State Geological Survey Contributions to the National Geothermal Data System

DOE Award Number: DE-EE0002850

Final Technical Report

EXECUTIVE OFFICE OF THE STATE OF ARIZONA (ARIZONA GEOLOGICAL SURVEY)

3/13/2015

This material is based upon work supported by the Department of Energy Geothermal Technologies Office under Award Number DE-EE0002850.

Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Contents

DOE Award Number: DE-EE0002850 1
Appendix Table of Contents 4
Appendix A: Overview of State Contributions to NGDS 4
Appendix B: Advisory Boards 4
Appendix C: USGIN Schema Development 4
Appendix D: NGDS Catalog System 5
Appendix E: Establishment and Maintenance of Nodes and Hubs 5
Appendix F: Subrecipient Training, Outreach, and Support 6
Appendix G: Outreach Materials 6
Appendix H: Sustainability Plan 6
Section 1. Grant Identification 6
Section 2: Notice of Distribution Limitations 8
Section 3: Executive Summary 8
Section 4: Goals and Objectives 11
Phase 1: Preparation and Background 11
Task 1. RDC and QA vision 11
Task 2. Data retrieval, collection, and development 12
Task 3. Quality and integrity analysis of the data format 15
Task 4. Establish regional technical resource centers 17
Phase 2: Transfer and Validation of Information to Data System 17
Task 5.0 Data integration into NGDS 17
Task 6.0 “Uploading” data to the NGDS 18
Task 7.0 Establishing links and/or virtual portals to the retrieved/collected/developed data 19
Task 8.0 System testing in conjunction with NGDS 19
Task 9.0 Assurance of the integrity of the original data set 20
Phase 3: Execution of Data Maintenance Sustainability Plan 20
Task 10.0 Formulation of the data maintenance and sustainability strategy 20
Task 11.0 Adding new technical or institutional data 21
Task 12.0 Publicizing addition of the new data 22
Task 13.0 Project Management and Reporting 22
Section 5: Summary of Project Objectives 25

Hypotheses 25
Approach 26
Preparation 26
Data Acquisition 30
Figure 5-1. Data Development Cycle 31
Production System 31
Technical Details 34
Problems and Issues 38
Quality control and validation 38
Well headers and well logs 39
Accessing across distributed services 39
Regional resource centers 39
Section 6: Products* 39
Websites and media 39
Websites and Internet Locations that Reflect the Results of this Project 39
Webinars, Videos, and Products to Promote and Explain NGDS Data 41
Webinars documenting how to use, submit and host NGDS data 41
Tools for Data Discovery and Publication 42
Excel to NGDS Services Tool 42
USGIN models API 42
Data Archiver 43
Workflow Management Tools 43
State Geothermal Data submission site 43
USGIN Task Management site 43
State Geothermal Data Tracking site 44
AASG Geothermal Data Repository database 44
Model Manager 45
Data replication process documentation 45
Section 7: Computer Modeling 45
Section 8: References 45

Appendix Table of Contents

Note: Each Appendix is paginated separately from the main report body. If viewing the report as a PDF document, note that 'bookmarking' is enabled.

Appendix A: Overview of State Contributions to NGDS

1. Inventory of participants and summary of contributions
    a. Data submissions by state (Excel spreadsheet): submitted separately as Excel file through E-link 241.3 (report number- DOE-AZGS-2850-5). Also available online at https://github.com/ngds/documents/blob/master/AASG-State-Contributions/A1-a-State-submissions_by_date.xlsx?raw=true
    b. Data contributions by services (Excel spreadsheet): submitted separately as Excel file through E-link 241.3 (report number- DOE-AZGS-2850-6). Also available online at https://github.com/ngds/documents/blob/master/AASG-State-Contributions/A1-a-State-submissions_by_date.xlsx?raw=true
    c. Data Contribution Summary of Work Accomplished by Subcontractors
2. Report on Sub-contractor Collaborations
3. Work performed against accepted Statements of Work (State Submissions Repository Site: https://github.com/usgin/aasgtrack) - submitted to OSTI as software

Appendix B: Advisory Boards

1. Science Advisory Board Framework
2. Technical Advisory Board Framework
3. Management Advisor Board Framework

Appendix C: USGIN Schema Development

1. Content Models List
2. Development Tutorials
    a. Content Model Guidelines
    b. Define New Information Exchange
    c. Content Model Workbook
3. Tools-submitted to OSTI as software
    a. Log ASCII standard Format Tools: Available to public at http://schemas.usgin.org/tools/ under Well Log Tools and https://github.com/usgin-models/BoreholeLASLogData - Submitted to OSTI
4. Websites – submitted to OSTI as software
    a. Content model development site: https://github.com/usgin-models- Submitted to OSTI
    b. Metadata and documents capture and management: http://repository.stategeothermaldata.org/repository/- Submitted to OSTI
    c. Content model management site: http://schemas.usgin.org/admin/- Submitted to OSTI

Appendix D: NGDS Catalog System

1. Data Validation Tutorials and Tools
    a. QA/QC Documentation
        i. GeoSciMLPortrayalCookbook
        ii. Data Development Cycle (graphic)
    b. QA/QC tools
        i. USGIN Models API: https://github.com/usgin/usginmodels- Submitted to OSTI
        ii. Content model Spreadsheet Validator: http://schemas.usgin.org/validate/cm- Submitted to OSTI
        iii. Web Feature Service validator at http://schemas.usgin.org/validate/wfs- Submitted to OSTI
        iv. Excel to NGDS Services Tool: https://github.com/usgin/ExcelToNGDSServiceTool- Submitted to OSTI
2. Conformance Tutorials, Documentation, and Tools
    a. Tutorials & Documentation
        i. USGIN metadata profile documentation (USGIN_ISO_Metadata.pdf)
        ii. XML Tutorial
        iii. XML Schema Profile for Simple Feature WFS Services
        iv. Import of Metadata to NGDS catalog (CatalogDocumentation.pdf)
        v. USGIN Identifier (URI) Conventions Tutorial
        vi. USGIN Lab Website Development Discussions
        vii. Metadata repository: https://github.com/usgin/metadata-repository- Submitted to OSTI
        viii. metadata compilation: (http://schemas.usgin.org/models/#Metadata)- Submitted to OSTI
    b. Tools
        i. Django, Node.js, and CouchDb metadata standardization and staging system: https://github.com/usgin/metadata-server- Submitted to OSTI
        ii. Data Archiver: https://github.com/ngds/data-archiver- Submitted to OSTI
        iii. CSV to XML python tool: https://github.com/usgin/usgincsv- Submitted to OSTI

Appendix E: Establishment and Maintenance of Nodes and Hubs

1. Hub/Node Selection
    a. Node Criteria and Participants (spreadsheet)
    b. Node Specification Criteria
    c. Node Map Graphic
    d. Hub Deployment Timeline (Excel spreadsheet)- submitted separately as Excel file through E-link 241.3 (report number- DOE-AZGS-2850-7). Also available online at https://github.com/ngds/documents/blob/master/AASG-State-Contributions/E-1-d-HubDeploymentTimeline.xlsx?raw=true
2. Establishment Instructions
    a. NGDS Installation Instructions
    b. Reviewing, Deploying, Validating Instructions

c. NGDS Software Requirements Summary
d. Geoserver Installation Instructions

3. Node Maintenance
    a. Data Replication Process Documentation: Hyper-V 2012 Replication on Workgroup Servers
4. Hub Testing
    a. Python Script: http://resources.arcgis.com/en/help/main/10.1/index.html#/Example_Check_a_folder_for_stopped_services/0154000005tr000000/ - Submitted to OSTI
5. Data Quality Reporting
    a. Issue Tracker: https://github.com/ngds/metadata-issues- Submitted to OSTI

Appendix F: Subrecipient Training, Outreach, and Support

1. Workflow Tools at State Geothermal Data site: http://www.stategeothermaldata.org/- Submitted to OSTI
2. Tools and Tutorials
    a. Repository Tutorial
    b. Workflow Tools Description
    c. Deduplication Workflow
3. Subrecipient regional workshop materials
    a. List of training webinars and PowerPoint presentations

Appendix G: Outreach Materials

1. List and Sample of Publications, Presentations, Media Coverage, and Social Media Sites with selected posters and abstracts
2. NGDS Atlas
3. Access Tools and Tutorials
    a. Scenario card tutorial videos
    b. NGDS Data Explorer: http://data.geothermaldatasystem.org/ - Submitted to OSTI

Appendix H: Sustainability Plan

1. Sustainability plan draft (previously submitted to DOE)
2. USGIN Foundation development
    a. USGIN articles of Incorporation
    b. USGIN bylaws
3. USGIN website: http://www.usgin.org/

Section 1. Grant Identification

DOE Award Number: DE-EE0002850

Name of Recipient: Arizona Geological Survey, on behalf of the Association of American State Geologists

Project Title: State Geological Survey Contributions to the National Geothermal Data System

Project Period: 1/8/2010 to 12/31/2014

Name of Project Director/PI: M. Lee Allison

CO-PI: Stephen M. Richard

Consortium Teaming Members:

Geological Survey of Alabama
Alaska Division of Geological & Geophysical Surveys
Arizona Geological Survey
Arkansas Geological Survey
Colorado Geological Survey
Delaware Geological Survey
Florida Geological Survey
University of Hawaii
Idaho Geological Survey
The Board of Trustees of the University of Illinois, Illinois Geological Survey
Trustees of Indiana University, Indiana Geological Survey
Iowa Department of Natural Resources – Geological & Water Resources Survey
University of Kansas Center for Research, Inc., Kansas Geological Survey
University of Kentucky Research Foundation, Kentucky Geological Survey
Board of Supervisors of Louisiana State University and A&M College, Louisiana Geological Survey
Maine Geological Survey
Massachusetts Geological Survey
Michigan
Mississippi Department of Environmental Quality
Missouri Geological Survey, Department of Natural Resources, Division of Geology and Land Survey
Montana Bureau of Mines & Geology
Board of Regents, NSHE, obo University of Nevada, Reno, Nevada Bureau of Mines and Geology
New Hampshire Geological Survey
New Jersey Geological Survey
New Mexico Institute of Mining & Technology
New York State Geological Survey
University of North Dakota
Ohio Department of Natural Resources, Division of Geological Survey
The Board of Regents of the University of Oklahoma, Oklahoma Geological Survey
Oregon Department of Geology and Mineral Industries
Pennsylvania Geological Survey
Rhode Island Geological Survey
South Carolina Geological Survey
Sinte Gleska University, South Dakota
Tennessee Division of Geology
Texas Bureau of Economic Geology
Utah Geological Survey
Vermont Office of Department of Environmental Conservation

Virginia Division of Geology and Mineral Resources
Washington State Department of Natural Resources
West Virginia Geological and Economic Survey
Wisconsin Geological and Natural History Survey
Wyoming State Geological Survey

In-Kind Partners:

Connecticut Geological Survey (through Massachusetts Geological Survey)
Delaware, Georgia, and Maryland Geological Surveys (through the Virginia Division of Geology and Mineral Resources)
Nebraska Geological Survey (through the University of North Dakota)
North Carolina Geological Survey (and through the Virginia Division of Geology and Mineral Resources)
North Dakota Geological Survey
South Dakota Geological Survey

Cost Share Partners:

Alabama Geological Survey
Illinois Geological Survey
Massachusetts Geological Survey

Section 2: Notice of Distribution Limitations

No patents were produced; however the data content models produced are under the intellectual property of the Arizona Geological Survey (AZGS) and licensed under a Creative Commons Attribution 3.0 Unported License with copyright of Arizona Geological Survey. As part of the sub-recipient agreement with the U.S. Department of Energy (DOE), AZGS provides DOE with an irrevocable, royalty-free, non-transferable, non-exclusive right and license to reproduce the content models.

Section 3: Executive Summary

The State Geological Survey Contributions to the National Geothermal Data System project is built on the work of the project managed by Boise State University to design and build the National Geothermal Data System, by deploying it nationwide and populating it with data principally from State Geothermal Surveys through collaboration with the Association of American State Geologists (AASG). This project subsequently incorporated the results of the design-build and other DOE-funded projects in support of the NGDS. The NGDS (*www.geothermaldata.org*) provides free open access to millions of data records, images, maps, and reports, sharing relevant geoscience, production, and land use data in 30+ categories to propel geothermal development and production in the U.S. NGDS currently serves information gathered from hundreds of the U.S. Department of Energy sponsored development and research projects and geologic data feeds from 60+ data providers throughout all 50 states. These data are relevant to geothermal energy exploration and development, but also have broad applicability in other areas including natural resources (e.g., energy, minerals, water), natural hazards, and land use and management.

The NGDS was formally launched by U.S. Secretary of Energy Ernest Moniz on May 28, 2014 at the White House Energy Datapalooza in Washington, DC. OneGeology, an international initiative of geological surveys, designated the NGDS project with a Four Star Accreditation rating in May, 2014.

Geothermal Energy Association awarded NGDS the Special Recognition Award for its outstanding achievement in the geothermal industry (presented at the National Geothermal Summit in Reno, Nevada, August 5, 2014).More details are available in the final technical report on the Design and Build portion of the National Geothermal Data System, which can be accessed at the following Web location: *http://www.osti.gov/scitech/biblio/1163354*.

The State Geological Survey Contributions project met or exceeded all deliverables and requirements, was on time, and within budget. The project received good to excellent scores in all categories for each of the annual Peer Reviews of DOE-funded geothermal awards, and was regularly among the highest rated among all 200+ projects.

As of January 2015, users can access approximately 10 million data records in the NGDS central catalog, including:

- 1.85 million well headers (all wells)
- 520,000 water wells (supply, monitoring, irrigation)
- 666,800 borehole geophysical logs (oil and gas, water, other wells)
- 2 million borehole lithology data records (intervals and intercepts)
- 815,600 digital geologic map features (contacts, units, active faults, volcanic vents, geothermal areas)
- 538,800 borehole temperatures (wells and springs)
- 661,500 well fluid production records
- 387,800 aqueous chemistry records (wells and springs)
- 73,300 drill stem/well test records
- 53,800 geochemistry records (physical sample, thermal conductivity, and geochemistry)
- 14,000 direct use and heat pump facilities

The records listed above are available as data that are interoperable ("Tier 3") and can be seamlessly mapped across all states. Also included are data available as scanned documents (Tier 1) and in structured databases or spreadsheets but not interoperable (Tier 2) See Section 5: Approach for detailed information on data tiers):

- 66,600 all other types of data (seismic, gravity, heat flow, hydraulics properties, geothermometry, power plants production)
- 59,300 documents, reports, and papers, including 6,300 geological maps
- 3 million data points compiled by Southern Methodist University

Detailed state contribution data totals can be found in Appendix A.

Four NGDS Regional Server Hubs (IL, KY, AZ, and NV) have been tasked with hosting web services for their surrounding data providers (other state surveys or institutions) who lacked institutional permission or the technical capability to host their own data in web service format. Experience has dictated that server outages can and will occur; the potential loss of data from a distributed network, while not as great as "closing the doors" on a traditional database, may occur. In order to preserve the vast amount of data digitization and preservation, the project conducted a one-time data harvest to capture the data available within the NGDS and place it as a back-up and disaster recovery mechanism on both cloud storage and

external hard drive. The project instituted a process and capability to make a copy of a virtual machine at a different geographic location for disaster recovery and data replication purposes.

In addition to collecting and formatting data for the NGDS central catalog, the State Contributions Project provided training and technical support to enable state geological surveys to self-host their own data. A major part of that task included assisting states in adopting and deploying the CKAN-based software stack to enable participating agencies to serve their own data as an NGDS node. The software stack, formerly known as "Node-in-a-Box" (NIAB), was developed under the Architecture, Design, and Maintenance award to Boise State University. The NIAB has been revamped and re-designated "GIN-stack" as part of the current project. A no-cost extension through December 31, 2014 was granted to allow at least six network data providers to deploy GIN-stack or another independent free-and-open-source server.

The NGDS is built using the Geoscience Information Network (USGIN) data integration framework, a data management framework that ensure that the system is compliant with international standards and protocols, scalable, extensible, and deployable throughout the world. USGIN is a federated information-sharing framework that uses free and open-source technology, capable of interacting with similarly configured data-sharing networks. The system accommodates resources in various forms, from unstructured text and images to documented, community Web services and interchange formats.

USGIN was developed over the past seven years through a partnership of AASG and the U.S. Geological Survey. Per the requirements of the NGDS Design-Build Project, the software produced in the project is under an open source license variant allowing users to copy, distribute, and transmit the software; adapt software for other applications; and make commercial use of the software under the condition that the attribution of the software is included in any copy or derived work. The distributed nature of the system means that stewardship of resources is determined by the resource owner. Participation in the network requires that a resource provider create metadata that conforms to the profile, and make the metadata and the described resource available. The network is open; anyone can deploy new nodes and components that implement one or more USGIN specifications, without requiring approval. New specifications can be introduced for service protocols, interchange formats, or vocabularies. Similarly, the system back end is decoupled from the front end, meaning anyone conforming to system configuration can build a portal into the system, consuming the data feeds and adding their own applications, visualizations, or other value-added functions. The International Renewable Energy Agency is already taking advantage of this, preparing to consume key data sets live for inclusion in the Global Atlas of Renewable Energy, accessible via the IRENA website, *www.irena.org*. Keeping resources under the stewardship of the parties responsible for the information promotes system sustainability because the stewards have a direct connection with the quality of the product rather than submitting the data to a centralized database and related data manager.

The use of USGIN standards and protocols to shape the contribution format, metadata, and exchange specifications for the State Contributions Project aligns with the goals put forth in 2007 at the inception of the USGIN project and with the goals of NGDS, to create a free and open data exchange network that promotes universal access to geoscience data.

Wherever possible, the system employs and encourages the use by data providers of free-and-open-source software to ensure continued open access and to encourage development of apps and programs using

NGDS data. NGDS contributors expose their data to a federated system where data owners maintain control of their data. This approach offers a long-term solution to data storage and sharing without any proprietary software. Developer specifications for USGIN are available via the USGIN GitHub site at *https://github.com/usgin*.

The NGDS vision is that the functions, accessibility, and availability of interoperable data will encourage continued participation within the NGDS from all system users – data consumers, data providers, and application developers. As each data node adds to its data repositories, the system-wide NGDS functions become increasingly valuable. Each data provider will have created a value-added service that is transportable and scalable to cover all data in its possession. An example of such is the recent decision by the State of Arizona to deploy a network among state agencies, built on the USGIN-NGDS underlying framework for natural resource, environmental, and transportation data. Thus, there are benefits to each participant to continue to add data to the system and maintain it. The long term goal is that the data network reaches a 'tipping point' at which it becomes a data equivalent to the World Wide Web – where everyone will maintain the function because it is expected by their clientele and it fills critical needs. Applying this vision to the NGDS opens the door for additional data providers to geothermal development, thus increasing the value of NGDS and its underlying data integration platform, USGIN.

The NGDS and the core principles of USGIN are fully compliant with the 2013 Presidential Executive Order 13642 (Making Open and Machine Readable the New Default for Government Information) and Office of Management and Budget (OMB) Memorandum M-13-13 regarding open data access. To address project sustainability, the team determined that a stand-alone organization offered the greatest independence and likely adherence to the organizations vision, goals, and missions and thus the longevity of a system such as the NGDS. In the fall of 2013, the Arizona Geological Survey successfully incorporated the USGIN Foundation, Inc. as a non-profit entity (IRS 501(c)3 tax status). This step in alignment with the recommendations of the Hutchison and Richard 2011 "Recommendations for the Future of the U.S. Geoscience Information Network," joint report to the Association of American State Geologists and U.S. Geological Survey, and has been incorporated into the State Contributions project as the development and execution of a sustainability plan.

Section 4: Goals and Objectives

This section reviews the project tasks from the Statement of Work (SOW), with the Purpose and Approach statements copied from that original plan. For each task or sub-task, the actual accomplishments are summarized. Details on technical process and workflow are described in Section 5, and complete product listings and links are found in Appendices A-G.

Phase 1: Preparation and Background

Task 1. RDC and QA vision

Purpose: Specify how this project will advance and dovetail with ongoing NGDS design, testing, and management.

Approach: Meet with NGDS Boise State project team to plan coordination between the data acquisition activities of this project and NGDS system development.

Actual accomplishments: The AZGS team was a subcontractor to the Boise State NGDS Design-Build project, with the task of developing system architecture. As a result, there is very close alignment of

the NGDS design developed by the Design-Build project and the AASG geothermal data acquisition processes. The AASG project work was the de facto system testing for the concurrent design-build activities, and the content compiled and acquired in the course of this project is a major part of the content available through the deployed NGDS.

Task 2. Data retrieval, collection, and development

Purpose: Consistently identify, collect, document, and add data from participants to the NGDS.

Approach: The project Science Advisory Board (SAB, described in further detail in Subtask 13.2) will serve as a peer review panel to review and approve the statement of work itemizing the types, amounts, and priorities of data deliverable on an annual basis for each project participant. The state subcontractors will provide the data based on the approved statement of work. At the close of each year, the state's deliverables will be reviewed for completion by the SAB prior to issuing the following year's contract. (Data submission information and processes are documented in Appendix A. Framework for Science Advisory Board, Technical Advisory Board, and Management Advisory Board is available in Appendix B.)

Actual accomplishments: Listed with Subtasks.

Subtask 2.1 Determining amounts and types of data that will be made available

Purpose: To identify the types of data most relevant to the geothermal resources in each state, prioritize data to be added to NGDS, and establish annual deliverables and milestones for subcontractors.

Approach: Each state will identify the types of data most in demand or most critical to identifying and characterizing its geothermal resources by cataloguing its data resources, and submitting a SOW to the SAB for review and approval at regular intervals. This SOW will explicitly identify deliverables and milestones for the subsequent work period.

Actual accomplishments: As outlined in section 5, data acquisition was conducted in an approximately annual cycle that began with the data provider assessing their holdings and the requirements of the geothermal community in their state to develop a Statement of Work (SOW) for the acquisition cycle. This SOW explicitly enumerates the data items to be compiled and the quantity of each that the provider expects to make available during that acquisition cycle. Each SOW was reviewed and revised according to recommendations by the SAB before approval to proceed with data acquisition.

Subtask 2.2 Collection, digitization, and indexing of "legacy" data

Purpose: Digitize existing priority data to add to state-based databases.

Approach: Each state will digitize data, beginning with their highest priorities, populating databases or tabular datasets and creating metadata for each product as part of the workflow. AZGS will coordinate data structures to promote compatibility between data from different states. (Data structure guidelines, models, and tools are listed in Appendix C.)

Actual accomplishments: SOWs developed under Subtask 2.1 were reviewed by the SAB to assess priorities and focus work on the most useful kinds of information. Metadata for approximately 5861 documents were created and added to the NGDS catalog. This total includes documents digitized by scanning paper copies on file at the surveys, as well as documents and files already available in digital form.

To assist state sub-recipients in digitizing and indexing legacy data, workshops and webinars were held and documentation developed and placed online. As the NGDS processing team at AZGS worked with project partners to develop USGIN information exchanges (*http://schemas.usgin.org/models*) used to make data compiled from different states interoperable. A metadata worksheet template was developed for the compilation of standardized metadata

describing available information items for inclusion in the NGDS catalog. The information exchanges were updated as new requirements emerged. The state recipients' SOWs were closely monitored by the NGDS data processing and management team to assess the progress of states' deliverables, prioritizing the most valuable given the time and funding constraints.

A more formal process for information exchange development was developed (described further in Section 5: Content Models). It was obvious from the initial information resource inventories that one of the most abundant entities of interest would be boreholes, so the first information exchange developed was for a 'well header'-- a description of the key properties of a well bore associated with the location at which the hole entered the Earth (the collar or header). The Well Log Observation information exchange was created to index electronic geophysical logs from boreholes, either as scanned paper logs or digitized log trace data. The team developed two software tools for working with digital versions of temperature-depth well logs in LAS ASCII text format, including an LAS converter, an automated well log conversion tool, and a tool to import digitized well log data into the Excel workbook content model (*https://github.com/usgin-models/BoreholeLASLogData*). The code for the programs (in Python) is available online at *https://github.com/usgin/WellLogDataTools*. The software implements two different conversions for well log data: 1) the BoreholeLASLogData content model conversion to an LAS 2.0 file, and 2) an LAS 2.0 file record data transfer into the WellLog content model feature fields. This process enables digitized well log data to be accessible online as discoverable records for interoperable web services.

Subtask 2.3 Documenting and adding digital databases to the NGDS

Purpose: Add existing digital data to the NGDS with appropriate metadata.

Approach: Expose existing digital databases to the NGDS or convert obsolete digital databases to data schema and formats suitable for incorporation in to NGDS, based on priorities established by the Science Advisory Board.

Actual accomplishments: 627 Tier 3 (interoperable and digital) datasets using NGDS content models are registered in the catalog. Statistics for Tier 2 (digital structured) datasets are lacking because in many cases, Tier 2 type data stored in workbooks or database files (e.g. dbf, mdb) are compressed into Zip archives before being placed online, so they cannot be simply quantified from the catalog.

Subtask 2.4 Collection of new data

Purpose: Make new measurements calculations, and interpretations, and collect samples.

Approach: Each state will assess the extent, usefulness, and nature of data to determine what gaps are most critical to fill.

Actual accomplishments: The original project scope was to work with existing data to make it more accessible online, but supplemental funding was obtained early in the project to support acquisition of new data. Proposals were solicited from the states and reviewed by the SAB. The following projects produced new observation data related to geothermal resources:

Arizona collected new data as water chemistry from springs and wells, and detailed new mapping of geology and structure associated with three known thermal springs areas.

Colorado collected new chemistry data/analyses and obtained new thermal conductivity measurements on existing core.

Idaho drilled four (4) new geothermal exploration wells, obtained lithology logs, thermal conductivity measurements, and heat flow data.

Indiana excavated pits and installed monitor systems that continuously collect shallow near-surface temperatures and thermal properties.

Massachusetts collected surface rock samples, and measured thermal conductivity and U, Th, K concentrations to calculate radiogenic heat production.

Maine collected new temperature-depth logs from existing deep wells.

New Jersey collected new temperature-depth logs from existing deep water wells.

New Mexico collected new temperature-depth logs from existing wells, new aqueous chemistry samples for analysis, and rock samples for geochronology analysis.

Nevada drilled two (2) new geothermal wells and obtained logs, thermal conductivity measurements and aqueous geochemistry data.

Oklahoma measured thermal conductivity and radiogenic heat production data from existing drill core samples and newly collected surface samples.

Oregon drilled three (3) new geothermal exploration wells and obtained logs, aqueous geochemistry and thermal conductivity data.

Pennsylvania collected well logs in new deep wells drilled by industry in 2014.

Utah drilled twelve (12) new geothermal exploration wells and obtained well logs, aqueous geochemistry and thermal conductivity data.

Vermont collected new temperature-depth logs and other geophysical logs from a set of existing deep wells without temperature logs. In addition, new rock samples were collected for thermal conductivity measurements.

Washington collected new aqueous chemistry samples for analyses, completed new temperature-depth logs in existing holes, conducted new geologic mapping related to geothermal activity.

Wisconsin drilled new wells as part of the primary project (not Supplemental) and obtained well logs and thermal conductivity data.

West Virginia collected new thermal conductivity measurements from existing core and cuttings samples.

Subtask 2.5 Creation of metadata catalogs

Purpose: Provide for easy user discovery of and access to data in the NGDS.

Approach: Set standards and requirements for metadata content and procedures for entry and preservation of metadata for all information resources exposed to the NGDS. Metadata will be made available through the standard Open Geospatial Consortium Catalog Services for the Web (CSW) interface. AZGS will work with each data provider to determine the most efficient procedure to create metadata and enter it into the NGDS catalog. (NGDS Catalog development is documented in Appendix D.)

Actual accomplishments: The project adopted the USGIN recommendations for metadata content, and the USGIN profile for ISO19115/19139 metadata. The catalog system uses the Open Geospatial Consortium Catalog Service for the Web for software access to the catalog. Metadata for resources produced by the states has been aggregated in catalog deployed as an Esri Geoportal instance at *http://catalog.stategeothermaldata.org/*. Metadata are harvested from this geoportal into the NGDS central aggregating catalog that is accessible for human users through the portal at *http://geothermaldata.org*.

Creation and maintenance of metadata emerged as one of the key project activities. Two major workflows were developed for obtaining metadata from the states. The first was to create individual metadata records using a form interface. Most of the metadata were compiled by data providers in the Excel Worksheet template (*http://schemas.usgin.org/models/#Metadata*), and converted to XML for harvest into the catalog (see workflow description in Section 5: Data Acquisition)

Through the course of the project, several iterations of web-based metadata forms were used, including the Metadata Wizard (Drupal, now disabled), a Drupal-based repository, and currently a Django-CouchDb repository application developed by project developer Ryan Clark and available at *http://repository.stategeothermaldata.org*. The NGDS CKAN-based GIN-stack user interface was updated to allow editing of the full USGIN ISO metadata content.

The project utilizes and has extended the portal web application developed by the Design-Build project, which is deployed at *http://geothermaldata.org*. This catalog component of the portal application uses a slightly modified version of the NGDS GIN-stack to implement metadata harvesting, form-based update, text and facet-based search and map-based search. The Esri Geoportal (*http://catalog.stategeothermaldata.org*) free-and-open-source application is used to aggregate metadata includes a transformation script to ensure that harvested records conform to USGIN/NGDS requirements. The catalog in the portal harvests metadata for resources registered in the State Geothermal project, as well as all other NGDS projects, to provide a single point of entry for finding information in the NGDS.

Task 3. Quality and integrity analysis of the data format

Purpose: Ensure the quality and integrity of the data provided to NGDS meets the system standards.

Approach: The project will follow the guidelines of the EPA "Guidance for Quality Assurance Plans." The project will also use guidelines in place and under development for the USGS National Geological and Geophysical Data Preservation Program (NGGDPP) National Digital Catalog that are being used by all the State Geological Surveys (SGSs) currently, and guidelines developed for the US Geosciences Information Network (USGIN). In addition, AZGS will employ standard measures of quality using statistical tools expected to be provided by the NGDS to assess variability, bias, systematic error, imprecision, accuracy, precision, reproducibility, etc.

Actual accomplishments: The project disseminated guidelines for data content, format, and quality to participants at the start of the project and requested that participants address them in the SOW for each Data Acquisition Cycle. (More details are available in Section 5: Content Models.)

Tools and automation. All Tier 3 (digital and interoperable) data sets were validated against the XML schema for its content mode, and inspected for field-specific out of range values (dates, spatial data, geometry, attributes, source), well-formed HTTP URIs, location accuracy, HTTP URL validity, and contact metadata. Metadata were reviewed under the following criteria:

- Contact and source data completeness and accuracy
- Consistent URI formats (human recognizable URI data chain if possible)
- Field filters show appropriate range and type of data
- Spatial data points within boundaries of project area
- Attributes listed within appropriate fields and using acceptable data valid terms

Metadata requiring revisions were returned to the provider with comments for guidance. Resubmitted metadata versions are re-inspected and validated. The completed and approved metadata in the Excel worksheet compilation tables were subject to additional automated validation to assure the metadata conforms to the USGIN profile before importing converting to XML for import into the NGDS catalog.

In order to ensure that content model data are schema valid, a Python API called usginmodels (*https://github.com/usgin/usginmodels*) was created. This library exposes functions that return individual models, URIs, layers and service names. It validates data in CSV format for conformance with the schema for a given content model. The content model schema is pulled from *http://schemas.usgin.org/contentmodels* in JSON format. CSV data are inspected for proper field names, field types, well-formed URIs, the uniqueness of primary URIs, unit and spatial reference system consistency, valid dates and range values for fields such as Latitude and Longitude. In the

cases where auto-correction is possible, the library applies the correction, which makes the data schema valid and returns the new data to the requesting agent. It also returns a list of warnings and errors for data which needs to be manually fixed.

This API is used within ArcGIS by the *Excel to NGDS Services Tool*, in the web-based interface at *http://schemas.usgin.org/validate/cm* and in the CKAN NGDS GIN-stack environment at *www.geothermaldata.org* for Tier 3 data upload. The *Excel to NGDS Services* software tool is used by hubs and states to process their own data, especially those who do not operate a GIN-stack software node and manually create web services from data in the Excel distribution of a given information exchange. The tool documentation is available *at https://docs.google.com/document/d/ - 1H08dObu5pWze7g3sSap3uDcI4DwYzuIEMfcP0V7Xj9s*.

The *Excel To NGDS Services Tool* takes as input an Excel or CSV file, validates the data, correcting the data if possible or returning errors, then creates an ArcGIS geodatabase which is ready to be deployed as a service. The web-based validation interface at http://schemas.usgin.org/validate/cm is part of a Django site for content models information. Given a CSV uploaded by the user, either a corrected data file or a list of errors is returned.

The NGDS data processing team created a Google Document for ongoing service validation and conformance monitoring, available at *https://docs.google.com/spreadsheet/ccc?key=0ArweWT - YyllxYdFppX1B2RVVWTVFfSlJjTWpsSkk2X0E*. This document tracks conformance to the following:

- service and layer names
- version of the information exchanges (and hence the appropriate namespace) implemented by the service
- appropriate and adequate metadata in the web service capabilities files and service rest pages
- specified projection (Spatial Reference System)
- rewrite rules, which redirects URI URLs
- schema validation

This conformance monitoring was kept for every state sub-recipient web service published and available through NGDS. Population of the online Google document worksheet content was done manually by AZGS data reviewers.

Workshops. NGDS workshops held in 2012 and 2013 at the Nevada survey, Illinois survey, and Virginia survey helped participants to learn how to validate data and integrate them into NGDS. These events were important for training state sub-recipients to process their own data, reducing much of the review necessary by the AZGS data processing team. In each workshop, the processing team was able to give states tools to transform their existing data and metadata into the NGDS specifications and create databases structured for web service publication, as well as workflows to manage GIS-enabled database management systems and publish the data as OGC-compliant, NGDS schema-conformant web services. Materials from the workshop such as PowerPoint slide sets and example data were made available at *https://github.com/usgin/Workshops*. These workshops are described in more detail in Section 5: Training.

Task 4. Establish regional technical resource centers

Purpose: Provide training, guidance, and assistance to sub-recipients in developing and implementing data quality and integrity guidelines, developing metadata, implementing data services, and configuring their servers to link seamlessly with the NGDS.

Approach: The project will establish multiple regional technical resource centers with a programmer/developer in each one, to serve that region. These centers will utilize existing facilities at the host agency. (Additional documentation on hubs and nodes is available in Appendix E.)

Actual accomplishments: Hubs are operational in Nevada, Arizona, Illinois and Kentucky. It was not in the scope of the project funding to employ web developers at the regional hubs. These hubs are staffed by state survey employees who are geologists, but also understand data and metadata management, server and database maintenance, and web services for maps and data. The hubs are a technical resource for the regions, publishing and hosting web services and any necessary files on Web Accessible Folders (WAFs) for online availability, transformed Tier 2 and Tier 3 maps and data into web services at the request of the states, and created various Python scripts to make the hub run smoothly and efficiently. For details about hub sustainability, see Task 10.1 below.

Workshops (described in Task 3 above and in Section 5: Training) were held at the Nevada survey, Illinois survey, and Virginia survey to engage state personnel and promote the regional centers.

Phase 2: Transfer and Validation of Information to Data System

Task 5.0 Data integration into NGDS

Purpose: Enable catalogs and databases to be exposed to the network for discovery, access, and retrieval

Approach: Develop, document, implement, and deploy web services to enable open access to information resources of the NGDS.

Actual accomplishments: Listed with subtasks, below.

Subtask 5.1 Registration of resources

Purpose: Enable data providers to register their resources by publishing metadata in the NGDS catalog. Verify that metadata meets requirements and that registered resources exist and can be accessed using information in metadata.

Approach: Acceptance and publication of metadata in the NGDS catalog will be the formal mechanism for making a resource part of the NGDS. The publication process requires a validation process to determine that metadata and the described resource are conformant with system requirements. This is a technical validation, not a scientific review. Project technical staff will work with data and application providers to efficiently produce metadata meeting system requirements and publish it to be harvested by the NGDS system catalog, and to develop automated processes for data and metadata validation. Policies for this validation process developed by Project technical staff and the NGDS core team, reviewed for approval by the project Technical Advisory Board (TAB), and once approved, will be made publically available.

Actual accomplishments: Three basic approaches to metadata creation have been implemented and utilized to populate the NGDS catalog. During registration, Globally Unique Identifiers (GUIDs) were created for each record. After a record is approved, metadata are saved as a comma-delimited text (CSV file and imported into the AASG Geothermal Data repository. Details are delineated in Chapter 5: Metadata Acquisition and Subtask 2.5: Creation of Metadata Catalogs.

Subtask 5.2 Implement data exchange software 'wrappers' to provide for interoperability of databases

Purpose: Provide tools and guidance to map data into interchange formats used by web services to make data available and interoperable.

Approach: Project technical staff will prepare training manuals ('cookbooks') and offer workshops on how to encode information for transmission and integration into NGDS. The technical staff will consult by phone and online with data providers. The AZGS project staff will implement the interchange mapping for providers lacking the technical staff to do it within their respective agencies. Schema and vocabulary mapping necessary to produce the interchange formats may be done using queries in a database, through XML transformations, or using custom software. Standardization of the interchange formats will reduce the amount of development required by enabling reuse of components and workflows.

Actual accomplishments: The project developed the *Excel to NGDS Services tool*, (described in Task 3) to assist data providers in validating data against an NGDS content model, and automate processes to load the data into an ArcGIS geodatabase that is ready for NGDS web service deployment on an ArcGIS Server. The geodatabase format can also easily be exported to other formats for use in open-source GIS web servers like GeoServer.

Project workshops held in 2012 and 2013 at the Nevada survey, Illinois survey, and Virginia survey (described in Task 3 and in Section 5: Training) explained the requirements, processes, and use of the software and tools.

NGDS is using the OneGeology Portrayal schema for geologic map data delivery. A webinar in August 2013 followed the 2012 OneGeology workshop held at USGS in Denver, to address the integration and registration of NGDS OneGeology web services with the OneGeology online portal. The NGDS championed integration of data services (WFS) with the map services (WMS) in the OneGeology architecture, and created documentation for use of GeoSciML portrayal and GeoServer application schemas implementation. (For detail see *http://repository.stategeothermaldata.org/resources/documents/GeoserverDocumentation.docx*). The webinar materials are available online for everything sub-recipients need to create metadata, transform data into content model standards, including GeoSciML/OneGeology, publish data services, and upload them to the NGDS.

The AZGS project team has also made itself available through telephone access to assist state data providers with any NGDS data processing issues.

Task 6.0 "Uploading" data to the NGDS

Purpose: Make each provider's data accessible to the NGDS through standard web service interfaces.

Approach: Providers will publish their data via web services, which they may implement themselves, or work through arrangements with other providers or the NGDS core. Services will follow documented NGDS service conventions.

Actual accomplishments: Tier 1 and Tier 2 data are being served from servers at nine participating agencies. Currently, 370 Tier 3, content model-based services are available online and registered in the catalog.

One of the goals of the NGDS is to provide an open-source software stack for releasing open data on the World Wide Web that is sustainable and a cost-effective option for data providers. The CKAN GIN-stack provides a tool allowing data providers to create metadata data for tabular datasets conforming to a registered NGDS content model and release it for file-based access as well as NGDS WMS and WFS services. The stack includes metadata management components providing an OGC Catalog Service for the Web (CSW) with USGIN-conformant ISO 19139 metadata. The NGDS CKAN GIN-stack (formerly known as Node in a Box, or NIAB) is the final, version 2 release of this compiled software stack.

In late summer 2014, unspent funds returned from the state's various data collection efforts were reallocated to eleven states to install GIN-stack or another catalog service and metadata management configuration to increase the number of states hosting their own services. Seven states

(Utah, North Dakota, Wisconsin, Delaware, Michigan, Indiana, and Alaska) opted to install Geoportal as a catalog server, coupled with GeoServer or ArcGIS Server for web services. Four states (Montana, Alabama, Hawaii, Washington) decided to install GIN-stack software. In both cases, data services will be transferred from the regional hubs back to these states to host, and the metadata for all resources submitted to NGDS hosted at the AASG Geothermal Data Repository will be extracted and transferred to local catalogs. In the future, the NGDS central aggregator will harvest metadata directly from these providers. Six additional states so (Arizona, New Hampshire, Kansas, Minnesota, Oregon, and West Virginia) were already serving their own web services using ArcGIS Server and continue to do so. (Details on node capabilities and software are found in Appendix E1a: NGDS Node Criteria Analysis).

Task 7.0 Establishing links and/or virtual portals to the retrieved/collected/developed data

Purpose: Make data and service products available through the NGDS Geothermal Desktop and other web portals.

Approach: Project technical staff will work with the NGDS Geothermal Desktop developers and external groups to create necessary web services, ontologies, and interface standardization to link the aggregated data resources of the project to NGDS web portal and enable other, independent portals to interface with the system. Staff will also work with other client application developers to implement connectors to NGDS catalog and data services. The objective is to enable widely used, off the shelf, low-cost or free software to access data provided by NGDS.

Actual accomplishments: The geothermal desktop component envisioned as part of the original Design-Build project was never completed. The AASG project has been working to link NGDS data with other online systems. A link to display NGDS WMS services in the NREL Geothermal Prospector web application has been implemented in the NGDS portal (this feature has not been implemented in the new version of the NGDS CKAN extensions). NGDS metadata is being harvested into *http://www.data.gov/* and the project team has been working to resolve problems with harvesting NGDS service records to the IRENA Global Atlas for Renewable Energy (*http://irena.masdar.ac.ae/*). Work with the NREL Geothermal Prospector team to improve connection between the NGDS catalog and the Prospector is ongoing.

Task 8.0 System testing in conjunction with NGDS

Purpose: Ensure that NGDS will be able to continue development and testing of key components of the system concurrently with the registration of data services with catalogs and implementation of data exchange standards.

Approach: Outside participants will work with the core development team at AZGS to test services and data formatting. NGDS core has stated intention to implement service status monitoring, and AZGS hopes to include conformance testing of data services against technical specifications. An operational network will require ongoing monitoring, and standardized performance and reliability testing. In collaboration with NGDS core developers, define standard tests of NGDS services that can be run to determine if service instances are functioning properly.

Actual accomplishments: The integration of the AASG data project with the last phase of the Design-Build project, working on the NGDS CKAN GIN-stack application has been very close. Metadata upload and data deployment to alpha and beta versions of the GIN-stack has been ongoing, testing the catalog harvest, indexing, and search capabilities, the USGIN URI redirection components, and the Tier 3 web service deployments.

The Arizona and Kentucky Hubs have implemented a Python script for monitoring services provided by Esri. This script monitors web services on a specified web server to ensure that they stay up and run without interruption. It is available at *http://resources.arcgis.com/en/help/main/10.1/index.html#/ - Example_Check_a_folder_for_stopped_services/0154000005tr000000/*.

Task 9.0 Assurance of the integrity of the original data set

Purpose: Establish criteria that can be used to filter data and categorize them according to established and user-defined quality levels.

Approach: Research will be necessary to identify automated processes that may be used to assess quality of data offered by a service. Automation can be used to assess conformance to usage of controlled vocabularies and schema cardinality requirements, the percentage of data fields that are null or have invalid data type values, and other similar metrics, as described in Task 3. **More subjective judgment of data quality can be made by the data provider and provided in metadata data quality elements, but these will not be mandatory in the near term because the information is commonly not available**. A system for user feed-back on data quality will be investigated. Implementation of such complex advanced features is not within the scope of the current project, but AZGS will work with NGDS core developers to identify approaches that can be implemented with resources available. Available quality information will be reported with each metadata search. When data with incomplete metadata is identified the entity responsible for the metadata will be notified.

Actual accomplishments: The only sure way to validate the accuracy and quality of incoming data is to conduct field work with spot checking to verify observation data and duplicate datasets to assess repeatability, but this was not part of the project work plan. Project data review was focused on syntax and inspection for typographic errors that produce wildly out of range values or obviously incorrect strings for dates or spelling in text.

The current workflow for reporting a metadata or data quality issues is through the NGDS metadata GitHub issues tracker at *https://github.com/ngds/metadata-issues*. This is a repository for tracking issues, alerting interested parties (e.g., metadata stewards, data providers), and documenting issues resolution. Interested parties can subscribe to the NGDS notifications RSS feed or follow the repository using a GitHub account. If an NGDS user at *www.geothermaldata.org* finds data or metadata that are incorrect or otherwise needs attention at a Data Details page for any resource, the user can use the "Report an issue," which creates an issue in the metadata tracker for the data provider or NGDS metadata administrators to address.

The *USGINmodels API* (see Task 3) can execute basic data quality tests for out of range values, but the tool is not designed to store or calculate metrics about the quality of the original data.

Phase 3: Execution of Data Maintenance Sustainability Plan

Task 10.0 Formulation of the data maintenance and sustainability strategy

Purpose: Ensure availability, reliability, and update of data and services in the NGDS during and beyond the project.

Approach: Educate system data providers and users through workshops at the regional technical resource centers. Data provided by state geological surveys can be hosted by the originating survey, at one of the regional centers, or by the NGDS core, and depending on policies developed by NGDS may be mirrored by other NGDS system servers. Regional centers will mirror and backup data for their region, and provide redundant service instances to prevent a single point of failure. Develop a business model to sustain system maintenance for data and services by engaging stakeholders who have financial resources available and an interest in continuing system operation.

Subtask 10.1 Implementation of regional aggregation services and back-up of data

Purpose: Ensure data backup, continuity of service, and the potential for a permanent repository.

Approach: Each regional center will host a data server that provides multiple functions: mirrored services for state servers in that region to ensure data backup, continuity of service, and the potential for a permanent repository.

Actual accomplishments: The project team has developed a scheme for replicating data between NGDS servers using the Microsoft Hyper-V 2012 Replication on Workgroup Servers. Hyper-V is a *Windows Server Role* that can be installed on Windows Server 2012 platforms to virtualize and replicate servers on x86 and x64 workgroup servers, backing up entire servers and their contents. If one server fails, the location of the replicated back-up can boot up with an exact replica of the server (with a server snapshot no greater than 15 minutes), which will use the same URLs so the user is unaware of any changes. This software is included with any Windows machine so no additional costs are incurred with this replication method.

Documentation on the replication process was provided to hubs in Virginia, Kentucky and Illinois and presented in the 2013 Western Regional Workshop for state sub-recipients, and is available online at *https://github.com/usgin/Workshops/blob/master/WesternRegionalWorkshop/Hyper-V%202012%20Replication%20on%20Workgroup%20Servers.docx* (Replication process documentation is available in Appendix E-3-a). This document has a step-by-step description of the process to implement replication and should be used by all those hosting data, files, metadata, and/or web services. The hubs are currently replicated regionally; the Illinois hub and Kentucky hubs have replicated and backed up one another's servers, as have the Arizona and Nevada hubs.

The NGDS development team developed an NGDS Data Archiver (*https://github.com/ngds/data-archiver*) application that scrapes the central catalog (or any CSW catalog), finds the URLs for each resource if there is an online distribution, gets the document or file, and saves to a local directory. If a WFS distribution is available for a dataset, the application accesses the service and creates an XML file containing all the features offered by the service that are placed in the local directory. Each resource in the local directory is packaged in a zip archive with the metadata record for the resource and a file containing the resource.

Development of the Sustainability Plan in 2013 was based on the idea of designating the NGDS structure and delivery system, USGIN, as a non-profit foundation to ensure the continued use, maintenance, and growth of the standards and protocols on which NGDS content models are based. USGIN Foundation, Inc. was incorporated in 2014 (See Appendix H for bylaws and Articles of Incorporation) as a legal entity to support NGDS. The USGIN website (*www.usgin.org*) has been updated to explain and promote USGIN capabilities and principles.

Task 11.0 Adding new technical or institutional data

Purpose: Add new data from participants and add participants.

Approach: Participants will revise their SOWs regularly in accord with guidance from the Scientific Advisory Board (SAB). Project management will be seeking participation by additional data and service providers throughout the project.

Actual accomplishments: The data acquisition workflow (see Section 5: Data Acquisition) consisted of a cyclic process that started with each data provider evaluating their data holdings that had not been registered in NGDS, along with the interests of the geothermal community in their region, to construct a SOW outlining the top priority data items to acquire for NGDS during the following year. These SOWs were reviewed by the SAB, and adjustments to the work plans were made based on the board's recommendations before work plans were approved.

Over the course of the project, in collaboration with the Design-Build project, the metadata has been harvested from Southern Methodist University (which includes all Geothermal Resource Council (GRC) library and Stanford geothermal program library holdings), the USGS Geothermal Energy program, and the University of Utah Energy and Geoscience Institute Geothermal Group.

Task 12.0 Publicizing addition of the new data

Purpose: Keep providers and users aware of new data available.

Approach: The project will implement the Technology Transfer plan (Subtask 13.4) and provide notification and syndication features in the servers and portals so that users can receive automatic notification when new data are available. Features will include the ability to select data from certain areas or of certain type.

Actual accomplishments: NGDS portal announces latest data acquisition. A Rich Site Summary (RSS) feed enables the publisher to syndicate data automatically. The RSS feed also benefit users who want to receive timely updates from favorite websites or to aggregate data from many sites. Subscribing to the NGDS website RSS removes the need for the user to manually check the website for new content as their browser constantly monitors the site and informs the user of any updates. The browser can also be commanded to automatically download the new data for the user.

Task 13.0 Project Management and Reporting

Purpose and Approach: Reports and other deliverables will be provided in accordance with the Federal Assistance Reporting Checklist following the instructions included therein. Comprehensive reports at the end of each project phase will serve as the basic milestones for each of the tasks and their subtasks. These reports will serve as the basis for go / no-go decisions for the project as a whole and for the various tasks. It is anticipated that the review of the reports will also lead to revisions of the tasks and subtasks as necessary.

Actual accomplishments: Listed with subtasks. The project received good to excellent scores in all categories for each of the annual Peer Reviews of DOE-funded geothermal awards. The project was regularly among the highest rated among all 200+ projects.

Average Rating (range)	2011 (0-4)	2012 (0-4)	2013 (0-10)
Relevance/Impact of Research	-	4.0	8.5
Scientific/Technical Approach	-	3.5	8.0
Accomplishments, Results, Progress	-	3.3	8.75
Project Management & Coordination	-	3.7	8.25
Total	3.5	3.6	8.4

Subtask 13.1 Implementation of organization structure

Purpose: To achieve project goals on budget and on schedule

Approach: Establish a management team at AZGS for the project; outline duties of the Management and Scientific Advisory Boards and appoint members; coordinate with the external Technical Advisory Board; implement a project communications plan; develop a technical training plan.

Actual accomplishments: The project was managed by a Project Manager at AZGS who reported to the project PI, Lee Allison. This position was held by Catherine Martinez at project start-up, by Diane Love as interim project manager in YR-1, and by Kim Patten as project coordinator thereafter, who coordinated subcontracts, reporting requirements, meeting logistics and scheduling. Randi Bellasai served as Project Manager for the final 6 months of the project. Diane Love supervised and provided guidance for most of the data compilation and review workflow. Celia Coleman and Christy Caudill provided lead for the technical services and web services. Project Co-PI Stephen Richard supervised the technical development.

Subtask 13.2 Coordination among participants, collaborators, NGDS, and stakeholders

Purpose: Ensure the project meets the needs of the NGDS, as well as project participants, collaborators, and stakeholders.

Approach: The team will establish advisory boards (described in Tasks 2, 3, and 13.1); integrate NGDS key personnel into the project management and technical processes; hold regular meetings with NGDS participants; establish formal affiliations among participants and collaborators; develop regular working relationships among the parties.

Actual accomplishments: The project assembled a Science Advisory Board (SAB) and Management Advisory Board (MAB) to provide input on priorities for data acquisition and evaluate results over the course of the project.

The SAB was comprised of 5 members appointed from funded participants in the project, along with DOE Program Officer Arlene Anderson. Dr. Rick Allis, State Geologist & Director, Utah Geological Survey, Utah was appointed Chair. Each appointed member has technical expertise or knowledge in one or more areas of geothermal energy. The other members are Dr. David Norman, Washington Geological Survey, Dr. Lisa Shevenell, Nevada Bureau of Mines & Geology, Dr. John Costain, Virginia Geological Survey, and Dr. Ed Deal, Montana Bureau of Mines & Geology.

Duties of the SAB included:

- Review priorities and guidelines for populating the NGDS with state-specific data
- Review the project solicitation to the participants, that is prepared by the Project Management
- Advise project staff on preparation of SOWs by participants
- Review the annual SOWs submitted by participating states
- Recommend actions to the Project Management to revise SOW deliverables (priorities, amounts, types of data)
- Review prior year's accomplishments (re: expectations and obligations) and make recommendations for continued funding to the Project Management

The SAB met in person each year to review and evaluate annual SOWs from participants. Because the SAB members were each actively engaged in meeting the project deliverables, they fully understood the requirements and needs for meeting deliverables. They provided frank and demanding guidance on deliverables from each of the participants and were rigorous in holding each participant to their commitments before endorsing approval of funding for subsequent years in the project. This internal peer review accountability was key to providing guidance to assure geothermally-relevant data and fully-executed delivery at each year end by all state subcontractors.

The Management Advisory Board was comprised of the Association of American State Geologists (AASG) President, who served as Chair, along with the AASG Vice President, AASG Past President, Dr. Walter Snyder, Principal Investigator, National Geothermal Data System, Boise State University, Boise, Idaho, and a representative from the DOE Geothermal Technologies Program in Golden CO (initially Ava Coy).

The MAB duties include:

- Review progress and advise Project Management team on goals, priorities, deliverables, and accomplishments
- Review and advise on Project Management's actions in response to SAB and Technical Advisory Board recommendations
- Review or recommend changes to meet AASG goals and project obligations
- Advice and consent on members of the project SAB

The MAB met in person at least twice yearly usually in conjunction with AASG meetings, which reduced travel costs and provided opportunities to engage other State Geologists who are the main subawardees on the project.

There were multiple positive outcomes of the MAB membership, structure, and activities, including helping create a highly successful national collaboration environment and active involvement by the participants, a non-trivial achievement. Having a cohesive national coalition that worked together seamlessly from day one was instrumental in keeping the project on task, on schedule, and on budget.

The project initially planned on using the previously established AASG-USGS Geoscience Information Network (GIN, now USGIN) Steering Committee as a Technical Advisory Board (TAB) to the project to review and make recommendations on data and metadata standards, protocols, services, and priorities. The TAB is intended to promote agreement and consensus on technical issues facing the data integration framework and system design of the NGDS.

Duties of the TAB were to promulgate plans and make recommendations to project managers on technical aspects of NGDS, including:

- Use cases to guide system development
- Priorities for technical development based on adopted use cases
- Adoption of specifications, including data and metadata content models, interchange formats, and service profiles
- Adoption of controlled vocabularies and ontologies for data integration
- Adoption of conformance and quality assurance tests and metrics
- Identify possible duplication of effort or incompatible component development and recommend solutions to maximize development efficiency and effectiveness
- Identify emerging best practices

Subsequently, we encouraged the establishment of a technical working group for the NGDS (GDSDPWG- Geothermal Data System Development and Population Technical Working Group), which included representatives from all the NGDS-related DOE funded projects.

A major advantage of this was to develop and promulgate common technical standards, guidelines, and requirements across all of the NGDS projects.

Subtask 13.3 Milestones

Purpose: Coordinate the project deliverables with the NGDS.

Approach: Management and technical staff from the project and from NGDS are working collaboratively and will be formally integrated into the respective projects.

Actual accomplishments: The project standards and protocols were developed during Phase I and implemented for project data concurrently. Networking functions were implemented during Phase II, on time and on budget. NGDS standards and protocols were available to project data providers at the appropriate time to contribute to the NGDS but were not implemented by the Design/Build project due to internal problems. This project subsequently assisted the Design/Build sub-recipients to adopt and implement technical standards and procedures.

Subtask 13.4 Technology transfer

Purpose: Reach out beyond the project participants and collaborators

Approach: NGDS team will make presentations at geothermal and geosciences meetings and conferences; exhibit at technical and industry conferences (in partnership with NGDS core, AASG, and others as opportunities permit); maintain and contribute to website(s); hold workshops and seminars (for data providers) opened to non-SGS's; prepare and distribute training materials ("cookbooks" and online tutorials) freely to anyone. (Outreach materials listed in Appendix H.)

Actual accomplishments: Outreach and technology transfer activities include 108 presentations at technical meetings, 31 briefings, 20 webinars, 22 exhibits at professional conferences 18 news media interviews, 5 non-technical articles, 72 sub-recipient presentations, news interviews, or other outreach events. Subjects included cyberinfrastructure, distributed networking, geoscience data exchange, geoinformatics, enhanced geothermal systems, the needs of the geothermal community, and open data. Talks and presentations took place before bodies and conferences organized by such as the American Institute of Professional Geologists; the European Geophysical Union; the Geothermal Resources Council; the Geothermal Energy Association; the American Association of Petroleum Geologists; and the Geological Society of America, with audiences including end users, data providers, and developers across the country and internationally. Demonstrations of NGDS capabilities and functionality were given at exhibit booths in venues such as summits, annual meetings, workshops and exhibitions. Social media outlets, including Facebook, Twitter, GitHub, LinkedIn and YouTube, were created to provide ongoing avenues for communication, promotion, and product updates. Multimedia products include tutorial videos, webinar recordings, data access scenarios, and lightning talks from the final presentation of the SAB. Webinars have been offered at no cost to potential users and contributors on various aspects of the system. A series of twelve access scenario video tutorials to enable users to search for and reformat data in map and database applications has been produced.

Some notable achievements were the articles about the project in the online and print edition of Scientific American and the cover story in EARTH magazine. A highlight of the project was the formal launch of NGDS by Secretary of Energy Ernest Moniz at the White House Energy Datapalooza, in Washington, DC, in May, 2014 which drew widespread news media coverage.

Section 5: Summary of Project Objectives

The project, to promote digital accessability to existing geothermal-relevant data, was executed in 3 phases. An initial planning and preparation phase was followed by the data acquisition phase. The third phase is deployment and operation of a sustainable production system, and has been an ongoing activity during the latter part of the project and will continue forward.

Hypotheses

Our starting hypotheses included:

1. State geological surveys are the stewards of a large body of information that would be useful in reducing the uncertainty and risk inherent in geothermal energy resource exploration and development.
2. The best way to increase the utility of this information is to convert it to digital formats and make it discoverable and accessible to the public on the World Wide Web.
3. Standardization of the formats and services used to provide access to the data is the most efficient way to provide access.
4. An architecture consisting of a distributed system of loosely-coupled data providers and a central catalog for search across this system is the best approach to integrating data from the large group of independently-governed data providers.

The initial proposal for the project identified the US Geoscience Information Network (USGIN) as the platform of choice for implementing the system. Work on this data compilation project was initiated and largely conducted in parallel with the NGDS Design and Build project based at Boise State University

(referred to as the 'Design and Build Project'). The Arizona Geological Survey was a sub-contractor on the Design and Build Project, so the integration between the projects was very close.

Based on the USGIN framework, the NGDS adopted the USGIN ISO19115/19139 metadata content model and XML encoding (USGIN, 2011b) profile for the description of NGDS resources in the system catalog. USGIN also provided a specification for the minimum and recommended content (USGIN, 2011a) to include in metadata records describing NGDS resources. The USGIN platform also includes the use of Open Geospatial Consortium (OGC) web service specification for map services (Web Map Service, WMS), feature services (Web Feature Service, WFS), and catalog services (Catalog Service for the Web, CSW).

Approach

Preparation

The project team implemented a Task management application for tracking review processing, and an Administrative application for tracking and reporting data submissions. Both applications were accessed via the web. The Task management application ('Task site') was implemented using Drupal v6, and the Administrative application ('Admin site') was implemented using Django and Python. These utility applications were designed for internal project management, not for public distribution. The system allowed the team to log new data submissions, assign review tasks to team members, track the status of submissions, count the numbers of submitted items to compare with sub-recipients' SOWs, and enabled state subcontractors to quickly review their progress.

Resource Classification: Tiers of NGDS data delivery

To plan for the data acquisition phase of the project, the team reviewed and synthesized the inventories of resources (files, databases, etc.) that the states had prepared as part of the response to the DOE funding opportunity. Analysis of the inventory resulted in a three-tier scheme to categorize data resources delivered during the course of the project. Tier 1 includes file-based resources as documents useful to people, but not designed for straightforward automated processing. Tier 2 resources are those that are in structured formats (tables, databases) designed for machine processing and access, but not organized in standardized formats. Tier 3 resources are data in an NGDS standardized interchange format documented in an NGDS information exchange template designed for interoperability. Tier 3 data use standardized schema with specific field names, data types and domain constraints designed as interchange formats that can be used for interoperable data services.

Tier 1

The simplest and most common access to resources is provided by Web links that result in a file download. Information contained in files can be accessed by users who have software that can recognize and open these files. This is the standard model for files accessible on the web, supported by HTTP servers and desktop web browser software.

Unstructured data requires user interpretation before it can be used for analysis. Users can utilize the information if they understand the encoding and language, but the system provides no support for this understanding, and little or no automation is possible across various sources of data. Audio files must be transcribed; text files must be parsed and mined for data that are then broken down and structured in ways

that can be processed by computers; images must be scanned, interpreted, and often georeferenced. Preparing Tier 1 data for analysis can be a painstaking and time-consuming process.

Tier 2

Tier 2 interoperability indicates that information content is structured (consistently organized) in a consistent pattern that is amenable to computer processing. However, Tier 2 data does not use a shared, documented data structure. Data in this tier must be transformed by the data consumer on a case-by-case basis for integration with other datasets, requiring them to study each new data source to discern how to extract the information they need. Numerical data tables and databases must be field matched and manipulated before consolidating. Obtaining data in a structured format is a step towards interoperability because once the format is understood, computer programs can be implemented to extract the desired information.

Tier 3

Tier 3 data are structured data that conform to an NGDS information exchange. Data that are published according to the exchange specification (content model, interchange format, service protocol) are interoperable with any other data published using that exchange. These are the most valuable data in the system, as they allow end users (researchers or computer programs) to retrieve and manipulate data from any source in a predictable and expedient way. When a data file (CSV or XML, for example) is "schema-valid," it means that the names of fields in the data structure and the data type and content of those fields conforms to an information exchange specification.

The Tier 3 data delivery approach used by NGDS is similar to, and informed by, the National Information Exchange Model (NIEM) (Anonymous, 2007-01-09) used by a number of U.S. federal agencies that deal with principally legal data. The core concept is the information exchange, a community agreement on conventions necessary for the interoperable exchange of some particular information. The exchange specification documents a scope, the content model (descriptive components) for the data items of interest, interchange formats for encoding and transmitting information electronically, and the protocols used to request information. By using an information exchange specification, a data provider can publish data that will be available to any client that implements the exchange, and a client application can access data from any provider publishing data according to the exchange conventions. This allows data in a variety of formats, organization, and structure to be integrated without having to manually transform it. For more information on information exchanges for data delivery see Section 5: Content Models.

USGIN Information Exchanges

Tier 3 data delivery requires the definition of the information-exchange entities of interest (feature types), and the information about those entities. Based on previous experience with standardization of data interchange formats, the project team determined that Tier 3 information exchanges would use 'simple feature' or 'flat-file' schemes, in which entities are modeled with a collection of single number- or string-valued properties. This requires de-normalizing normalized database content, but makes the interchange documents simple, and accessible via OGC Simple Feature Web Feature Services (WFS) (van den Brink et al., 2011). Commonly used Geographic Information System (GIS) software such as Esri ArcGIS or MapInfo, and the free, open-source Quantum GIS software can access data from these services out of the box.

The concept of an information exchange emerged as a very useful framework for structuring workflows to release data in a distributed system like NGDS. In the USGIN framework, an information exchange is a specification for how to expose a particular kind of data through a particular service protocol in a particular interchange format. The specification includes 1) an abstract content model for the information that defines the entity or entities of interest, their properties, and the logical data type to specify property values, possibly including controlled vocabularies; 2) and implementation of the content model using a specific encoding scheme to define the information exchange format; 3) the service protocol or protocols that will be used for accessing information. The process for developing information exchange specifications developed by the project and outlined here is based on operational systems of the Open Geospatial Consortium (OGC, *http://www.opengeospatial.org/ogc/process*), International Organisation for Standardisation (ISO, *http://www.iso.org/iso/home/standards_development.htm*), and the EPA Environmental Information Exchange Network (http://www.exchangenetwork.net/data-exchange/).

Content model

Content models are abstract specifications that define a feature for information interchange, along with the properties associated with the feature and relationships to other features. The content model in its simplest form is a statement that some particular entity (feature or observation) will have some list of properties. For example a content model record describing a book might include: title, author, publisher, publication date, publication place, citation, and an ISBN number (a URI).

Content models were developed through an organic process during the first year of the project, evolving rapidly as production-scale data compilation commenced. WFS services were deployed as the models evolved, and iteration of model versions and XML schema for corresponding WFS features resulted in some discrepancies between the Excel Workbooks and corresponding XML documents. During early September 2011, the AZGS team reviewed and synchronized the deployed services and XML schema to produce a collection of schema that will validate the deployed services. For currently deployed NGDS services, interchange document formats are implemented in XML, with document structure specified by an XML schema. The use of XML schema allows data documents delivered by services to be validated to ensure that they conform to the syntax dictated by the information exchange.

To avoid recurrence of discrepancies, a more formal process for information exchange development was instituted, using GitHub version tracking. The information content in the interchange documents was determined empirically based on examination of sample data obtained from data providers. This 'bottom-up' approach was favored in order to avoid inclusion of fields for properties for which none of the providers actually had data. Content models for NGDS information exchanges were assembled using Microsoft Excel workbooks. The use of well-known software made it easier for project participants to review content models and for the project to get more input. Each model is managed in its own GitHub repository (see *https://github.com/usgin-models*), with edits under version control.

New content models were developed when new data types emerged from the data provider network. The initial draft of a new content model is developed by an individual or a working group. Formation of the workgroup to propose new models is intentionally left very informal to reduce the barriers to participation. Content models may be developed as Excel workbooks, text descriptions, or UML models. Documents used in the development process are hosted in a version-controlled (VC) repository. Documents defining each content model are managed in one repository. The development documents

describe and specify the content model, but do not implement the content model in a computer-processable form. These documents are intended to help human users understand the content model.

When a draft is deemed ready for review (by the submitting group), the content model specification documents are made available in the GitHub repository, and availability of a content model for review is publicized in a request for comment broadcast to the intended user community by e-mail and RSS feed (http://notifications.usgin.org/). The request for comments includes instructions for commenting, and a time frame during which comments will be accepted. Comments are collected by e-mail and compiled in a comments worksheet that is maintained in the content model VC repository. At the end of the review period, the workgroup incorporates comments to produce a final specification document.

When a model is accepted for production work, a copy of the workbook defining the model is moved to a 'tag' branch in the repository (e.g. *https://github.com/usgin-models/WellFluidProduction/tree/v1.1.1*), and an XML schema is generated from the model. XML schema are currently being hand-crafted, a process that is tractable because we are keeping the structure of the interchange documents simple. The workbook and XML schema in the tag branch in the repository are locked so that they may not be modified, and copies are made into the public schema repository at *http://schemas.usgin.org/models/*.

The public model-repository site (*http://schemas.usgin.org*) provides various resources for users of the content model, including downloads to obtain the content model workbook, the XML schema, text explanation of the model, documentation for fields in the model, and Esri layer (.lyr) or OGC Styled Layer Descriptor (SLD) files that can be used to consistently symbolize data provided through the NGDS information exchange. The repository application also provides a service for validating a deployed WFS service for conformance with the NGDS information exchange, and a service that will determine if a comma-delimited (csv) tabular text file conforms to the field definitions in a stated content model.

XML schemas that implement the content models are versioned, and the namespace for the schema elements is unique to that schema version. Thus, namespace-aware client applications can determine if an instance document is using a known schema version. This enables greater user-application reliability.

The content model development process and conventions are documented in USGIN specification documents in the USGIN specification GitHub repository (*https://github.com/usgin/usginspecs*, available in Appendix C).

Content model workbook

An Excel workbook has been used for generating most of the current content models, and use of this workbook is recommended because it calls out important information that should be included, provides a starting point useful for most content models, and promotes following naming patterns used in other content models. Full description of the content model workbook is available in Appendix C.

Interchange format

In order to actually use a content model, it must be implemented using an encoding scheme that is computer processable. The eXtensible Markup Language (XML) is typically used to implement the content models, but other schemes (e.g. JSON) may be adopted as technology evolves. The important thing is that the information is encoded consistently in a structured, well documented format. This enables

computer programs to parse the interchange documents and extract desired information, and as newer encodings are adopted, conversion between formats can be automated using software.

Versioning and URIs

Test implementation of a content model should be done during the review of the abstract model, because this provides an excellent review of the model. When the specification is finalized ('tagged'), a final implementation can be developed. The implementation is specified by a separate document. In the case of XML implementations, the normative specification is an XML schema; other encoding schemes will have other normative implementation specifications. The implementation encoding is identified by a URI (a Namespace URI in the case of an XML implementation). This URI should be included in any instance document based on that interchange format implementation such that software consumers of the document content can determine that the document is encoded using a known format.

Each implementation must define a validation process that will determine if an instance document is conformant with the implementation.

Data Acquisition

Figure 1 below displays the data acquisition cycle. The cycle starts with development of a SOW by the data provider agency, outlining the information resources that in their judgment are the highest priority for online access, based on knowledge of current local conditions affecting geothermal energy development and the information holdings of their agency. The SOWs were reviewed by the SAB to evaluate the proposed work for geothermal relevance, technical soundness, and alignment of proposed budgets with proposed work. The board provided feedback and recommendations on each proposed SOW, in some cases requiring revision of the SOW. Plans were approved to proceed by a majority vote of the board, after revisions, if those were requested.

After acceptance of the SOW for an acquisition cycle, each participant was requested to prepare a small test data set to prototype the data delivery content and format for each deliverable product. The NGDS development team at AZGS reviewed prototype data that was provided to evaluate content consistency with the related information exchange. Debug iterations were typically required between the NGDS project team and each data producer to refine the data to conform to the NGDS information exchange requirements. In many cases, the data providers did not provide test data early in the process, and this iterative debugging did not occur until a near complete data product had been compiled. This resulted in additional work that might have been avoided in some cases.

When the data producer completed data compilation and formatting, they made final data submissions for technical review and approval, which was recorded in the internal tracking system. After approval by the review team and release of that year's proposed data deliverables through NGDS web services, the data producer could submit their SOW for the next data development cycle. Three annual cycles were conducted by most of the sub-contracting data providers over the course of the project.

Science Advisory Board process

The SAB was established to ensure alignment of project work with priorities for geothermal information requirements. The board includes technical experts with experience in geothermal energy exploration and development, selected from members of the AASG or project sponsors. The SAB provided oversight on proposals, ensuring their feasibility and relevance to geothermal energy exploration and development.

The SAB reviewed each project's annual SOW and voted on final approval for each year's work effort by subcontracting data providers. SAB members were required to abstain from comment on projects related to their home state.

For year one and year two data-acquisition cycles, the SAB held a face-to-face meeting with follow up teleconferences to review revisions and SOWs received after the face to face meeting. In year three, SAB meetings were held by teleconference.

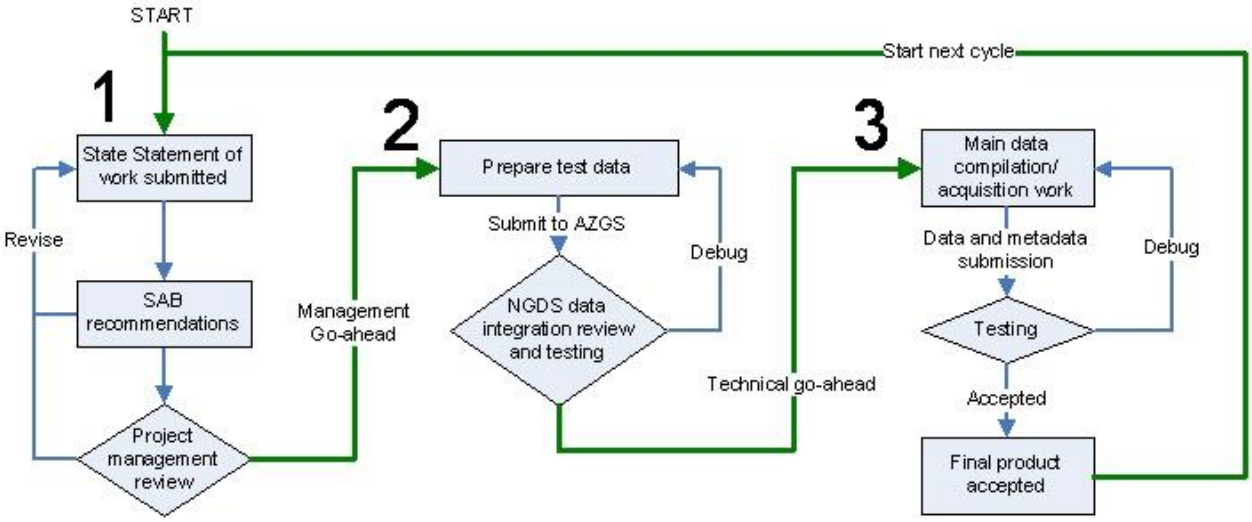


Figure 5-1. Data Development Cycle

Production System

One of the goals of the project was to deploy a sustainable, community-maintained information system. The approach to this goal was to engage as many of the project participants as possible in all aspects of data compilation and review, resource cataloging, and resource delivery on the Web. The approach was implemented through a series of workshop and webinars, as well as one-on-one communication between the project team at AZGS and other project participants. State subcontractors who required guidance through the data information exchange formatting process or expressed interest in deploying their own services were targeted, in view of the ultimate goal of having the data providers take ownership of their information resources and the provision of those resources.

Hub locations were selected and established with the states that had existing information technology capabilities sufficient to understand and deploy the NGDS technology. The hubs were geographically distributed with the goal of distributing expertise widely to promote diffusion of know-how to all project participants.

NGDS file repositories and data management

The focus of development for NGDS has been on the information exchange definitions--standardizing the service protocols and data interchange formats. The underlying assumption is that data providers in the system will be implementing and using their own data management practices and infrastructure, and NGDS does not mandate any particular approach. However, because many of the states decided they do not have the technical capability or resources to host their own repository, the NGDS hubs have been

made available to provide repository services. This is a very simple workflow for states, but is less sustainable than having states create their own node and managing their metadata locally.

The hub repositories are backed up according to the local back up practices for servers at each hub. In addition, we have implemented a replication system between pairs of hubs for disaster recovery backup. A third level of backup has been implemented using a software tool we developed that harvests all the resources listed in the catalog and stores them in an Amazon S3 bucket.

Hub deployment

Recognizing that many state geological surveys lack the technical resources to host USGIN/NGDS data services or catalogs in-house, the project established four hubs that were made available for hosting data from states in their region. Hubs were located at Arizona Geological Survey, the Nevada Bureau of Mines and Geology, the Illinois Geological Survey, and the Kentucky Geological Survey, and each hub was grouped with states in their region (See Figure 5-2 map). Hubs received extra project funding to support acquisition of new servers and extra personnel to provide assistance to states utilizing the hub's hosting resources. As a disaster recovery backup, the hub servers were virtualized during the third year of the project and virtual machine replication was established between Arizona and Nevada and between Illinois and Kentucky. The replication uses Microsoft Hyper-V technology, and allows recovery of the hub server at its replicated location almost immediately, with the lag in service mostly determined by the time for the host-name mapping to be updated in the Domain Name System for the internet. (See Appendix E for Hub/Node tools and processes).

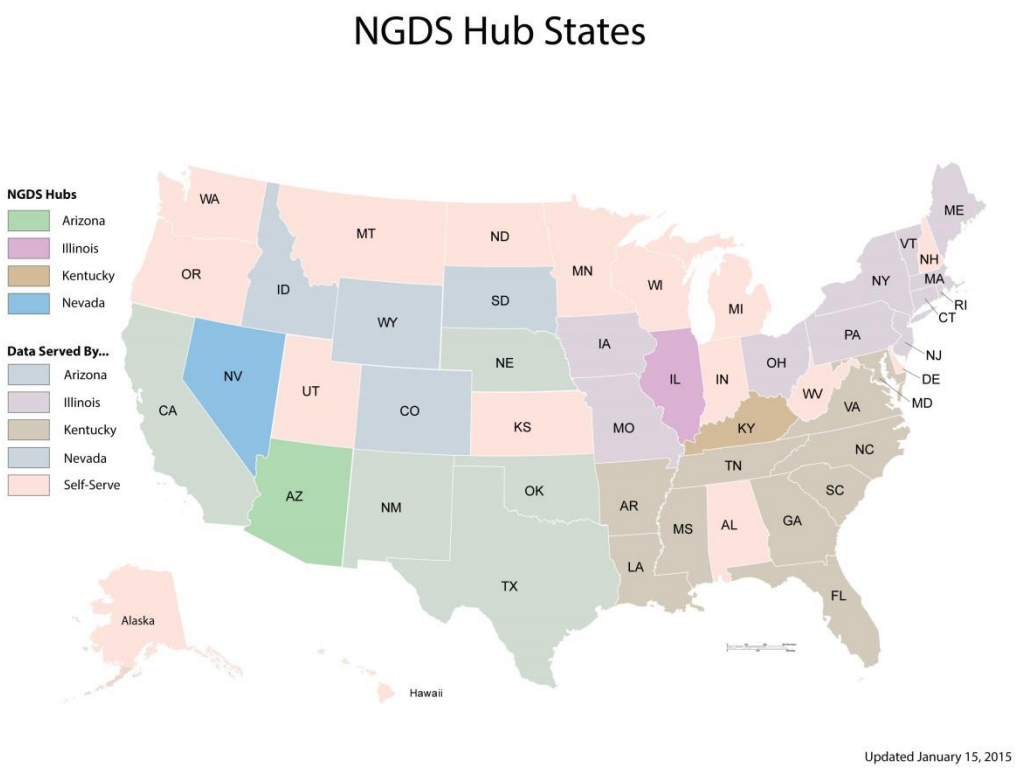


Figure 5-1. NGDS Hub Graphic.

Node deployment by state surveys

One of the most important moves toward a model to sustain system maintenance for data and services is to move web service management from the hubs and metadata management from the AASG Geothermal Data Repository back to the states themselves. This makes the system more distributed, sharing the associated costs in personnel and servers and time spent on data and metadata management for each node. The effort to move more states to their own node (hosting their own state's NGDS data, web services, and metadata through a catalog server) was a focus of activity near the end of the project term in October 2014 – December 2014 (see Task 5.1 *Integrating Nodes* for details about node installation and integration into the system). Remaining funding that had been returned from the state's various data collection efforts was reallocated to 12 states to install GIN-stack or another catalog service and metadata management configuration. The solicitation for funding to set up self-managed nodes was opened up to all state sub-recipients in October 2014 through Google Forms (Metadata Management Proposal and Metadata Management Financial Survey). As described in Task 6.0, seven states (Indiana, Utah, North Dakota, Wisconsin, Delaware, Michigan, and Alaska) opted to install Geoportal as a catalog server, coupled with GeoServer or ArcGIS Server for web services. Four states (Montana, Alabama, Hawaii, Washington) decided to install GIN-stack software. (One contracted state, TX, was not able to create a functional node). In both cases, data services have been transferred from the regional hubs back to these states to host, and the metadata hosted at the NGDS aggregating catalog have been extracted and also transferred to the local catalog for maintenance. These metadata are now harvested to the aggregator from their originating catalog. Six additional states so (Arizona, New Hampshire, Kansas, Minnesota, Oregon, and West Virginia) were already serving their own web services using ArcGIS Server and continue to do so.

Training

Regional Workshops

Throughout 2012 and 2013, the NGDS processing team planned and conducted as series of regional workshops to train project sub-recipients on NGDS specifications, practices, and operation to integrate their data into the system and develop capabilities to host their own nodes (pertinent to Tasks 3, 4, and 5, respectively). The 2013 workshops updated states on the tools and methods for Extract-Transform-Load (ETL) operations to prepare data for Tier 3 delivery. In each workshop, the processing team was able to give states more automated processes with tools to transform their existing data and metadata into the NGDS standards and prepare databases for web service publication, as well as workflows to manage GIS-enabled databases and to publish the data as OGC-compliant, NGDS schema-conformant web services. These workshops walked through the entire ETL process, starting with raw data and metadata, showing web service publication of that data, then explaining the importance and process of service validation, as well as hosting metadata and making it discoverable in the system. The materials from the workshops, including PowerPoint presentations and example data, are available at *https://github.com/usgin/-Workshops*.

Online Training

Training webinars were conducted to supplement face-to-face workshops and provide guidance as processing tools and methods evolved, and to inform project participants of new software developments

like the CKAN GIN-stack, and the new versions of the main NGDS aggregator UI. The NGDS championed integration of data services (WFS) along with the map services (WMS) which was previously the only service type specified by that organization. Additional considerations were needed for OneGeology WFS than was provided in the OneGeology documentation, which was the focus of the webinar.

Technical Details

Project submission workflow

Each sub-recipient received an account to log in at *http://stategeothermaldata.org*. Sub-recipients submitted products by logging in at this site to access project functions. They created a new submission by populating the information form 'Creating an Upload for File Submission' identifying the originating data compiler and deliverable item, and then uploading files containing content for review. The AZGS reviewers were automatically notified when a submission was received. The submissions were logged into the internal tracking system and a task was created for the review process and preparing the data for services (Subtask 2.3). The processing work flow is outlined in Figure 5-3.

Metadata acquisition

In order to be accessible through NGDS, a resource (data set, file, data service) must be registered by making NGDS-conformant metadata accessible to the system (see NGDS catalog, below, or Subtask 2.5). The resource must be at a web location with a URL provided in the metadata that will allow access to the resource. Project approaches to create or update metadata included providing web-based forms (the Metadata Wizard and NGDS repository publisher), metadata compilation spreadsheets, and mapping from existing database metadata using queries.

Several iterations of forms for entering metadata records individually were used during the project. The first interface was built using Drupal v.6, accessed either through a repository interface that allowed uploading files with the metadata record, and a 'metadata wizard' web site that simply created a USGIN ISO metadata record and made it available for harvesting via a web-accessible folder. The second generation of metadata forms was created with the new *http://repository.stategeothermaldata.org/repository* web site, using node.js, Django, and JavaScript. The new repository was implemented to support a staging database for better management of submitted metadata. Finally, the CKAN-based NGDS GIN-stack includes a metadata-creation form interface that was adapted to generate USGIN-conformant metadata. Each of these tools functioned similarly, with tabs for various sections of the metadata record, use of background user log-in information to auto-populate some of the metadata, as well as providing pick lists and autocomplete functions in the data entry fields. The date and timestamp of publication can be recorded, saving the data provider from having to manually create dates for each submission. The forms also provide some basic validation to assure that required content is entered.

Form-based metadata creation was most useful for registering single resources or small collections of resources. One common scenario was registration of Tier 3 datasets. The NGDS GIN-stack software stack enables generation of the metadata for the dataset using tabbed forms, upload of the tabular data as a CSV file, validation of the tabular data for conformance to a selected content model, and if valid, deployment of OGC WMS and WFS services for the dataset.

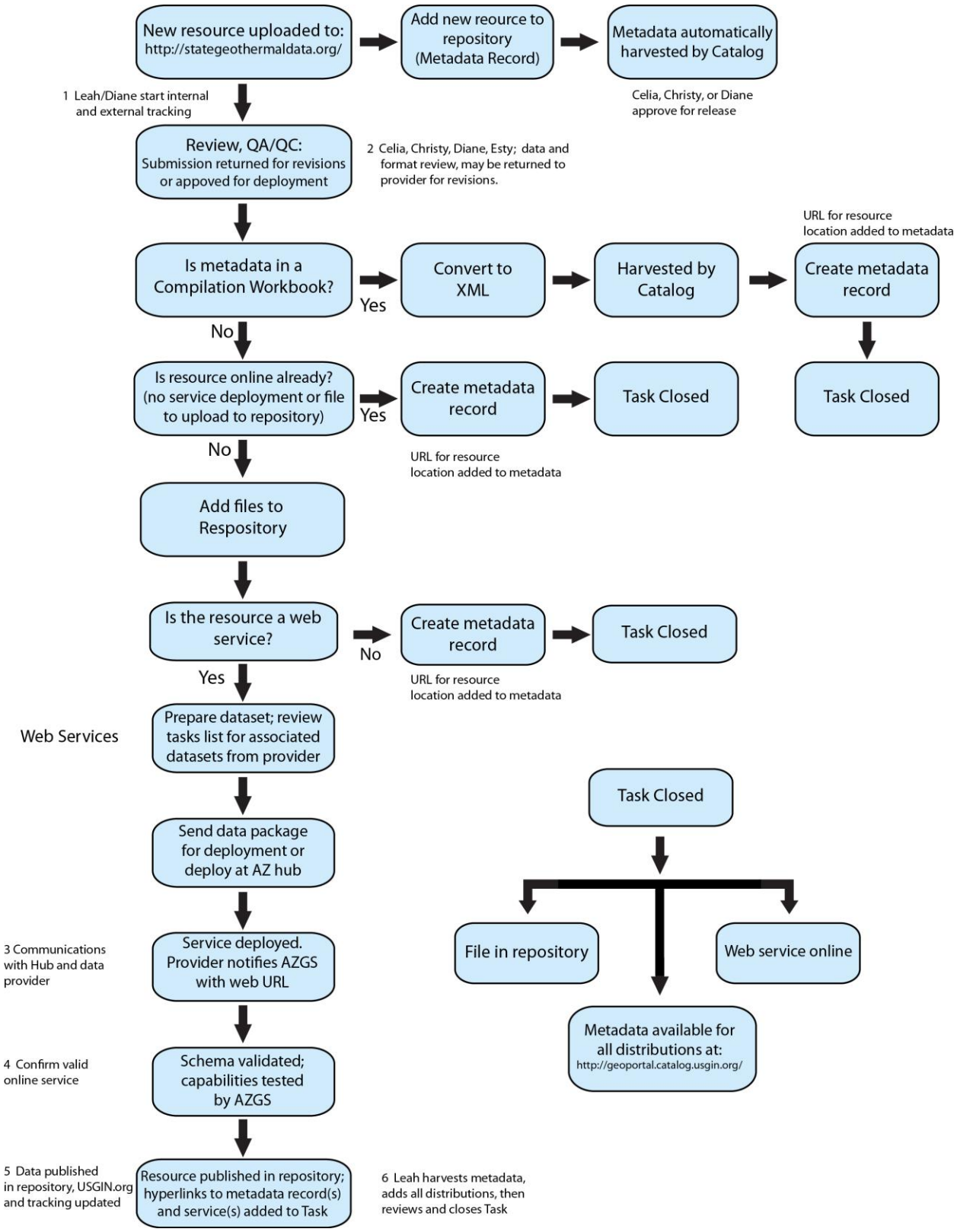


Figure 5-2: Data submission and processing work flow.

Many resources compiled by the project were already documented using various metadata content schemes from the internal resource management systems of the project participants. Existing metadata included lists of files compiled in a text document or worksheet, various databases constructed by organizations to manage their library holdings, and formal metadata conformant in varying degrees to Federal Geographic Data Committee (FGDC) or ISO standards, constructed according to locally varying interpretation and practice. In some cases, the available metadata collected was not sufficient to conform to USGIN recommendations, or did not exist, in which case metadata had to be created or updated.

To assist with capturing and updating this existing metadata, a workbook template (*http://schemas.usgin.org/models/#Metadata*) known as the compilation worksheet was created to assist users with extract-transform-load processing using familiar cut, paste, search/replace, and fill-down operations supported by the workbook software. The worksheet metadata compilation fields are mapped to the USGIN ISO metadata profile, and metadata entered in each record can automatically be converted to an XML record for import into the NGDS catalog. The software that does the conversion operates on a comma-delimited text (CSV) formatted table, exported from the worksheet or created by a variety of other workflows.

Metadata submissions in table format were required be in the Metadata Compilation content model Excel file (*http://schemas.usgin.org/models/#Metadata*). If the metadata records in the compilation worksheet did not have unique identifiers, a Globally Unique Identifier (GUID) was created for each record to allow unique identification in the NGDS system. The script for generating these unique GUIDs is available at *https://github.com/usgin/csvtometadata/wiki/Creating-Unique-GUIDs-for-Metadata*. These identifiers stay with the individual records through all subsequent review, updates, and processing to load into the NGDS catalog.

After review and approval for release, the metadata in the worksheet are saved as a comma-delimited text (CSV) file which is then imported into the AASG Geothermal Data repository application (*http://repository.stategeothermaldata.org*) for subsequent maintenance. This application also generates USGIN ISO 19139 XML documents for each published metadata record that are placed in a web-accessible folder for harvest into a catalog search application. A standalone CSV to XML Python tool is also available online at *https://github.com/usgin/csvtometadata*, to transform the encoding from NGDS metadata template CSV to USGIN ISO 1913 XML records.

Data providers with metadata contained in an existing relational database have a variety of options for publishing the metadata to the NGDS catalog. Structured Query Language (SQL) views that duplicate the table structure of the metadata compilation worksheet can be used to export CSV files that can be converted to XML.

Metadata staging database

The AASG Geothermal Data Repository (*http://repository.stategeothermaldata.org*) uses node.js to communicate to the Django-driven user interface component, and uses CouchDb (a free, open-source document database) as a staging database to store metadata records as JSON documents for internal review and editing. This staging system stores metadata in an internal JSON document that is designed against the USGIN metadata content requirements independent of the interchange formats, providing a flexible mechanism to accommodate different input metadata and allowing more options for editing and updating metadata. Bulk updates to metadata record sets were done using CouchDb instead of editing

individual records in the forms-based user interface. Bulk updates were used to change items such as contact information for a particular organization, adding keywords, or even modifying URLs. When edits are made to resources in CouchDb, these changes are immediately reflected in the repository. The code base for this system is onlined at *https://github.com/usgin/metadata-server*. Access to the AZGS network is necessary to view the CouchDb staging database. Bulk updates to records in CouchDb require a JavaScript script (analogous to an SQL script in a relational database) that applies the changes or additions to the records.

Validation

Most validation processes are listed in Task 3. The *usginmodels Application Programming Interface* (API) is used to run various validation tests as part of the file upload process for Tier 3 data (see Appendix E for validation documentation and tools). The *usginmodels API* was created to allow the content models to be made available to software for automated validation of data, available at *https://github.com/usgin/usginmodels*. Various NGDS tools have been developed to use this validation package The Excel to NGDS Services tool is deployed as an ArcMap toolbox, (see *https://github.com/usgin/ExcelToNGDSServiceTool*) reads data from an Excel worksheet or comma-delimited text (csv) file, runs the validation tests against a selected content model and version, and if the dataset validates, pushes the data to a geodatabase, read to deploy the Tier 3 service using ArcGIS server. The validation API is also accessible via a web page (*http://schemas.usgin.org/validate/cm*) that has the user select the content model and version, then upload a file (less than 15 Mb) to run the validation processes and return the results on a web page.

Using the NGDS GIN-stack, the user registers a resource as a structured dataset and selects an NGDS content model, then uploads a comma-delimited text formatted file containing the data table. The NGDS-extended CKAN interface uses the *usginmodels API* to validate the file against the selected schema. If the file is not valid, errors are shown to the user indicating necessary modifications to the data to make it conformant to the schema and valid tier 3 data and the user is redirected back to the Upload page of the CKAN interface so that a corrected file can be uploaded. This provides a mechanism to ensure that datasets are schema-valid before publishing the metadata.

Tier 3 Service Deployment

Nodes must install a web-facing, geographic data server software to publish data services. Through the NGDS GIN-stack, the user can deploy a validated Tier 3 dataset as an OGC WMS and WFS using the GeoServer component of the stack by clicking a button. Because the NGDS GIN-stack has only become operational during the last phase of the project, current production NGDS services have all been deployed using ArcGIS server (commercial) or GeoServer (FOSS). Each of the Hubs was associated with a university having access to an Esri ArcGIS site license. It was decided that the benefits of all the hubs using similar software configurations that they were already familiar outweighed the desire to use FOSS as much as possible. No matter what platform is used to deploy Tier 3 data services, the data provider must ensure that the names of layers (or feature Types) in the service are as specified in the schema for the NGDS data type, also given at *https://github.com/ngds/documents/wiki/NGDS-Tier-3-Service-Names,-Layer-Names,-and-URI-tokens*).

ArcGIS Server web service publishing

The AZ hub uses ArcCatalog to set up a database connection to an internal PostgreSQL SDE database ('staging'), and data for the services is deployed to a PostgreSQL database on a public-facing ArcGIS Server ('production'). Schema-valid data tables are edited in a 'staging database' and pushed to a 'production database' for service deployment.

GeoServer web service publishing

GeoServer is a free, open-source geographic data server software application that has been used to deploy some NGDS services. GeoServer uses PostgreSQL with PostGIS for back-end data management systems. Data have been deployed using GeoServer, by importing shapefiles containing Tier 2 or validated Tier 3 data into PostGIS to create a table.

The NGDS catalog

The foundation of the NGDS is the catalog of resources registered in the system using the system metadata conventions. From a user perspective, the face of the system is the NGDS portal at *http://geothermaldata.org*, which provides a search interface for searching the NGDS aggregated catalog, maintained at the AZGS hub. The aggregated catalog harvests metadata from all NGDS nodes. Registration for harvest is currently done manually through e-mail to the catalog manager at metadata@usgin.org.

Metadata for state contributions to this project are harvested into the State Geothermal Data Catalog (*http://catalog.stategeothermaldata.org/geoportal/catalog/main/home.page*) to provide a public-facing endpoint for all project metadata. Metadata from this project repository are harvested into the NGDS Portal at *http://geothermaldata.org*, into the USGIN Catalog (*http://catalog.usgin.org/geoportal/catalog/main/home.page*) as part of the larger USGIN, and by Data.gov to include in that national catalog. The State Geothermal Data Catalog is hosted in an Esri Geoportal instance, and the NGDS Portal catalog is hosted by an NGDS CKAN node.

Problems and Issues

Quality control and validation

Due to the large volume of data, the project was dependent for much of the quality control on the data providers. The collaborating agencies displayed widely varying levels of technical competence and attention to detail.

Additional desired work: To automate as much as possible QA/QC of submitted data, the project team as begun editing the XML schema implementation of the information exchange (XSD) to include data constraints on double (number) fields based on known geologic and natural constraints (e.g., temperature, depth, latitude/longitude, heat flow, thermal conductivity). This will not eliminate the need for data providers to thoroughly QA/QC their submitted data, but will greatly reduce the occurrence of typos and other obvious yet overlooked mistakes in data entry. This is a function that could increase the reliability of the data by adding another layer of quality control into the information exchanges themselves.

Current and proposed future efforts prioritize more in-depth data validation checks. These would be warnings given during processing or any schema validation (post-processing) to the user to ensure correct

data values based on natural known constraints built into the schema (XSD) implementations of the information exchanges.

Well headers and well logs

Initially the plan was to index all scanned well logs in the central aggregating catalog. As soon as large well-log index datasets were started to be loaded, it was found that search results became less efficient. The number of well logs so far exceeded the number of other kinds of documents that most free-text searches with fairly general search terms would yield results mostly consisting of singular well logs, making it hard to find the actual document resources of interest in the results.

To address this issue, the project team has deployed 'sampling feature services', which allow a user to search for the borehole or specimen (the sampling feature) that is the immediate subject of a well log or chemical analysis, and provide the binding between the analysis and the sample location and sampled geologic unit. The central aggregating catalog lists the sampling feature services, requiring a second search within in interface specific to the sampling feature service. With more resources, search could be chained from the catalog to the sampling feature service, and even to produce results across multiple well header services (see next section). This chained-search approach requires a more sophisticated search client than the basic catalog search provided by CKAN, Geoportal, or other off-the-shelf catalog solutions; see next item for a proposed solution.

Accessing across distributed services

The data getter application provides a mechanism to download datasets extracted from multiple services, but there is still no way to search on feature level properties across multiple services. We have prototyped a solution that scrapes the catalog to find instances of feature services offering a particular data type (e.g. borehole temperatures) and creating a SOLR index that allows seamless search of data offered across the service instances as though they were in a single database.

Regional resource centers

The project has not succeeded in getting stand-alone resource centers operational and robustly sustainable to the degree we would have liked. Success has been achieved when individuals take an interest in the system (e.g. Denise Hills, Alabama), but sustainability has been a problem. Engaged developers and tech people at Arizona, Arkansas, and Illinois have moved on to other projects or jobs. However, perseverance and maintained engagement has led to six independent data hosting nodes in addition to the eleven states contracted to become nodes.

Section 6: Products*

*All online products, including code bases, websites, and URLs, have been uploaded to OSTI.

Websites and media

Websites and Internet Locations that Reflect the Results of this Project

The following websites display, share, promote or otherwise make note of NGDS products or materials.

Primary NGDS websites

- National Geothermal Data System; *http://geothermaldata.org/*

    - Includes individual pages dedicated to Map, Library, Resources, Content
- State Geothermal Data; *http://www.stategeothermaldata.org/*
- NGDS Data Explorer; *http://data.geothermaldatasystem.org/*
- SMU Geothermal Data Aggregation – NGDS Node; *http://geothermal.smu.edu/gtda/*
- AASG National Geothermal Data System – Western Regional Data Hub; *http://web2.nbmg.unr.edu/GeoT/*
- GeoFRAT Geothermal Financial Risk classification and Assessment Tool; *http://geofrat.egi.utah.edu*

Websites with capabilities to serve NGDS data

- Data.gov- National federal data website; *https://www.data.gov/*
- National Map – USGS website of online maps and mapping products; *http://nationalmap.gov/*
- OneGeology – International map initiative of world geological surveys; *http://www.onegeology.org/*

USGIN-related websites

- USGIN Lab; *http://lab.usgin.org/tools* (with pages dedicated to standards, profiles, applications, implementations, community and tools)
- USGIN Document Repository
- USGIN informational site (www.usgin.org)
    - How USGIN Works; *http://usgin.org/page/how-usgin-works*
    - USGIN in Action: NGDS *http://usgin.org/page/usgin-action-ngds*
- Search USGIN; *http://search.usgin.org/*
- USGIN GitHub; *https://github.com/usgin*
- USGIN Technical Information (tutorials); *http://tech.usgin.org/tutorials*
- Delaware Geological Survey USGIN; *http://www.dgs.udel.edu/projects/united-states-geoscience-information-network-usgin*
- USGIN Geoscience Content Models; *http://schemas.usgin.org/home/*

Promotional or informational sites

- What is the National Geothermal Data System; *https://www1.eere.energy.gov/geothermal/pdfs/ngds_factsheet.pdf*
- Energy.gov - *http://energy.gov/search/site/national%20geothermal%20data%20system?gid=859446* Energy.gov includes a number of online documents providing information on NGDS, e.g., Energy Department Announces National Geothermal Data System to Accelerate Geothermal Energy Development (*http://energy.gov/eere/articles/energy-department-announces-national-geothermal-data-system-accelerate-geothermal*)
- An Open Source Stack for the National Geothermal Data System; *http://lanyrd.com/2013/foss4gna-2013/schqtt/*
- The National Geothermal Data System – Free online access to America's Digital Geothermal Resources; *https://www.isgs.illinois.edu/national-geothermal-data-system-free-online-access-america-s-digital-geothermal-resources*
- AASG NGDS Project – Powerpoint Presentation; *http://www.stategeologists.org*

Social media sites

- National Geothermal Data System – Facebook; *https://www.facebook.com/NationalGeothermalDataSystem*
- Geothermal Data-NGDS –Twitter; *https://twitter.com/geothermaldata*

- National Geothermal Data System – YouTube; *https://www.youtube.com/user/geothermaldata*
- NGDS Outreach – LinkedIn; *https://www.linkedin.com/pub/ngds-outreach/66/a6b/17b*
- USGIN Foundation – Twitter; *https://twitter.com/USGINFoundation*
- USGIN – Facebook; *https://www.facebook.com/USGIN*

Webinars, Videos, and Products to Promote and Explain NGDS Data

12 "Scenario Card" NGDS Access Tutorial Videos – These videos, available on the NGDS YouTube channel and on the main NGDS website, walk the user through use scenarios such as viewing data in GeoServer, and importing to a WFS or database program.

- NGDS Data Access Scenario 01: Viewing a Well Log Service with the NGDS Data Explorer

- NGDS Data Access Scenario 02: Wildcard Search
- NGDS Data Access Scenario 03: Merging Datasets from Different Sources
- NGDS Data Access Scenario 04: Basic Measurements with the NREL Geothermal Processor
- NGDS Data Access Scenario 05: Creating an OpenLayers Web Application
- NGDS Data Access Scenario 06: Creating a Leaflet Web Application
- NGDS Data Access Scenario 07: Filtering NGDS Catalog Results
- NGDS Data Access Scenario 08: Performing Analyses with NGDS Data
- NGDS Data Access Scenario 09: Performing Geospatial Analyses with the USGS National Map Viewer
- NGDS Data Access Scenario 10: Performing Geospatial Analyses with ArcMap
- NGDS Data Access Scenario 11: Discovering Related Datasets in uDIG
- NGDS Data Access Scenario 12: OneGeology Maps and URIs
- NGDS Data Access Scenario 13: WFS Symbology in QGIS

NGDS Atlas – This atlas, accessible online on the main NGDS website and available for print, showcases the state contributions of data and derivative products to the NGDS and describes positive outcomes of geothermal exploration made possible by the NGDS product.

Webinars documenting how to use, submit and host NGDS data

- AZGS Webinar: "Metadata tools and workflows for AASG Geothermal Data," February 15, 2011 [32 attendees]
- AZGS Webinar: "Metadata tools and workflows for AASG Geothermal Data," February 16, 2011 [35 attendees]
- AZGS Webinar: "Web Services for Bottom-Hole Data," March 16, 2011 [36 attendees]
- AZGS Webinar: "AZGS HUBS – Data and Serving Data," April 12, 2011 [6 attendees]
- AZGS Webinar: "AZGS HUBS – Data and Serving Data," April 26, 2011 [4 attendees]
- AZGS Webinar: "Reviewing Data for Online Delivery," May 17, 2011
- AZGS Webinar: "Reviewing Data for Online Delivery," May 17, 2011 *https://www.youtube.com/watch?v=IUVA7TKCUJk&index=3&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx*
- Webinar: Lehnert, Kerstin. "Introduction to IGSN and SESAR," March 8, 2012 [31 attendees] *https://www.youtube.com/playlist?list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx*
- Webinar: Kay, Martin, Gosnold, W., & Lovenitti, J., "Geothermal Data: How's it Used & Why it's Important". April 26, 2012. [37 Registered Participants]

https://www.youtube.com/watch?v=CRPfSyjX6cs&index=2&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx

- AZGS Webinar: "Updated Data Delivery Tools for AASG NGDS," October 4, 2012.
- USDOE Virtual Workshop: Allison, M. Lee, "The National Geothermal Data System and the Energy-Water Nexus," Energy/Water Datasets and Models Workshop, U.S. DOE, November 28, 2012.
- Allison, M. Lee, moderator, "Web Accessibility Strategies," Association of American State Geologists Annual Meeting, Breakout Session, Deadwood, SD, June 10, 2013. LiveStream & YouTube. http://www.youtube.com/watch?v=_rggaq-hCs8&feature=share&list=UU8wP7yOz8_ebyNkSoUGmdSg

Tools for Data Discovery and Publication

Excel to NGDS Services Tool

Task 5.2 Workflow and tool for data providers to encode data into information exchanges describes the workflow for using the *Excel to NGDS Services tool* and accompanying documentation, which has some quality control parameters in the code. These parameters, as outlined in Task 5.2, include constraints on number fields such as location information, temperature, depth, and others. This tool was used by the data processing team after state data submissions were entered and tracked through the system, (workflow described in Task 2.3). In addition, the processing team included multiple layers of manual quality checks for outliers possibly indicating incorrect data in number fields, as well as many comparisons of text fields, embedded links, URIs, geospatial locations, and other components indicated in the data. As any inconsistencies were found, the dataset was returned to the state data provider for corrections or to ensure that outliers are indeed correct. These problems are indicated in the ReviewersComments worksheet (see Task 5.2 for information on specific worksheets) of the content model workbook. After the data provider corrects or otherwise validates the data and addresses the issues identified by the data processing team, the file is once again submitted to *stategeothermaldata.org* to go through the review processing system again.

Tool documentation is available at *https://docs.google.com/document/d/1H08dObu5pWze7g3sSap3uDcI4DwYzuIEMfcP0V7Xj9s/edit?usp=sharing*.

https://github.com/usgin/ExcelToNGDSServiceTool - Tool to validate and convert the data worksheet from an Excel content model to a geodatabase, ready to be deployed as a NGDS Service.

USGIN models API

NGDS CKAN USGIN stack uses the *usginmodels API* to validate data. This API uses the same validation code base as Excel to NGDS Services, but is accessible as a service and through web-based interface at *http://schemas.usgin.org/validate/cm.*

A more updated implementation of validation of data against a given schema was written into the file upload process of the CKAN user interface at geothermaldata.org, which will likely be the method utilized by State sub-recipients and other data providers moving forward, as more GIN-stack (see Task 5) catalog and node solutions are used. As part of the file upload process in the CKAN code for the main

aggregator at geothermaldata.org and the GIN-stack software, the CKAN stock code of the datapusher extension has been extended to run the usginmodels Application Programming Interface (API) validation as part of the asynchronous file upload process. The *usginmodels API* was created to allow the content models to be made available to software for automated validation of data, available at *https://github.com/usgin/usginmodels*. When a user uploads files through the CKAN interface to create metadata in the system and valid Tier 3 web services, hooks to the *usginmodels API* have been implemented, where the file is validated against the specified schema. If the file is not valid, errors are shown to the user indicating necessary modifications to the data to make it conformant to the schema and valid Tier 3 data, and the user is redirected back to the Upload page of the CKAN interface so that a corrected file can be uploaded. This does not in any way remove the need for the data provider to manually check the data for quality control, but does provide a mechanism to ensure that schema-valid data is uploaded to the system. The *usginmodels API* does include the same data quality checks as in the *Excel to NGDS Services* tool, as the code was replicated from this tool to *usginmodels API*.

https://github.com/usgin/usginmodels - Python API for interacting with USGIN Content Models

Data Archiver

The Data Archiver *(https://github.com/ngds/data-archiver)* is a program was developed to scrape metadata from a CSW catalog and copy all files and OGC-WFS-based datasets to an archive on Amazon S3. It is a Node.js command line utility for:

- Scraping XML metadata from CSW and WFS
- Downloading linkage data from CSW and WFS
- Pinging hosts and returning status codes
- Pinging data linkages and returning status codes
- Writing ping status to CSV files
- Uploading data to Amazon S3

Workflow Management Tools

As part of state data collection and documentation, a workflow for managing state deliverables was implemented for use through the entirety of state data processing over the life of the project funding. This workflow was enabled by creation of multiple project administration sites which are outlined in more detail below. More distributed and sustainable methods of data management were developed and are in use moving forward, which are addressed in Task 5.1 *Integrating Nodes* below.

State Geothermal Data submission site

The workflow for state data submissions management begins by states uploading submissions at *http://stategeothermaldata.org/node/add/member-submitted-files* (available for those with site logins). This Drupal site has the specific state deliverables as indicated on their SOWs at a drop-down list. The user must choose from this list to upload files which allows for better management of incoming deliverables for the user and submission processing team. After submission, the files are tracked internally.

USGIN Task Management site

The Task Management site at *http://task.usgin.org/* is the processing management and tracking site where members of the data processing team are assigned one task per file submission. File name, location,

instructions, information about the deliverable, and contact information are indicated. The processing team preforms QA/QC on the data, ensures conformance to content model standards, and processes Tier 3 data to produce web services (see Task 3 and Task 5.2 below). Any necessary files are made available online, generally hosted at the hubs (see Task 4). Through this process, each step is detailed in a comment to follow the exact steps or actions preformed on any submission. This assists in tracking any edits or changes made, with a timeline and reasons for changes, and helps to ensure the integrity of the original data. It became apparent that this detail in data or metadata processing was essential to understand the deliverables especially if the data was significantly transformed in format, distribution, or transformation into a data standard.

State Geothermal Data Tracking site

Submissions were entered in the internal administration tracking system that associated the submission to the proposed deliverable from the SOW, accounting for the number of records submitted, file name, State, submission date, and comments as the files progressed through the review process. The data from the administration tracking system was also provided as a Drupal report interface at *http://repository.stategeothermaldata.org/track/report/xx (xx being the state abbreviation)* for sub-recipients to check their submissions against each year's proposed deliverables and provides a summary report of the QA/QC data process and the submission status in the review process at any given time. This gives the state data providers an online location to track submissions as they go through review and processing and enter the NGDS. For each state, the deliverables items are categorized by project Year SOWs with associated submissions listed beneath the SOW item, date submitted, status and number of records. These submissions and deliverables are summarized by data type at the top of the tracking report page showing percent complete, number of records submitted, and other information. (these final totals per state are also indicated in Task 2.4 (the Supplemental project for NEW data below). All online data URLs for file downloads and WMS services are tracked in a window in the upper right side of the tracking report. These links direct the provider to the repository data locations and WMS services. The code base for this is available at *https://github.com/usgin/aasgtrack*.

AASG Geothermal Data Repository database

The AASG Geothermal Data Repository (repository.stategeothermaldata.org) uses node.js to communicate to the Django-driven user interface component, and uses CouchDB (a free, open-source document database) to store metadata records as JSON documents. The code base for this system is online at *https://github.com/usgin/metadata-server*.

The administration site, available to users with logins at *http://repository.stategeothermaldata.org/admin/*, is a site used for internal tracking of state submissions and deliverables. Specific deliverables from the SOWs are entered here to create lists just as those in the State Geothermal Data submission site. Once users upload files there, the processing team puts the file on a local network server in specific directories and makes an entry into this administration site. The administrator indicates the date submitted, the status (submitted, in review, returned, approved, online), the deliverable from the SOW list, the number of records, and any comments. Once the file is tracked, a task is created and a data processing team member is assigned.

Model Manager

https://github.com/usgin/modelmanager - Django app used to manage content models which are available to the public at schemas.usgin.org. These are uploaded here after adoption from the community development GitHub repositories at *https://github.com/usgin-models*.

Data replication process documentation

Documentation on the Hyper Vreplication process was provided to hubs and presented in the 2013 Western Regional Workshop for state sub-recipients, and is available at *https://github.com/usgin/-Workshops/blob/master/WesternRegionalWorkshop/Hyper-V%202012%20Replication%20on%20Work-group%20Servers.docx*.

Section 7: Computer Modeling

No computer models were generated or run as part of this project.

Section 8: References

Anonymous, 2007-01-09, NIEM Concept of Operations Version 0.5: accessed at http://reference.niem.gov/niem/guidance/concept-of-operations/0.5/

USGIN, 2011a, Metadata Recommendations for Geoscience Resources: U.S. Geoscience Information Network, document USGIN2011-002 accessible at http://repository.usgin.org/uri_gin/usgin/dlio/335.

USGIN, 2011b, Use of ISO metadata specifications to describe geoscience information resources: U.S Geoscience In-formation Network, document USGIN2010-09, accessible at http://repository.usgin.org/uri_gin/usgin/dlio/337.

van den Brink, Linda, Portele, Clemens, Vretanos, Panagiotis A., editors, 2011-05-11, Geography Markup Language (GML) simple features profile (with technical note) v2.0: Open Geospatial Consortium Inc., Document 10-100r3.

Appendix A: Overview of State Contributions to NGDS

1. List of Participants and Inventory of Contributions
    a. State data submissions by state (Excel spreadsheet): submitted separately as Excel file through E-link 241.3 (report number- DOE-AZGS-2850-5). Also available online at https://github.com/ngds/documents/blob/master/AASG-State-Contributions/A1-a-State-submissions_by_date.xlsx?raw=true
    b. Data Contributions by services (spreadsheet) – submitted separately as Excel file through E-link 241.3 (report number- DOE-AZGS-2850-6). Also available online at https://github.com/ngds/documents/blob/master/AASG-State-Contributions/A1-b-State_contributions.xlsx?raw=true
    c. Data Contribution Summary of work accomplished by subcontractors……2
2. Report on Sub-contractor Collaborations
    a. Original letters of commitment……51
    b. Comparison spreadsheet: stated collaborations with actual collaborations ..119
    c. State Responses……..121
3. Work performed against accepted Statements of Work (State Submissions Repository Site: https://github.com/usgin/aasgtrack) - submitted to OSTI as software

State Data Contribution Summary

This appendix summarizes the data contributions from all state sub-recipients. The Statements of Work (SOW) for each acquisition cycle were closely monitored and compared with progress from the states. As states begin acquiring and compiling the actual data, the scope of work became better defined; consequently, some states failed to deliver 100% of their initial data assessments as indicated in their SOWs due to data unavailability, initial overestimates, or other issues associated with locating and digitizing data. Just as often as finding a lack of originally indicated data, states were reporting additional data which could be added to their SOWs. In this case the SOWs were consequently modified along with contracts to better align actual deliverables with allocated funding, ensuring that all possible data available could be integrated into NGDS. Due to the contract modifications, no states were shown to deliver greater than 100% as indicated by their SOWs. Most states indeed did deliver exactly 100% of their deliverables. Those brief figures are outlined in the subtasks below.

Subtask 2.4 (A) Alaska: See general Subtask 2.4 description above.

The following table shows the final deliverables from Alaska. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 17,595 records in the NGDS, inclusive of data and metadata for other resources.

Bottom hole temperature YR1	Temperature & Heat flow
Document - metadata for CMs YR1	Metadata and Documents
Hot spring description YR1	Temperature & Heat flow
Borehole Physical Samples YR2	Other
Document metadata OSTI YR2	Metadata and Documents
Metadata - Datasets YR2	Metadata and Documents
Temperature Depth Log YR2	Temperature & Heat flow
Volcanic Vents YR2	Recent Tectonics
Active Faults YR3	Recent Tectonics
Aqueous chemistry YR3	Water Chemistry
Borehole Physical Samples YR3	Other
Earthquake Hypocenter YR3	Recent Tectonics
Extra: Direct Use Facilities	Metadata and Documents
Geothermal Resources Map YR3	Map Data
Metadata for CMs YR3	Metadata and Documents

Pilgrim HS MT Report	Metadata and Documents

Subtask 2.4 (B) Alabama: See general Subtask 2.4 description above.

The following table shows the final deliverables from Alabama. They completed 93.75% of the deliverables as initially outlined in their original SOWs, with 54,763 records in the NGDS, inclusive of data and metadata for other resources.

Core analysis data YR1	Well Logs
Develop GIS Data/Metadata YR1	Metadata and Documents
Digital Well Logs YR1	Well Logs
Earthquake epicenter YR1	Recent Tectonics
Fault information YR1	Recent Tectonics
Formation Temperatures YR1	Temperature & Heat flow
Geochemical Borehole Fluid data YR1	Water Chemistry
References YR1	Metadata and Documents
Fault Information YR2	Recent Tectonics
Formation Temperatures YR2	Temperature & Heat flow
Geochemical data YR2	Water Chemistry
Core and Sample Analysis Data YR3	Rock Chemistry
Formation Temperatures YR3	Temperature & Heat flow
Production Information YR3	Well Logs
Quad Maps YR3	Map Data
State Geologic Map YR3	Map Data

Subtask 2.4 (C) Arkansas: See general Subtask 2.4 description above.

The following table shows the final deliverables from Arkansas. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 6,845 records in the NGDS, inclusive of data and metadata for other resources.

Active Fault Map YR1	Recent Tectonics
Borehole lithology log dataset YR1	Well Logs
Bottom hole temperature YR1	Temperature & Heat flow
Digital well log YR1	Well Logs
Document - AGS Pubs YR1	Metadata and Documents
Document - metadata for CMs YR1	Metadata and Documents
Earthquake Hypocenter YR1	Recent Tectonics
Temperature depth log YR1	Well Logs
Active Fault Map YR2	Recent Tectonics
Borehole Lithology Log Dataset YR2	Well Logs
Bottom Hole Temperature YR2	Temperature & Heat flow
Digital Well Logs YR2	Well Logs
Document - metadata for CMs YR2	Metadata and Documents
Documents - AGS YR2	Metadata and Documents
Earthquake Hypocenter YR2	Recent Tectonics
Temperature Depth Log YR2	Temperature & Heat flow
YR2 Thermal Conductivity	Temperature & Heat flow
Aeromagnetic Survey Maps YR3	Map Data
Brine Chem info South AR YR3	Rock Chemistry
Conduct Sample Calibration YR3	Rock Chemistry
Geologic maps YR3	Map Data
Geothermal Gradient Map YR3	Map Data
Heat Flow Map YR3	Map Data

Subtask 2.4 (D) California: See general Subtask 2.4 description above.

The following table shows the final deliverables from California. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 79,068 records in the NGDS, inclusive of data and metadata for other resources.

Active faults YR1	Recent Tectonics
Active Volcanic Vents YR1	Recent Tectonics
Bottomhole Temperature YR1	Temperature & Heat flow
Documents - CGS YR1	Metadata and Documents
DOGGR Geosteam Data YR1	Temperature & Heat flow
Geothermal Energy Documents YR1	Metadata and Documents
Geothermal Fields YR1	Map Data
Hot Springs Features YR1	Temperature & Heat flow
O&G Borelogs YR1	Well Logs
State Map / Geothermal Data YR1	Map Data
Status Idle Wells YR1	Water Chemistry
Temp Grads Sediment Basin YR1	Metadata and Documents
Thermal Aqueous Chemistry YR1	Water Chemistry
Thermal Wells - DOGGR YR1	Well Logs
BHT - Oil and Gas Well Data YR2	Temperature & Heat flow
Docs Research and References YR2	Metadata and Documents
Docs Scans New GeoT Data YR2	Metadata and Documents
Geological Maps YR2	Map Data
Geological / Regional Maps YR2	Map Data
Volcano-hydrothermal systems YR2	Metadata and Documents
BHT O&G Data YR3	Temperature & Heat flow
CA Geothermal Power Plants YR3	Other

California GEA Documents YR3	Metadata and Documents
Production and Injection Data YR3	Well Logs
Well Logs O&G	Well Logs

Subtask 2.4 (E) Delaware: See general Subtask 2.4 description above.

The following table shows the final deliverables from Delaware. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 34,274 records in the NGDS, inclusive of data and metadata for other resources.

Bottom hole temperature YR1	Temperature & Heat flow
Digital Well Logs, geothermal test wells YR1	Well Logs
Digital well log, water wells YR1	Well Logs
Document - Metadata for CMs YR1	Metadata and Documents
Document - research VTech YR1	Metadata and Documents
Geologic maps YR1	Map Data
Geologic Unit -geot-charac-YR1	Metadata and Documents
Heat flow measurement YR1	Temperature & Heat flow
Lith Interval Logs-Geot Wells YR1	Well Logs
Lith interval logs-water wells YR1	Well Logs
Temperature depth log YR1	Well Logs
Thermal Conductivity YR1	Temperature & Heat flow
Bottom Hole Temperature YR2	Temperature & Heat flow
Digital Well Log, Water Wells YR2	Well Logs
Document - Research VTech YR2	Metadata and Documents
Geologic Maps DE YR2	Map Data
Geologic Unit -geot-charac-YR2	Metadata and Documents

Heat Flow Measurement YR2	Temperature & Heat flow
Lith Interval Logs-Geot Wells YR2	Well Logs
Lith interval logs-water wells YR2	Well Logs
Temperature Depth Logs YR2	Well Logs
Thermal Conductivity YR2	Temperature & Heat flow
Well Log Water Wells (Docs) YR2	Metadata and Documents
BHT Observations YR3	Temperature & Heat flow
Docs-VA Tech 1969-1989 YR3	Metadata and Documents
Digital Well LOgs O&G wells YR3	Well Logs
Digital Well Log, Water Wells YR3	Well Logs
Geologic Map modern/historic YR3	Map Data
Lithology logs, water wells YR3	Temperature & Heat flow

Subtask 2.4 (F) Florida: See general Subtask 2.4 description above.

The following table shows the final deliverables from Florida. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 8,605 records in the NGDS, inclusive of data and metadata for other resources.

Borehole Data-artesian wells YR1	Metadata and Documents
Doc-scanned log metadata YR1	Metadata and Documents
Docs-Temp Log LAS Files YR1	Temperature & Heat flow
Hot springs YR1	Temperature & Heat flow
O&G Borehole Header Features YR1	Well Logs
O&G Bottomhole Temp YR1	Well Logs
UIC Wells Borehole Data YR1	Well Logs
FLGS BHT YR2	Temperature & Heat flow
Geology Map Service	Map Data

Metadata O&G and State BHT YR2	Metadata and Documents
Metadata-O&G geophysl logs YR2	Metadata and Documents
Metadata UIC Geophysical Logs YR2	Metadata and Documents
Metadata-WMD Geophysl Logs YR2	Metadata and Documents
UIC Borehole Temperatures YR2	Temperature & Heat flow
WMD Borehole Temperatures YR2	Temperature & Heat flow
WMD Geophysical Logs YR2	Well Logs
Oil and gas database BHT data YR3	Temperature & Heat flow
Oil and Gas Database Metadata YR3	Metadata and Documents
Scanned UIC Logs Metadata YR3	Metadata and Documents
UIC Borehole Temperatures YR3	Temperature & Heat flow
Water Management District BHT YR3	Temperature & Heat flow

Subtask 2.4 (G) Hawaii: See general Subtask 2.4 description above.

The following table shows the final deliverables from Hawaii. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 12,115 records in the NGDS, inclusive of data and metadata for other resources.

Document - Metadata (reports) YR1	Metadata and Documents
Geologic map YR1	Map Data
Hot Springs Features YR1	Temperature & Heat flow
Temperature observation YR1	Temperature & Heat flow
Volcanic rift systems YR1	Map Data
Well Header and Logs YR1-EXTRA SUBMISSION	Well Logs
Docs-Well Drill Records YR2	Metadata and Documents
Metadata (cored intervals) YR2	Metadata and Documents

Metadata (reports) YR2	Metadata and Documents
Temperature Observation YR2	Temperature & Heat flow
Water Chemistry Observation YR2	Water Chemistry
Bibliogrphic Data YR3	Metadata and Documents
Geothermal Reports YR3	Metadata and Documents
Well Data YR3	Well Logs

Subtask 2.4 (H) Illinois: See general Subtask 2.4 description above.

The following table shows the final deliverables from Illinois. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 157,231 records in the NGDS, inclusive of data and metadata for other resources.

Bottom hole temperature YR1	Temperature & Heat flow
Document - metadata for CMs YR1	Metadata and Documents
Documents - References YR1	Metadata and Documents
Drill stem test YR1	Temperature & Heat flow
Temperature depth log YR1	Well Logs
Bottom hole temperature YR2	Temperature & Heat flow
Drill Stem Test YR2	Temperature & Heat flow
Bottom Hole Temperatures YR3	Temperature & Heat flow
Drill Stem Test YR3	Other

Subtask 2.4 (I) Indiana: See general Subtask 2.4 description above.

The following table shows the final deliverables from Indiana. They completed 97.44% of the deliverables as initially outlined in their original SOWs, with 45,944 records in the NGDS, inclusive of data and metadata for other resources.

Bedrock Geology Map 1:500K YR1	Map Data
BHT (Petroleum Logs) YR1	Temperature & Heat flow

Bottomhole Temperature YR1	Temperature & Heat flow
Geol Map Elevation Tops Fms YR1	Map Data
Geologic map Bedrck Elevation YR1	Map Data
Geologic map Bedrock 1:24K YR1	Map Data
Geologic map (Core Samples) YR1	Map Data
Geologic map Depth to Bedrock YR1	Map Data
Geologic map (Petroleum Logs) YR1	Map Data
Geologic map Soil Characteris YR1	Map Data
Geologic map (Well Samples) YR1	Map Data
Geophysical Data Index Maps YR1	Map Data
Permeability YR1	Other
Seismic Depth to Bedrck Index YR1	Map Data
Water Chemistry TDS YR1	Map Data
Water Chemistry UCMs YR1	Map Data
Water Source Feature UCMs YR1	Map Data
Water Source Mineral Quarries YR1	Map Data
Develop GIS Data/Metadata YR2	Metadata and Documents
Geologic Map Bedrock topo YR2	Map Data
Geologic Map depth-to-bedrock YR2	Map Data
Geologic Map IN YR2	Map Data
Geologic Unit GeoT Character YR2	Other
Geothermal System Feature YR2	Other
Bedrock Geology Maps YR3	Map Data
Bedrock Wells Aqueous Chem YR3	Water Chemistry

Bottom Hole Temperatures YR3	Temperature & Heat flow
Depth-to-bedrock Map YR3	Map Data
Digital Elevation Map YR3	Map Data
Geologic map YR3	Map Data
Geothermal Map YR3	Map Data
Temp-depth logs YR3	Well Logs
Unconsol. Aquifer Aq Chem YR3	Water Chemistry
Geologic Unit -In-situ Heat SUP	Map Data
Geologic Unit-physical property SUP	Map Data
Geologic Unit-vol. moisture SUP	Map Data
Temperature SUP	Temperature & Heat flow
Thermal conductivity SUP	Temperature & Heat flow
Thermal Diffusivity SUP	Temperature & Heat flow

Subtask 2.4 (J) Iowa: See general Subtask 2.4 description above.

The following table shows the final deliverables from Iowa. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 298,662 records in the NGDS, inclusive of data and metadata for other resources.

Bedrock depth Quat-raster YR1	Map Data
Bedrock depth Quat-vector YR1	Map Data
Bedrock Geologic map YR1	Map Data
Bedrock Surface Elevatn-raster YR1	Map Data
Bedrock Surface Elevatn-vector YR1	Map Data
Geologic Sampling Points YR1	Map Data
Geothermal system features YR1	Well Logs
Resource Suitability Map YR1	Map Data

First Aquifer (300 ft. depth) YR2	Well Logs
Geolgc Sample Pts update2011 YR2	Map Data
Geology Map Service	Map Data
GeoT systems update 2011 YR2	Well Logs
Karst probability YR2	Map Data
Known GW Contaminants YR2	Other
Known Underground Mining YR2	Other
Potential Well Yield (aquifers) YR2	Well Logs
Pumped GW Temperatures YR2	Temperature & Heat flow
Resource Suitability Map YR2	Map Data
Geologic Sampling Pts YR3	Map Data
Geothermal systems YR3	Other
Iowa Well Logs Year 3	Well Logs
Karst probability YR3	Map Data
Known GW contamin. YR3	Other
OneGeology map YR3	Map Data
Resource Suitability YR3	Map Data
Shallow well yield YR3	Other
Top aquifer YR3	Map Data
Underground mining YR3	Map Data

Subtask 2.4 (K) Kansas: See general Subtask 2.4 description above.

The following table shows the final deliverables from Kansas. They completed 95.24% of the deliverables as initially outlined in their original SOWs, with 2,217,730 records in the NGDS, inclusive of data and metadata for other resources.

BHT (Logs on File) YR1	Temperature & Heat flow

BHT (Scanned E-logs) YR1	Temperature & Heat flow
Digital well log (O&G wells LAS) YR1	Well Logs
Digital well O&G E-logs scans YR1	Well Logs
Document - metadata for CMs YR1	Metadata and Documents
Drill stem test - digital YR1	Temperature & Heat flow
DST (Scanned DTS Reports) YR1	Metadata and Documents
Geologic map (County Maps) YR1	Map Data
BHT Observation Service YR2	Other
Document - Metadata for CMs YR2	Metadata and Documents
Drill Stem Test, Western Kansas YR2	Temperature & Heat flow
DST (Digital DTS Reports) YR2	Temperature & Heat flow
Geologic map (County Maps) YR2	Map Data
Max Temp Data, East Kansas YR2	Temperature & Heat flow
Well Header Service YR2	Other
BHT Obs. Service YR3	Temperature & Heat flow
Drillers logs and reports YR3	Well Logs
Drill Stem Tests YR3	Temperature & Heat flow
Geologic map YR3	Map Data
KS Geothermal Direct Use Report	Metadata and Documents
Well Header Service YR3	Map Data

Subtask 2.4 (L) Kentucky: See general Subtask 2.4 description above.

The following table shows the final deliverables from Kentucky. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 1,633,259 records in the NGDS, inclusive of data and metadata for other resources.

BHT (Observation) YR1	Temperature & Heat flow

Digital Well Log (Coverage) YR1	Well Logs
Temp-Depth Log (Coverage) YR1	Well Logs
Temp Log image (Document) YR1	Well Logs
BHT (Observation) YR2	Temperature & Heat flow
Borehole Cores /Rock Cuttings YR2	Other
Borehole Lithology Log Dataset YR2	Well Logs
Digital Well Log (Coverage) YR2	Well Logs
Document - Geologic Reports YR2	Metadata and Documents
Drill Stem Tests (Document) YR2	Temperature & Heat flow
Geologic Map (digital) YR2	Map Data
Temp Log Image (Document) YR2	Well Logs
Whole-Rock Chemistry YR2	Rock Chemistry
3-D Stratigraphic Surface YR3	Map Data
Fault Features - 1:24,000 scale YR3	Recent Tectonics
Geologic Contacts 1:24K YR3	Map Data
Geologic Unit Features 1:24K YR3	Map Data
Geologic Units 1:500K YR3	Map Data
Oil and Gas Well Headers YR3	Well Logs
Stratigraphic Top of Formation YR3	Well Logs
Water Well Headers YR3	Well Logs

Subtask 2.4 (M) Louisiana: See general Subtask 2.4 description above.

The following table shows the final deliverables from Louisiana. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 16,515 records in the NGDS, inclusive of data and metadata for other resources.

Bottom hole temperature YR1	Temperature & Heat flow

Document - metadata for CMs YR1	Metadata and Documents
Documents - LGS YR1	Metadata and Documents
Gravity Map (Aero magnetic) YR1	Map Data
Gravity map (Bouguer) YR1	Map Data
Bottom Hole Temperature YR2	Temperature & Heat flow
Geologic Map of Louisiana YR2	Metadata and Documents
Metadata for LA Documents YR2	Metadata and Documents
Bottom Hole Temperatures YR3	Temperature & Heat flow
Map SW LA Max Fresh Water YR3	Map Data
Metadata YR3	Metadata and Documents
Oil and Gas Map of Louisiana YR3	Map Data

Subtask 2.4 (N) Maryland: See general Subtask 2.4 description above.

The following table shows the final deliverables from Maryland. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 9,738 records in the NGDS, inclusive of data and metadata for other resources.

BHT Geothermal test wells YR1	Temperature & Heat flow
BHT Other wells YR1	Temperature & Heat flow
Digital well log, GeoT wells YR1	Well Logs
Digital well log, O&G wells YR1	Well Logs
Digital well log, water wells YR1	Well Logs
Document - published reports YR1	Metadata and Documents
Geologic maps YR1	Map Data
Geologic Unit GeoT charact YR1	Metadata and Documents
Heat flow measurement YR1	Temperature & Heat flow
Lithology logs, GeoT wells YR1	Well Logs

Lithology logs, O&G Wells YR1	Well Logs
Lithology logs, water wells YR1	Well Logs
Temperature depth log YR1	Well Logs
Thermal conductivity YR1	Temperature & Heat flow
Thermal Springs Data YR1	Temperature & Heat flow
BHT Other Wells YR2	Temperature & Heat flow
Digital well log, GeoT wells YR2	Temperature & Heat flow
Digital Well Log, O&G Wells YR2	Well Logs
Digital Well Log, Water Wells YR2	Well Logs
Document - Published Reports YR2	Metadata and Documents
Geologic Maps MD YR2	Map Data
Geologic Unit GeoT charact YR2	Metadata and Documents
Gravity Map Data MD YR2	Map Data
Heat Flow Measurements YR2	Temperature & Heat flow
Lithology Logs, GeoT Wells YR2	Well Logs
Lithology logs, O&G Wells YR2	Well Logs
Lithology Logs, Water Wells YR2	Well Logs
Temperature Depth Log YR2	Well Logs
Thermal Conductivity YR2	Temperature & Heat flow
BHT Other wells YR3	Temperature & Heat flow
Digital Well Log, O&G Wells YR3	Well Logs
Digital Well Log, Water Wells YR3	Well Logs
Document 1969-189 YR3	Metadata and Documents
Intrusive Body with Heat YR3	Temperature & Heat flow

Lithology logs, O&G Wells YR3	Well Logs
Lithology logs, water wells YR3	Well Logs

Subtask 2.4 (O) Michigan: See general Subtask 2.4 description above.

The following table shows the final deliverables from Michigan. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 73,752 records in the NGDS, inclusive of data and metadata for other resources.

Bibliographic Metadata YR1	Metadata and Documents
Bottom Hole Temperature YR1	Temperature & Heat flow
Drill stem test YR1	Temperature & Heat flow
Formation water chemistry YR1	Water Chemistry
Metadata CM's YR1	Metadata and Documents
Temperature Depth Log YR1	Well Logs
Bibliographic Metadata YR2	Metadata and Documents
Bottom Hole Temperatures YR2	Temperature & Heat flow
Drill Stem Tests YR2	Temperature & Heat flow
Formation Water Chem Data YR2	Water Chemistry
Geology Map Service	Map Data
Lithology Interval Data YR2	Well Logs
Metadata CM's YR2	Metadata and Documents
Well Header Data YR2	Well Logs
Bottom Hole Temperatures YR3	Temperature & Heat flow
Direct Use GeoT Installs YR3	Other
Drill Stem Test YR3	Temperature & Heat flow
Lithology Interval Data YR3	Temperature & Heat flow
Shallow Water Well Lithology YR3	Water Chemistry

Well Header Data YR3	Well Logs

Subtask 2.4 (P) Minnesota: See general Subtask 2.4 description above.

The following table shows the final deliverables from Minnesota. They completed 83.33% of the deliverables as initially outlined in their original SOWs, with 1,477,724 records in the NGDS, inclusive of data and metadata for other resources.

Borehole lithology log dataset YR1	Well Logs
Document - metadata for CMs YR1	Metadata and Documents
Heat flow measurement YR1	Temperature & Heat flow
Borehole Lithology Log Dataset YR2	Well Logs
Geologic Map - 3D YR2	Map Data
Heat Flow Measurement YR2	Temperature & Heat flow

Subtask 2.4 (Q) Mississippi: See general Subtask 2.4 description above.

The following table shows the final deliverables from Mississippi. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 10,028 records in the NGDS, inclusive of data and metadata for other resources.

Bottom Hole Temperatures YR1	Temperature & Heat flow
Document - Metadata for CM's	Metadata and Documents
Document (OF1 report) YR1	Metadata and Documents
Map publications (OF1 maps) YR1	Map Data
Oil and Gas Field Map YR1	Map Data
Water Coprod O&G Production YR1	Other
Bottom Hole Temperatures YR2	Temperature & Heat flow
Bottom Hole Temperatures YR3	Temperature & Heat flow
Geothermal Website/Database YR2	Metadata and Documents
Geothermal Website YR3	Other

Subtask 2.4 (R) Nebraska: See general Subtask 2.4 description above.

The following table shows the final deliverables from Nebraska. They completed 88.89% of the deliverables as initially outlined in their original SOWs, with 18,413 records in the NGDS, inclusive of data and metadata for other resources.

Bottom Hole Temperatures YR2	Temperature & Heat flow
Geothermal Map YR2	Map Data
Heat Flow from BHT Logs YR2	Temperature & Heat flow
Lithologic Sections of Heat Flow Sites YR2	Well Logs
Lithologic Units YR2	Well Logs
Temperature Depth Log YR2	Temperature & Heat flow
Temperature Depth Plots YR2	Temperature & Heat flow
Bottom Hole Temperatures YR3	Temperature & Heat flow
Heat Flow Data YR3	Temperature & Heat flow

Subtask 2.4 (S) New Hampshire: See general Subtask 2.4 description above.

The following table shows the final deliverables from New Hampshire. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 16,128 records in the NGDS, inclusive of data and metadata for other resources.

Bottom Hole Temperature YR1	Temperature & Heat flow
Geologic Map (1:24,000 scale) YR1	Map Data
Geologic Map (1:250,000 scale) YR1	Map Data
Geologic map SE NH YR1	Map Data
Geologic Unit GeoT Charact YR1	Metadata and Documents
Metadata - Subsurface Borings YR1	Metadata and Documents
Metadata Unweathered Rock YR1	Metadata and Documents
Metadata Well Compl Reports YR1	Metadata and Documents
Temperature Measurement YR1	Temperature & Heat flow

Whole-rock Chemistry YR1	Rock Chemistry
Whole-rock Chem Samples YR1	Rock Chemistry
Geologic map Central NH YR2	Map Data
Geologic Unit GeoT Charact YR2	Metadata and Documents
Geology Map Service YR2	Map Data
Well Logs YR2	Well Logs
Aeromag Maps YR3	Map Data
Bedrock Geologic map YR3	Map Data
BHT and Profiles YR3	Temperature & Heat flow
Geologic Map GIS YR3	Map Data
Gravity Map YR3	Map Data
Petrology Rock Chem YR3	Rock Chemistry
Spring Temp & Flow YR3	Temperature & Heat flow
Thermal Conductivity YR3	Temperature & Heat flow
Whole-rock Chemistry YR3	Rock Chemistry

Subtask 2.4 (T) New York: See general Subtask 2.4 description above.

The following table shows the final deliverables from New York. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 286,302 records in the NGDS, inclusive of data and metadata for other resources.

Bottom hole temperature YR1	Temperature & Heat flow
Earthquake Hypocenter YR1	Recent Tectonics
Fault, Fracture, Lineament Map YR1	Map Data
Geologic Unit Description YR1	Map Data
Isopach-Structure Contour Maps YR1	Map Data
Rock samples - Permeability YR1	Rock Chemistry

Temperature depth log YR1	Temperature & Heat flow
Bottom Hole Temperature YR2	Temperature & Heat flow
Digital Well Log YR2	Metadata and Documents
Earthquake Epicenter YR2	Recent Tectonics
Fault, Fracture, Lineament Map YR2	Map Data
Geologic Unit Description YR2	Map Data
Isopach-Structure Contour Maps YR2	Map Data
Temperature Depth Log YR2	Temperature & Heat flow
Core Inventory YR3	Rock Chemistry
Cuttings Inventory YR3	Rock Chemistry
Data Mining Previous Studies YR3	Metadata and Documents
Production Data YR3	Well Logs
Rock samples - Permeability YR3	Rock Chemistry
Temperature Depth Log YR3	Temperature & Heat flow

Subtask 2.4 (U) North Carloina: See general Subtask 2.4 description above.

The following table shows the final deliverables from North Carolina. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 9,472 records in the NGDS, inclusive of data and metadata for other resources.

Documents (reports) YR3	Metadata and Documents
Borehole lith logs GeoT test wells YR3	Well Logs
Borehole lith logs, O&G wells YR3	Well Logs
Borehole lith logs, water wells YR3	Well Logs
Bottom hole temperature YR3	Temperature & Heat flow
Digital well log, O&G wells YR3	Well Logs
Digital Well Logs, GeoT Test wells YR3	Well Logs

Digital well log, water wells YR3	Well Logs
Geologic map YR3	Map Data
Geologic Unit YR3	Map Data
Gravity map data YR3	Map Data
Gravity Station data YR3	Map Data
Hot Spring Description YR3	Temperature & Heat flow
Intrusive body with heat YR3	Temperature & Heat flow
Thermal conductivity YR3	Temperature & Heat flow
Whole-rock Chemistry YR3	Rock Chemistry

Subtask 2.4 (V) North Dakota: See general Subtask 2.4 description above.

The following table shows the final deliverables from North Dakota. They completed 75% of the deliverables as initially outlined in their original SOWs, with 15,713 records in the NGDS, inclusive of data and metadata for other resources.

Bottom Hole Temperatures YR2	Temperature & Heat flow
Drill Stem Test YR2	Temperature & Heat flow
Geothermal Map YR2	Metadata and Documents
Temperature Depth Log YR2	Temperature & Heat flow
Thermal Conductivity Data YR2	Temperature & Heat flow
Thermal Stratigraphy YR2	Temperature & Heat flow
Borehole Temperature Data YR3	Temperature & Heat flow
Heat Flow Data YR3	Temperature & Heat flow

Subtask 2.4 (W) Ohio: See general Subtask 2.4 description above.

The following table shows the final deliverables from Ohio. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 13,651 records in the NGDS, inclusive of data and metadata for other resources.

Bottom Hole Temperature YR1	Temperature & Heat flow
Digital Well Logs YR1	Well Logs
Document - metadata for CMs YR1	Metadata and Documents
Earthquake epicenter YR1	Recent Tectonics
Geologic maps YR1	Map Data
Geologic unit feature YR1	Map Data
Lithology Log WellHeaders YR1	Well Logs
Preliminary Geothermal Map YR1	Map Data
Temperature Depth Log YR1	Well Logs
AUMs - GeoT Reservoirs YR2	Temperature & Heat flow
Borehole Temperatures YR2	Temperature & Heat flow
Geology Map Sevrice YR2	Map Data
Gravity and magnetic maps YR2	Map Data
GW Source GeoT Energy Prodct YR2	Map Data
Heat Flow model YR2	Temperature & Heat flow
Map of deep, subsurface faults YR2	Map Data
Metadata for CMs and Maps	Metadata and Documents
Thermal Conductivity YR2	Temperature & Heat flow
Abandoned UnderGrnd Mines YR3	Other
Abnd UnderGrd Mines IMS YR3	Map Data
Final Report	Metadata and Documents
Geochem from deep wells YR3	Water Chemistry
GW Inventory AqGeochem YR3	Water Chemistry
GW temperature map YR3	Map Data

Known and Probable Karst YR3	Map Data
ODNR Water Well Header YR3	Well Logs
ODNR Water Well Log YR3	Well Logs
Potentiometric Surface maps YR3	Map Data
Quaternary Geologic Map YR3	Map Data
Revised gravity map YR3	Map Data
Revised magnetic maps YR3	Map Data
Shaded Bedrock Topo Map YR3	Map Data
Shaded Drift Thickness Map YR3	Map Data
State Aquifer Maps YR3	Map Data
Surficial Geology Stack Maps YR3	Map Data

Subtask 2.4 (X) Rhode Island: See general Subtask 2.4 description above.

The following table shows the final deliverables from Rhode Island. They completed 93.75% of the deliverables as initially outlined in their original SOWs, with 8,789 records in the NGDS, inclusive of data and metadata for other resources.

Borehole lithology logs YR1	Metadata and Documents
Digital boring logs - lithology YR1	Metadata and Documents
Digital well logs -water wells YR1	Well Logs
Document - metadata for CMs YR1	Metadata and Documents
Geologic map - bedrock YR1	Map Data
Geologic map Surficial 1:24K YR1	Map Data
Geologic map Surficial 1:31K YR1	Map Data
Geologic map - water resources YR1	Map Data
Borehole Lithology Log Dataset YR2	Well Logs
Digital Boring Log, Lithology YR2	Well Logs

Digital Well Log, Water Wells YR2	Well Logs
Geo & Aeromag Maps YR2	Map Data
Geologic Map Digital Bedrock YR2	Map Data
Geologic map-Quat surficial YR2	Map Data
Geologic maps Bedrock 1:100K YR2	Map Data
Geologic Map simplified 1:24K YR2	Map Data
Wells GeoT Heat Pump Ground YR2	Well Logs
Bedrock Map 100K, 125K YR3	Map Data
Bedrock Map, 24, 31, 100K - YR3	Map Data
Borehole Lith Interval YR3	Metadata and Documents
Borehole Lith Log YR3	Temperature & Heat flow
Digital Lith log YR3	Well Logs
Digital water well log YR3	Well Logs
Geochem and structure data YR3	Rock Chemistry
Geology/Aeromag Maps YR3	Map Data
Groundwater Map 24K - YR3	Map Data
GSHP wells YR3	Well Logs
Heat flow data YR3	Temperature & Heat flow
Radiometric Age Bedrock YR3	Map Data
RI and MA co-op Field Data YR3	Other
Scanned Groundwater Map 24K - YR3	Metadata and Documents
Surficial Map, 24, 31, 100K - YR3	Map Data

Subtask 2.4 (Y) South Carolina: See general Subtask 2.4 description above.

The following table shows the final deliverables from South Carolina. They completed 84.62% of the deliverables as initially outlined in their original SOWs, with 6,113 records in the NGDS, inclusive of data and metadata for other resources.

Bottom hole temperature YR1	Temperature & Heat flow
Document - metadata for CMs YR1	Metadata and Documents
Geophysical Well Logs YR1	Well Logs
Borehole Data YR2	Well Logs
Borehole Temperatures YR2	Temperature & Heat flow
Document -VPI YR2	Metadata and Documents
Earthquake Hypocenter YR2	Recent Tectonics
Basement Surface Contour Map YR3	Map Data
Bouguer Gravity Data YR3	Map Data
CO2 Sequestration Prjt YR3	Other
Earthquakes YR3	Recent Tectonics
Geologic maps YR3	Map Data
Vectorized Aeromag Map YR3	Map Data

Subtask 2.4 (Z) South Dakota: See general Subtask 2.4 description above.

The following table shows the final deliverables from South Dakota. They completed 83.33% of the deliverables as initially outlined in their original SOWs, with 4,050 records in the NGDS, inclusive of data and metadata for other resources.

Bottom hole temperature YR1	Temperature & Heat flow
Geologic maps YR1	Map Data
Documents YR2	Metadata and Documents
Geologic map YR2	Map Data
Temperature YR2	Temperature & Heat flow
Flow Rate YR2	Temperature & Heat flow

Subtask 2.4 (AA) Tennessee: See general Subtask 2.4 description above.

The following table shows the final deliverables from Tennessee. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 12,219 records in the NGDS, inclusive of data and metadata for other resources.

Digital Well Logs YR1	Well Logs
Doc-GeoT Installation Reports YR1	Metadata and Documents
Document - DOE ORNL Report YR1	Metadata and Documents
Document - metadata for CMs YR1	Metadata and Documents
Thermal Conductivity TVA YR1	Temperature & Heat flow
Bottom Hole Temperature YR2	Temperature & Heat flow
Doc-GeoT Installation Reports YR2	Metadata and Documents
Hot Spring Data YR2	Temperature & Heat flow
Temp-Depth Digital Log YR2	Temperature & Heat flow
Temperature Depth Log Well Logs YR2	Well Logs
Well Header Data YR2	Well Logs
Bedrock Core Samples YR3	Rock Chemistry
Bottom Hole Temperature YR3	Temperature & Heat flow
Depth to Bedrock Map YR3	Map Data
Geologic Map 1:24K YR3	Map Data
Geo-T Installation Document YR3	Metadata and Documents
Hot Spring Data YR3	Temperature & Heat flow
Temp-Depth Log (well header) YR3	Well Logs
Temp-Depth Well Logs YR3	Well Logs
Well Cutting Samples YR3	Rock Chemistry

Subtask 2.4 (AB) Texas: See general Subtask 2.4 description above.

The following table shows the final deliverables from Texas. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 488,090 records in the NGDS, inclusive of data and metadata for other resources.

Borehole lithology log dataset YR1	Well Logs
Bottom Hole Temperature YR1	Temperature & Heat flow
Digital well log YR1	Well Logs
Document Geothermal Energy YR1	Metadata and Documents
Existing BHT Dataset YR2	Temperature & Heat flow
Expanding BHT Dataset YR2	Temperature & Heat flow
Well Log Observation Data YR2	Well Logs
Documents - GeoThermal YR3	Metadata and Documents
Exisiting BHT temp dataset YR3	Temperature & Heat flow
Expanding BHT temp dataset YR3	Temperature & Heat flow
Reservoir Characteristics YR3	Temperature & Heat flow
TX Active Faults -ADD	Recent Tectonics
Well Log Metadata YR3	Well Logs

Subtask 2.4 (AC) Virginia: See general Subtask 2.4 description above.

The following table shows the final deliverables from Virginia. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 44,870 records in the NGDS, inclusive of data and metadata for other resources.

Bottom hole temperature YR1	Temperature & Heat flow
Digital well log, GeoT wells YR1	Metadata and Documents
Digital well log, O&G wells YR1	Metadata and Documents
Digital well log, water wells YR1	Metadata and Documents
Document - metadata for CMs YR1	Metadata and Documents
Doc - VA Tech publish reports YR1	Metadata and Documents

Geologic maps YR1	Map Data
Geologic Units YR1	Metadata and Documents
Heat flow measurement YR1	Temperature & Heat flow
Hot Spring Description YR1	Temperature & Heat flow
Lithology logs GeoT wells YR1	Well Logs
Lithology logs O&G wells YR1	Well Logs
Lithology logs, water wells YR1	Well Logs
Temperature depth log YR1	Well Logs
Thermal Conductivity YR1	Temperature & Heat flow
Bottom Hole Temperature YR2	Temperature & Heat flow
Digital well log, GeoT wells YR2	Metadata and Documents
Digital well log, O&G wells YR2	Metadata and Documents
Digital Well Log, Water Wells YR2	Metadata and Documents
Geologic Units Radioactive YR2	Temperature & Heat flow
Gravity map data YR2	Map Data
Heat Flow Measurement YR2	Temperature & Heat flow
Hot Spring Description YR2	Temperature & Heat flow
Lithology Logs GeoT Wells YR2	Well Logs
Lithology Logs, O&G Wells YR2	Well Logs
Lithology logs, water wells YR2	Well Logs
Temperature Depth Logs YR2	Temperature & Heat flow
Thermal Conductivity YR2	Temperature & Heat flow
Whole rock Chemistry YR2	Rock Chemistry
Bottom Hole Temperature YR3	Temperature & Heat flow

Digital well log, GeoT wells YR3	Well Logs
Digital Well Log, O&G Wells YR3	Well Logs
Digital Well Log, Water Wells YR3	Well Logs
Document (Various Vtech) YR3	Metadata and Documents
Geologic Map (1:24,000) YR3	Map Data
Geothermal Map YR3	Well Logs
Heat Pump Facility YR3	Other
Hot Spring Description YR3	Temperature & Heat flow
Lithology logs GeoT wells YR3	Temperature & Heat flow
Lithology logs, O&G Wells YR3	Temperature & Heat flow
Lithology logs, water wells YR3	Temperature & Heat flow
Resource Suitability Map YR3	Map Data

Subtask 2.4 (AD) Wyoming: See general Subtask 2.4 description above.

The following table shows the final deliverables from Wyoming. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 19,067 records in the NGDS, inclusive of data and metadata for other resources.

Basin GeoT Resource Reports YR1	Temperature & Heat flow
Bottom Hole Temperature YR1	Temperature & Heat flow
Hot Springs/Water Chemistry YR1	Temperature & Heat flow
Bottom Hole Temperature YR2	Temperature & Heat flow
Bottom Hole Temperature YR3	Temperature & Heat flow
Digital maps with BHT	Map Data
Digitize GeoT Basin Studies YR3	Metadata and Documents
Metadata YR3	Metadata and Documents

Subtask 2.4 (AE) Wisconsin: See general Subtask 2.4 description above.

The following table shows the final deliverables from Wisconsin. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 1616 records in the NGDS, inclusive of data and metadata for other resources.

Borehole lithology log dataset YR1	Well Logs
Bottomhole Temperature YR1	Temperature & Heat flow
Crustal Stress data YR1	Recent Tectonics
Digital Well Logs YR1	Well Logs
Documents - Pub YR1	Metadata and Documents
Permeability YR1	Other
Borehole lithology log dataset YR2	Well Logs
Bottomhole Temperature YR2	Temperature & Heat flow
Digital Well Logs YR2	Well Logs
Permeability YR2	Rock Chemistry
Rock Chemistry YR2	Rock Chemistry
Borehole Lithology Log Dataset YR3	Temperature & Heat flow
Bottom Hole Temperature YR3	Temperature & Heat flow
Digital Well Log YR3	Well Logs
Permeability YR3	Rock Chemistry
Rock Chemistry YR3	Rock Chemistry

Subtask 2.4 (AF) New Mexico: See general Subtask 2.4 description above.

The following table shows the final deliverables from New Mexico. They completed 96.43% of the deliverables as initially outlined in their original SOWs, with 163,164 records in the NGDS, inclusive of data and metadata for other resources.

Active Fault Map YR1	Map Data
Borehole Lithology YR1	Metadata and Documents
Bottom hole temperature YR1	Temperature & Heat flow

Temperature-depth logs YR1	Temperature & Heat flow
Thermal conductivity data YR1	Temperature & Heat flow
Volcanic centers map YR1	Map Data
Aquifer-Reservoir Hydr Data YR2	Water Chemistry
BHT NMBGMR O&G Records YR2	Temperature & Heat flow
Borehole Lithology YR2	Well Logs
Digital NMBGMR O&G Records YR2	Temperature & Heat flow
GeoT Electric Geophy Studies YR2	Other
Hot Spring Description YR2	Temperature & Heat flow
NMBGMR O&G Porosity-Perm YR2	Rock Chemistry
Scan geophysical well logs - YR2	Well Logs
Well-Spring Temperature YR2	Temperature & Heat flow
Additional Data: One Geology Map Data	Map Data
BHT from NMBGMR YR3	Temperature & Heat flow
Data Logger temps YR3	Well Logs
Drill Stem tests Petro wells YR3	Well Logs
Geocronology volc ctrs YR3	Other
Lith Logs YR3 NMBGMR	Well Logs
Temp-depth logs YR3	Well Logs
Thermal Conductivity YR3	Temperature & Heat flow
Well/spring BHT YR3	Temperature & Heat flow
Well/spring Chemistry YR3	Water Chemistry
40Ar/39Ar Ages SUP	Rock Chemistry
Geothermal Waters SUP	Water Chemistry

Temperature Depth Logs SUP	Temperature & Heat flow

Subtask 2.4 (AG) Washington: See general Subtask 2.4 description above.

The following table shows the final deliverables from Washington. They completed 96.55% of the deliverables as initially outlined in their original SOWs, with 863,869 records in the NGDS, inclusive of data and metadata for other resources.

Active fault YR1	Recent Tectonics
Bottom Hole Temperature YR1	Temperature & Heat flow
Bottom Hole Temperature YR1	Temperature & Heat flow
Document (Geology) YR1	Metadata and Documents
Geologic maps YR1	Map Data
Geothermal maps YR1	Metadata and Documents
Lithology log - geotech YR1	Well Logs
Lithology log - O,G & G wells YR1	Well Logs
Temperature-depth logs YR1	Well Logs
Volcanic vent feature YR1	Recent Tectonics
Borehole Lithology Logs YR2	Well Logs
Bottom Hole Temperature YR2	Temperature & Heat flow
Earthquake hypocenter YR2	Recent Tectonics
EQ Epicenter/Hypocenter YR2	Recent Tectonics
Geology Map Service	Map Data
Gravity map data YR2	Map Data
Temperature Depth Logs YR2	Temperature & Heat flow
Volcanic vent feature YR2	Recent Tectonics
Age Dates YR3	Rock Chemistry
Bottom Hole Temperatures YR3	Temperature & Heat flow

Direct Use YR3	Other
Geochemistry YR3	Rock Chemistry
Geologic Maps (DGER) YR3	Map Data
Thermal Spring/Water Source YR3	Water Chemistry
Well Log Data YR3	Well Logs
Hot spring geochem/thermometers SUP	Temperature & Heat flow
Temperature Depth Log SUP	Temperature & Heat flow
Temperature Gradient Holes SUP	Temperature & Heat flow
Wind River Valley Report SUPP	Metadata and Documents

Subtask 2.4 (AH) Indiana: See general Subtask 2.4 description above.

The following table shows the final deliverables from Indiana. They completed 97.44% of the deliverables as initially outlined in their original SOWs, with 45,944 records in the NGDS, inclusive of data and metadata for other resources.

Bedrock Geology Map 1:500K YR1	Map Data
BHT (Petroleum Logs) YR1	Temperature & Heat flow
Bottomhole Temperature YR1	Temperature & Heat flow
Geol Map Elevation Tops Fms YR1	Map Data
Geologic map Bedrck Elevation YR1	Map Data
Geologic map Bedrock 1:24K YR1	Map Data
Geologic map (Core Samples) YR1	Map Data
Geologic map Depth to Bedrock YR1	Map Data
Geologic map (Petroleum Logs) YR1	Map Data
Geologic map Soil Characteris YR1	Map Data
Geologic map (Well Samples) YR1	Map Data
Geophysical Data Index Maps YR1	Map Data

Permeability YR1	Other
Seismic Depth to Bedrck Index YR1	Map Data
Water Chemistry TDS YR1	Map Data
Water Chemistry UCMs YR1	Map Data
Water Source Feature UCMs YR1	Map Data
Water Source Mineral Quarries YR1	Map Data
Develop GIS Data/Metadata YR2	Metadata and Documents
Geologic Map Bedrock topo YR2	Map Data
Geologic Map depth-to-bedrock YR2	Map Data
Geologic Map IN YR2	Map Data
Geologic Unit GeoT Character YR2	Other
Geothermal System Feature YR2	Other
Bedrock Geology Maps YR3	Map Data
Bedrock Wells Aqueous Chem YR3	Water Chemistry
Bottom Hole Temperatures YR3	Temperature & Heat flow
Depth-to-bedrock Map YR3	Map Data
Digital Elevation Map YR3	Map Data
Geologic map YR3	Map Data
Geothermal Map YR3	Map Data
Temp-depth logs YR3	Well Logs
Unconsol. Aquifer Aq Chem YR3	Water Chemistry
Geologic Unit -In-situ Heat SUP	Map Data
Geologic Unit-physical property SUP	Map Data
Geologic Unit-vol. moisture SUP	Map Data

Temperature SUP	Temperature & Heat flow
Thermal conductivity SUP	Temperature & Heat flow
Thermal Diffusivity SUP	Temperature & Heat flow

Subtask 2.4 (AI) West Virginia: See general Subtask 2.4 description above.

The following table shows the final deliverables from West Virginia. They completed 94.12% of the deliverables as initially outlined in their original SOWs, with 61,182 records in the NGDS, inclusive of data and metadata for other resources.

Aeromagnetic map YR1	Map Data
Basement Map YR1	Map Data
Bottom Hole Temperature YR1	Temperature & Heat flow
Digitized LAS Well Logs YR1	Well Logs
Gravity map YR1	Map Data
Scanned Well Logs and Data YR1	Well Logs
Bottom Hole Temperature YR2	Temperature & Heat flow
Digitized LAS Well Log YR2	Metadata and Documents
Earthquake Epicenter YR2	Recent Tectonics
Faults in West Virginia YR2	Recent Tectonics
Scanned Well Logs and Data YR2	Temperature & Heat flow
Spring/Water source feature YR2	Well Logs
Bottom Hole Temperature YR3	Temperature & Heat flow
Digitized LAS Well Logs YR3	Well Logs
Geologic map (Coverage) YR3	Map Data
Scanned Well Logs and Data YR3	Temperature & Heat flow
Thermal conductivity SUP	Temperature & Heat flow

Subtask 2.4 (AJ) Oklahoma: See general Subtask 2.4 description above.

The following table shows the final deliverables from Oklahoma. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 15,211 records in the NGDS, inclusive of data and metadata for other resources.

Arbuckle digital logs YR1	Temperature & Heat flow
Depth to bedrock YR1	Map Data
Gravity maps YR1	Map Data
Magnetic map YR1	Map Data
Active Faults YR2	Recent Tectonics
Arbuckle digital logs YR2	Temperature & Heat flow
Arbuckle well penetrate YR2	Well Logs
Depth to bedrock YR2	Map Data
Gravity maps YR2	Map Data
Magnetic maps YR2	Map Data
Arbuckle well penetrate YR3	Well Logs
Gravity Maps of OK YR3	Map Data
Magnetic Maps YR3	Map Data
Radiogenic heat production SUP	Temperature & Heat flow

Subtask 2.4 (AK) Massachusetts: See general Subtask 2.4 description above.

The following table shows the final deliverables from Massachusetts. They completed 95.83% of the deliverables as initially outlined in their original SOWs, with 2,222 records in the NGDS, inclusive of data and metadata for other resources.

Chemical analysis (whole-rock) YR1	Rock Chemistry
Extra Hot Springs Aqueous Chemistry	Water Chemistry
Hot Spring Description YR1	Temperature & Heat flow
Temperature depth log YR1	Well Logs
Temperature (Direct Use & GSHP data) YR1	Temperature & Heat flow

Thermal Conductivity YR1	Temperature & Heat flow
Trace constituent chemistry YR1	Rock Chemistry
Heat Flow Measurement YR2	Temperature & Heat flow
Temperature Observations YR2	Temperature & Heat flow
Thermal conductivity lith basin YR2	Temperature & Heat flow
Trace Chemistry data YR2	Rock Chemistry
Trace chem lith basin YR2	Rock Chemistry
Whole-Rock Chemistry YR2	Rock Chemistry
Whole-rock chem - lith basin YR2	Rock Chemistry
Heat Flow Measurement YR3	Temperature & Heat flow
Heat Production Map YR3	Map Data
Inferred Heat Flow Map YR3	Map Data
Inferred Temp Map By Depth YR3	Map Data
Metadata Maps YR3	Metadata and Documents
Soil Thermal Conductivity Map YR3	Map Data
Thermal Conductivity Map YR3	Map Data
Metadata SUP	Metadata and Documents
Thermal conductivity Bedrock SUP	Rock Chemistry
Thermal conductivity Soils SUP	Rock Chemistry

Subtask 2.4 (AL) Connecticut: See general Subtask 2.4 description above.

The following table shows the final deliverables from Connecticut. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 613 records in the NGDS, inclusive of data and metadata for other resources.

Aeromagnetic Map CT YR1	Map Data
Aeroradioactivity Maps NY, NE YR1	Map Data

Lithologic Logs YR1	Well Logs
Metadata for Documents and Maps	Metadata and Documents
Trace Chemistry Data YR1	Rock Chemistry
Whole-rock chemistry YR1	Rock Chemistry
YR1_Aeroradioactivity Map Parts of NY, CT, RI, MA YR1	Map Data
Aeromagnetic Map CT YR2	Map Data
Whole-rock chem-Granitoids YR2	Rock Chemistry
Glacial Deposit Thickness YR2	Rock Chemistry
Metadata for Documents YR2	Metadata and Documents
Radon Potential Map CT YR2	Map Data
Thermal Conductivity YR2	Temperature & Heat flow
Trace Chemistry Lith basin YR2	Rock Chemistry
Trace Chemistry Whole-Rock YR2	Rock Chemistry
Whole Rock Chem Lithified YR2	Rock Chemistry
Heat Production Map YR3	Map Data
Inferred Heat Flow Map YR3	Map Data
Inferred Temp Map By Depth YR3	Map Data
Metadata Maps and documents YR3	Metadata and Documents
Moodus Seismic Report YR3	Metadata and Documents
Soil Thermal Conductivity Map YR3	Map Data
Thermal Conductivity Map YR3	Map Data
Thermal Conductivity YR3	Temperature & Heat flow
Thermal conductivity (bedrock) SUPP	Temperature & Heat flow
Thermal conductivity (soils) SUPP	Temperature & Heat flow

Subtask 2.4 (AM) Vermont: See general Subtask 2.4 description above.

The following table shows the final deliverables from Wisconsin. They completed 88.89% of the deliverables as initially outlined in their original SOWs, with 3069 records in the NGDS, inclusive of data and metadata for other resources.

Bedrock, Surf, Hydrogeo Maps YR1	Map Data
Borehole Temperatures YR1	Temperature & Heat flow
Doc Metallic/NonMetal minerals YR1	Metadata and Documents
Doc Mineral Resource Data YR1	Metadata and Documents
Doc Natural Gas Explore Well YR1	Metadata and Documents
Docs - Non-VGS publications YR1	Metadata and Documents
Docs VTGS Pubs: Report/maps YR1	Metadata and Documents
Document - Masters theses YR1	Metadata and Documents
Document - Misc documents YR1	Metadata and Documents
Borehole Data YR 2	Well Logs
Document - Map Images YR2	Map Data
Surficial/Bedrock Digitial Maps YR2	Map Data
Borehole Temperatures YR3	Temperature & Heat flow
Heat Pump Systems Dataset YR3	Temperature & Heat flow
Well Header Data YR3	Well Logs
Borehole Lithology Log SUP	Temperature & Heat flow
Chemical analysis (whole-rock) SUP	Rock Chemistry
Thermal conductivity SUP	Temperature & Heat flow

Subtask 2.4 (AN) Colorado: See general Subtask 2.4 description above.

The following table shows the final deliverables from Colorado. They completed 70.37% of the deliverables as initially outlined in their original SOWs, with 571,279 records in the NGDS, inclusive of data and metadata for other resources.

Active faults YR1	Recent Tectonics
BHT Corrections YR1	Temperature & Heat flow
BHT data YR1	Temperature & Heat flow
BHt maps YR1	Map Data
Document - Bibliography YR1	Metadata and Documents
Document - References YR1	Metadata and Documents
Geologic Map - Part 1 YR1	Map Data
Heat flow measurement - map YR1	Temperature & Heat flow
Hot Spring Description - basic YR1	Temperature & Heat flow
Hot Spring / geochemistry YR1	Water Chemistry
Geologic Map - Part 2 YR2	Map Data
Gradient BHT Data YR2	Temperature & Heat flow
Heat Flow Measurement - Data YR2	Temperature & Heat flow
Surface Temperature Data YR2	Temperature & Heat flow
Direct Use Facilities YR3	Metadata and Documents
Flow Rate/Co-prod. Fluids YR3	Other
Geothermal Map YR3	Map Data
Radiogenic Heat Prod. YR3	Temperature & Heat flow
Sedimentary Basin Map YR3	Map Data
Temp-depth Logs YR3	Well Logs
Thermal Conductivity YR3	Temperature & Heat flow
Drill Stem Test Data SUP	Well Logs
TSK1 HeatFlow Measurement SUP	Temperature & Heat flow
TSK2 Thermal Conductivity SUP	Temperature & Heat flow

TSK3 Radiogenic Heat Production SUP	Temperature & Heat flow
TSK4 Trace constituent Water-Chem SUP	Water Chemistry
TSK5 Thermal conductivity Cuttings SUP	Temperature & Heat flow

Subtask 2.4 (AO) Idaho: See general Subtask 2.4 description above.

The following table shows the final deliverables from Idaho. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 11,447 records in the NGDS, inclusive of data and metadata for other resources.

Active Fault YR1	Recent Tectonics
BHT - additional well log data	Temperature & Heat flow
Digital well log - DWR YR1	Well Logs
Digital well log - IGS YR1	Well Logs
Docs-Geothermal relavent YR1	Metadata and Documents
Docs-metadata for CMs YR1	Metadata and Documents
Docs-Thermal Water Source YR1	Metadata and Documents
Whole-rock chemistry YR1	Rock Chemistry
Active Fault YR2	Recent Tectonics
BHT - additional to well logs data	Temperature & Heat flow
Digital Log-GeoT Test Wells YR2	Well Logs
Digital Logs-Oil and Gas Wells YR2	Well Logs
Digital well log, Thermal Wells YR2	Well Logs
Doc-Contact Info YR2	Metadata and Documents
Doc-Geothermal Metadata YR2	Metadata and Documents
Doc-Thermal Springs/Wells YR2	Metadata and Documents
Geology Map Service	Map Data
Whole-Rock Chemistry YR2	Rock Chemistry

Active Fault Linked Document YR3	Metadata and Documents
Chemical analysis (whole-rock) YR3	Rock Chemistry
Digital Well Logs YR3	Well Logs
Drill stem test YR3	Well Logs
Formation Tops YR3	Well Logs
Geologic Map YR3	Map Data
Digital Well Log SUP	Well Logs
Document SUP	Metadata and Documents
Heat Flow Estimate SUP	Temperature & Heat flow

Subtask 2.4 (AP) Utah: See general Subtask 2.4 description above.

The following table shows the final deliverables from Utah. They completed 97.22% of the deliverables as initially outlined in their original SOWs, with 12,934 records in the NGDS, inclusive of data and metadata for other resources.

Active Faults Quaternary faults YR1	Recent Tectonics
BottomHole Temparture YR-1	Temperature & Heat flow
Document (bibliography) YR1	Metadata and Documents
Document (Quaternary Faults) YR1	Metadata and Documents
Doc - Volcanic Rocks & Vents YR1	Metadata and Documents
Drill stem tests YR1	Well Logs
Geologic maps YR1	Map Data
Hot spring descriptions YR1	Temperature & Heat flow
Temperature-depth logs YR1	Temperature & Heat flow
Thermal well descriptions YR1	Temperature & Heat flow
Volcanic Vents (Quat rocks) YR1	Map Data
Active Faults YR2	Recent Tectonics

Borehole Temperatures YR2	Temperature & Heat flow
Drill Stem Test YR2	Well Logs
Geologic Maps YR2	Map Data
Hot spring descriptions YR2	Temperature & Heat flow
Thermal-gradient dataset YR2	Temperature & Heat flow
Thermal well descriptions YR2	Temperature & Heat flow
Volcanic Vents & Flows YR2	Recent Tectonics
BHT O&G wells YR3	Temperature & Heat flow
Drill stem tests YR3	Temperature & Heat flow
Geologic maps YR3	Map Data
Heat-flow Map of Utah YR3	Temperature & Heat flow
Temp/Wire Logs Water Wells YR3	Temperature & Heat flow
Thermal Conductivity YR3	Metadata and Documents
Thermal-gradient dataset YR3	Temperature & Heat flow
Borehole Lithology Log SUP	Well Logs
Chemical Analyses-Exploratory SUP	Water Chemistry
Chemical Analyses-water wells SUP	Water Chemistry
Cutting Samples SUP	Rock Chemistry
Geo- Interpretations SUP	Metadata and Documents
Geophysical Logs SUP	Well Logs
Temperature-Depth Logs Borehole SUP	Temperature & Heat flow
Temperature-Depth Log (water) SUP	Water Chemistry
Temperature Gradient Wells SUP	Temperature & Heat flow
Thermal-conductivity SUP	Temperature & Heat flow

Subtask 2.4 (AQ) Oregon: See general Subtask 2.4 description above.

The following table shows the final deliverables from Oregon. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 46,769 records in the NGDS, inclusive of data and metadata for other resources.

Aquifer BHT - Task 1 YR1	Temperature & Heat flow
Digital well log - Task 4 YR1	Well Logs
Document - metadata for CMs YR1	Metadata and Documents
Document - Task 2 YR1	Metadata and Documents
Hot springs - Task 3 YR1	Temperature & Heat flow
Docs -GeoT Research 69-89 YR2	Metadata and Documents
Document - metadata for CMs YR2	Metadata and Documents
Geologic Maps YR2	Map Data
Radiometric age, Task 3 YR2	Rock Chemistry
Temp-depth log Task 1 YR2	Temperature & Heat flow
Volcanic vent feature, Task 2 YR2	Recent Tectonics
Active Faults, Task 4 - YR3	Recent Tectonics
Aqueous Chemistry, Task 1 - YR3	Water Chemistry
Direct Use Features, Task 2 - YR3	Other
GeoT Power Plant Task 3 -YR3	Other
Metadata Task 5 - YR3	Metadata and Documents
Borehole lithology log dataset SUP	Temperature & Heat flow
Chemical Analysis (rock chem) SUP	Rock Chemistry
Document (SUP)	Metadata and Documents
Geologic Maps SUP	Map Data
Geologic Unit feature SUP	Map Data
Sample Cuttings SUP	Rock Chemistry

Temperature Depth Log SUP	Temperature & Heat flow
Temperature Gradient Wells SUP	Well Logs
Thermal Infrared Imagery SUP	Map Data

Subtask 2.4 (AR) Nevada: See general Subtask 2.4 description above.

The following table shows the final deliverables from Nevada. They completed 74.19% of the deliverables as initially outlined in their original SOWs, with 158,119 records in the NGDS, inclusive of data and metadata for other resources.

Geophyscial logs and well records YR1	Well Logs
Metadata Digital Logs, Samples, Production Records YR1	Temperature & Heat flow
Metadata Exist Digital records YR1	Metadata and Documents
Paper Files Scanned YR1-3	Metadata and Documents
Power Plants Production Statistics YR1	Other
Rock Samples from Core YR1	Temperature & Heat flow
Thermal Wells Aqueous Chemistry	Water Chemistry
Water chemistry	Water Chemistry
Well Headers YR1	Well Logs
Borehole Temperatures YR2	Temperature & Heat flow
Geologic Maps YR2	Map Data
NBMG Online Items Metadata YR2	Metadata and Documents
Power Plant Features Y2	Other
Update Well Header Records YR2	Other
Update Well Headers w/ Related ResourcesYR2	Metadata and Documents
Well Header Data YR2	Well Logs
Additional Services YR3	Other

Core Photograph Data YR3	Well Logs
Core Samples YR3	Well Logs
Direct Use Sites YR3	Temperature & Heat flow
GeoT Areas Shape File YR3	Map Data
Metadata 3 Core Holes YR2	Metadata and Documents
Misc Additional Data Delivery YR3	Other
Power Plant Annual Data YR 3	Other
Scientific Datasets YR3	Metadata and Documents
Volcanic Vents YR3	Map Data
TSK10 Shallow-temperatures SUP	Temperature & Heat flow
TSK11 Large-scale Drilling A-F SUP	Water Chemistry
TSK12 Drilling temp-gradient NV SUP	Temperature & Heat flow
TSK8 Support Drilling -Permitting SUP	Water Chemistry
TSK9 Thermal Springs SUP	Water Chemistry

Subtask 2.4 (AS) Maine: See general Subtask 2.4 description above.

The following table shows the final deliverables from Wisconsin. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 2,925 records in the NGDS, inclusive of data and metadata for other resources.

Groundwater Temperature YR1	Temperature & Heat flow
Heat flow measurement YR1	Temperature & Heat flow
Aqueous Chemistry YR1	Water Chemistry
State geologic map YR1	Map Data
Thermal Conductivity YR1	Temperature & Heat flow
SUPP Borehole Observations Data	Temperature & Heat flow
SUPP Temperature Depth Logs	Temperature & Heat flow

Subtask 2.4 (AT) Arizona: See general Subtask 2.4 description above.

The following table shows the final deliverables from Arizona. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 50,760 records in the NGDS, inclusive of data and metadata for other resources.

Active faults YR1	Recent Tectonics
Bottom hole temperatures YR1	Temperature & Heat flow
Documents YR1	Metadata and Documents
Drill stem tests YR1	Water Chemistry
Geological Maps AZ YR2	Map Data
Hot spring descriptions YR1	Temperature & Heat flow
Hot Springs Water Chemistry YR1	Water Chemistry
Lithologic logs YR1	Well Logs
References YR1	Metadata and Documents
Temp-depth profiles YR1	Temperature & Heat flow
Temperature-depth logs YR1	Temperature & Heat flow
Thermal Well Header YR1	Well Logs
Well chemistry dataset YR1	Water Chemistry
Arizona Rock Chemistry YR2	Rock Chemistry
Heat Flow (Documents) YR2	Temperature & Heat flow
Isotopic Data (Documents) YR2	Metadata and Documents
Temperature Gradient Data YR2	Temperature & Heat flow
Water Chemistry Dataset YR2	Water Chemistry
Geothermal Direct Use YR3	Other
Lithologic Intervals ADWR Wells YR3	Well Logs
Temp-Depth ADWR Wells YR3	Temperature & Heat flow

Temperature Gradient Data YR3	Temperature & Heat flow
Volcanic Vents YR3	Recent Tectonics
Water Well Temperature Data YR3	Temperature & Heat flow
Document (SUP)	Metadata and Documents
Geochronologic analysis (SUP)	Recent Tectonics
Geology Map (SUP)	Map Data
Geothermometry (SUP)	Water Chemistry
Hot Springs Water Chemistry (SUP)	Water Chemistry

Subtask 2.4 (AU) New Jersey: See general Subtask 2.4 description above.

The following table shows the final deliverables from New Jersey. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 14,874 records in the NGDS, inclusive of data and metadata for other resources.

Borehole lithology logs YR1	Well Logs
Document - Metadata (logs) YR1	Metadata and Documents
Documents Public School GHPS YR1	Metadata and Documents
Documents YR1	Metadata and Documents
GeoT Heat Pump Systems YR1	Temperature & Heat flow
3D Analysis GeoT Parameters YR2	Temperature & Heat flow
3D Mapping Newark Basin YR2	Map Data
GW Heat Pump Study YR2	Temperature & Heat flow
Metadata - 3D spatial Analysis YR2	Metadata and Documents
Metadata GW HeatPump Study YR2	Metadata and Documents
Metadata Temp Profiles YR2	Metadata and Documents
Borehole Studies YR3	Metadata and Documents
Lithology Interval Logs YR3	Well Logs

Well Headers YR3	Well Logs
Down Hole Temperature Profiles SUP	Temperature & Heat flow
Metadata SUP	Metadata and Documents

Subtask 2.4 (AV) Pennsylvania: See general Subtask 2.4 description above.

The following table shows the final deliverables from Pennsylvania. They completed 100% of the deliverables (or more of the initially indicated deliverables) as outlined in their SOWs, with 109,148 records in the NGDS, inclusive of data and metadata for other resources.

Bottom HoleTemperature YR1	Temperature & Heat flow
Documents YR1	Metadata and Documents
Earthquake epicenter YR1	Recent Tectonics
Geologic map (Basement) YR1	Map Data
Geologic map of PA YR1	Map Data
Borehole lithology log Dataset YR2	Well Logs
Bottom Hole Temperature YR2	Temperature & Heat flow
Map Data (Fold Map) YR2	Map Data
Water Chemistry Dataset YR2	Water Chemistry
Bottom Hole Temperatures YR3	Temperature & Heat flow
GeoT Heat Pump facility YR3	Other
Gravity Bouger map data YR3	Map Data
Seismic Profiles YR3	Recent Tectonics
Water Chemistry YR3	Water Chemistry
Whole-rock Geochemistry YR3	Rock Chemistry
Geophysical well logs SUP	Well Logs

Janice K. Brewer
Governor

State of Arizona
Arizona Geological Survey
416 W. Congress, Suite 100
Tucson, Arizona 85701
(520) 770-3500

M. Lee Allison, Ph.D., R.G.
Director and State Geologist

July 22, 2009

To Whom It May Concern:

The Arizona Geological Survey (AZGS) hereby provides assurance to the U.S. Department of Energy that funds awarded to AZGS from DE-FOA-0000109 will be used in such a manner that any laborers on projects funded directly or indirectly with such funds shall be paid at rates equal to or greater than the applicable prevailing wage. While it is not expected that this project will include public works construction, AZGS will ensure that any applicable work will indeed be in compliance with the Davis-Bacon Act.

M. Lee Allison
State Geologist & Director

Letters of Commitment and Support

Association of American State Geologists

Boise State University

Energy Industry Metadata Standards Work Group

U.S. Geological Survey

State of Alaska, Department of Natural Resources, Division of Geological and Geophysical Surveys

Geological Survey of Alabama

Arkansas Oil & Gas Commission

USGS Arkansas Water Science Center

Colorado Geological Survey

Connecticut Geological Survey

Florida Geological Survey

Iowa Geological and Water Survey

Idaho Geological Survey

Illinois State Geological Survey

Indiana Geological Survey

Kansas Geological Survey

Maine Office of Energy Independence and Security

Kentucky Geological Survey

Louisiana State University (Louisiana Geological Survey)

Massachusetts Clean Energy Center

Massachusetts Department of Energy Resources

University of Massachusetts and the Office of the State Geologist

Maine Geological Survey

Michigan Geological Repository for Research and Education

Minnesota Geological Survey

Missouri Department of Natural Resources – Division of Geology & Land Survey

Mississippi Office of Geology

North Dakota Geological Survey

Minnesota Geological Survey

New Hampshire Geological Survey

New Jersey Geological Survey
New Mexico Bureau of Geology and Mineral Resources
Nevada Bureau of Mines and Geology
Great Basin Center for Geothermal Energy
New York State Geological Survey
Ohio Division of Geological Survey
Oregon Department of Geology and Mineral Resources
Pennsylvania Bureau of Topographic and Geological Survey
South Carolina Geological Survey
State Geological Survey of Texas
Rhode Island Geological Survey
South Dakota Geological Survey
Utah Geological Survey
Vermont Geological Survey
Washington Geological Survey
Weyerhaeuser Corporation
Pacific Northwest National Laboratory
Snohomish County Public Utility District (Washington State)
Skamania County Board of Commissioners (Washington State)
USGS at University of Washington
Wisconsin Geological and Natural History Survey
West Virginia Geological and Economic Survey
West Virginia Division of Energy
D. H. Jones & Son Heating and Cooling
Wyoming State Geological Survey

♦ PRESIDENT
David R. Wunsch
New Hampshire Geological Survey
Department of Environmental Services
29 Hazen Drive
Concord, New Hampshire 03302-0095
Phone: 603-271-6482
FAX: 603-271-3305
E-mail: david.wunsch@des.nh.us

♦ PAST-PRESIDENT
Berry H. (Nick) Tew, Jr.
Geological Survey of Alabama
P O Box 869999
Tuscaloosa, AL 35486-6999
Phone: 205-247-3679
FAX: 205-349-2861
Email: ntew@gsa.state.al.us

♦ PRESIDENT-ELECT
James Cobb
Kentucky Geological Survey
228 Mining and Mineral Resources
Lexington, KY 40506-0107
Phone: 859-257-5500
FAX: 859-257-1147
Email: cobb@uky.edu

♦ VICE-PRESIDENT
Vicki S. McConnell
Oregon Dept. of Geology and Mineral Industries
800 NE Oregon St., Suite 965
Portland, OR 97232
Phone: 971-673-1550
Fax: 971-673-1562
Email :vicki.mcconnell@dogami.state.or.us

♦ SECRETARY
Robert Swenson
Alaska Division of Geological & Geophysical Surveys
3354 College Rd.
Fairbanks AK 99709-3707
Phone: 907-451-5001
Fax: 907-451-5030
Email; bob.swenson@alaska.gov

♦ TREASURER
Harvey Thorleifson
Minnesota Geological Survey
University of Minnesota
2642 University Avenue W., Room 104
St. Paul, MN 55114-1057
Phone: 612-627-4780
FAX: 612-627-4778
E-mail: thorleif@umn.edu

♦ Executive Committee Members

EDITOR
Karl W. Muessig
New Jersey Geological Survey
Division of Land Use Management
Department of Environmental Protection
P O Box 427
Trenton, NJ 08625
Phone: 609-2921185
FAX: 609-633-1004
Email: Karl.Muessig@dep.state.nj.us

HISTORIAN
William M. Kelly
New York State Geological Survey
NYSM
3140 CEC
Albany, NY 12230
Phone: 518-474-7559
FAX: 518-486-2034
Email: wkelly@mail.nysed.gov

STATISTICIAN
Rick Allis
1594 West North Temple
P.O. Box 146100
Salt Lake City, UT 84114-6100
Phone: 801-537-3300
Fax;:801-537-3400
Email: rickallis@utah.gov

AASG

Association of American State Geologists

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, AZ 85701

Dear Dr. Allison:

The Association of American State Geologists (AASG) strongly supports the efforts of the Arizona Geological Survey in taking the lead in the preparation of a proposal on the Association's behalf to seek funding from the U.S. Department of Energy (DOE) Geothermal Technology Program, Topic Area 3, Geothermal Data Development, Collection, and Maintenance. This cooperative effort among the cooperating state geological surveys will support the creation and enhancement of the National Geothermal Data System (NGDS). The individual state geological surveys have extensive geological, geophysical, and geothermal data and samples, as well as extensive archives of ancillary materials that are critical for geothermal exploration and development. Thus, the state surveys collectively contain a wealth of information that should be incorporated in the data system for the benefit of the nation.

Previously, the AASG heartily endorsed the report of the NSF-sponsored workshop, *"Role of State Geological Surveys and USGS in a Geological Information System for the Nation: A report to the AASG and USGS on the results of a workshop conducted February 21-22, 2000"*, and senior USGS management has agreed to create a joint Steering Committee to pursue implementation of this network. Subsequently, the proposal you helped develop for the creation of the National Geothermal Data System is entirely compatible and supportive of this agreement and will benefit immeasurably from building on this existing national program.

As you know, at its recent annual meeting, the AASG unanimously supported you and the Arizona Geological Survey in your roles as co-chair of the AASG-USGS GIN Steering Committee, and in your cooperative efforts with the Universities and other geosciences entities to build the NGDS. The recent news that the Energy Industry Metadata Standards Working Group has formally committed to collaborate with AASG on our geothermal data proposal is yet another milestone towards this cooperative effort to build the NGDS. The Association fully supports the consortium's goals to achieve this community vision.

AASG will commit to providing support to the project through the engagement of the Geothermal Steering Committee for this effort; and will encourage AASG members to fully participate in the project, provide access for member states to participate, and to collectively serve as a forum for making data content available to the NGDS.

Sincerely,

David R. Wunsch
President

BOISE STATE UNIVERSITY

Walter S. Snyder
Director, National Geothermal Data System

Department of Geosciences
1910 University Drive Boise, Idaho 83725-1535
phone: 208-426-5905
fax: 208-426-4061
wsnyder@boisestate.edu
www.ngdsystem.org

July 17, 2009

M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, AZ 85741

Lee:

We have communicated extensively about the Association of American State Geologist proposal "State Geological Survey Contributions to the National Geothermal Data System." Since the Geoscience Information Network (GIN) is already part of the National Geothermal Data System (NGDS), and you and Steve Richard are PI and Co-PI, I can say with confidence that the interface between the AASG's network (developed by GIN) and the NGDS Network will be a success. As we've discussed, the NGDS will be a federated network and GIN is already committed to a lead role in the development of the NGDS Network, and thus extending that network, though difficult, is quite feasible.

The state geological surveys clearly are important data providers for the NGDS. Your proposal provides a way to capture those data and make them available to the NGDS. As a member of your management board I will be able to work closely with your members and help to foster the success of the effort. In summary, we look forward to working with you for the success of the project and the growth of the NGDS.

Sincerely

Walter S. Snyder

Walter S. Snyder
Director, National Geothermal Data System
-and-
Department of Geosciences
Boise State University
Boise, Idaho 83725
Phone: (208) 426-5905
Fax: (208) 426-4061
Cell: (208) 863-9260

July 8, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, AZ 85701

Dear Dr. Allison:

On behalf of the Energy Industry Metadata Standards Work Group, I would like to thank you and accept your invitation for the Work Group to collaborate with the GIN Project on the proposed effort to link the National Geothermal Data System into the Geoscience Information Network.

As an ad-hoc group, you are aware that we have no resources to commit beyond those focused on the Work Group goals. In addition, we expect to become a formal Energistics Work Group near the end of August, at which time Energistics will presumably need to renew this commitment. I expect Energistics to conclude the same as the current Work Group, however: that this opportunity is a natural extension of our current collaboration with the GIN Project, as manifest by Steve Richard's participation on our 7-member Steering Team. And we see that extension not only as highly aligned with our goals and vision, but as a strategic opportunity to participate in a visible pilot project.

We see the potential impact of a successful, visible, and well conceived pilot project as considerable. It will be the first opportunity most in the energy community have to appreciate the vision toward which the Work Group is driving, a vision the essence of which I know is shared with the GIN Project:

> *... one in which all stakeholders in the energy community exploit a common set of metadata standards and guidelines to enable efficient cataloging, discovery, evaluation, and retrieval of available information resources, regardless of whether those resources are hosted internally or externally to their organization.*
>
> *"Standards for Energy Industry Use of Metadata" position paper (in prep.)*

Perhaps more importantly, however, the community will have the opportunity to appreciate the feasibility of that vision. We expect this appreciation will encourage adoption of the standards and guidelines, and thus gradual realization of the vision.

We also see this invitation as fostering alignment at multiple levels, something my experience suggests is a prerequisite to interoperability and needed to avoid inadvertent fragmentation of the community. Collaboration on the proposed NGDS work will encourage alignment between the Work Group and the GIN Project at both the strategic level where visions are shared, for example, and at the tactical level where shared understanding of deployment issues produces design choices that enable interoperability.

We look forward to collaborating with the GIN Project on this important effort.

Sincerely,

Scott Hills, Chevron

<u>On behalf of the ad-hoc Energy Industry Metadata Standards Work Group Steering Team</u>:

- Alan Doniger, Energistics
- Dave Danko, ESRI
- Grant Tucker, Shell
- Lisa Derenthal, Gimmal Group
- Robert Graham, BHP Billiton
- Steve Richard, GIN Project/Arizona Geological Survey

United States Department of the Interior

U.S. GEOLOGICAL SURVEY
Reston, Virginia 20192

July 20, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, AZ 85701

Dear Dr. Allison:

The U.S. Geological Survey (USGS) is pleased that the Arizona Geological Survey is proposing, on behalf of the Association of American State Geologists, and with the support of a number of other cooperators, that the U.S. Department of Energy fund the enhancement of the National Geothermal Data System (NGDS). The USGS has a long history of cooperation and collaboration with the Association of American State Geologists, participates in established forums, and maintains working relationships in the areas that are covered in the project and with many of the organizations involved. If this proposal is funded, the USGS will, to the extent allowable by the availability of agency funds and personnel, collaborate with your NGDS project team through the Geothermal Resource Studies Project, the Geoscience Information Network Steering Committee, the National Data Catalog, the National Geologic Map Database, and other USGS groups as appropriate, enhancing coordination, contributing data and sharing expertise on data systems.

We look forward to working with you and the other collaborators on the "State Geological Survey Contributions to the National Geothermal Data System."

Sincerely,

Linda Gundersen
Chief Scientist for Geology

STATE OF ALASKA

DEPARTMENT OF NATURAL RESOURCES

DIVISION OF GEOLOGICAL & GEOPHYSICAL SURVEYS

SARAH PALIN, GOVERNOR

☐ 3354 COLLEGE ROAD
FAIRBANKS, ALASKA 99709-3707
PHONE: (907) 451-5000
FAX: (907) 451-5050

☐ GEOLOGIC MATERIALS CENTER
18225 FISH HATCHERY ROAD
PO BOX 772805
EAGLE RIVER, ALASKA 99577-2805
PHONE: (907) 696-0079
FAX: (907) 696-0078

July 10, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Alaska Division of Geological & Geophysical Surveys is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

Alaska Division of Geological & Geophysical Surveys is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Alaska Division of Geological & Geophysical Surveys is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Robert F. Swenson
State Geologist and Division Director

GEOLOGICAL SURVEY OF ALABAMA

Berry H. (Nick) Tew, Jr.
State Geologist

420 Hackberry Lane
P.O. Box 869999
Tuscaloosa, Alabama 35486-6999
Phone (205)349-2852
Fax (205)349-2861
www.gsa.state.al.us

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison:

The Geological Survey of Alabama (GSA) is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The GSA is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The GSA is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely yours,

Berry H. (Nick) Tew, Jr.
State Geologist and
Oil & Gas Supervisor

itb

Science and Service for the People of Alabama

ARKANSAS OIL AND GAS COMMISSION

Director's Office:
301 Natural Resources Drive
Suite 102
Little Rock, AR 72205
Phone: (501) 683-5814
Fax: (501) 683-5818
http://www.aogc.state.ar.us

Mike Beebe
Governor

Lawrence Bengal
Director

El Dorado Regional Office:
2215 West Hillsboro
El Dorado, AR 71730
Phone: (870) 862-4965
FAX: (870) 862-8823

Fort Smith Regional Office:
3309 Phoenix Avenue
Fort Smith, AR 72903
Phone: (479) 646-6611
FAX: (479) 649-7656

July 6, 2009

Bekki White, Director
Arkansas Geological Survey
3815 West Roosevelt Road
Little Rock, Arkansas 72204

Re: DOE Geothermal Data Collection Project

Dear Ms. White:

The Arkansas Oil and Gas Commission supports your efforts in seeking funds to participate in the Department of Energy Geothermal Data Collection Project and to that end will assist in providing data contained in the Commission files. The Commission also supports the effort from the standpoint of studying the potential of generating electricity from geothermal sources, which may allow operators to lower the cost of production in the oil producing area of Arkansas, which is primarily from stripper wells.

Given the need to expand and optimize domestic energy resources and to provide Arkansas oil and gas operators with alternative electricity sources, the Commission fully supports and encourages the Geological Survey to pursue this project.

Sincerely,

Lawrence Bengal
Director

LB/sh

COMMISSION MEMBERS
Chad White, Chairman, Magnolia
W. Frank Morledge, Vice-Chairman, Forrest City
Charles Wohlford, Fort Smith • Bill Poynter, Texarkana
Mike Davis, Magnolia • Kenneth Williams, Jersey
William L. Dawkins, Jr., Fort Smith • Jerry Langley, Smackover
Chris Weiser, Magnolia

An equal opportunity employer

United States Department of the Interior

U.S. GEOLOGICAL SURVEY
Water Resources Division
401 Hardin Road
Little Rock, Arkansas 72211

July 2, 2009

U.S. Department of Energy

The U.S. Geological Survey (USGS) Arkansas Water Science Center (WSC) fully supports the Geothermal Data Collection Project that the Association of American State Geologist (AASG) is proposing. The AASG is ideally situated to perform an outstanding job in compiling and analyzing the geothermal data necessary to conduct this project. The State Geologists and their staff have a very good working relationship with other State and Federal environmental agencies that will ensure project success.

The Arkansas WSC has a long history of working with the Arkansas Geological Survey (AGS) related to groundwater, water quality, and streamflow data collection and interpretative projects. Our office will assist AGS in providing groundwater (well and spring) temperature data and borehole geophysical log information. We have conducted several projects and published reports related to the geothermal resources of Hot Springs National Park. In addition, the Arkansas WSC has more than 7,000 groundwater temperature measurements in Arkansas stored in our computer database, and an additional tens of thousands of temperature measurements from continuous (15 minute and hourly) monitoring sites of springs and wells. The USGS will benefit from the proposed study by gaining an improved understanding of the distribution and source of thermal and non-thermal groundwater through the compilation and interpretation of these data.

We look forward to assisting in any way we can.

Sincerely,

John E, Terry
Director
USGS Arkansas Water Science Center

STATE OF COLORADO

COLORADO GEOLOGICAL SURVEY

Department of Natural Resources
1313 Sherman Street, Room 715
Denver, CO 80203
Phone: (303) 866-2611
Fax: (303) 866-2461

COLORADO
DEPARTMENT OF NATURAL RESOURCES

Bill Ritter
Governor

Harris D. Sherman
Executive Director

Vincent Matthews
Division Director and
State Geologist

July 8, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Colorado Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Colorado Geological Survey has actively worked on gathering and developing geothermal data in Colorado over the past three years. We want to advance our efforts to collect, develop, and maintain geothermal data to the maximum extent possible, making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Colorado Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Vince Matthews
State Geologist and Director

STATE OF CONNECTICUT
DEPARTMENT OF ENVIRONMENTAL PROTECTION

OFFICE OF INFORMATION MANAGEMENT
CONNECTICUT GEOLOGICAL SURVEY

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Connecticut Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Connecticut Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Connecticut Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Steven O. Fish
Director, Office of Information Management
steve.fish@ct.gov

Florida Department of Environmental Protection

Florida Geological Survey
903 West Tennessee Street • Gunter Building
Tallahassee, FL 32304-7716
Phone: (850) 488-4191 Fax: (850) 488-8086

Charlie Crist
Governor

Jeff Kottkamp
Lt. Governor

Michael W. Sole
Secretary

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Florida Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Florida Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Florida Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Jon Arthur, Ph.D., P.G.
Director and State Geologist

Fields of Opportunities

STATE OF IOWA

CHESTER J. CULVER, GOVERNOR
PATTY JUDGE, LT. GOVERNOR

DEPARTMENT OF NATURAL RESOURCES
RICHARD A. LEOPOLD, DIRECTOR

July 10, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Iowa Geological and Water Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Iowa Geological and Water Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development. It is also in-line with the state of Iowa's goals for green energy development and carbon emission control.

The Iowa Geological and Water Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Robert D. Libra

Robert D. Libra
State Geologist of Iowa
Iowa Geological and Water Survey

Iowa Geological and Water Survey / 109 Trowbridge Hall / Iowa City IA 52242-1319
319-335-1575 FAX 319-335-2754 www.igsb.uiowa.edu

University of Idaho

Office of Sponsored Programs
PO Box 443020
Moscow ID 83844-3020

Phone: 208-885-6651
Fax: 208-885-5752
osp@uidaho.edu

07/09/09

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The University of Idaho, Idaho Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

University of Idaho, Idaho Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The University of Idaho, Idaho Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

for Polly Knutson
Interim Director, Office of Sponsored Programs
University of Idaho

To enrich education through diversity, the University of Idaho is an equal opportunity/affirmative action employer.

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Institute of Natural Resource Sustainability
Illinois State Geological Survey

615 East Peabody Drive
Champaign, Illinois 61820

July 8,2009

Dr. M. Lee Allison

State Geologist & Director

Arizona Geological Survey

416 W. Congress, #100

Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Illinois State Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Illinois State Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Illinois State Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Beverly Seyler

Senior Geologist and Head Oil and Gas Section

telephone 217-333-4747 • *fax* 217-244-7004 • www.isgs.uiuc.edu

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Office of Sponsored Programs
and Research Administration
1901 South First Street, Suite A
South Research Park
Champaign, IL 61820

July 14, 2009

AZ Geological Survey
Attn: L. Christian Hanson, CAE
416 W. Congress, #100
Tuscon, IL. 85701

U of I REF. NO. 2010-00183
TITLE: Illinois Geothermal Database Development Project
AMOUNT: $299970
PERIOD: 10/1/09-9/30/12
PRINCIPAL INVESTIGATOR(s): Huff, Bryan G.
DEPARTMENT: IL State Geological Survey
TYPE OF REQUEST: New Request

Enclosed are copies of the above referenced proposal. This proposal has been approved for submission by the proper University administrative official(s).

Your consideration will be appreciated. Any contract or grant supporting the above described project must be issued in the University's corporate name, The Board of Trustees of the University of Illinois, Urbana, Illinois 61801.

Any questions of a non-technical nature regarding this proposal should be addressed to the individual below at (217) 333-2187:

Kim Walsh

Sincerely,

Kathy Young, Director
Office of Sponsored Programs and Research Administration

KY:jge

Enclosure

cc: Lisa Young

ATTACHMENT TO PROPOSAL TRANSMITTAL LETTER

(The following General Information is provided to assist potential Sponsors. It is recognized some information may not be applicable to this specific proposal and, if inappropriate, should be disregarded.)

1. **The University of Illinois reserves the right to negotiate the terms and conditions of any definitive Contract/Grant which may result from this proposal application. UIUC is a public research university subject to an increasing number of state and federal regulations that are unique to higher education. As a result, most contracts provided by our sponsors require minor revisions before we can legally sign them.**

2. **Any resulting Contract/Grant should be made in the University's legal corporate name, "The Board of Trustees of the University of Illinois", c/o Office of Sponsored Programs & Research Administration, at the address listed below in item 3.**

3. **All contractual correspondence should be mailed to:**

University of Illinois
Office of Sponsored Programs
& Research Administration (OSPRA)
1901 South First Street, Suite A
Champaign, IL 61820
E-mail: gcoaward@uillinois.edu

Contractual Signature Authority:

Walter K. Knorr, Comptroller
ATTEST:

Michele M. Thompson, Secretary

4. **General Information, Mailing Instructions, Representations/Certifications, etc: Jodi Bear (217) 265-7794**

Proposals		Contracts/Grants	
Joe Evans	(217) 333-5897	Kamil Tamimie	(217) 244-0315

5. **University Contacts related to Proposal Review PHONE (217) 333-2187 FAX #(217) 239-6830**

Kim Darnell	(217) 244-4762	Amber Floyd	(217) 265-7687
Carol Stickrod	(217) 265-7689	Deborah Gorvin	(217) 265-7680
Tim Tufte	(217) 265-7708	Kim Walsh	(217)265-7889
Julie McCabe	(217) 244-9029		

6. **Cognizant Federal Admin. Agency:**
Office of Naval Research
230 South Dearborn Avenue, Rm. 380
Chicago, IL 60604-1595
Attn: Administrative Contact
(312) 886-5423; E-Mail: ONR_Chicago@onr.navy.mil

7. **Contract/Grant payments should be mailed to:**
University of Illinois at Urbana-Champaign-Grants & Contracts
PO Box 4610
Springfield, IL 62708-4610

DUNS # 04-154-4081
FEIN # 37-6000.511
Cage Code: 4B808

8. **Authorized Institutional Officials for Submitting Proposal Applications:**

Administrative:	Business:
R.K. Iyer	Kathy Young, Director
Research Board	Office of Sponsored Programs & Research Administration

9. **The following research indirect cost rates have been currently negotiated with the Office of Naval Research:**

MTDC Indirect Cost Rate	Predetermined Graduate Asst. Tuition	Period
58.5%	56.0%	7/1/08 – 6/30/11

Illinois Geothermal Database Development Project

For Submission To:

Arizona Geological Survey
416 W. Congress, #100
Tucson, AZ 85701

By

Bryan G. Huff
Illinois State Geological Survey
615 E. Peabody Drive
Champaign, IL 61820

Make Grant to: The Board of Trustees of the University of Illinois
c/o Office of Sponsored Programs & Research Administration
1901 S. First Street, Suite A
Champaign, Illinois 61820-7406
Ph: (217)333-2187 Fax: (217)239-6830

Requested: $299,970

Project Period: 10/01/2009 – 09/30/2012

Bryan G. Huff, Principal Investigator
Illinois State Geological Survey

R. K. Iyer, Chair
Research Board
University of Illinois

E. Donald McKay, Interim Director
Illinois State Geological Survey

Kathy Young, Director
Office of Sponsored Programs &
Research Administration
University of Illinois

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Institute of Natural Resource Sustainability
Illinois State Geological Survey

615 East Peabody Drive
Champaign, Illinois 61820

July 8,2009

Dr. M. Lee Allison

State Geologist & Director

Arizona Geological Survey

416 W. Congress, #100

Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Illinois State Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Illinois State Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Illinois State Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Beverly Seyler

Beverly Seyler

Senior Geologist and Head Oil and Gas Section

R. K. Iyer

R. K. Iyer
Chair, Research Board

telephone 217-333-4747 • *fax* 217-244-7004 • www.isgs.uiuc.edu

ISGS Scope

Title: Illinois Geothermal Database Development Project
PI Bryan G. Huff

The Illinois State Geological Survey (ISGS) proposes to examine, evaluate, and enter data currently in non-digital paper format into its Enterprise basic well database. There are numerous wells drilled for petroleum exploration and other purposes that have Bottom Hole Pressure and Drill Stem Test data available in paper format. There are currently approximately 200,000 wells with paper records located in the Geologic Records Unit of the Illinois State Geological Survey. The Geologic Records Unit is the location of the state mandated repository for well data. The ISGS also maintains a digital Enterprise Basic Well Database. The ISGS digital well database is incomplete and data fields may lack entries particularly for wells drilled prior to the inception of our modern digital relational database in the 1980s. Many older wells have data pertinent to the state DOE RFP devoted to development, collection, and maintenance of geothermal database. Tasks to be included for Illinois' proposal are: (1.) entry of data in missing fields in the Enterprise Well Database. The focus will be on entry of geothermal related data such as bottom hole temperature and drill stem test data, (2.) extraction of existing pertinent well data into a separate geothermal database, (3.) The bottom hole pressure, drill stem test, and other data added to the ISGS Enterprise database will be made available via the ISGS ILoil and ILwater websites and (4.) geothermal data will also be made into GIS 3-D and contoured temperature data maps for identification of potential anomalies and correlations.

The primary focus of this project is data entry. The ISGS plans on hiring student assistants to work under the supervision of highly experienced and dedicated Geologic Records Unit staff. Professional staff will train, supervise, and conduct quality control for the data entry phase of this project. The ISGS has two professional Oracle programmers that possess the expertise to guide the data entry and data extraction phases of this project, to ensure the integrity of the ISGS digital database, and to ensure extraction of pertinent data for a national database.

INDIANA GEOLOGICAL SURVEY
611 N. Walnut Grove Ave., Bloomington, IN 47405-2208 · (812) 855-7636
http://igs.indiana.edu · IGSinfo@indiana.edu

July 8, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Indiana Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

As a member of the AASG, the Indiana Geological Survey supports the proposal to establish a comprehensive national inventory of geothermal data, and to making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Indiana Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, the AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative. We look forward to the establishment of the National Geothermal Data System and see its role as an integral component in advancing geothermal technology and creating associated economic development opportunities in the United States.

Sincerely,

John C. Steinmetz
Director and State Geologist

INDIANA UNIVERSITY

KU KANSAS GEOLOGICAL SURVEY
The University of Kansas

July 8, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear ~~Dr. Allison,~~ Lee

The Kansas Geological Survey (KGS) is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

KGS is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The KGS is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

W.E. Harrison

Dr. William E. Harrison
Director, Kansas Geological Survey
The University of Kansas

1930 Constant Avenue, Lawrence, KS 66047-3726, (785) 864-2965, Fax (785) 864-5317, www.kgs.ku.edu

STATE OF MAINE
OFFICE OF THE GOVERNOR
1 STATE HOUSE STATION
AUGUSTA, MAINE
04333-0001

JOHN ELIAS BALDACCI
GOVERNOR

JOHN M. KERRY
DIRECTOR
OFFICE OF ENERGY
INDEPENDENCE AND SECURITY

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Maine Office of Energy Independence and Security supports the proposed collaborative project that you are submitting in cooperation with the Association of American State Geologists (AASG) and the Maine Geological Survey to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

Driven in part by the desire to break Maine's overwhelming reliance on fossil fuels for winter heating, Maine Governor John Baldacci has promoted an energy future for Maine that will be lower cost, less dependent on imported fossil fuels, cleaner, and better for the state's economy. The Maine Office of Energy Independence and Security provides leadership in this effort. The Office monitors energy trends, provides information for consumers, and creates energy policies that enable the State of Maine to be a leader in energy conservation and efficiency. The Office also has worked to ensure that Maine State Government is a leader by example on sustainable energy practices, including the use of biodiesel to heat state offices, the adoption of LEED building standards, improved fuel economy and reduced travel for the state vehicle fleet, and energy efficiency improvements at state buildings.

Geothermal resources are an important element of the renewable energy future for our nation. It is important to develop a comprehensive understanding of this resource in order to implement the most efficient strategies for its use. While in Maine the primary geothermal application will be ground-coupled heat-pumps, it will be valuable to this expanding industry to understand the variable distribution of natural heat flow in order to implement the most efficient commercial systems.

The Maine Office of Energy Independence and Security is pleased to support in this project and awaits the results of this important initiative.

Sincerely,

John M. Kerry
Director
Governor's Office of Energy Independence and Security

UNIVERSITY OF KENTUCKY

July 8, 2009

Kentucky Geological Survey
Research
504 Rose Street
228 Mining & Mineral Resources Bldg.
Lexington, KY 40506-0107
Phone: (859) 257-5500
Fax: (859) 257-1147
www.uky.edu

Dr. M. Lee Allison
Director and State Geologist
Arizona Geological Survey
416 West Congress Street, Suite 100
Tucson, AZ 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Kentucky Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

Kentucky Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Kentucky Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

James C. Cobb

Jim Cobb
Director and State Geologist

KGS

An Equal Opportunity University

LOUISIANA STATE UNIVERSITY

Office of Sponsored Programs

Louisiana State University
202 Himes Hall
Baton Rouge, LA 70803

O 225-578-2760
F 225-578-2751
www.lsu.edu/osp
osp@lsu.edu

July 14, 2009

M. Lee Allison, PhD, pG
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, AZ 85701

Re: LSU proposal # 35157

Dear M. Lee Allison:

Louisiana State University (LSU) agrees to participate as a subawardee in your geothermal data collection project to DOE in response to FOA # DE-FOA-0000109. The LSU portion of the work as described in the attached proposal will be under the primary direction of Dr. Chacko John, Director of Louisiana Geological Survey.

The amount requested for LSU's participation in this research project is $299,979 over three years. Please see the attached budget and budget narrative for a detailed explanation of costs.

If your proposal is funded by DOE, LSU requests a subaward with terms and conditions applicable to an R&D project conducted by a state funded institution of higher education. We look forward to a rewarding and productive research effort.

Sincerely,

Karen Baumonte for
James L Bates
Executive Director

Enclosures

MASSACHUSETTS
CLEAN ENERGY
CENTER

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison:

The Massachusetts Clean Energy Center (CEC) fully supports the Office of the State Geologist's participation in the Association of American State Geologists' collaborative proposal to the U.S. Department of Energy to collect, develop and maintain geothermal data through Funding Opportunity Announcement DE-FOA-0000109, *Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

As the state's quasi public organization tasked to promote the clean energy industry and technologies, we recognize that new capabilities and technologies are creating options in clean energy. We welcome your efforts to compile and make available data critical to the potential development of geothermal resources – a critical part of bringing new geothermal technologies to the commercial marketplace.

For the proposed effort, the Office of the State Geologist is ideally situated as our State's lead agency. Its resources, mission, goals and affiliation with a major research institution are aligned to maximize participation and learning in this critical project.

The CEC, in coordination with the state energy agencies, commits to supporting the Office of the State Geologist in this effort by providing connections with industrial partners and others who may have geothermal data valuable to this project.

Sincerely,

Patrick Cloney
Executive Director

COMMONWEALTH OF MASSACHUSETTS
EXECUTIVE OFFICE OF
ENERGY AND ENVIRONMENTAL AFFAIRS
DEPARTMENT OF ENERGY RESOURCES
100 CAMBRIDGE ST., SUITE 1020
BOSTON, MA 02114
Internet: www.Mass.Gov/DOER
E-mail: Energy@State.MA.US

Deval L. Patrick
Governor

Timothy P. Murray
Lieutenant Governor

Ian A. Bowles
Secretary, Executive Office of Energy and Environmental Affairs

Philip Giudice
Commissioner

TELEPHONE
(617) 626-7300

FACSIMILE
(617) 727-0030
(617) 727-0093

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison:

The Massachusetts Department of Energy Resources endorses the Office of the State Geologist's participation in the Association of American State Geologists' collaborative proposal to the U.S. Department of Energy to collect, develop and maintain geothermal data through Funding Opportunity Announcement DE-FOA-0000109, *Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

We recognize that our energy future is on the threshold of major change and that building an energy portfolio that is diverse is essential to meeting the challenges ahead. We welcome your efforts to compile and make available data critical to the potential development of geothermal resources.

The Office of the State Geologist is ideally situated to be our State's lead agency for this proposal. Its resources, mission, goals and affiliation with a major research institution are aligned to maximize participation in this critical project.

The Massachusetts Department of Energy Resources will fully commit to supporting the Office of the State Geologist in this effort by providing connections with industrial partners and others who may have geothermal data that is valuable to this project. We look forward to working with you in this important endeavor.

Sincerely,

Philip Giudice
Commissioner, Department of Energy Resources

UMASS geosciences

DEPARTMENT OF GEOSCIENCES
Programs in Geology
Geography
Earth Systems
Tel: 413.545.0745
Fax: 413.545.1200

July 8, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Department of Geosciences and the Office of the Massachusetts State Geologist are committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Department and Office of the Massachusetts State Geologist are dedicated to advancing the efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Department and Office of the Massachusetts State Geologist are pleased to participate in this project as a subawardee and are looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Michael L. Williams

Dr. Michael L. Williams
Chair, Department of Geosciences

Department of Geosciences, 233 Morrill Science Center, University of Massachusetts, Amherst MA 01003

UNIVERSITY OF MASSACHUSETTS
AMHERST

Research Administration Building
70 Butterfield Terrace
Amherst, MA 01003-9242

Office of Grant and
Contract Administration

voice: 413.545.0698
fax: 413.545.1202

July 13, 2009

Arizona Geological Survey
c/o Lee Allison
416 W. Congress St., Suite 100
Tucson, AZ 85701-1381

RE: Proposed Subcontract to University of Massachusetts
UM Referenced No. 110-0054
Entitled: Geothermal Technologies Program – MA and CT

Dear Lee:

Attached is the subject proposal submitted on behalf of Professor Stephen B. Mabee of the Geosciences Department.

It is our understanding that this proposal will be included in a prime proposal which you are submitting to the U.S. Department of Energy.

The University is looking forward to participating in this project, subject to the execution of a mutually acceptable subcontract.

If you have questions on the technical aspects of the proposal, please contact Professor Mabee at (413) 545-4814. Administrative concerns may be directed to John Fillio, Grant and Contract Administrator, at (413) 545-0698. In all future correspondence about this proposal, please refer to UMass Proposal No. 110-0054.

Sincerely,

Jennifer Donais

Jennifer Donais, CRA
Associate Director

JD/keg
enclosure
cc: Stephen B. Mabee, P.I.

The University of Massachusetts is an Affirmative Action/Equal Opportunity Institution

Printed on Recycled Paper

STATE OF MAINE
DEPARTMENT OF CONSERVATION
22 STATE HOUSE STATION
AUGUSTA, MAINE
04333-0022

JOHN ELIAS BALDACCI
GOVERNOR

PATRICK K. MCGOWAN
COMMISSIONER

14 July 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Maine Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

Maine Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Maine Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Robert G. Marvinney, Ph.D.
State Geologist and Director

MAINE GEOLOGICAL SURVEY
ROBERT G. MARVINNEY, DIRECTOR AND STATE GEOLOGIST

PHONE: (207) 287-2801
FAX: (207) 287-2353
TTY: (207) 287-2213

Michigan Geological Repository for Research and Education
Department of Geosciences
Western Michigan University
Kalamazoo, MI 49008-5241
Fax: (269) 387-8669
Phone: (269)387-8633

William B. Harrison, III
Curator and Professor Emeritus
David A. Barnes
Robb Gillespie
G. Michael Grammer
Professors
Linda K. Harrison
Niah B. H. Venable
Administrators

July 16, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Michigan Geological Repository for Research and Education is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Michigan Geological Repository for Research and Education is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Michigan Geological Repository for Research and Education is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG,

the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Dr. William B. Harrison, III
Curator,
Michigan Geological Repository for Research and Education
Department of Geosciences
Western Michigan University

UNIVERSITY OF MINNESOTA

Twin Cities Campus

Minnesota Geological Survey
Newton Horace Winchell School of Earth Sciences
Institute of Technology

2642 University Avenue
St. Paul, MN 55114-1057
612-627-4780
Fax: 612-627-4778
thorleif@umn.edu

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

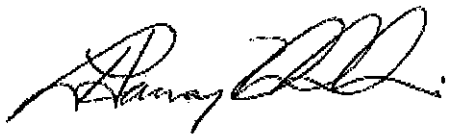
Dear Dr. Allison,

The Minnesota Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

Minnesota Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Minnesota Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,



Harvey Thorleifson Ph.D., P.Geo. D.Sc., Director, Minnesota Geological Survey, State Geologist of Minnesota, Professor, Department of Geology and Geophysics University of Minnesota; 2642 University Ave W, St Paul, MN 55114-1057 USA; Telephone 612-627-4780 ext 224; Fax 612-627-4778; thorleif@umn.edu

STATE OF MISSOURI
Jeremiah W. (Jay) Nixon, Governor • Mark N. Templeton, Director
DEPARTMENT OF NATURAL RESOURCES
www.dnr.mo.gov

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison:

The Missouri Department of Natural Resources' Division of Geology and Land Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Division of Geology and Land Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

As Missouri's State Geologist, I can assure you that we are pleased to participate in this project as a subawardee and we are looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

DIVISION OF GEOLOGY AND LAND SURVEY

Joe Gillman, RG
Director and State Geologist

JG:ta

Recycled Paper

STATE OF MISSISSIPPI
HALEY BARBOUR
GOVERNOR
MISSISSIPPI DEPARTMENT OF ENVIRONMENTAL QUALITY
TRUDY D. FISHER, EXECUTIVE DIRECTOR

July 10, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison:

The Office of Geology of the Mississippi Department of Environmental Quality is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Mississippi Office of Geology is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Mississippi Office of Geology is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Trudy D. Fisher
Executive Director

TDF:jar

POST OFFICE BOX 2261 • JACKSON, MISSISSIPPI 39225-2261 • TEL: (601) 961-5000 • FAX: (601) 961-5794 • www.deq.state.ms.us
AN EQUAL OPPORTUNITY EMPLOYER

North Dakota Geological Survey

Edward C. Murphy - State Geologist
Department of Mineral Resources
Lynn D. Helms - Director
North Dakota Industrial Commission
https://www.dmr.nd.gov/ndgs/

July 13, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The North Dakota Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The North Dakota Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The North Dakota Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Edward C. Murphy
State Geologist

600 E Boulevard Ave - Dept 405, Bismarck, North Dakota 58505-0840 Phone (701)328-8000 Fax (701)328-8010

UNIVERSITY OF UND NORTH DAKOTA

GEOLOGY AND GEOLOGICAL ENGINEERING
LEONARD HALL ROOM 101
81 CORNELL STREET STOP 8358
GRAND FORKS ND 58202-8358
(701) 777-2811
FAX (701) 777-4449
www.geology.und.edu
www.und.edu/dept/sem

Harvey Thorleifson Ph.D., P.Geo., D.Sc.
Director, Minnesota Geological Survey
Professor, Department of Geology and Geophysics
University of Minnesota; 2642 University Ave W, St Paul, MN

Dear Dr. Thorleifson,

The University of North Dakota (UND) contribution to the National Geothermal Database compilation includes subsurface temperature data, thermal conductivity data, and radioactive heat production heat data. We will record these data and the metadata in digital format for the NGDB.

We anticipate that about 100 temperature vs. depth measurements can be made in existing wells. We also will analyze Precambrian crystalline rocks for radiogenic heat production and thermal conductivity in collaboration with MN NRRI and the MN Geological Survey.

The field and laboratory work will be done by W. Gosnold and a graduate student research assistant.

Proposed timeline:

Year 1 Begin temperature, thermal conductivity, and radiogenic heat production data collection.

Years 2 and 3 Continue data collection and heat flow determinations

I enthusiastically support this application and look forward to working with the Minnesota Geological Survey on this project.

Sincerely yours,

Will Gosnold

Will Gosnold
Distinguished Professor
Chair, Geology and Geological Engineering

UND is an equal opportunity/affirmative action institution

JOHN H. LYNCH
GOVERNOR

STATE OF NEW HAMPSHIRE
OFFICE OF ENERGY AND PLANNING
4 Chenell Drive
Concord, NH 03301-8501
Telephone: (603) 271-2155
Fax: (603) 271-2615

OEP
www.nh.gov/oep

July 14, 2009

Ms. Genevieve Wozniak
Contracting Officer
US Department of Energy
Golden Field Office
Mail Stop 1501
1617 Cole Boulevard
Golden, Colorado 80401

Dear Ms. Wozniak:

I am writing this letter in support of the New Hampshire Geological Survey's (NHGS) partnership with other state geological surveys for their proposal in response to U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance*. The NHGS is a proposed sub-awardee in this proposal, which is being prepared in coordination with the Association of American State Geologists (AASG), which represents the State Geologists and Directors of the geologic bureaus of the 50 states.

The State of New Hampshire has a goal of producing 25 percent of its basic energy needs from renewable energy sources by 2025, and geothermal energy will play a part in our energy portfolio to meet this goal. Currently, geoexchange heating and cooling systems are very popular with New Hampshire homeowners and businesses for providing a green and economical heating and cooling option. However, direct-heat geothermal, and perhaps electricity generation from enhanced geothermal systems may be in our future as well. The Conway Granite in New Hampshire has long been recognized as a potential source of geothermal heat, and was explored previously with support from DOE in 1978 (Redstone Exploration Hole). It is still perhaps the best potential location for enhanced geothermal energy production east of the Mississippi River, according to the recently published 2007 MIT report: *"The Future of Geothermal Energy."* With new technologies and access to archival, as well as new data, the tools will be available for the next steps with regard to permitting and exploration of this green energy resource. Recently, state law created by the passage of HB0415 requires the New Hampshire state geologist to collect geologic information reported by persons installing geothermal systems or exploring hydrothermal or geothermal resources, and to prepare a revised geothermal assessment by 2011. Thus, New Hampshire is committed to the exploration of geothermal resources within our state.

The NHGS is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System currently in development. The grant proposal described herein for which New Hampshire wishes to take part would provide the resources and support to make enhanced geothermal data widely available. Moreover, the assets created through this geothermal program will assist the

Office of Energy and Planning to effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development. Collectively, the information and operational database proposed by the AASG will be an asset for the entire nation, and I strongly encourage you to fund the entire proposal.

Sincerely,

Amy L. Ignatius
Director
New Hampshire Office of Energy and Planning

State of New Jersey

DEPARTMENT OF ENVIRONMENTAL PROTECTION
Land Use Management
New Jersey Geological Survey
29 Arctic Parkway
PO Box 427
Trenton, NJ 08625-0427
Tel. # (609) 292-1185 – Fax (609) 633-1004 – Home Page: http//www.njgeology.org/

JON S. CORZINE
Governor

MARK MAURIELLO
Acting Commissioner

July 10, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The **New Jersey Geological Survey (NJGS)** is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

NJGS is dedicated to advancing its efforts to collect, develop and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

NJGS is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Finally, NJGS is looking forward to hosting AASG's 102nd Annual Meeting next year, at which time we anticipate discussions on the progress of this important geothermal project with all of the participating state geological surveys.

Sincerely,

Karl W. Muessig, Ph.D.
State Geologist

New Jersey is an Equal Opportunity Employer Printed on Recycled Paper and Recyclable

New Mexico Bureau of Geology & Mineral Resources
A DIVISION OF
NEW MEXICO INSTITUTE OF MINING & TECHNOLOGY
801 Leroy Place
Socorro, NM 87801-4796

Director: 575-835-5302
FAX: 575-835-6333
E-mail: scholle1@nmt.edu

Information: 575-835-5420
Publications: 575-835-5410
http://geoinfo.nmt.edu/

9 July 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The New Mexico Bureau of Geology & Mineral Resources is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance*.

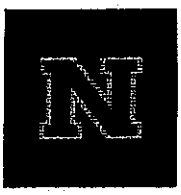
The New Mexico Bureau of Geology & Mineral Resources is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The New Mexico Bureau of Geology & Mineral Resources is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Peter A. Scholle
Director & State Geologist

NEW MEXICO TECH IS AN EQUAL OPPORTUNITY/AFFIRMATIVE ACTION INSTITUTION



University of Nevada, Reno

July 10, 2009

To: DOE GTP review panel

From: Lisa Shevenell
Director, Great Basin Center for Geothermal Energy
Faculty, Nevada Bureau of Mines and Geology
On behalf of
Jonathan G. Price, Ph.D.
State Geologist and
Director, Nevada Bureau of Mines and Geology

Subject: Letter of Commitment

The personnel at the Nevada Bureau of Mines and Geology and Great Basin Center for Geothermal Energy (GBCGE) are committed to working with the AASG group on this data gathering proposal in response to Funding Opportunity Announcement Number: DE-FOA-0000109. If the proposal is funded, we will provide the technical assistance indicated in the proposal for the state of Nevada, as well as technology transfer of the information developed in the course of performing the work outlined in the proposal. The GBCGE is also a founding member of the National Geothermal Database (NGDB) led by Boise State, and hence the data produced in this project will feed directly into the NGDB.

Thank you for your consideration.

Nevada Bureau of Mines and Geology
Mail Stop 178
Reno, Nevada 89557-0178
Telephone: (775) 784-6691
FAX: (775) 784-1709
www.nbmg.unr.edu

University of Nevada, Reno
Statewide • Worldwide

nbmg

Jonathan G. Price, Ph.D.
State Geologist and Director
Nevada Bureau of Mines and Geology
1664 North Virginia Street
University of Nevada, Reno/Mail Stop 178
Reno, Nevada 89557-0178

Office telephone: 775-784-6691 extension 5
Home telephone: 775-329-8011
Fax: 775-784-1709
E-mail: jprice@unr.edu

13 July 2009

To: DOE GTP review panel

Subject: Letter of Commitment

The personnel at the Nevada Bureau of Mines and Geology and Great Basin Center for Geothermal Energy (GBCGE, whose Director, Dr. Lisa Shevenell, is a member of the faculty of the Nevada Bureau of Mines and Geology) are committed to working with the Association of American State Geologists on this data gathering proposal in response to Funding Opportunity Announcement Number: DE-FOA-0000109. If the proposal is funded, we will provide the technical assistance indicated in the proposal for the State of Nevada, as well as technology transfer of the information developed in the course of performing the work outlined in the proposal. The GBCGE is also a founding member of the National Geothermal Database (NGDB) led by Boise State, and hence the data produced in this project will feed directly into the NGDB.

Thank you for your consideration.

Sincerely,

Jonathan G. Price

The Nevada Bureau of Mines and Geology (NBMG) is a research and public service unit of the University of Nevada, Reno and is the State geological survey. Established by the Nevada Legislature as a department within the public service division of the Nevada System of Higher Education, NBMG is part of the Mackay School of Earth Sciences and Engineering within the College of Science and one of the Statewide Programs at the University of Nevada, Reno. NBMG's mission, to provide the State's needs for geological and mineral-resource information and research, is defined in its enabling legislation. NBMG scientists conduct research and publish reports that focus on the economic development, public safety, and quality of life in urban and rural areas of Nevada.

NEW YORK STATE GEOLOGICAL SURVEY

3140 Cultural Education Center

Albany, NY 12230

518/474-5816 • Fax 518/473-8724

New York State Education Department – Office of Cultural Education – New York State Museum

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The New York State Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The New York State Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The New York State Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

William Kelly

William Kelly
State Geologist and Director

Ohio Department of Natural Resources

TED STRICKLAND, GOVERNOR

SEAN D. LOGAN, DIRECTOR

Division of Geological Survey
Lawrence H. Wickstrom, *Chief*

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Ohio Division of Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Ohio Division of Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Ohio Division of Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Larry Wickstrom
Division Chief & State Geologist
614-265-6598 Direct
614-268-3669 Fax

Oregon
Theodore R. Kulongoski, Governor

Department of Geology & Mineral Industries
Administrative Office
800 NE Oregon Street #28, Suite 965
Portland, OR 97232
(971) 673-1555
Fax: (971) 673-1562
www.oregongeology.com

July 10, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Oregon Department of Geology and Mineral Industries is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Oregon Department of Geology and Mineral Industries is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Oregon Department of Geology and Mineral Industries is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Don Lewis
Assistant Director

Pennsylvania Department of Conservation and Natural Resources

3240 Schoolhouse Road
Middletown, PA 17057-3534
July 20, 2009

Bureau of Topographic and Geologic Survey

(717) 702-2017
FAX (717) 702-2065

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Pennsylvania Bureau of Topographic and Geologic Survey (BTGS) is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

BGTS is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

BGTS is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Jay Parrish

Jay Parrish, Ph.D., P.G.
BGTS Director and State Geologist

South Carolina Department of
Natural Resources

John E. Frampton
Director

Don Winslow
Chief of Staff

July 21, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The South Carolina Department of Natural Resources, Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The South Carolina Department of Natural Resources, Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The South Carolina Department of Natural Resources, Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Don Winslow
Chief of Staff

1000 Assembly St • P.O. Box 167 • Columbia, S.C. 29202 • winslowd@dnr.sc.gov • Telephone: 803/734-9102
EQUAL OPPORTUNITY AGENCY www.dnr.sc.gov PRINTED ON RECYCLED PAPER

BUREAU OF ECONOMIC GEOLOGY, *OFFICE OF THE DIRECTOR*
THE UNIVERSITY OF TEXAS AT AUSTIN

John A. & Katherine G. Jackson School of Geosciences • University Station, Box X • Austin, Texas 78713-8924
10100 Burnet Road, Bldg. 130 • Austin, Texas 78758-4445 • (512) 471-0209 • FAX (512) 471-0140

July 10, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The State Geological Survey of Texas (Bureau of Economic Geology) is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The State Geological Survey of Texas is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The State Geological Survey of Texas is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,
Warm regards,

Scott W. Tinker, Director
State Geologist of Texas

University of Rhode Island
Department of Geosciences
Kingston, RI 02881-0807

Rhode Island
Geological Survey
Jon C. Boothroyd, State Geologist

Phone: (401) 874-2265
Fax: (401) 874-2190
email: RIGSurv@etal.uri.edu

July 10, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Rhode Island Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

Rhode Island Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Rhode Island Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Jon C. Boothroyd, PhD
Rhode Island State Geologist and Professor of Quaternary Geology

South Dakota Department of
Environment and Natural Resources
GEOLOGICAL SURVEY

AKELEY-LAWRENCE SCIENCE CENTER, USD
414 EAST CLARK STREET
VERMILLION, SD 57069–2390
605–677–5227
FAX 605–677–5895
INTERNET *www.state.sd.us/denr*

July 14, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Geological Survey Program, South Dakota Department of Environment and Natural Resources, is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Geological Survey Program, South Dakota Department of Environment and Natural Resources, is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Geological Survey Program, South Dakota Department of Environment and Natural Resources, is pleased to participate in this project and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Derric L. Iles

Derric L. Iles
State Geologist

JON M. HUNTSMAN, JR.
Governor
GARY R. HERBERT
Lieutenant Governor

State of Utah
DEPARTMENT OF NATURAL RESOURCES
MICHAEL R. STYLER
Executive Director
Utah Geological Survey
RICHARD C. ALLIS
State Geologist/Division Director

July 8, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Utah Geological Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance*.

The Utah Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Utah Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely

Rick Allis
Director/State Geologist

1594 West North Temple, Suite 3110, PO Box 146100, Salt Lake City, UT 84114-6100
telephone (801) 537-3300 • facsimile (801) 537-3400 • TTY (801) 538-7458 • *geology.utah.gov*

VERMONT

Vermont Department of Environmental Conservation
Commissioner's Office
103 South Main Street, 1 South [phone] 802-241-3808
Waterbury, VT 05671-0401 [fax] 802-244-5141

Agency of Natural Resources

July 13, 2009

Laurence R. Becker
Vermont State Geologist
Vermont Geological Survey
Department of Environmental Conservation
103 South Main Street
Waterbury, VT 05671

Dear Mr. Becker:

The Association of American State Geologists (AASG) is responding to the U.S. Department of Energy's (DOE) Funding Opportunity Announcement for *Topic Area 3: Geothermal Data Development, Collection, and Maintenance* under the *Recovery Act: Geothermal Technologies Program.* The Vermont Geological Survey (Division of Geology and Mineral Resources) in the Vermont Department of Environmental Conservation is submitting a VT state proposal to be included as part of the AASG multi-state proposal going to DOE.

DOE is seeking to develop, collect, and maintain technical geothermal and institutional data for the U.S. in order to make the data publically available and to reduce the risks associated with the initial stages of geothermal project development. This additional funding will accelerate data system development and populate critical geothermal site attribute information, such as temperature at depth, selsmicity/microseismicity, fracture maps, drilling data, permeability data, well logs, and geophysical surveys.

As part of this effort DOE has established the National Geothermal Data System (NGDS) to serve all publically accessible geothermal data. The AASG-USGS Geoscience Information Network (GIN) is chosen by DOE to power the NGDS.

As state geological surveys are the repositories for publically available geological information in their states, and Vermont is no exception, I strongly encourage the Vermont Geological Survey to take the state lead in developing existing and new information in this important energy area. Vermont is already seeing strong growth in the use of groundwater heat exchange systems. Enhanced Geothermal Systems (deeper resources) will eventually be tested in the eastern United States. With testing and application, economic viability of will improve and heat at depth in Vermont could well be a target for energy production in the future. This data development project will assist any future exploration efforts.

Sincerely,

Justin Johnson
Acting Commissioner

WASHINGTON STATE DEPARTMENT OF
Natural Resources
Peter Goldmark - Commissioner of Public Lands

Caring for
your natural resources
... now and forever

July 14, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, AZ 85701

Subject: Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison:

I am writing to express support for the Department of Natural Resources (DNR) Division of Geology and Earth Resources (Washington Geological Survey) proposal to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The DNR-Washington Geological Survey is regarded as the primary source of geological products and information in the state to support decision making by Washington's government agencies, its business, and the public. As the state's land management agency, the DNR is in a unique position to help lease and develop geothermal energy. Furthering our understanding of geothermal energy in Washington is important and this proposal will provide technology and information that can be readily transferred to others in our state who are interested in geothermal energy development.

There is a strong possibility for successful geothermal development in Washington because of the high heat flow that is obviously demonstrated by our many volcanoes. Despite our high heat flow, there is a lack of data on geothermal energy and we simply need more information to make good exploration decisions. The development of a geothermal system in the high heat flow areas of the Cascade Range and the Columbia Basin in Washington would have far reaching implications for renewable energy in our state.

As the Commissioner of Public Lands, it is one of my top priorities to support development of renewable energy in Washington. This project is essential to developing Washington's geothermal energy potential and I urge your support.

Thank you for your time and consideration.

Sincerely,

Peter Goldmark
Commissioner of Public Lands

1111 WASHINGTON ST SE ▮ MS 47001 ▮ OLYMPIA, WA 98504-7001
TEL (360) 902-1000 ▮ FAX (360) 902-1775 ▮ TTY (360) 902-1125 ▮ TRS 711 ▮ WWW.DNR.WA.GOV
EQUAL OPPORTUNITY EMPLOYER

RECYCLED PAPER

SNOHOMISH COUNTY PUD
PUBLIC UTILITY DISTRICT NO. 1

Your Northwest renewables utility invites you to be a Conservation Sensation.

July 13, 2009

Dave Norman
State Geologist
Department of Natural Resources
1111 Washington St. SE
Olympia, WA 98504

Dear Dave,

I am writing to express our strong support for the Geothermal Data Development, Collection, and Maintenance proposal that the Washington Department of Natural Resources is preparing in response to the U.S. Department of Energy's Geothermal Technologies Program Funding Opportunity Announcement DE-FOA-0000109.

Snohomish County PUD has taken a very proactive approach in the exploration of new renewable energy resource development in our region and believes that geothermal energy in particular has enormous potential to help ensure future energy needs are met in a sustainable way. The baseload characteristics of geothermal energy represent tremendous value relative to the intermittent and unpredictable nature of other renewable resources such as wind and solar. If our region is to be successful in achieving its vision for a clean, safe, reliable and sustainable energy future geothermal energy must be a prominent element of the resource mix.

Our ongoing geothermal exploration and development activities would be greatly facilitated by the activities described in your proposal. Lack of relevant data, effectively organized in an accessible and useful way, substantially increases the risk associated with the initial stages of geothermal project development. Snohomish County PUD, as well as other potential geothermal developers in our region, would greatly benefit from the results gained from your proposed activities.

Best of luck with your proposal. Snohomish County PUD looks forward to a successful collaboration with DNR resulting in the notable advancement of geothermal energy development in our region.

Sincerely,

Craig W. Collar
Senior Manager, Energy Resource Development
Snohomish County Public Utility District

2320 California Street • Everett, WA • 98201 / *Mailing Address:* P.O. Box 1107 • Everett, WA • 98206-1107
425-783-1000 • Toll-free in Western Washington at 1-877-783-1000 • www.snopud.com

SKAMANIA COUNTY
BOARD OF COMMISSIONERS

Skamania County Courthouse
Post Office Box 790
Stevenson, Washington 98648

(509) 427-3700 FAX: (509) 427-3708
TDD Relay Service (800) 833-6388

PAUL J. PEARCE
District 1

JIM RICHARDSON
District 2

JAMIE TOLFREE
District 3

July 9, 2009

Dave Norman
State Geologist
Department of Natural Resources
Division of Geology and Earth Resources
Washington Geological Survey
360.902.1439

Subject: Letter of Support for Washington Department of Natural Resources- Division of Geology and Earth Resources application for DE-FOA-0000109

Dear Mr. Norman:

We are writing to express our strong support for Recovery Act: Geothermal Technologies Program Topic Area 3: Geothermal Data Development, Collection, and Maintenance proposal that Washington State Department of Natural Resources is submitting to United States Department of Energy, funding opportunity announcement DE-FOA-0000109.

Skamania County is a natural resource-rich and diverse county, including geothermal resources. The majority of the land within Skamania County's boarders has potential for low and moderate geothermal energy production but there is also excellent high temperature geothermal potential as we have the highest recorded temperature and geothermal gradient in a well drilled by the Department of Natural Resources in the 1980's. While other natural resource economies have changed significantly in our county, we recognize the need to explore new opportunities that suit our resource strengths, such as geothermal energy production.

Skamania County Economic Development Council's Community Task Force has identified "geothermal energy feasibility study" as its ninth priority out of fifty. Fulfilling your proposed objectives to collect and develop data on geothermal energy sources in Skamania County would greatly help our efforts to focus our attention, resources and development in geothermal technologies. Having publically available data would also help our efforts to attract interested businesses and utilities to invest in geothermal technologies in Skamania County.

We look forward to hearing that you have been successful in obtaining support from the Department of Energy for your Geothermal Technologies Program proposal. Upon receiving the support, you can expect active participation in your efforts by our communities.

Sincerely,

Jamie Tolfree, Vice Chair

Paul Pearce, Commissioner

Weyerhaeuser
The future is growing™

Corporate Headquarters
32125 32nd Avenue S. (98001)
P.O. Box 9777, NB2 1D4
Federal Way, WA 98063-9777
Tel (253)924-4094
Fax (253)924-3862

July 15, 2009

David K Norman
State Geologist
Washington Geological Survey
PO Box 47007
Olympia WA 98501

Re: Weyerhaeuser Support for Geothermal Proposal Submittal for DE-FOA-0000109

Dear Mr. Norman,

I am writing to express our support for the Department of Natural Resources-Washington Geological Survey proposal to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Washington Geological Survey is regarded as the primary source of geological products and information in the state to support decision making by Washington's government agencies, its business, and the public. Because the Washington Geological Survey has done extensive research in geothermal energy and has a long continuing interest in the development of geothermal energy in our state Weyerhaeuser Corporation has been able to obtain important geological information to guide important business decisions. Furthering our understanding of geothermal energy in Washington is important and this proposal will provide technology and information that can be readily transferred to others in our state who are interested in geothermal energy development.

There is a strong possibility for successful geothermal development in Washington because of the high heat flow that has been documented but under explored. We simply need more information to make good exploration decisions. The development of a geothermal system in the high heat flow areas of the Cascade Range in Washington would have far reaching implications for energy development in our state and other states as well. This is a project that should go forward and it is the type of proposal that the Department of Energy should fund.

As a major land owner in Washington, Weyerhaeuser is supportive of the development of geothermal energy and we have leased extensive land holdings for exploration. New data and information developed by the Washington Geological Survey would be invaluable in guiding future exploration and developing new geothermal resources. We will be able to offer our expertise and support throughout the project and we will be able to attend progress meetings over the life of the project as needed.

We look forward to hearing that you have been successful in obtaining support from the Department of Energy for your proposal.

Sincerely,

Gary R. Helnemeyer *(WA State RPG#2294)*
Project Geologist and Business Development Manager

Pacific Northwest
NATIONAL LABORATORY

Tel: (509) 372-4548
Fax: (509) 372-6710
MSIN: K9-46
terry.walton@pnl.gov

July 21, 2009

David K Norman
State Geologist
Washington Geological Survey
PO Box 47007
Olympia WA 98501

Dear Mr. Norman:

Re: Geothermal Proposal Submittal for DE-FOA-0000109

I am writing to express our enthusiastic support for the Department of Natural Resources-Washington Geological Survey proposal to collect, develop, and maintain Geothermal data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

Our organization looks forward to contributing to the State of Washington's efforts to assess the viability of Geothermal Energy in the State and the Region. For this proposal, the Pacific Northwest National Laboratory is pleased to provide expertise in subsurface science and in particular our knowledge of the Columbia Basin along with access to a significant source of unpublished geological data that has been collected by the Department of Energy for the last several decades on the region.

The high heat flow of the Columbia Basin indicates our region has geothermal potential and the data generated from this study will help evaluate that potential for commercial geothermal development. The development of a geothermal systems in the Cascade Range and the Columbia Basin could have far reaching implications for renewable energy development in Washington State.

The proposal by the Washington Geological Survey is the first step in evaluating and developing the geothermal resources of the State and we look forward to working collaboratively with the State on this proposal.

Sincerely,

Terry Walton
Director of Energy and Environmental Programs

TLW/TJG:bjw

United States Department of the Interior

U.S. GEOLOGICAL SURVEY

July 20, 2009

Dr. David K. Norman
State Geologist
Washington Geological Survey
Department of Natural Resources
Olympia, Washington

Re: Washington Geothermal Proposal Submittal to Department of Energy DE-FOA-0000109

Dear Dave,

The United States Geological Survey has strong interest in your proposal that you are submitting to the U.S. Department of Energy's Funding Opportunity Announcement DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program—Topic Area 3: Geothermal Data Development, Collection, and Maintenance.* I am particularly enthusiastic to see your proposal to move existing data relevant to understanding geothermal systems to an accessible, digital platform. This is long overdue, not only for geothermal exploration and studies but for a whole host of earth science and, increasingly, biological applications as well.

As you are keenly aware, Washington's geothermal potential for generating electricity is not well studied, but without greatly improved access to key data it will be difficult to get significant interest in taking the first steps toward new assessments. Thus, your first phase is critical to generating interest in pursuing geothermal research and updated assessments in Washington. Having the data readily available for the initial stages of geothermal project development, will is an effective step toward promoting the development of geothermal energy in Washington.

Your second phase, to collect new data needed for assessments follows directly from creating the databases of existing observations. It's clear that lidar has become a key tool for virtually all exploration in earth sciences, and I'm delighted to see a significant component of the funding slated for more acquisition of this data. And there is no question that additional temperature gradient holes are needed. The USGS is eager to assist you by coordinating our planned lidar acquisitions with yours and discussing our continuing expansion of high-resolution aeromagnetic data collection in Washington.

Sincerely,

Craig S. Weaver
Pacific Northwest Earthquake Program Coordinator
United States Geological Survey
@ University of Washington
Seattle, WA 98195

UW Extension WISCONSIN GEOLOGICAL AND NATURAL HISTORY SURVEY

3817 Mineral Point Road
Madison, WI 53705-5100
Tel 608/262.1705
Fax 608/262.8086
www.uwex.edu/wgnhs/

James M. Robertson
Director and State Geologist

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Wisconsin Geological and Natural History Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

The Wisconsin Geological and Natural History Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Wisconsin Geological and Natural History Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

James M. Robertson
Director and State Geologist
Wisconsin Geological and Natural History Survey

UW-Extension provides equal opportunities in employment and programming, including Title IX and ADA.

July 9, 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Wisconsin Geological and Natural History Survey is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance*.

The Wisconsin Geological and Natural History Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The Wisconsin Geological and Natural History Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

David Hart

David Hart
Geophysicist/Hydrogeologist

Joe Manchin III
Governor

Kelley Goes
Secretary, WV Department of Commerce

WEST VIRGINIA DIVISION OF
ENERGY

Jeff Herholdt
Director

July 10, 2009

West Virginia Geological and Economic Survey
Michael Ed. Hohn, director and state geologist
1 Mont Chateau Road
Morgantown, WV 26508-8079

Dear Mike:

The West Virginia Division of Energy endorses the West Virginia Geological and Economic Survey's taking the lead for West Virginia in the Association of American State Geologists proposal for seeking funding from the U.S. Department of Energy to advance the nation's geothermal technology program by developing, collecting, and maintaining geothermal data development, collection, and maintenance.

The Division of Energy and the WVGES are sister agencies under the West Virginia Department of Commerce. A better appreciation of West Virginia's geothermal resources will aid the division in its support of energy-related economic development.

WVGES is ideally situated as our State's lead agency. Its resources, mission, and goals are aligned to maximize participation in this critical project.

The West Virginia Division of Energy will fully commit to supporting the West Virginia Geological and Economic Survey in this effort. We look forward to working with you in this important endeavor.

Sincerely,

John F. "Jeff" Herholdt Jr., director
West Virginia Division of Energy

WASHINGTON STATE DEPARTMENT OF
Natural Resources
Peter Goldmark - Commissioner of Public Lands

Caring for
your natural resources
... now and forever

July 9, 2009

Dr. M. Lee Allison
Director and State Geologist
Arizona Geological Survey
416 W. Congress St., Suite 100
Tucson, Arizona 85701-1381

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The Washington Geological Survey has a long-continuing interest in the development of geothermal energy, and is committed to participating in the collaborative project being proposed by the Arizona Geological Survey in cooperation with the Association of American State Geologists (AASG) through the U.S. Department of Energy's Funding Opportunity Announcement DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program—Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

As a repository for information about Washington's geology and resources, the Washington Geological Survey is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible, and making the entire body of information widely and easily accessible through the National Geothermal Data System. Washington's geothermal potential for generating electricity is not well studied, and furthering our understanding of geothermal energy in Washington is important to the development of this energy resource. This proposed project will use modern information technology to readily transfer relevant data to those interested in geothermal energy development. This approach of making geothermal data broadly available will reduce the risks associated with the initial stages of geothermal project development, thereby effectively promoting the development of geothermal energy in Washington.

The Washington Geological Survey is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this vital program.

Sincerely,

David K. Norman
Washington State Geologist

GEOLOGY AND EARTH RESOURCES DIVISION ▮ 1111 WASHINGTON ST SE ▮ MS 47007 ▮ OLYMPIA, WA 98504-7007
TEL (360) 902-1450 ▮ FAX (360) 902-1785 ▮ TTY (360) 902-1125 ▮ TRS 711 ▮ WWW.DNR.WA.GOV
EQUAL OPPORTUNITY EMPLOYER
RECYCLED PAPER

WEST VIRGINIA
GEOLOGICAL AND ECONOMIC SURVEY

1 Mont Chateau Road
Morgantown, WV 26508-8079

Joe Manchin, III, *Governor*
Kelley Goes, *Secretary, Department of Commerce*
Michael Ed. Hohn, *Director and State Geologist*

Phone: (304) 594-2331
Fax: (304) 594-2575
E-mail: info@geosrv.wvnet.edu
Web Site: http://www.wvgs.wvnet.edu

9 July 2009

Dr. M. Lee Allison
State Geologist & Director
Arizona Geological Survey
416 W. Congress, #100
Tucson, Arizona 85701

Re: AASG Geothermal Proposal Submittal for DE-FOA-0000109

Dear Dr. Allison,

The West Virginia Geological and Economic Survey (WVGES) is committed to participating in the proposed collaborative project that the Arizona Geological Survey is submitting in cooperation with the Association of American State Geologists (AASG) to collect, develop, and maintain data through the U.S. Department of Energy's Funding Opportunity Announcement #DE-FOA-0000109, *Recovery Act: Geothermal Technologies Program, Topic Area 3: Geothermal Data Development, Collection, and Maintenance.*

WVGES is dedicated to advancing its efforts to collect, develop, and maintain geothermal data to the maximum extent possible and making the entire body of information widely and easily accessible to the public through continued data submission to the National Geothermal Data System. This approach will effectively promote the development of geothermal energy applications and reduce the risks associated with the initial stages of related project development.

The WVGES is pleased to participate in this project as a subawardee and is looking forward to collaborating with you, AASG, the other state geological surveys, the U.S. Geological Survey, and other partners in this important initiative.

Sincerely,

Michael Ed. Hohn
Director and State Geologist

D. H. Jones & Son Heating and Cooling

100 Stuart Street St. Albans, WV 25177
(304) 727-4811 Fax (304) 727-0329

July 16, 2009

Dr. Michael Ed Hohn, Director and State Geologist
West Virginia Geological and Economic Survey
1 Mont Chateau Road
Morgantown, WV 26508-8079

Dear Dr. Hohn,

D. H. Jones and Son Heating and Cooling would be honored to endorse West Virginia Geological and Economic Survey taking the lead for West Virginia in the Association of American State Geologists proposal for seeking funding from the U.S. Department of Energy to advance the nation's geothermal technology program by developing, collecting and maintaining geothermal data development, collection, and maintenance.

I, David A. Jones, in 1993 attended (IGSHPA #ID11247-1193) schooling in Florida at the FHP (Florida Heat Pump) program. The class was by the International Ground Source Heat Pump Association. I am proud to say that D.H. Jones and Son have installed approximately three hundred (300) Geothermal Units in the southern parts of West Virginia. We have installed many different types of applications, from closed loop systems to ponds and open water wells or "pump and dump" systems, and in many different types of soils. In the world of Geothermal heating and cooling, I have been introduced to many different types of installations, from a home built into a cave, to a home that we installed a 19-Ton Cooling unit in, just to name a few. We would also like to address the importance of the duct system, which plays a big part of the efficiency of any heating and cooling equipment, forced air units as well as Geothermal. For example, you can have a perfect heart, but if your arteries are not in good shape, your heart will not function properly.

Geothermal systems represent a savings to homeowners of around 70% in the heating mode, and up to 50% in the cooling mode compared with conventional systems. Conservative estimates show that for every unit of energy into a geothermal system, 3 additional free units of energy are transferred from the ground -- 400% verses a high efficiency gas furnace at around 90%. Geothermal systems are the most environmentally clean and cost-effective space conditioning systems. Geothermal qualifies as a renewable energy - a typical two bedroom house-sized installation is equivalent to taking two cars off the road or planting one acre of trees. Geothermal systems work toward the preservation of the environment by minimizing present environmental problems like acid rain, air pollution, and the destruction of the ozone layer.

Sub-Contractor Collaboration

State	Listed Collaborators	Commitment	Summary of Collaborator Support
Arkansas	USGS AR Water Science Center	Assist AGS in providing groundwater (well and spring) temperature data and borehole geophysical log information	USGS was willing to submit data but was told these data would be redundant, as they had already been collected. Submitted instead some brine inndustry data and thermal conductance readings performed on core stored at the AGS.
Arkansas	AR Oil and Gas Commission	general support	Project received much of its data from AR Oil and Gas Commission well records, which formed the bulk of the NGDS data set contributed by AR.
Maine	Maine Governor's Office of Enery Independence and Security	general support	No action mentioned.
Massachusetts	MA Clean Energy Center	provide connections with industrial partners and others who may have geothermal data	
New Hampshire	State of NH Office of Energy and Planning	general support	
Washington	US Geological Survey	coordinating USGS planned lidar acquisitions with the Survey's and exploring continuing expansion of high-resolution aeromagnetic data collection in WA	Collaborated on Lidar acquisition and discussed expansion of aeromag but no funding was acquired for this
Washington	Pacific Northwest National Laboratory	provide expertise in subsurface science and knowledge of Columbia Basin; access to a significant source of unpublished geological data from the DOE over several decades	Collected subsurface data in the Columbia Basin as indicated, and many people have used the data through the NGDS
Washington	Weyerhaeuser	general expertise and support	Ongoing meetings and sharing of knowledge and expertise
Washington	Skaminia County Board of Commissioners	active participation and cooperation	Currently undertaking an active geothermal evaluation project in Skamania County
Washington	Snohomish County Public Utility District No 1	general support	Past general support and cooperation; no recent conversation with PUD.

Sub-Contractor Collaboration

State	Listed Collaborators	Commitment	Summary of Collaborator Support
West Virginia	D.H. Jones and Son Heating and Cooling	general endorsement of heating exchange systems	Interest in project waned when focus turned from low to high -temp potential geothermal resources to EGS enterprise. No contact since Y1.
West Virginia	WV Division of Energy	commitment to support and partner, especially to improve public access to data	After funding to build a local website to display temperature at depth mapping analyses was denied by NGDS, a separate proposal to the WV Division of Energy to fund the data analyses and website, encouraged by the agency, was funded to perform those tasks independent of NGDS data submission. (West Virginia Division of Energy, Award No. DE-EE0000838). The website was finished in 2012 (http://ims.wvgs.wvnet.edu/WVGeothermal/viewer.htm)

Arkansas

From: Bill Prior [mailto:Bill.Prior@arkansas.gov]
Sent: Tuesday, December 02, 2014 12:03 PM
To: Randi Bellassai
Subject: RE: Arkansas letter of collaboration

Dear Mr. Bellassai

The short answer to your 2 comments is that there has been no further responses from either the U.S.G.S Water Division or from the Arkansas Oil and Gas Commission concerning the NGDS or of geothermal energy in general here in Arkansas. During the NGDS project the Arkansas Geological Survey (AGS) did received much of its data from the Arkansas Oil and Gas Commission well records. These files form the bulk of the NGDS data set for Arkansas.

Although the U.S.G.S. was a strong supporter of the NGDS Project and would have been more than willing to supply their data we at the AGS were informed shortly after the projects start date that providing U.S.G.S. data would be redundant to submit U.S.G.S. data. So no U.S.G.S. data was submitted instead some brine industry data was allowed to be submitted instead along with thermal conductance readings we performed on core stored here at the AGS.

Since the end of 2013 we have received a handful of inquiries from the public and industry about the NGDS and of the deep brines in southern Arkansas.

To date this all that I am aware of that has happen here in Arkansas. The use of geothermal heat pumps is something that is promoted by the Arkansas Electric Cooperatives.

I hope I have answered your questions.

Regards

Bill Prior

From: Marvinney, Robert G. [mailto:Robert.G.Marvinney@maine.gov]
Sent: Thursday, December 11, 2014 1:54 PM
To: Randi Bellassai
Subject: RE: Maine Letter of Collaboration

Dear Randi,

With regard to the Maine Geological Survey's contributions to the NGDS, I really have nothing to report. In Maine, use of geothermal resources is exclusively through the development of ground-coupled heat pumps for residential and commercial buildings. This activity is conducted through private businesses, which may or may not access the NGDS, but I have no information on that. My guess is that they haven't accessed the database.

We are still in the process of creating a new heatflow map for Maine, incorporating our new data. When that is complete, I'm sure there will be more research interest in the broader resources of the region.

Best wishes for the Holidays.

Bob

Robert G. Marvinney, Ph.D.
State Geologist, Maine Geological Survey
Director, Bureau of Resource Information and Land Use Planning
Maine Department of Agriculture, Conservation, and Forestry
93 State House Station
Augusta, ME 04333

Direct line: 207-287-2804
Main line: 207-287-2801
Fax: 207-287-2353
robert.g.marvinney@maine.gov
http://www.maine.gov/dacf/mgs/index.shtml

WASHINGTON STATE DEPARTMENT OF
Natural Resources

December 9,2014
Randi S. Bellassai
Project Manager
Arizona Geological Survey
416 W. Congress St.
Tucson, AZ 87501

Dear Randi S. Bellassai:

The National Geothermal Data System was a boon for digital data compilation in Washington State, not only making it easier for those interested in geothermal exploration to access critical data, but for the rest of the geologic community as well, as many of the datasets created during the project have direct application in many other geologic disciplines.

Thermal and mineral spring chemistry data collected for the NGDS project has been used recently by scientists with the USGS studying noble gas isotopes in the Cascadia Forearc. Other USGS scientists used heat flow data and geothermal gradients from our NGDS data to study heat flow loss in the Columbia Basin. New geologic mapping in the Wind River valley near the Columbia River recently facilitated the siting of wells that have elevated warm water temperatures. Geochronology and rock geochemistry data is currently being harvested from the NGDS to help compile the 1: 100,000-scale geologic map of the Tacoma quadrangle and several 7.5-minute quadrangles. Private geothermal exploration companies also use our data to identify potential resources in their geothermal lease holdings in the state as well. Additionally the NGDS is being used

extensively in a government-industry-academia a Geothermal Play Fairway Analysis to evaluate geothermal potential in Washington.

All in all we think the NGDS is a very worthwhile and has been extremely successful.

Sincerely,
David K. Norman
State Geologist/Division Manager

West Virginia

From: Ronald McDowell [mailto:mcdowell@geosrv.wvnet.edu]
Sent: Tuesday, December 09, 2014 9:50 AM
To: Randi Bellassai
Cc: 'Mike Hohn'
Subject: RE: West Virginia letters of collaboration

Ms. Bellassai,

I spoke with the original PI for the Geothermal Project regarding parties who wrote letters of support during the proposal phase. In the case of the first correspondent, D. H. Jones and Son, their expressed interest in the Project related specifically to low-temperature geothermal applications. As a result, our Agency's proposal and initial workplan addressed the full spectrum of potential geothermal resources from low- to high-temperature. However, during the first year of the Project, it became clear that focus had shifted to Enhanced Geothermal resources and interest from low-temperature commercial enterprises rapidly waned. We have had no contact with D. H. Jones and Son since the first year of the Project.

In the case of the second correspondent, Mr. Jeff Herholdt of the West Virginia Division of Energy, a major concern was always that data generated by our Agency would become available in a fashion that was readily accessible and readily usable by the general public. To this end, our Agency proposed to create a "local" website that would display the results of data analyses in the form of temperature-at-depth mapping based on modeling of data we were submitting to NDGS. The Project's Scientific Advisory Board (SAB) rejected funding for the creation of such a website and attempts to perform our own analyses of the data. These tasks were removed from our revised workplan. However, Mr. Herholdt continued to be interested in this aspect of the Project and requested that our Agency prepare a separate proposal to the West Virginia Division of Energy to fund the data analyses and website creation. We did so and were given State funding (West Virginia Division of Energy, Award No. DE-EE0000838) to perform these tasks independently of NGDS data submission. The website was finished in 2012 (http://ims.wvgs.wvnet.edu/WVGeothermal/viewer.htm).

I hope this is the information you were looking for. I would be glad to answer additional questions if I am able.

Sincerely,

Ronald McDowell

Appendix D: NGDS Catalog System

1. Data Validation Tutorials and Tools
    a. QA/QC documentation
        i. GeoSciMLPortrayalCookbook……2
        ii. Data Development Cycle (graphic)……39
    b. QA/QC tools- Uploaded separately to OSTI
        i. USGIN Models API: https://github.com/usgin/usginmodels
        ii. Content model Spreadsheet Validator: http://schemas.usgin.org/validate/cm
        iii. Web Feature Service validator at http://schemas.usgin.org/validate/wfs
        iv. Excel to NGDS Services Tool: https://github.com/usgin/ExcelToNGDSServiceTool
        v. NGDS data archiver: https://github.com/ngds/data-archiver
2. Conformance Tutorials, Documentation, and Tools
    a. Tutorials & Documentation
        i. USGIN metadata profile documentation……40
        ii. XML Tutorial……153
        iii. XML Schema Profile for Simple Feature WFS Services……163
        iv. Import of Metadata to NGDS catalog……168
        v. USGIN Identifier (URI) Conventions Tutorial……185
        vi. USGIN Lab Discussions……192

        Uploaded separately to OSTI:

        vii. Metadata repository: https://github.com/usgin/metadata-repository
        viii. Metadata compilation: (see http://schemas.usgin.org/models/#Metadata)
    b. Tools
        i. Django, Node.js, and CouchDb metadata standardization and staging system: https://github.com/usgin/metadata-server
        ii. CSV to XML python tool: https://github.com/usgin/usgincsv

GeoSciML-Portrayal

An introduction to a simple feature schema for supporting map services and guidelines for service deployment

A product of the Arizona Geological Survey

Table of Contents

Table of Contents 1
1 Introduction 3
1.1 What is GeoSciML? 3
1.1.1 What is GeoSciML-Portrayal? 3
1.1.2 GeoSciML Specifications 3
1.2 What is USGIN? 4
1.3 Cookbook Prerequisites 4
1.3.1 Prerequisite Notes 5
1.4 Cookbook Workflow 5
2 GeoSciML-Portrayal Feature Types 6
2.1 ContactView Features 6
2.1.1 ContactView Feature Specifications 7
2.2 ShearDisplacementView Features 7
2.2.1 ShearDisplacementStructureView Feature Specifications 9
2.3 GeologicUnitView Features 9
2.3.1 GeologicUnitView Feature Specifications 11
3 Integration of Web Map Services 12
4 Deployment Cookbook 12
4.1 Mapping from source data 12
4.1.1 Interchange formats 13
4.1.2 Schema mapping in ESRI ArcGIS 13
4.1.3 Schema mapping in SQL 15
4.1.4 Notes on schema mapping in SQL 16
4.2 Vocabulary mapping 17
4.3 Excel Workbook template 18
4.4 Configuring OGC services 18
5 Styling 19
5.1 How to Create a Styled Layer Descriptor (SLD) using Arc2Earth 19
6 Deploying Your Web Service 20
7 Testing Your Web Service 21
7.1 Web Service Requests 21
7.2 Adding a Web Service in ArcCatalog 22
7.3 Adding a Web Service in uDig 22
1

8 Registering with OneGeology and Submitting your Service 23
9 Appendix A: Deploying GeoSciML-Portrayal Web Services in GeoServer 23
9.1 Logging in 23
9.2 Service-Level Metadata 24
9.3 Creating a Workspace 24
9.4 Connecting to your PostGIS Database 25
9.5 Adding Layers to a Workspace 26
9.5.1 The Data tab of the Edit Layer page 27
9.5.2 The Publishing tab of the Edit Layer page 27
9.5.3 Importing Layer Styles from an SLD File 27
9.5.4 Finishing Up 27
9.6 GeoServer Troubleshooting 27
10 Appendix B: Deploying GeoSciML-Portrayal Web Services in ArcGIS Server 29
10.1 Connecting to your ArcGIS Server instance 29
10.2 Create an ArcMap Project for your web service 29
10.3 Publishing on ArcGIS Server 29
10.4 Troubleshooting ArcGIS Server 30
11 Appendix C: OneGeology Web Service Requirements 30
11.1 Service-Level Metadata Requirements 30
11.2 Service Title and Service Name Requirements 31
11.2.1 OneGeology Service Title Examples: 31
11.3 Layer Naming and Title Conventions 31
11.3.1 Layer Title Examples 32
11.3.2 Layer Name Examples 32
11.3.3 A sample WMS GetCapabilities document demonstrating Layer Names and Layer Titles 32
11.4 Additional Layer-level Metadata Requirements 32
12 Glossary 34

1 Introduction

This document is a cookbook designed to facilitate the deployment and maintenance of geoscience information as interoperable, publicly available OGC web services in GeoSciML-portrayal in accordance with USGIN and OneGeology specifications.

To facilitate interoperability, it is possible to perform all of the steps indicated in this cookbook with free-and-open-source software.

1.1 What is GeoSciML?

GeoSciML stands for Geoscience Markup Language.

GeoSciML was developed to support interoperable map services, these services delivering geologic map unit, contact, and shear displacement structure (fault and ductile shear zone) portrayals. Ideally, use of standard vocabulary enables map portrayal using shared legends to achieve basic visual harmonization of maps, even when these maps are provided by different services.

In other words: GeoSciML is a specialized markup language by which geoscience data can be shared and portrayed interoperably, as web services, using shared legends.

1.1.1 What is GeoSciML-Portrayal?

GeoSciML-portrayal is a simplified **view** of GeoSciML (Figure 1).

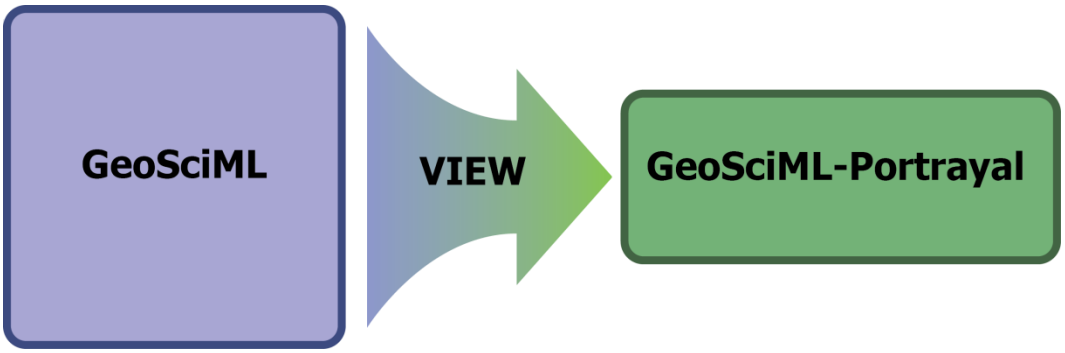


Figure 1: A visual representation of the relationship between GeoSciML and GeoSciML-Portrayal

In the context of a database, a **view** is a selection of fields. For example: given a database with twelve fields, an arbitrary grouping of four such fields would constitute a **view** of the database. A more concrete example is a card catalog: if one chose to look *only* at the **Title** field of each card in the catalog, this choice would constitute a discrete **view** of the catalog.

Markup languages can and do function similarly to databases, wherein markup language tags are analogous to database fields. Consequently, it is possible to create the equivalent of a **view** for a markup language.

The **GeoSciML-portrayal** is the equivalent of a **view** of the GeoSciML markup language: a specialized, distilled form of the full GeoSciML markup language.

Consequently, GeoSciML-portrayal can be considered a derivative markup language. As such, documents written using the GeoSciML-portrayal markup language can be structured by schemas. In part, the GeoSciML-portrayal cookbook will describe the GeoSciML-portrayal schema that structures documents written using the GeoSciML-portrayal markup language.

1.1.2 GeoSciML Specifications

GeoSciML is based on XML and developed as a GML v3.2 application for encoding a wide variety of geoscientific information (Richard et al., 2007); GeoSciML defines a collection of complex GML features for describing geoscience entities. Use of standard vocabulary enables map portrayal using shared legends to achieve some basic visual harmonization of maps provided by different services. In addition the scheme includes text information for human users when browsing a geologic map, a link to a full GeoSciML feature element if available, and a symbol identifier field to enable a user-defined symbolization scheme in each map service. Most WMS implementations simplify deployment of WFS service, as the same schema is used. Linking the simple feature WMS and WFS allows clients to acquire basic text geologic feature descriptions that can be used in web-mapping applications to construct custom legends. Linking to full GeoSciML features allows the portrayal scheme to be used in a map browsing and query interface to identify and select features for further processing that can be acquired as highly structured, information-rich, complex GML features.

The GeoSciML-portrayal schema is separate from, but harmonized with, GeoSciML v3.0. GeoSciML-portrayal conforms to the level 0 of the Simple Features Profile for GML (OGC 10-100r3 - van den Brink et al., 2011; OGC 06-049). The simple features profile supports only a limited subset of possible GML geometry types that may be used to describe feature geographic location and shape. For the purposes of GeoSciML simple features, these include gml:Point, gml:LineString, gml:Curve, gml:Polygon, gml:Surface, gml:MultiPoint, gml:MultiCurve, gml:MultiSurface and multi-geometry types consisting of collections of these base types. For a useful discussion of GML simple vs. complex features, see this Geoserver documentation page.

GeoSciML-Portrayal features are analogous to GeoSciML mapped features, with additional text attributes for human consumption, a flatted-relation view of the age, and assignment to a single lithology. The portrayal scheme consists of 'free-text' fields and identifier fields. In robust services the free-text fields will contain well-structured summaries of complex GeoSciML data in a format suitable for reading by the intended users. Identifier fields should contain identifiers for concepts in a controlled vocabulary (for example CGI Simple Lithology) that specify representative thematic properties. Inclusion of these standardized identifiers enables interoperability across services. Ideally these should be URIs that can be dereferenced to obtain machine-processable or human-readable representations of the identified concepts.

In addition, each GeoSciML-portrayal feature includes an (optional) identifier for a specification, which is a resource containing a description of that particular feature. In many cases, the descriptions will be the same for all polygons assigned to the same map unit or classified as the same kind of contact or structure. If more complete information is available, different descriptions may be associated with subsets of features of the same type that are portrayed with the same symbol. In the most extreme case, each feature might have a unique description that captures the full spatial variability of a geologic unit or structure. Following the standard patterns of web architecture, the specification_uri should be dereferenceable to obtain one or more representations of that description. For maximum interoperability, one of these representations should be a GeoSciML encoded description of the feature, but other encodings might also be available, for example html web pages, other XML schema, or JSON. For those familiar with full GeoSciML v3.0, the specification_uri property is equivalent to the specification association from MappedFeature to GeologicFeature.

Deployment of an interoperable dataset requires first that the interchange format used for interoperability is well understood. The following sections are intended to provide the necessary background understanding of GeoSciML-portrayal, and should be studied carefully. The deployment process consists of determining how the information in the dataset to be published can best be represented using the elements in the GeoSciML-portrayal model, populating the feature collections that will be served, and configuring a WMS server to display the data.

1.2 What is USGIN?

The United States Geoscience Information Network (USGIN) initiative is the product of a partnership between the Association of American State Geologists (AASG) and the United States Geological Survey (USGS) created to facilitate discovery of, and access to, geoscience information provided by state and federal geological surveys of the United States.

USGIN is:

- conceptual framework for sharing geoscience data
- a distributed data-sharing network
- a collection of open-source applications, standards, procedures, and protocols for sharing geoscience data
- web-based, distributed, open-source, and interoperable

In other words: USGIN is a distributed data-sharing network that uses open-source software and existing World Wide Web infrastructure and browsers. Anyone can access the data shared over USGIN, and any data that is shared according to USGIN specifications is automatically a part of the network.

For more information about USGIN, see the USGIN website (http://usgin.org/).

1.3 Cookbook Prerequisites

To comply with this cookbook, you will need the following:

1. Geologic data in tabular form; valid data formats include:
    a. File geodatabases
    b. Microsoft Access databases
    c. Personal geodatabases
    d. PostGIS databases

    e. SDE databases
    f. Shapefiles
2. Server software capable of fielding web service requests; examples include:
    a. GeoServer and accompanying software
    b. ESRI ArcGIS Server (proprietary)
3. GIS software; examples include:
    a. ArcGIS
    b. Grass GIS
    c. gvSIG
    d. OpenLayers
    e. PostGIS plugin for PostgreSQL DBMS
    f. Quantum GIS
    g. uDIG
4. GeoSciML-portrayal cookbook (this document)
5. An Internet connection to deploy your data as a publicly accessible web service

1.3.1 Prerequisite Notes

The following section provides supplementary background information about topics introduced in Section 1.3:

- As indicated in Section 1.3, Cookbook Prerequisites, **GeoServer** and **ESRI ArcGIS Server** both constitute server software capable of providing web services
    - **ESRI ArcGIS Server** is proprietary server software
    - **GeoServer** is actively maintained, free-and-open-source server software
    - **GeoServer** can be found at the following web locations:
        - http://geoserver.org/
        - http://opengeo.org/technology/GeoServer/
    - **GeoServer** can ingest a **PostgreSQL** database that is **PostGIS**-enabled
- **PostGIS** is a plugin for **PostgreSQL** database management system (DBMS)
- **PostgreSQL** DBMS is a free-and-open-source database management system
- A database management system is a program that controls the creation, maintenance, and use of a database; numerous examples exist, including:
    - Relational DBMS, such as MRDS, Rel, and Micro DBMS
    - SQL DBMS, such as INGRES, DB2, and Microsoft SQL Server
        - SQL DBMS are so named because they use Standardized Query Language (SQL)
    - NoSQL DBMS, such as memcached, Redis, and CouchDB
        - NoSQL DBMS do not use SQL as their primary **query** language (NoSQL means "not only SQL")

For back-end server management, the Arizona Geological Survey uses pgAdmin III (included in the software bundle from opengeo.org). pgAdmin III provides the user with an interface to create and edit databases as well as data tables. An introduction to pgAdmin as well as documentation for using the software can be found on their website here.

For information on PostGreSQL and deploying services using multiple geodatabases, click here.

1.4 Cookbook Workflow

This cookbook will encompass the following overall steps:

1. Map your data from its source schema to an appropriate GeoSciML-portrayal schema; this step can be further broken down into the following sub-steps
    a. Map your data from its source database to an interchange format structured according to the appropriate GeoSciML-portrayal schema
    b. Export your data from the interchange document to a database format appropriate for the server software you will use to deploy your web service
2. Define appropriate symbology for your web service
3. Deploy your web service and, if desired, register it with OneGeology

2 GeoSciML-Portrayal Feature Types

Within the scope of the GeoSciML-portrayal markup language, three feature classes are defined:

- **ContactView**: contains features that represent the mapped traces of boundaries between two geologic units
- **ShearDisplacementStructureView**: used to represent the mapped trace of any type of fault or shear zone that is treated as a single surface for map portrayal
    - The concept of 'Shear displacement structure' includes all fault types or shear zones along which displacement has occurred, from a simple, single 'planar' brittle or ductile surface to a fault system com-prising multiple strands of both brittle and ductile nature. This structure may have some significant thickness (a deformation zone) and have an associated body of deformed rock that may be considered a DeformationUnit. However ,because this feature class is constrained to have a linear geometry, representation is limited to shear displacement structures that are considered single surfaces at the scale of portrayal.
- **GeologicUnitView**: contains polygon features that represent the outcrop of a geologic unit mapped on some outcrop surface called the 'map horizon'. Example map horizons include:
    - Earth surface: the most common map horizon for a geologic map
    - Top of basement: bedrock on which a stratographic section of interest is deposited. For example, in the Grand Canyon, the Precambrian crystalline rocks in the inner gorge (e.g. Vishnu schist) would be considered the basement.
    - Mission-Pima pit, 6/20/1990

  Since the map horizon is not specified in each feature's properties, it must be described in the metadata for the feature collection. Overlapping polygons representing outcrops on different horizons (e.g. 'Earth surface' and 'bedrock surface') will be represented as distinct features in different GeoSciML-portrayal services. Each GeoSciML-portrayal service provides geologic unit outcrop polygons, associated contacts between the units, and fault traces on a single map horizon at any particular location in the extent of the feature collection.

Each of the above feature classes contains features that provide a simplified **view** of existing GeoSciML features. In other words: the above GeoSciML-portrayal feature classes contain features that are also contained by GeoSciML features; but since GeoSciML-portrayal is a simplified **view** of GeoSciML, GeoSciML-portrayal features are simpler than their GoeSciML counterparts.

For more information about the relationship between GeoSciML and GeoSciML-portrayal, refer back to the GeoSciML-Portrayal section.

2.1 ContactView Features

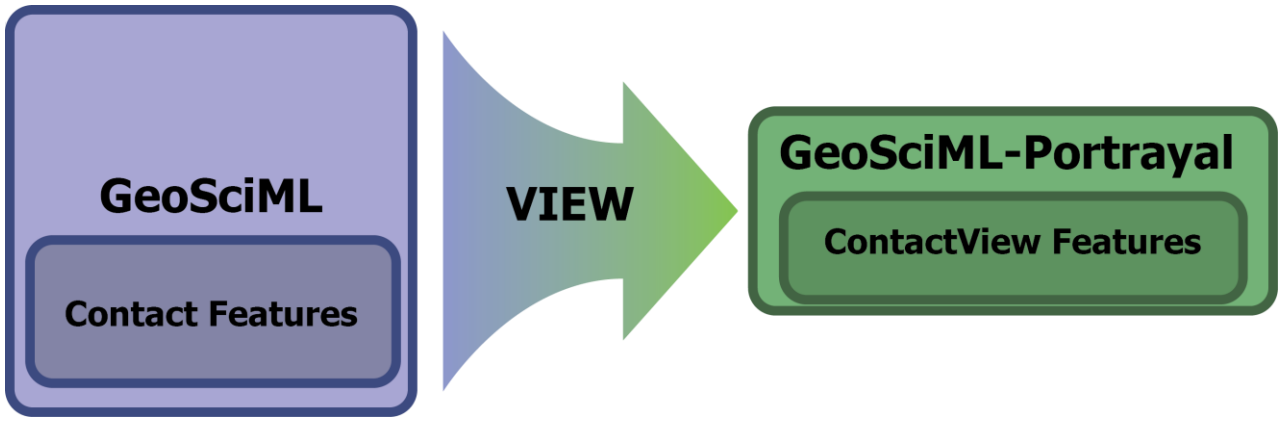


Figure 2: The relationship between GeoSciML Contact Features and GeoSciML ContactView Features

These features provide a simplified **view** of GeoSciML Contact Features (Figure 2).

Table 1 describes each of the elements that describe **ContactView Features**; in this table:

- **Name** is the name of an element that describes a given feature
- **Implementation data type** indicates the specifications for a given element:
    - AnyURI: This element will contain an appropriate URI
        - URIs can be used to refer to other resources, such as controlled concepts in published vocabularies

- String: This element will contain a free-text series of alphanumeric characters
- Decimal: This element will contain a numeric value

Note that these values are modified by the **xs:** prefix; the xs: prefix indicates conformity to the W3C XML Schema

- **Notes** contains a general description of a given element

Table 1: ContactView feature scheme

Name	Implementation data type	Notes
identifier	xs:AnyURI	Globally unique identifier for the individual feature. Recommended practice is that this identifier be derived from the primary key for the spatial objects in the source data in case information needs to be transferred from the interchange format back to the source database. This identifier is analogous to the identifier for a GeoSciML MappedFeature.
name	xs:string	Display name for the Contact. Examples: 'depositional contact', 'unconformity', 'Martin-Escabrosa contact.'
description	xs:string	Text description of the contact; may be a generic description of a contact type taken from an entry on a geological map legend or a more specific description of the particular contact.
contactType	xs:string	Text label specifying the kind of surface separating two geologic units including primary boundaries such as depositional contacts, all types of unconformities, intrusive contacts, and gradational contacts, as well as faults that separate geologic units. Ideally this would be the preferred label for the concept identified by contactType_uri.
observationMethod	xs:string	Metadata snippet indicating how the spatial extent of the feature was determined. ObservationMethod is a convenience property that provides a quick and dirty approach to observation metadata.
positionalAccuracy	xs:decimal	Quantitative values define the radius of an uncertainty buffer around a mappedFeature (e.g. a positionAccuracy of 100 m for a line feature defines a buffer polygon of total width 200 m centered on the line).
source	xs:string	Text describing feature-specific details and citations to source materials, and if available providing URLs to reference material and publications describing the geologic feature. This could be a short text synopsis of key information that would also be in the metadata record referenced by metadata_uri.
contactType_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the Contact types. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
specification_uri	xs:AnyURI	URI referring the GeoSciML Contact feature that describes the instance in detail. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
metadata_uri	xs:AnyURI	URI referring to a formal metadata record describing the provenance of data.
genericSymbolizer	xs:string	Identifier for a symbol from standard (locally or community defined) symbolization scheme for portrayal. There should be an accompanying SLD file. available defining the symbol associated with each genericSymbolizer value.
shape	GM_Object	Geometry defining the extent of the feature of interest. This is the only element with complex content, and must contain a GML geometry that is valid for the Geography Markup Language (GML) simple features profile (OGC 06-049r1.)

2.1.1 ContactView Feature Specifications

In GeoSciML terms, GeoSciML-portrayal ContactView Features will be an instance of a GeoSciML MappedFeature with key property values from the associated ContactFeature summarized in text (data type xs:string) fields as well as properties suffixed with '_uri' containing URIs referring to other resources (for example, controlled concepts in published vocabularies).

2.2 ShearDisplacementView Features

7

These features provide a simplified **view** of GeoSciML ShearDisplacementStructure Features (Figure 3).

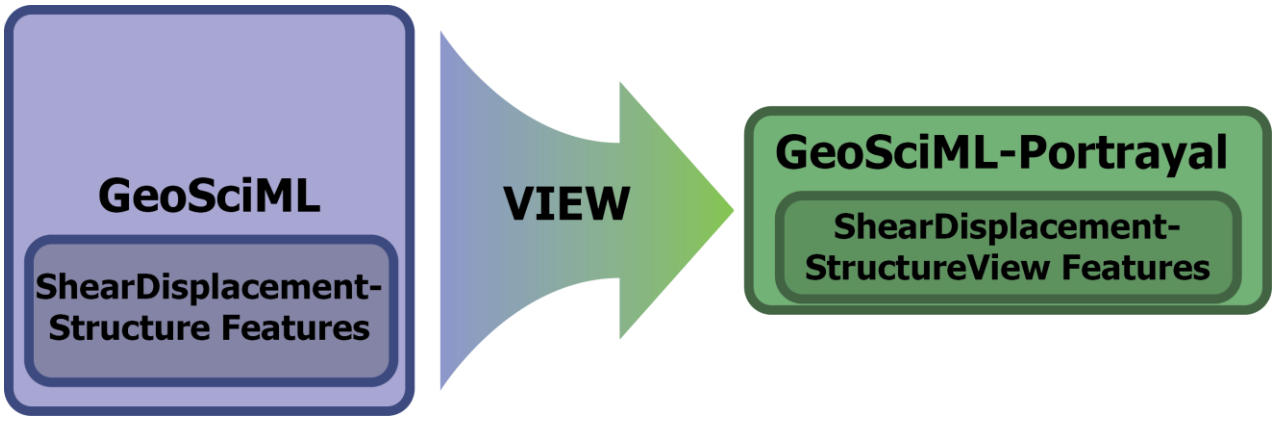


Figure 3: The relationship between GeoSciML ShearDisplacementStructure Features and GeoSciML ShearDisplacementStructureView Features

Table 2 describes each of the elements that describe **ShearDisplacementStructure Features**; in this table:

- **Name** is the name of an element that describes a given feature
- **Implementation data type** indicates the specifications for a given element:
    - AnyURI: This element will contain an appropriate URI
        - URIs can be used to refer to other resources, such as controlled concepts in published vocabularies
    - String: This element will contain a free-text series of alphanumeric characters
    - Decimal: This element will contain a numeric value

  Note that these values are modified by the **xs:** prefix; the xs: prefix indicates conformity to the W3C XML Schema
- **Notes** contains a general description of a given element

Table 2: Elements in ShearDisplacementStructureView feature scheme

Name	Type	Notes
identifier	xs:AnyURI	Globally unique identifier for the individual feature. Recommended practice is that this identifier be derived from the primary key for the spatial objects in the source data in case information needs to be transferred from the interchange format back to the source database. This identifier is analogous to the identifier for a GeoSciML MappedFeature.
name	xs:string	Display name for the the ShearDisplacementStructure. This may be a generic fault type (e.g. 'thrust fault', 'strike-slip fault') or a particular fault name (e.g. 'Moine thrust', 'San Andreas Fault').
description	xs:string	Text description of the ShearDisplacementStructure, typically taken from an entry on a geological map legend.
faultType	xs:string	Type of ShearDisplacementStructure (as defined in GeoSciML).
movementType	xs:string	Summary of the type of movement (e.g. dip-slip, strike-slip) on the ShearDisplacementStructure.
deformationStyle	xs:string	Description of the style of deformation (e.g. brittle,ductile etc) for the ShearDisplacementStructure.
displacement	xs:string	Text summary of displacement across the ShearDisplacementStructure.
geologicHistory	xs:string	Text (possibly formatted with formal syntax) description of the age of the ShearDisplacementStructure (where age is a sequence of events and may include process and environment information).
observationMethod	xs:string	Metadata snippet indicating how the spatial extent of the feature was determined. ObservationMethod is a convenience property that provides a quick and dirty approach to observation metadata when data are reported using a feature view (as opposed to observation view).

Name	Type	Notes
positionalAccuracy	xs:string	Quantitative values which define the radius of an uncertainty buffer around a mappedFeature (eg: a positionAccuracy of 100 m for a line feature defines a buffer polygon of total width 200 m centered on the line).
source	xs:string	Text describing feature-specific details and citations to source materials, and if available providing URLs to reference material and publications describing the geologic feature. This could be a short text synopsis of key information that would also be in the metadata record referenced by metadata_uri.
faultType_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the fault type (ShearDisplacementStructure). Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
movementType_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the ShearDisplacementStructure movement type. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
deformationStyle_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the ShearDisplacementStructure deformation style. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
representativeAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative stratigraphic age interval for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing the all or part of the feature's history.
representativeLowerAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative lower value in a range of stratigraphic age intervals for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing the all or part of the feature's history.
representativeUpperAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative upper value in a range of stratigraphic age intervals for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing the all or part of the feature's history.
specification_uri	xs:AnyURI	URI referring the GeoSciML ShearDisplacementStructure feature that describes the instance in detail. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
metadata_uri	xs:AnyURI	URI referring to a metadata record describing the provenance of data.
genericSymbolizer	xs:string	Identifier for a symbol from standard (locally or community defined) symbolization scheme for portrayal.
shape	GM_Object (GM_curve)	Geometry defining the extent of the feature of interest.

2.2.1 ShearDisplacementStructureView Feature Specifications

In GeoSciML terms this will be an instance of a MappedFeature with key property values from the associated ShearDisplacementStructure feature summarized in text fields (data type xs:string) as well as fields containing identifiers (URI) for fault type, deformation style, movement type, geologic age, and a formally-encoded (ideally in GeoSciML) specification for interoperability. The latter are the properties suffixed with '_uri' and will contain URIs referring to other resources, for example controlled concepts in published vocabularies.

2.3 GeologicUnitView Features

9

These features provide a simplified **view** of GeoSciML GeologicUnit Features (Figure 4), which are used to provide properties associated with a body of material in the earth.

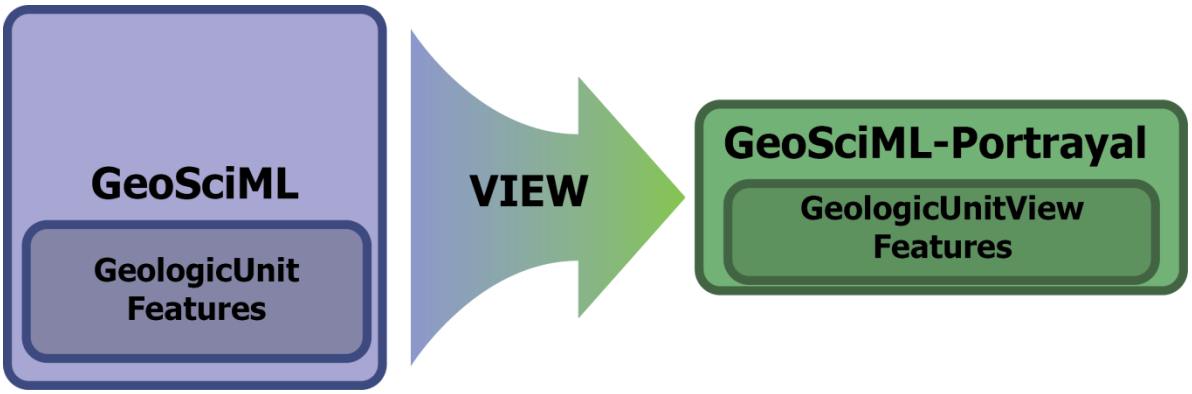


Figure 4: The relationship between GeoSciML GeologicUnit Features and GeoSciML GeologicUnitView Features

Table 3 describes each of the elements that describe **ShearDisplacementStructure Features**; in this table:

- **Name** is the name of an element that describes a given feature
- **Implementation data type** indicates the specifications for a given element:
  - AnyURI: This element will contain an appropriate URI
    - URIs can be used to refer to other resources, such as controlled concepts in published vocabularies
  - String: This element will contain a free-text series of alphanumeric characters
  - Decimal: This element will contain a numeric value

  Note that these values are modified by the **xs:** prefix; the xs: prefix indicates conformity to the W3C XML Schema
- **Notes** contains a general description of a given element

Table 3: Elements in GeologicUnitView feature class

Name	Type	Notes
identifier	xs:AnyURI	Globally unique identifier for the individual feature. Recommended practice is that this identifier be derived from the primary key for the spatial objects in the source data in case information needs to be transferred from the interchange format back to the source database. This identifier is analogous to the identifier for a GeoSciML MappedFeature.
name	xs:string	Display name for the GeologicUnit; this may be populated by a geologic unit name, or more likely, an abbreviation used to label outcrops of the unit in a map display.
description	xs:string	Text description of the GeologicUnit, typically taken from an entry on a geological map legend.
geologicUnitType	xs:string	Type of GeologicUnit (as defined in GeoSciML).
rank	xs:string	Stratigraphic rank of GeologicUnit (as defined in GeoSciML). Examples include formation, member, group, or supergroup.
lithology	xs:string	Text (possibly formatted with formal syntax) description of the GeologicUnit's lithology.
geologicHistory	xs:string	Text (possibly formatted with formal syntax) description of the age of the GeologicUnit (where age is a sequence of events and may include process and environment information).
observationMethod	xs:string	Metadata snippet indicating how the properties of the feature were determined. ObservationMethod is a convenience property that provides a quick and dirty approach to observation metadata. Example values might include 'field observation by author', 'compilation from published maps', 'air photo interpretation'.

Name	Type	Notes
positionalAccuracy	xs:string	Quantitative values define the radius of an uncertainty buffer around a mappedFeature (eg: a positionAccuracy of 100 m for a line feature defines a buffer polygon of total width 200 m centered on the line).
source	xs:string	Text describing feature-specific details and citations to source materials, and if available providing URLs to reference material and publications describing the geologic feature. This could be a short text synopsis of key information that would also be in the metadata record referenced by metadata_uri.
geologicUnitType_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the GeologicUnit types. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
representativeLithology_uri	xs:AnyURI	URI referring to a controlled concept specifying the characteristic or representative lithology of the unit. This may be a concept that defines the super-type of all lithology values present within a GeologicUnit or a concept defining the lithology of the dominant CompositionPart (as defined in GeoSciML) of the unit. Typically this identifier might be used as the symbol key for a lithologic map portrayal of the geologic unit features.
representativeAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative stratigraphic age interval for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing all or part of the feature's history. Typically, this identifier might be used as a symbol key for a geologic-age-based portrayal of the geologic unit features.
representativeLowerAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative younger value in a range of stratigraphic age intervals for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing all or part of the feature's history.
representativeUpperAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative older value in a range of stratigraphic age intervals for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing all or part of the feature's history.
specification_uri	xs:AnyURI	URI for a complete description of the geologic unit cropping out within the extent of the feature's geometry. Preferred representation is a GeoSciML GeologicUnit feature instance. Mandatory property - if no value is provided then a nil reason URI explaining why the value is nil must be provided.
metadata_uri	xs:AnyURI	URI referring to a metadata record describing the provenance of data.
genericSymbolizer	xs:string	Identifier for a symbol from standard (locally or community defined) symbolization scheme for portrayal.
shape	GM_Object, (GM_polygon)	Geometry defining the extent of the feature of interest.

2.3.1 GeologicUnitView Feature Specifications

In GeoSciML terms GeologicUnitView features are instances of MappedFeature with key property values from the associated GeologicUnit feature summarized in text fields (data type xs:string) for human data consumers, with standard identifiers for geologic unit type, representative lithology, and geologic age. The specification_uri identifies a description resource specific to the geologic unit cropping out in the extent of the polygon (or other) geometry of the feature. The specification_uri should dereference to yield a formally-encoded representation of the geologic unit, ideally in GeoSciML for interoperability. Properties populated by identifiers are suffixed with '_uri' and contain URIs referring to other resources, for example controlled concepts in published vocabularies.

3 Integration of Web Map Services

The key to interoperable map services is the ability to portray adjacent maps using the same symbolization scheme such that when the maps are displayed, the visual discontinuity at the boundary is minimized. Because of the discrepancies in mapping interpretation and intention, there are commonly differences in the definition of map units between geologic maps produced by different authors or at different times. Resolving such differences is a compilation process that often requires additional field work; this is outside the scope of service deployment.

What can be avoided is visual discontinuity due to different portrayal schemes that assign different colors to similar units on adjacent maps. GeoSciML-portrayal addresses this issue by requesting that geologic unit mapped features be categorized using the CGI Simple Lithology vocabulary (representativeLithology_uri) and the ICS stratigraphic time scale (representativeAge_uri) using registered identifiers.

The use of community portrayal schemes associated with these categorization schemes provides a basis for first order lithologic and age harmonization between maps provided by different services. The genericSymbolizer property is intended to enable preservation of the original legend units and portrayal from the source map.

4 Deployment Cookbook

Once geologic map data is automated into a structured digital form, publication of the data for access as an Open Geospatial Consortium (OGC) web map service (WMS) or web feature service (WFS) is a three-step process:

1. Extract the data from its source data set
2. Transform the data to an interchange format
    - In the context of this cookbook, the GeoSciML-portrayal schema constitutes the appropriate interchange format
3. Load the transformed data into whatever sort of data store is most convenient for the server implementation of the OGC WMS and WFS

This process is commonly called 'ETL'.

There exists a broad spectrum of approaches to the ETL process. We focus in this tutorial on what we deem to be a common situation in which the source data is in a database feature class (possibly in an ESRI geodatabase, shape file, or PostGIS table) with one record for each geometric feature (line or polygon), each feature has either a link to a map unit description or includes fields specifying the age and lithology of the unit. Source information for the map data and unit descriptions is also required.

The most difficult part of the ETL process is typically determining what representative lithology, age, younger age and older age categories from a standard vocabulary to assign to each unit using URIs. OneGeology conformance proscribes that these vocabularies should be the CGI SimpleLithology vocabulary for lithology and the International Stratigraphic Commission Geologic Time Scale 2009 for age.

4.1 Mapping from source data

The following outline describes the basic ETL workflow:

1. Determine which fields in the source data contain the information that is to be delivered in the interchange format fields. Multiple source fields may be combined into single interchange fields, and a single source field may impact values in multiple interchange fields.
2. Determine what steps are necessary to get the content into the interchange format. This may involve some calculation, such as concatenating text from multiple fields to populate requisite fields in the interchange format. Use of the standard vocabularies for interoperability will likely require mapping vocabulary terms in the source data to identifiers for concepts in the controlled vocabularies for the fields that require URIs (see next section).
3. Set up a query to generate a table with **field names exactly matching the field names in the interchange schema**. In some cases it may be convenient to generate the interchange schema table in several steps, populating subsets of the fields each time. It may also be useful to generate a table with unique combinations of contact, fault, or geologic unit properties from the source data, and map each combination to corresponding properties in the interchange format. Depending on individual situations, the unique descriptions can be identified using the specification_uri, name, or genericSymbolizer field; identifiers in this field can then be used to join the interchange format properties with individual features.

4. The table of unique descriptions can then be **joined** with the geometry elements to generate the final feature classes for the portrayal view service. ESRI Shapefiles and PostGIS tables are useful representations for the final feature class that the WMS server will use. Shapefiles will truncate field names longer than 10 characters, so in order to use them, the map server must support defining XML element names distinct from data table field names that are longer than 10 characters.
5. Set up the configuration for the WMS server. This procedure will be specific to the particular server implementation that is being used. Procedures for deployment using GeoServer and ArcGIS server are discussed in Appendices A and B respectively.
6. Create metadata for the dataset and its service distribution and load it into a catalog server to make the serviced discoverable.

4.1.1 Interchange formats

For a variety of reasons, this cookbook specifies GeoSciML-portrayal XML documents as interchange formats. These documents should be structured according to an appropriate GeoSciML-portrayal schema.

The GeoSciML-portrayal feature classes detailed in Section 2 constitute **representations** of three appropriate schemas.

- Table 1 represents the schema for the GeoSciML-portrayal ContactView schema by describing each element and the data that should be entered therein
- Table 2 represents the schema for the GeoSciML-portrayal ShearDisplacementStructureView by describing each element and the data that should be entered therein
- Table 3 represents the schema for the GeoSciML-portrayal GeologicUnitView by describing each element and the data that should be entered therein

By entering geoscience data into an XML document structured according to any of the above schemas, one is creating an interchange document.

4.1.2 Schema mapping in ESRI ArcGIS

This section provides an example of schema mapping in ESRI ArcGIS, using ArcMap, ArcCatalog, and ArcToolbox.

1. Open ArcCatalog
2. Locate or create a **spatial database** that will be used to host your web services in accordance with the GeoSciML-portrayal schema
    a. Sometimes it's easier to create a new database in accordance with the GeoSciML-portrayal schema; other times, it's better to create a new table within an existing database. The conditions needed to make this determination are multifarious, highly dependent on individual context, and therefore beyond the scope of this document.
3. Connect to the desired spatial database under **Database Connections**. To connect to a database, you will need to click **Add Spatial Database Connection** (Figure 5) in ArcCatalog (below **Database Connections**). This will open the **Spatial Database Connection** window.

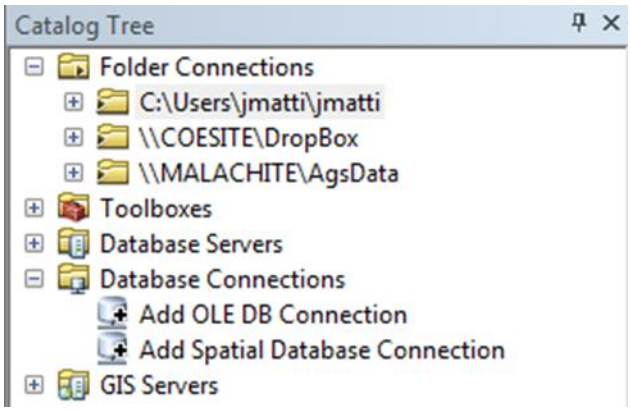


Figure 5: Adding a Spatial Database Connection in ArcCatalog

4. In the **Spatial Database Connection** window, specify the location and (if necessary) the credentials needed to access the desired spatial database. Click **OK** when finished.
5. After you have connected to the spatial database you intend to use for your services, right-click the desired database; in the context menu that appears, click **New** > **Feature Class**…

   This will bring up the **New Feature Class** dialogue box.

a. In the **New Feature Class** dialogue box, type the name for your feature class (this is *not* the name of the web service). Use the pulldown menu to select the appropriate **Type** of feature geometry for the data you are going to publish. When you are finished filling out the form, click **Next**.
b. Choose the **coordinate system** for your feature class.
    i. For the purposes of interoperability, WGS 1984 is the recommended coordinate system. To select the WGS 1984 coordinate system, navigate to the **Geographic Coordinate System** > **World** directory and click the WGS 1984 coordinate system.

When you have located the desired coordinate system, click **Next**.

c. Specify the resolution for your feature class by entering an appropriate amount of degrees into the **XY Tolerance** field.
    i. The XY tolerance defaults to 0.000000008983153, which is fine for most applications, but should be changed if necessary.

When finished, click **Next**.

d. Specify the **Configuration keyword** for your feature class.
    i. The default Configuration keyword is fine for most database storage architectures.

When finished, click **Next**.

e. Enter appropriate field headings under the **Field Name** column; specify the parameters for each field by entering appropriate values in the **Data Type** column and in the **Field Properties** section (Figure 6).
    i. **To comply with GeoSciML-portrayal, all field names and associated parameters must conform exactly to the GeoSciML-portrayal schema that is appropriate for your dataset**. This includes capitalization, spacing, underscores, field lengths, and cardinality.

    Appropriate values can be found in Section 4, GeoSciML-Portrayal Feature Types.

When done, click **Finish**.

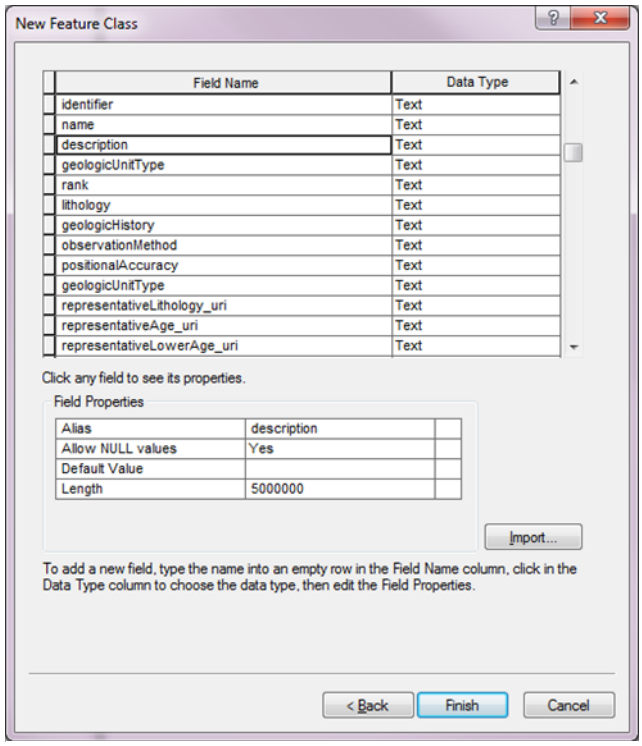


Figure 6: Step 5E - Entering GeoSciML-portrayal schema field names and parameters

Your new **Feature Class**, structured by the appropriate **GeoSciML-portrayal schema**, will appear in the spatial database you selected or created above.

Next, you will use **ArcToolbox** to import and map your existing data into the new feature class you just selected or created above.

1. Open **ArcCatalog**

2. On the menu bar of ArcCatalog, click the **ArcToolbox** button
3. In ArcToolbox, click **Data Management Tools > General > Append**
4. In the **Append** window (Figure 7), you need to perform a number of tasks:
    a. Specify the **Input Dataset**. This will be your source data that does not conform to the GeoSciML-portrayal schema
    b. Specify the **Target Dataset**. This will be the Feature Class you created above
    c. Change the default **Schema Type** to **NO_TEST**. This last instruction is very important; it allows you to perform the schema mapping
    d. For each field in the **Field Map (optional)** table, map your **Input Dataset** to the **Target Dataset**:
        i. Right-click a target field in the **Field Map (optional)** table
        ii. In the context menu that appears, click **Add Input Field**
        iii. In the dialog box that appears, select the corresponding field from your **Input Dataset**
    e. Repeat steps **i-iii** for every field in the **Target Dataset**
    f. When finished, click **OK**
5. Repeat Step 4 in this list for each dataset you wish to deploy as a [web service](#).

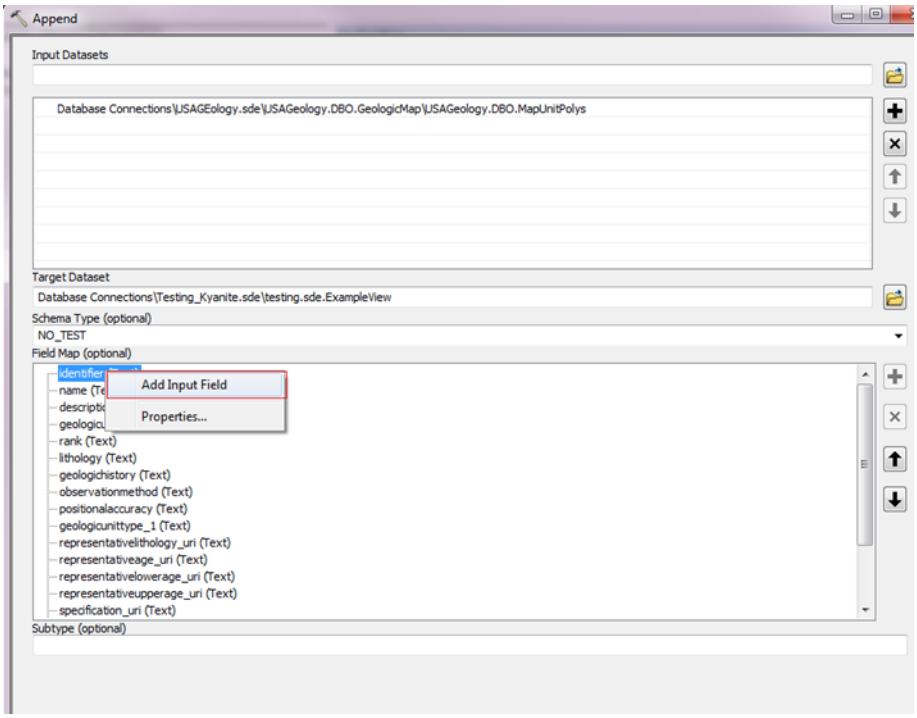


Figure 7: The Append window

You are now ready to deploy your services.

Note: If you have multiple tables that you need to map into a single schema, it may be easier to write a SQL query to do the mapping for you, as described in Section 4.1.3. For our GeologicUnitView features, we originally mapped the NCGMP-09 data into GeoSciML-portrayal in PostGIS using an SQL query by creating a view. For more information, see page XX of the accompanying document.

For an in-depth, technical description of how our tables were originally mapped, visit our lab.usgin blog by following this link.

4.1.3 Schema mapping in SQL

The following example constitutes a [schema](#) mapping query, written in Postgres-flavor SQL (minor modifications may be necessary for other SQL environments). The source schema is the USGS/AASG National Cooperative Geologic Mapping Program NCGMP09 database schema (Haugarud et al., 2010), and the target is a GeoSciML-portrayal GeologicUnitView. This is meant as an illustrative example of the kind of processing that may be required. The level of difficulty in the schema mapping is largely determined by how similar the internal data model schema is to the GeoSciML conceptual model, and how much content the data provider wishes to make available in the portrayal view.

Code Example 1: PostGIS SQL query to produce GeoSciML-portrayal view for map unit polygons.

```
1  CREATE TABLE sde.geologicunitview As
2  SELECT
3  mup.objectid as objectid,
4  mup.mup_id AS identifier,
5  dmu.description AS name,
6  mup.notes AS description,
7  'Geologic Unit'::text AS "geologicUnitType",
8  'Not Specified'::text AS rank,
9  polyextattr.lith6name AS lithology,
10 dmu.age AS "geologicHistory",
11 (datasources.source::text || ' '::text) || datasources.notes AS source,
12 'http://…/cgi/geologicunittype/0008'::text AS "geologicUnitType_uri",
13 polyextattr.lithuri AS "representativeLithology_uri",
14 polyextattr.ageuri AS "representativeAge_uri",
15 mapunitages.ageyoungerterm AS "representativeLowerAge_uri",
16 mapunitages.ageolderterm AS "representativeUpperAge_uri",
17 'http://www.opengis.net/def/nil/OGC/0/missing'::text AS specification_uri,
18 'http://catalog.usgin.org/geoportal/…'::text AS metadata_uri,
19 mup.mapunit AS "genericSymbolizer",
20 shape::geometry as shape
21 FROM mapunitpolys AS mup
22 LEFT JOIN polyextattr ON mup.mapunitpolys_id = polyextattr.ownerid
23 LEFT JOIN datasources ON mup.datasourceid = datasources.datasources_id
24 LEFT JOIN descriptionofmapunits as dmu ON mup.mapunit = dmu.mapunit
25 LEFT JOIN mapunitages ON mup.mapunit = mapunitages.mapunit;
```

This example SQL query creates a table that has fields conforming to the GeologicUnitView for the geosciml portrayal scheme. The SQL is a PostGIS dialect, so field names that are not all lower case are enclosed in quotes, and constant values are explicitly types (e.g. '::text'). 'mapunitpolys' (mup),' datasources', 'descriptionof-mapunits' (dmu) are tables from the source NCGMP09 database, abbreviated as indicated in the query. 'poly-extattr' and 'mapunitages' are Postgres views that aggregate lithology and age properties correlated with polygons or map units respectively. The source data is a regional dataset that correlates lithology and age categories at the polygon level to document lithologic and age variation within the map units. This information is contained in the 'extendedattributes ' and 'geologicevent' tables in the source dataset (See NGCMP, 2010 or http://ngmdb.usgs.gov/Info/standards/NCGMP09/ for details on the input data structure).

4.1.4 Notes on schema mapping in SQL

Line numbers refer to numbered lines of text in Code example 1.

- Line 1. Result table is placed in ESRI sde schema in PostGIS database.
- Line 3. For ESRI Spatial Database Engine (SDE) to recognize the table as a feature class that may be used to source a service, an integer unique-value field must be present. In this case the field from the source dataset is used because it serves the same purpose there. This field does not need to be set up as an autoincrement field because the resulting table is a read-only view of the data. The ObjectID field will likely be placed at the last column in the table for better interoperability with systems that do not require such a field.
- Line 4. Unique polygon identifiers from the source dataset are carried into the interchange format view.
- Line 7. Source data map units are not categorized into specific unit types (e.g. lithostratigraphic, chronostratigraphic, geophysical, etc.), so this field gets a constant value 'Geologic Unit', which is the preferred label for the corresponding concept in the CGI Geologic Unit Type vocabulary.
- Line 8. Likewise, rank is not assigned for the regional geologic units in the source dataset, so this is a constant that is the preferred label ('Not Specified') for the corresponding concept in the CGI Strati-graphic Rank vocabulary.

- Line 9, 10. Text descriptions of lithology and the age of the unit are derived from corresponding fields in the source data. The lithology description is mapped at the individual polygon level (via the polyextattr Postgres view), note the JOIN in line 22 that uses the .mapunitpolys_id as the foreign key. The age text description is mapped at the map unit level, note the JOIN in line 24 that uses the .mapunit field as the foreign key.
- Line 11. The text for the source field in the portrayal scheme is calculated by concatenating two fields (.source and .notes) from the datasources table in the input dataset.
- Line 12. Since all the map units are classified simply as 'Geologic Unit', this URI is also a constant value from the CGI Geologic Unit Type vocabulary.
- Line 13, 14. Representative lithology and age categories are assigned on a polygon by polygon basis, thus the polyextattr table is the source of these URIs. Note the JOIN in line 22 that uses the .mapunitpolys_id as the foreign key.
- Line 15, 16. Younger and older age bounds are assigned by map unit (not at the individual polygon level), so these fields are sourced from the mapunitages aggregation query in the database; this query uses the extendedattributes table to join map-units with geologicevent records that contain the lower and upper age bounds for the unit.
- Line 17. specification_uri is a link to a more complete description of the geologic unit that crops out in the polygon's extent. This description can be viewed as a resource in the context of web architecture. The intention of the GeoSciML design team was that this URI should dereference to return a full Geo-SciML GeologicUnit element instance. With the use of content negotiation on the web, this description might also have representations as a web page or other structured description (rdf, owl…). In the example instance, a more complete description is not available and an OGC nil URI is used to indicate that the resource is missing (does not exist).
- Line 18. metadata_uri identifies a metadata resource that contains more complete information on the provenance of the information in the feature element. This may be a metadata record that is scoped to the individual feature, or may identify a metadata record describing some collection of polygons that have similar enough provenance to document together. The text in the source field (Line 11) should be a succinct summary of the information in this metadata record pertinent to the containing feature.
- Line 19. The genericSymbolizer field should contain an identifier for the symbol used to display this poloygon in the default portrayal (legend) chosen by the data provider for the gsmlp:GeologicUnitView instance in the containing feature collection. This field can be used to capture the map unit assignment and portrayal color scheme from the original data from which the GeologicUnitView polygon was digitized. The legend may be encoded in an accompanying Styled Layer Descriptor (SLD) file, and if such an SLD exists it should be recorded as a related resource in the metadata record specified by metadata_uri.
- Line 20. The shape field contains a representation of the geometry of the outcrop area described by the GeologicUnitView instance. The content of this field will be managed by the GIS and WMS server, and thus will generally not need to be manipulated by users outside the GIS environment. The shape field may need to be cast into a recognized geometry field in order to get PostGIS and ESRI SDE to recognize the output as a feature class. In our configuration at AZGS, we ended up using this:

  'st_geometryfromtext(ST_AsText(shape)::text, 4326)::geometry as shape'

  More expert PostGIS users may have a better solution for this problem.

4.2 Vocabulary mapping

There are many possible approaches to mapping terms from one vocabulary to another.

One situation in which a standard process can be defined involves a controlled vocabulary used to populate a field in the source data, and that field maps directly to a field in the interchange format. For instance: if the source data contains a 'dominant lithology' field, the information in this field can be used to populate the 'representativeLithology_uri' for a GeologicUnitView feature.

Recommended procedure:

1. Produce a table of the unique values in the source data
2. Add a column to this table for the corresponding element in the target interchange scheme
3. Determine the best matching value from the interchange vocabulary for each term in the source vocabulary

In general: the most specific term from the interchange vocabulary that completely subsumes (encompasses) the meaning of the term in the source vocabulary should be used.

If the source vocabulary has terms that are more specific than the controlled vocabulary, there will be some information lost in this process, but the original source terminology can be preserved in free-text description fields in the interchange document. Sample free-text fields that may be used to preserve source vocabulary information are as follows:

- Lithology
- geologicHistory

Remember, the primary purpose of the controlled vocabulary fields is for data integration and use as search criteria, likely by non-expert users.

In some cases, unique values from a combination of multiple source data fields may be necessary to define mapping to interchange concepts, particularly for the following interchange fields:

- representativeAge_uri
- representativeLithology_uri
- genericSymbolizer

The procedure is the same, but any unique values query will involve more than one source field and multiple-field **joins** will be necessary to construct queries generating the output schema content.

Standard vocabularies are listed in the accompanying Excel Workbook.

4.3 Excel Workbook template

The accompanying Microsoft Excel Workbook provides spreadsheets containing the content models for the description properties associated with geologic contacts, faults or shear zones, and geologic units in appropriate GeoSciML-Portrayal schemas.

These spreadsheets do not include geometry fields that are necessary for an actual GIS feature class for the corresponding ContactView, ShearDisplacementStructureView, and GeologicUnitView features. Rather, they provide a template for compiling the necessary descriptions (combinations of attributes) that can then be **joined** with a GIS feature class to produce the dataset for feature web service deployment.

The specification_uri in each feature class provides a link to structured representation of the geologic feature that is intended to be a GeoSciML feature.

The metadata_uri provides a link to a structured metadata record maintaining complete provenance information for features.

See the notes tab in the workbook for additional information describing the GeoSciML Portrayal feature classes and the work-sheets included in this document.

4.4 Configuring OGC services

GeoSciML-Portrayal is intended to support map services for online viewing and exploration of geologic data, and the creation of mash-ups integrating data from different servers and possibly different thematic domains.

The OGC WMS specification includes a 'getFeatureInfo' operation that returns a description of a map feature at a user-identified point on the map. The operation returns a document determined by the server configuration. A variety of different formats may be offered, and specified by a request parameter. For GeoSciML-Portrayal services, an XML document conforming to the GeoSciML-Portrayal XML schema should be offered. Other getFeatureInfo response formats may also be offered.

Server software that implements the OGC Web Map Service will typically also allow parallel deployment of an OGC Web Feature Service with no additional effort. The XML provided by the feature service is the same as that provided in the WMS getFeatureInfo response formats.

The GML simple feature profile requires that each feature include a gml:id attribute. According to the GML specification, this identifier must be an alphanumeric string that has an alphabet letter as its first character. These ID's can be generated from the primary key in the data table used to compose the features for the map service. If data are coming from an ArcGIS environment, the ObjectID required for any ArcGIS feature class can be used to construct the gml:id. If no primary key is present in the feature class for the service, one will have to be added to source the gml:id property. Technically, the gml:id is meant to be unique within the scope of a feature collection

from a particular service to enable internal cross referencing within a single WFS response document or to allow queries to retrieve a feature using the gml:id.

Server configuration using GeoServer and ESRI ArcGIS Server is described in Appendices A and B as examples of the workflow necessary to set up a GeoSciML-Portrayal service. Other implementations exist, and the details of the procedure will change as new versions of the server software are released.

5 Styling

Part of the process of deploying a web service includes defining a map legend.

ESRI ArcGIS Server provides a number of tools to define a map legend, but if you are using GeoServer to deploy your data as a web service, you will need an OGC SLD file to define your map legend.

OGC Styled Layer Descriptors (SLDs) are XML documents that define a mapping from some property of an OGC feature to a symbol to use for portrayal. By mapping a property of a feature to a symbol that will be used to represent a feature with a given property, you are essentially defining a map legend.

SLD files for portrayals using the CGI simple lithology and ICS 2009 time scale URIs are available at the following web location:

> http://schemas.usgin.org/schemas/slds/

Aside from those SLDs made available at the above web location, it is possible to use a basic text editor to manually create an SLD. Unfortunately, manually creating an SLD for data with a complicated symbology scheme, such as a geologic map, is not a trivial proposition, for a variety of reasons. For example, colors must be converted from RGB to hexadecimal on a per-symbol basis, which is a labor-intensive process. Instructions for creating an SLD file manually are out of the scope of this document.

Alternatively, software exists that permits the user to create an SLD file based on symbology extant in ArcGIS. See the following section for an example.

5.1 How to Create a Styled Layer Descriptor (SLD) using Arc2Earth

Arc2Earth is a plugin for ArcGIS that is available in several different editions. The **Community** edition is free and provides access to the SLD generator. Arc2Earth is available at http://www.arc2earth.com/.

1. Install the Arc2Earth plugin
2. Open ArcMap
3. Open the ArcMap project containing the layer or layers to export as SLDs
4. Select the desured layer in the catalog tree
5. In the **Arc2Earth** toolbar, click **Export** > **Export Layer Style to SLD**
6. Navigate to the appropriate location
7. Click **Export**

After your SLD has been exported, open the .sld in any XML or text editor. Note that the Arc2Earth plug-in adds an outline (stroke) to polygon layers. If you are working with polygon layers, the stroke may need to be tailored or removed manually. Also, it may be necessary to customize the heading for your SLD, or change the schema location.

> **Note**: though .sld files, like any XML document, may be edited by a generic text editor, such as the **Notepad** or **Wordpad** text editors included with Windows operating systems, some text editors are better suited to working with XML documents than others. One example of a free-and-open-source text editor designed with XML support is Notepad++.

Below is a snippet of an SLD created using Arc2Earth (Code Example 2).

Code Example 2: Styling in XML

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<StyledLayerDescriptor xmlns:xlink='http://www.w3.org/1999/xlink' xmlns:gml='http://www.opengis.net/gml'
xmlns:ogc='http://www.opengis.net/ogc' xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance' version='1.0.0'
xsi:schemaLocation='http://www.opengis.net/sld StyledLayerDescriptor.xsd' xmlns='http://www.opengis.net/sld' >
        <NamedLayer>
                <Name><![CDATA[GeologicAge]]></Name>
                <UserStyle>
                        <FeatureTypeStyle>
                                <Rule>
                                        <Name><![CDATA[Holocene]]></Name>
                                        <Title><![CDATA[Holocene]]></Title>
                                        <ogc:Filter>
                                                <ogc:PropertyIsEqualTo>
                                                        <ogc:PropertyName>representativeage_uri</ogc:PropertyName>

        <ogc:Literal><![CDATA[http://resource.geosciml.org/classifier/ics/ischart/Holocene]]></ogc:Literal>
                                                </ogc:PropertyIsEqualTo>
                                        </ogc:Filter>
                                        <PolygonSymbolizer>
                                                <Fill>
                                                        <CssParameter name="fill" >#f9f97f</CssParameter>
                                                        <CssParameter name="fill-opacity" >1</CssParameter>
                                                </Fill>
                                                <Stroke>
                                                        <CssParameter name="stroke" >#f0f0f0</CssParameter>
                                                        <CssParameter name="stroke-width" >0.4</CssParameter>
                                                        <CssParameter name="stroke-opacity" >1</CssParameter>
                                                </Stroke>
                                        </PolygonSymbolizer>
                                </Rule>
                                <Rule>
                                        <Name><![CDATA[Quaternary]]></Name>
                                        <Title><![CDATA[Quaternary]]></Title>
                                        <ogc:Filter>
                                                <ogc:PropertyIsEqualTo>
                                                        <ogc:PropertyName>representativeage_uri</ogc:PropertyName>

        <ogc:Literal><![CDATA[http://resource.geosciml.org/classifier/ics/ischart/Quaternary]]></ogc:Literal>
                                                </ogc:PropertyIsEqualTo>
                                        </ogc:Filter>
                                        <PolygonSymbolizer>
                                                <Fill>
                                                        <CssParameter name="fill" >#f9f97f</CssParameter>
                                                        <CssParameter name="fill-opacity" >1</CssParameter>
                                                </Fill>
                                                <Stroke>
                                                        <CssParameter name="stroke" >#6e6e6e</CssParameter>
                                                        <CssParameter name="stroke-width" >0</CssParameter>

                                                        <CssParameter name="stroke-opacity" >1</CssParameter>
```

6 Deploying Your Web Service

If you have completed the following steps, you are ready to deploy your web service in a GeoServer environment.

1. You meet the prerequisites detailed in Section 1.3, Cookbook Prerequisites
2. You have successfully mapped your data to an appropriate GeoSciML-portrayal schema, as described in Section 4, Deployment Cookbook
3. You have symbolized your data, as described in Section 5, Styling.

If you have followed all previous steps in this cookbook, you should now be ready to deploy your GeoSciML-portrayal web service.

The overall process of web service deployment is as follows:

1. Place your data in a web-accessible folder
2. Set up and/or access server software that is capable of providing a web service, such as:
    a. GeoServer
    b. ArcGIS Server
3. Use the desired server software to deploy your data as a web service
4. Test your web service (see Section 7, Testing Your Web Service, for more details)

Unfortunately, there are as many ways to deploy a web service as there are individual contexts; to attempt to account for all possible scenarios is beyond the scope of this cookbook.

Instead, this cookbook will provide two generic examples of web service deployment:

- Web service deployment in a GeoServer environment (Appendix A)
- Web service deployment in an ArcGIS Server environment (Appendix B)

7 Testing Your Web Service

Testing your web service is very simple procedure: you attempt to access it using your preferred GIS application. If you can access your web service successfully, and if your web service displays properly, then nothing more need be done.

If, on the other hand, you cannot access the web service, or if it displays improperly, then you need to troubleshoot your web service.

7.1 Web Service Requests

In order to access your web service, you must perform a **GetCapabilities** request in your web browser. A **GetCapabilities** request is a request for data from a web service that returns an XML document describing the capabilities of the web service (Figure 8).

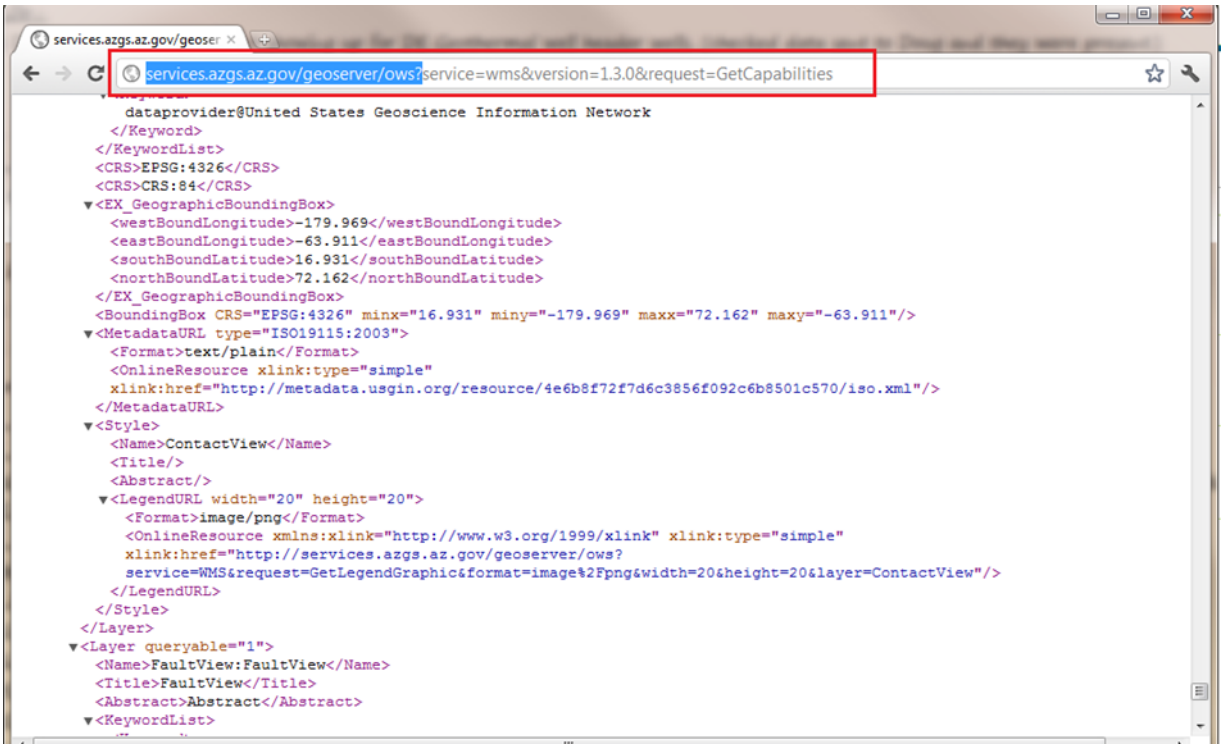


Figure 8: A typical capabilities document returned by a GetCapabilities request in a web browser

To perform a **GetCapabilities** request, simply use your web browser to navigate to the appropriate **service directory** and click your web service. This will return a **capabilities document** in your web browser; the web address of this capabilities document is your **service request**.

Sample GetCapabilities requests:

21

http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/ALBoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities&service=WFS

http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/ALBoreholeTemperatures/MapServer/WMSServer?request=GetCapabilities&service=WMS

In both examples, **bolded** text constitutes the **service endpoint**, whereas all text afterwards constitutes the **service request**. **In order to test your web service, you will need to select and copy the service endpoint of your web service and append a service request to it**.

Once you have accessed the capabilities document of your web service and copied the service endpoint, you can proceed to the Section 7.2.1 or 7.2.2.

Sections 7.2.1 and 7.2.2 contain instructions for accessing a web service via GIS applications. As 90% of the USGIN user base uses ArcGIS, Section 7.2.1 will describe the process of accessing a web service via ArcCatalog. Section 7.2.2, however, will describe the process of accessing a web service via the free-and-open-source desktop GIS application uDig.

If you are not sure how to access your web service, here is a short list of GIS software that supports web services:

- ArcGIS Desktop
- ArcGIS Explorer
- ArcGIS Online
- Grass GIS
- gvSIG
- OpenLayers
- Quantum GIS
- uDig

7.2 Adding a Web Service in ArcCatalog

1. Open **ArcCatalog**
2. In the Catalog Tree, expand **GIS Servers**
3. Click **Add WMS Server** or **Add WFS Server**, depending on the kind of web service you wish to access.
    - See the glossary for more information about WMS and WFS.
4. Paste the **service endpoint** (see Section 7, Testing) for your web service into the URL field
5. It may be necessary to navigate back to GeoServer for your URL. On the GeoServer Welcome page, Service Capabilities are listed on the right. Click on the latest version of WMS or WFS, whichever service you wish to use.
6. When redirected, copy the URL (from beginning up to the question mark) and paste into the URL field of Add WMS Server in ArcCatalog:
7. http://services.azgs.az.gov/geoserver/ows?service=wms&version=1.3.0&request=GetCapabilities
8. Click Get Layers
9. Click OK
10. Add the desired layer(s) to ArcMap

7.3 Adding a Web Service in uDig

1. Open uDig
2. Software available at: http://udig.refractions.net/
3. Create a new map, or open an existing map to which you would like to add the web service
4. On the upper menu bar, click Layer > Add…
5. In the window that appears, click Web Map Server or Web Feature Server, as appropriate; click Next
6. Paste the service endpoint in the URL field; click Next
7. In the Resource Selection window that appears, select all layers you wish to add. When you are done, click Finish

8 Registering with OneGeology and Submitting your Service

To submit your service to OneGeology, you will first need to register your organization on their website by filling out their registration form. If the data provider is able to host their own service, visit this link:

http://www.onegeology.org/technical_progress/data_coordination.cfm

If the data provider would prefer a 'buddy' serve their data, fill out this form:

http://www.onegeology.org/technical_progress/buddy_coordination.cfm

Next, you will need to send an email to onegeology@bgs.ac.uk with the draft URL of the proposed WMS service. Include this information: Name of Geographic Area, Name of Data Provider, Name of Service Provider.

You will be contacted by the OneGeology secretariat with feedback, and confirmation of receipt.

9 Appendix A: Deploying GeoSciML-Portrayal Web Services in GeoServer

This appendix provides a tutorial of GeoSciML-portrayal web service deployment in a GeoServer environment.

9.1 Logging in

To deploy a web service in the GeoServer environment, your first step is to log in to the **Web Administration Interface** for the GeoServer instance you will be using to deploy your data as a web service (Figure 8). The Web Administration Interface is, as the name suggests, a browser-based user interface capable of administering server software remotely via your Internet connection.

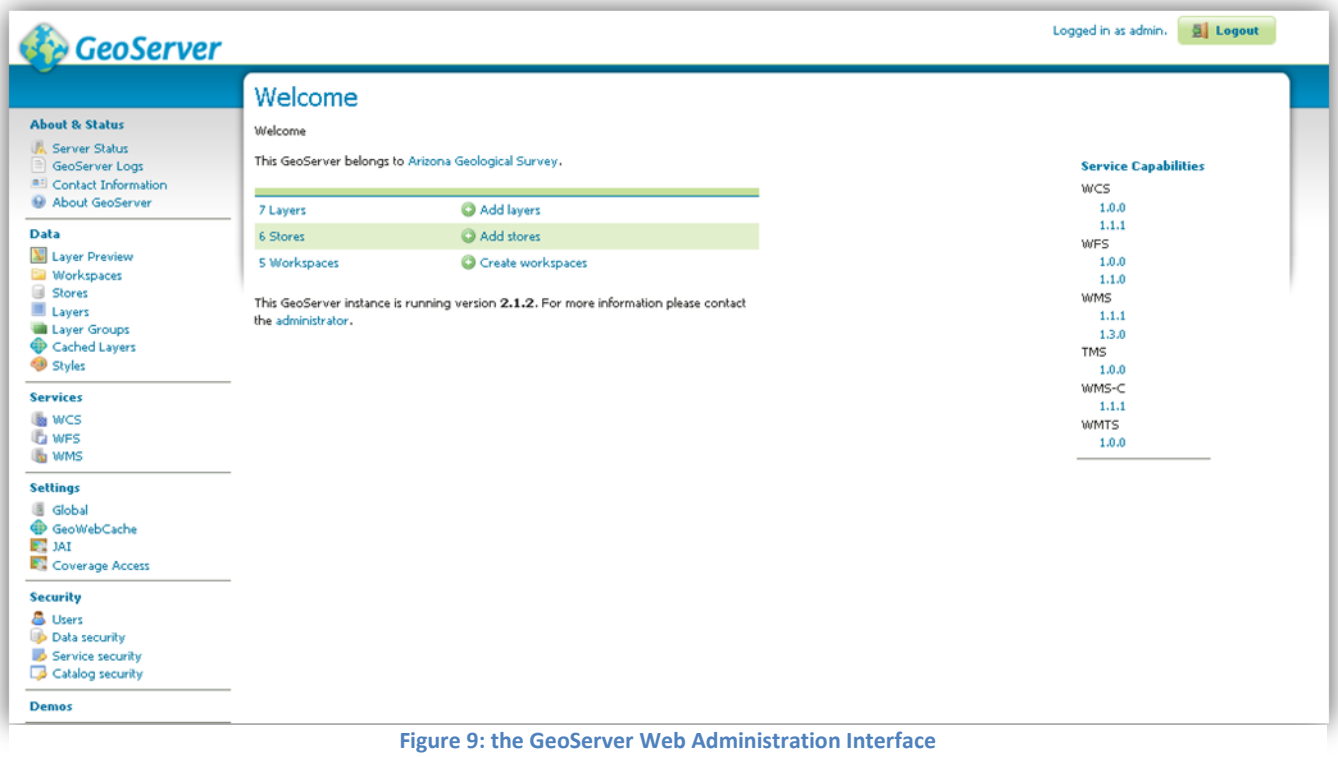


Figure 9: the GeoServer Web Administration Interface

To access the Web Administration Interface of a GeoServer instance, open your web browser and enter the appropriate web address into the navigation bar.

GeoServer instances located on your physical machine can be accessed by entering the following address into your web browser:

http://<**host**>:<**port**>/**geoserver/web/**, where <**host**> is usually **localhost** and <**port**> is usually **8080**

Unfortunately, remote users will be unable to access your GeoServer instance at this location via the World Wide Web. You will need to contact your network administrator to configure the host and port of your GeoServer instance in such a way that your GeoServer instance can be accessed remotely.

The default account settings for GeoServer are as follows:

Username: admin
Password: geoserver

For security reasons, it is recommended that you change your password to something more secure as soon as possible.

9.2 Service-Level Metadata

Within the GeoServer **Web Administration Interface** (Figure 8), click **Contact Information**, under **About & Status**.

This brings you to a **Contact Information** form (Figure 9) in which you can provide contact information for your GeoServer instance. The information entered here becomes service-level metadata for each web service you host from this GeoServer instance. Consequently, Contact Information entered here should be as precise and comprehensive as possible.

If you wish to deploy your service in compliance with OneGeology standards, your server-level metadata will need to meet very specific

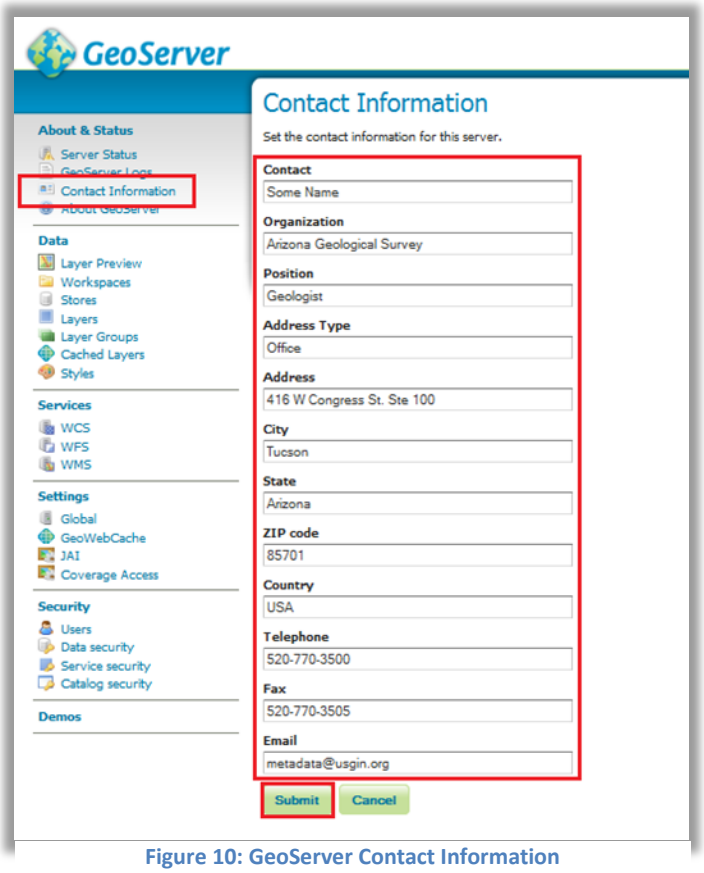


Figure 10: GeoServer Contact Information

requirements. See Appendix C.1 for details.

9.3 Creating a Workspace

After entering contact information for your GeoServer instance, you will need to create a workspace for your web service. The workspace is the equivalent of the **service name** when deploying a web service in the ArcGIS Server environment.

If you wish to deploy your service in compliance with OneGeology standards, your workspace will need to meet very specific requirements. See Appendix C.2 for details.

On the left side of the GeoServer **Web Administration Interface**, under **Data**, click **Workspaces**. This will bring you to the **Workspaces** page, wherein you can manage existing workspaces and create new workspaces.

Click **Add New Workspace**. This will bring you to the **Edit Workspace** page for your new workspace.

Two fields are present on the Edit Workspace page:

Name: The **service title**; may contain spaces or special characters; usually named according to the following convention:

GeographicArea_Theme, where GeographicArea is the overall region containing the map extent of the web service Theme is an indicator of the kinds of features

Namespace URI: A URI associated with your project

When you are finished, click **Save**.

9.4 Connecting to your PostGIS Database

Having created a new workspace, you will now create a **store**, or **database connection**, for your service.

On the left side of the GeoServer **Web Administration Interface**, under **Data**, click **Stores**. This will bring you to the **Stores** page. On the **Stores** page, click **Add New Store**.

This will bring you to the **New Data Source** page (Figure 10).

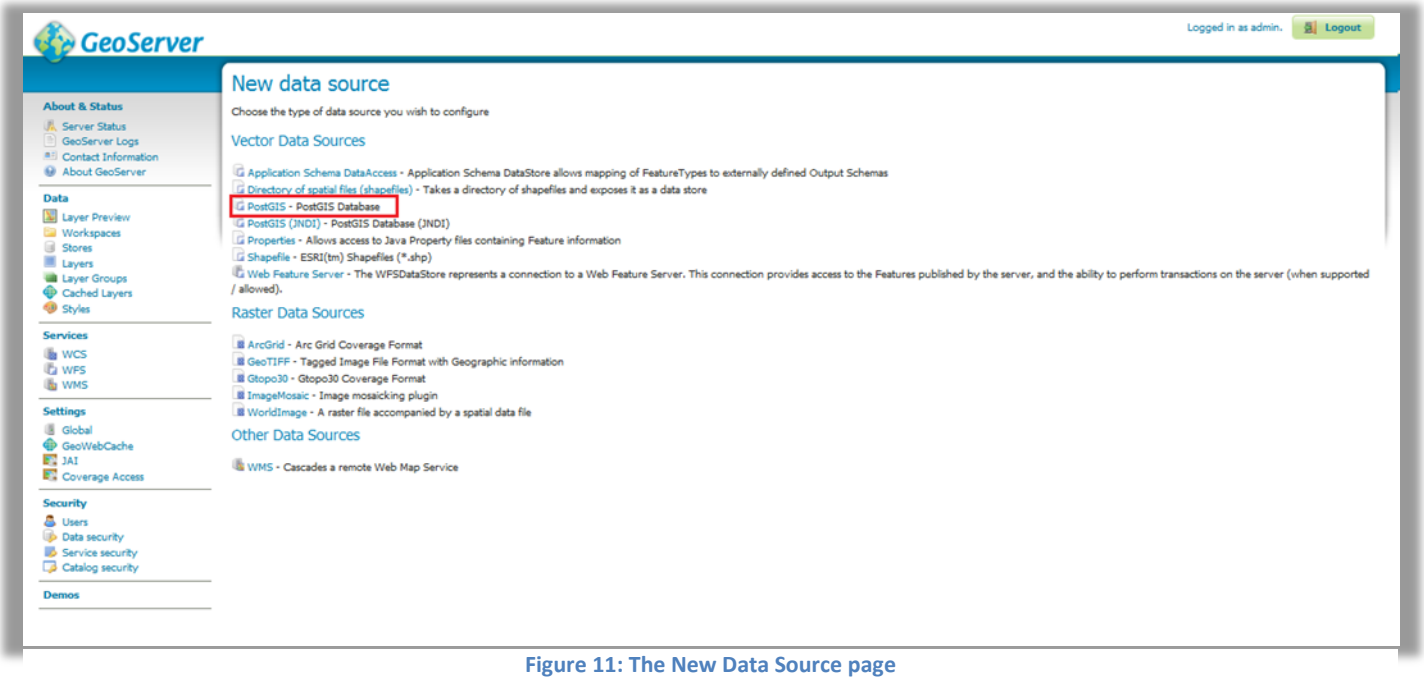


Figure 11: The New Data Source page

On the **New Data Source** page, choose **PostGIS** as the data source by clicking **PostGIS**. This will bring you to the **New Vector Data Source** page. Complete the following steps:

- Select a **Service Title** from the **Workspace** drop down menu. Select the workspace you created in Appendix A.3
- Type a name for your data store in the **Data Source Name field** (spaces are acceptable here); add a description if desired
- Make sure that the **Enabled** checkbox is checked
- Set the **Connection Parameters** for your PostGIS data source; if the PostGIS data source is located on a remote server, you will need to provide the appropriate host, port, database name, user name, and password to access it:
    - **Host**: use "localhost" if the PostGIS data source is on the same machine as your GeoServer instance; more specific host information will be necessary if your PostGIS data source is on a remote server
    - **Port**: default is 5432

- **Database name**: this information will depend on the PostGIS data source
- **Schema**: this information will depend on the PostGIS data source
- **User name**: this information will depend on the PostGIS data source
- **Password**: this information will depend on the PostGIS data source

When finished, click **Save**.

9.5 Adding Layers to a Workspace

Having created a workspace and specified a PostGIS data source for your web service, you will now populate your web service with data layers from your workspace.

On the left side of the GeoServer **Web Administration Interface**, under **Data**, click **Layers**. This will bring up the **Layers** page.

On the **Layers** page, click **Add a new resource**. This will take you to the **New Layer** page.

On the **New Layer** page, use the pulldown menu at the top of the page to select the data source you specified in Appendix A.4. Doing so will populate the **New Layer** page with a list of layers that may be published; click **Publish** to make the associated layer publicly accessible to anyone who connects to your web service.

Note: You may publish the same layer multiple times. To do so, click **Publish again.

After clicking **Publish**, the **Edit Layer** page for the corresponding layer automatically appears (Figure 11). The **Edit Layer** page contains two tabs, **Data** and **Publishing**.

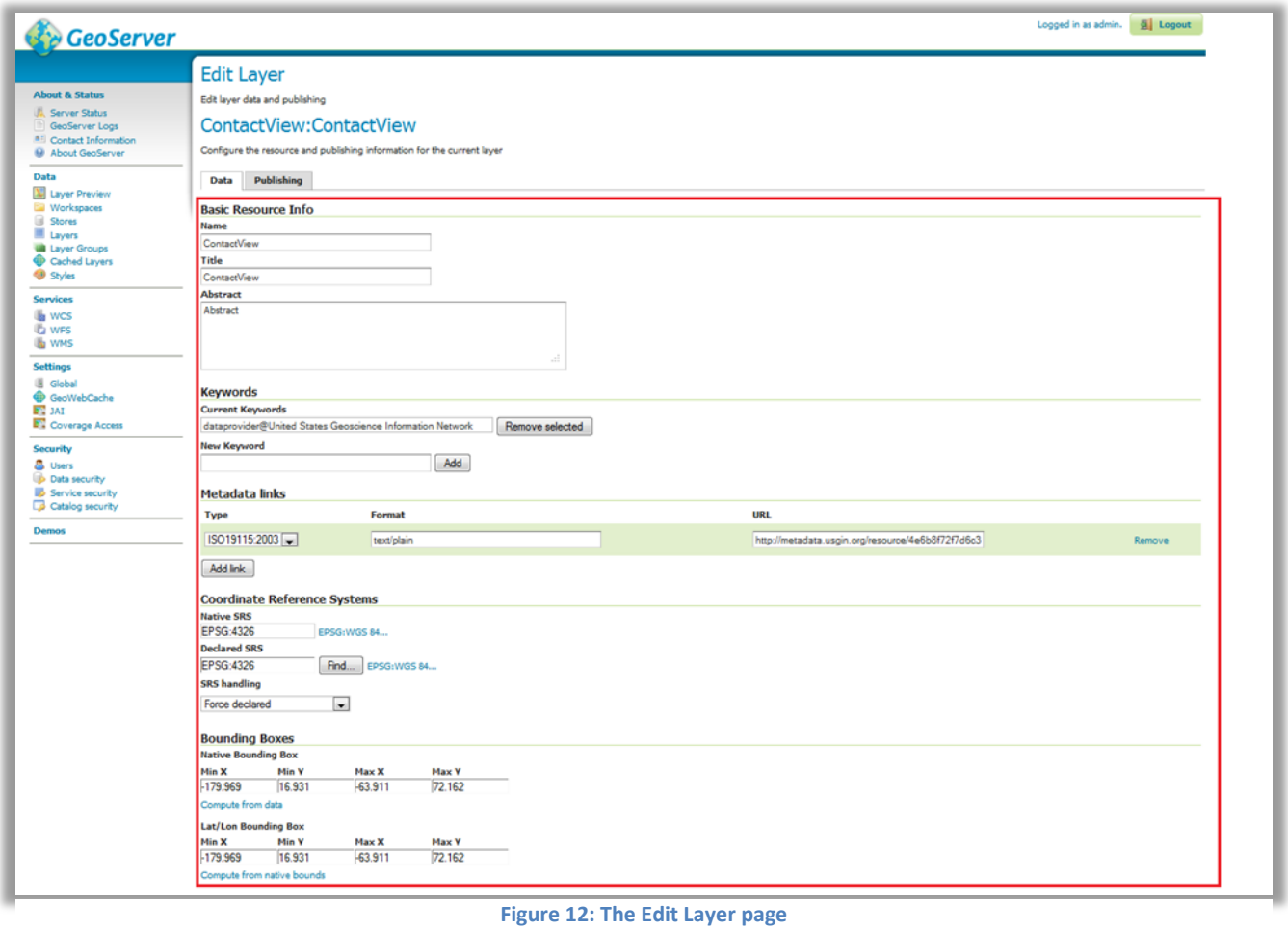


Figure 12: The Edit Layer page

9.5.1 The Data tab of the Edit Layer page

The **Data** tab contains fields within which you may specify the title, abstract, bounding box, spatial reference system, keywords, and metadata links for each layer in your web service. This information will be present within the **Capabilities document** produced by your web service in response to a **GetCapabilities** request, so it is very important to enter this information carefully for each layer in your web service.

If you wish to deploy your service in compliance with OneGeology standards, the information you enter for your layers will need to meet very specific requirements. See Appendix C.3 for details.

It is recommended to enter the bounding boxes for your service manually, as doing so permits you to provide a more useful bounding box for your web service.

**Note: GeoServer occasionally enters a loop of web errors. To fix this, we recommend restarting the service on Apache-Tomcat. It also helps to use Mozilla Firefox or Google Chrome for debugging (rather than Internet Explorer).

9.5.2 The Publishing tab of the Edit Layer page

After populating the fields in the **Data** tab, click the **Publishing** tab.

The **Publishing** tab contains **Layer Style** settings for the corresponding layer of your web service. **Layer Style** settings are dependent on the geometry of the layer (point, line, polygon). See Section 5, Styling, of this document for more details.

Generally, it is faster and more precise to import **Layer Style** settings from an existing style than it is to manually specify values for each field in the **Publishing** tab. To import **Layer Style** settings from an existing style, select the desired style from the **Available Styles** list.

To populate the **Available Styles** list, you might need to import layer styles from an SLD file. For further instructions, see Appendix A.5.3, Importing Layer Styles from an SLD File.

When you have populated the fields in the **Publishing** tab, click **Save**.

9.5.3 Importing Layer Styles from an SLD File

On the left side of the GeoServer **Web Administration Interface**, under **Data**, click **Styles**. This will open the **Styles** page.

If the Style you want is not listed on the **Styles** page, you will need to add it to the list by clicking **Add a New Style**. This brings you to the **Style Editor** page.

On the Style Editor page, you have the choice to copy/paste an SLD or upload a .sld XML document. The **Validate** button may be used to validate the SLD file against a schema before using it. Click **Submit** to add the SLD to the list on the **Styles** page.

9.6 Finishing Up

Repeat Appendix A.5 until all data you wish to publish has been added to the desired workspace. When finished, return to Section 7, Testing Your Web Service.

9.7 GeoServer Troubleshooting

Q: I made a change in the database on my server and now the service is not working.

A: Try clearing the cache and reloading GeoServer on the Server Status page (Figure 22). If that doesn't work, try hard restarting the service through Apache Tomcat.

Q: Can I use a replicated Feature Class to create a service on Geoserver?

A: No.

Q: I set up my services under three separate workspaces. When I connected to the WMS in ArcCatalog, all the layers appeared as one bundle. Is there a way to separate them out so I can add them individually?

A: Yes.

Though setting up your services under different workspaces seems to imply that they can be accessed as discrete services, Geoserver defaults to providing one capabilities document containing the information for all of the services set up on your instance of GeoServer. To access workspaces individually, you will need customize your Get request to specify the desired workspace.

For example: the Arizona Geological Survey runs three services on Geoserver (http://services.azgs.az.gov/geoserver/web/):

- GeologicUnitView
- FaultView
- ContactView

To perform a GetCapabilities request for GeologicUnitView, your GetCapabilities request will appear as follows:

http://services.azgs.az.gov/geoserver/GeologicUnitView/ows?service=wms&version=1.3.0&request=GetCapabilities

This URL opens the WMS **capabilities document** for the GeologicUnitView workspace only. A generic form of the service endpoint for the request is as follows:

http://[host server]/geoserver/[Workspace Name]/ows

Individual workspaces can be added in ArcGIS, as well.

Q: Is it possible to configure GeoServer so that I do not need to use PostGIS?

A: Try installing the ArcSDE plug-in for GeoServer. To do this, you will need to download the extension from GeoServer's website. Make sure to match the versions of the extension and GeoServer.

Q: All of my data are in Shapefiles. Can I deploy a shapefile as a GeoSciML-portrayal service?

A: The problem you will run into is the truncation of field names that occurs in shapefiles. Ideally you will have a full version of the data in PostGIS. As mentioned in the above document, to be compliant with GeoSciML-portrayal, you will need to make sure there is no truncation in field names; they must be an exact match for the GeoSciML-portrayal schema.

Q: I would like to deploy a service using full GeoSciML v.3.0 schema; where can I go?

A: Visit the GeoSciML website at http://www.geosciml.org/.

Q: Why would I want to create a WMS or WFS at all?

A: To display your data in a map format. Map format shows: data relative to its geographical surroundings, data relative to other data on the same map, data points relative to other data points in the same dataset (size, quantity, order, etc).

Q: How do I use the ArcCatalog Interoperability Extension?

A: Here is a brief tutorial on connecting to the WFS in ArcCatalog.

1. Open ArcCatalog
2. Enable the Interoperability Extension
    - For more information visit http://www.esri.com/software/arcgis/extensions/datainteroperability/common-questions.html
3. Click **Add Interoperability Connection**; this will bring up the **Interoperability Connection** dialog box:
    1. Click the '**…**'button next to **Format** and select the appropriate format
    2. Paste **the service endpoint** for the service you want to add into the appropriate field
    3. Click the **Parameters** button. In the window that appears, click the '**…**' button next to **Table List** and select the tables you wish to load. Confirm that the **Max Features** box is scaled appropriately
        - If you are loading a dataset with 10 million features, the Max Features box must be scaled to ten million. If you are loading a dataset with three (3) data points, the Max Features box can be scaled down quite a bit.
    4. Click **OK** in the Interoperability Connection dialogue box.
4. Depending on the size of the dataset, it may take some time to load. When it has loaded, expand the Connection (you can also rename the connection, which can be helpful).

5. Add your service to ArcMap

10 Appendix B: Deploying GeoSciML-Portrayal Web Services in ArcGIS Server

This appendix provides a tutorial of GeoSciML-portrayal web service deployment in a GeoServer environment.

10.1 Connecting to your ArcGIS Server instance

1. Open **ArcCatalog**
2. In the catalog tree, expand **GIS Servers** by clicking the nearby +
3. Double-click **Add ArcGIS Server**
4. Click the Manage GIS Services radio button in the **Add ArcGIS Server** window and then click **Next**.
5. Type in the URL for your services and the name of the server, then click **Finish**.

10.2 Create an ArcMap Project for your web service

1. Open an instance of ArcMap
2. Create a blank map
    a. Note: there is a one-to-one relationship between .mxd files and services: each .mxd file, along with all layers it contains, is equal to one web service
3. Add your source data to the map as a layer
    a. If you wish to deploy your service in compliance with OneGeology standards, your layers will need to meet very specific requirements. See Appendix C.3 for details
4. Save your map as a .mxd file

10.3 Publishing on ArcGIS Server

1. Open ArcCatalog.
2. Navigate to the database connection you set up in Appendix B.1, Connecting to your ArcGIS Server instance
3. In the catalog tree on the left side of the screen, expand the connection to ArcGIS Server
4. Right-click your server and select **Add New Service**; this opens a window containing fields for **Service Title** and **Service Description; if** you wish to deploy your service in compliance with OneGeology standards, your **Service Title** will need to meet very specific requirements. See Appendix C.2 for details
    a. Specify the appropriate **Service Title** based on content and map extent
    b. Using the pulldown menu, specify whether you are deploying a Web Map Service or a Web Feature Service
    c. Enter an appropriate description for your service in the **Description** field
    d. Click **Next**
5. Specify the capabilities of your web service
    a. Use the checkboxes to indicate which capabilities your web service will support
        i. OneGeology requires either WMS or WFS capabilities
            1. Specifying WMS or WFS will reveal a form in the **Properties** box; if you wish to deploy your service in compliance with OneGeology standards, this information must be provided in accordance with very specific requirements. See Appendix C for more details.
            2. Alternatively, if you know of a web service capabilities document that is already compliant with OneGeology standards, you may save this capabilities document as an XML file and import it
                a. To import an external XML document, click **Use external capabilities files** and then specify the location of the XML document you will use
                b. To save a capabilities document as an XML document, access the appropriate capabilities document (described in detail in Section 7) in your web browser and click **File > Save Page As...**
    b. Click the **Use layer names from the map document** checkbox
    c. Click **Next**
    d. Click **Next** again
    e. Click **Finish** to start your service
        i. Note: you will need to log in to your service and **Stop** your service in order to edit the capabilities document. Service title cannot be edited after publishing; you will need to create a new service

Proceed to Section 7, Testing Your Web Service.

10.4 Troubleshooting ArcGIS Server

Q: How do I connect to the WFS in ArcCatalog?

A: Here is a brief tutorial on connecting to the WFS in ArcCatalog.

1. Open **ArcCatalog**
2. Enable the **Interoperability Extension** in ArcCatalog.
    a. For more information visit http://www.esri.com/software/arcgis/extensions/datainteroperability/common-questions.html
3. Click **Add Interoperability Connection**
4. A Dialog box appears
    a. Click the '…' button next to the **Format** pulldown menu
        i. Select the appropriate type of web service from the dialog
        ii. Click **OK**
    b. In the **Dataset** field, paste the URL of the capabilities document for the service you wish to add
        i. For more information, see Section 7, Testing Your Web Service
    c. Click the **Parameters** button.
        i. Click the '…' button next to **Table List**
        ii. Select the tables you wish to load
        iii. Confirm that the **Max Features** box is scaled appropriately
            1. If you are loading a dataset with 10 million features, the "Max Features" box must be scaled appropriately to ten million
            2. If you are loading a dataset with three (3) data points, the "Max Features" can be scaled down quite a bit
        iv. Click **OK**
    d. Once the layers have been added under the interoperability connections, you will see something like this in ArcCatalog
    e. Click **OK**
5. Depending on the size of the dataset, it may take some time to load. When it is finished, expand the Connection (you can also rename the connection, which can be helpful).
6. Drag and drop the layer into your ArcMap Project.

Q: All of my data are in Shapefiles. Can I deploy a shapefile as a GeoSciML-portrayal service?

A: The problem you will run into is the truncation of field names that occurs in shapefiles. Ideally you will have a full version of the data in an sde databse (or file/personal geodatabase). As mentioned in the above document, to be compliant with GeoSciML-portrayal, you will need to make sure there is no truncation in field names; they must be an exact match for the GeoSciML-portrayal schema.

Q: I would like to deploy a service using full GeoSciML v.3.0 schema; where can I go?

A: Visit the GeoSciML website

Q: Why would I want to create a WMS or WFS service at all?

A: To display data in a map format. Map format shows: data relative to its geographical surroundings, data relative to other data on the same map, data points relative to other data points in the same dataset (size, quantity, order, etc).

11 Appendix C: OneGeology Web Service Requirements

If you wish to submit your web service to the OneGeology portal, your web service must comply with OneGeology standards.

This appendix summarizes the service-level and layer-level requirements that must be respected for compliance with OneGeology standards.

11.1 Service-Level Metadata Requirements

The following section describes OneGeology standards as they relate to service-level metadata.

When you submit your web service to OneGeology, service-level metadata is harvested into the portal in the GetCapabilities response document. The minimum requirements for inclusion with OneGeology (source: http://www.onegeology.org/wmscookbook/2_4_1.html)

Table 4: One-Geology Service-Level Metadata Requirements

Field	Description	Conditions
Version (1.3.0, 1.1.1)	This will be automatically populated by GeoServer	N/A
Title	Required for 2-star service (2-stars); see Section 7.2 for more information	Required for 2 Stars
Abstract	May include data owner organization, goals, scale of data and other information	Required for 2 Stars
Access Constraints	Definition of access constraints, otherwise "none" (2 to 3 Stars)	Required for 2-3 Stars
Keywords	"OneGeology", "MD_DATE@2012-02-02", "MD_LANG@ENG" (3 Stars)	Required for 3 Stars
Data (owner) provider	Full name of data owner organization (not service provider) (2 Stars)	Required for 2 Stars
Image Format	This will be automatically populated by GeoServer	N/A
On-line Resource	Required by the WMS specification (2 Stars)	Required for 2 Stars
Contact Person	Named Individual or contact within data organization (2 to 3 Stars)	Required for 2-3 Stars
Country	Country of the data owner organization (2 to 3 Stars)	Required for 2-3 Stars
Email Address	Email address of the named contact (2 to 3 Stars)	Required for 2-3 Stars
Fees	Any fees required to use the WMS Service; otherwise "none" (3 Stars)	Required for 3 Stars

11.2 Service Title and Service Name Requirements

This section describes the OneGeology standards that govern web service **Title** and **Name**.

Service Title: Conforming to the naming conventions for service title is a 2 star requirement with OneGeology. **Service Titles** should contain the following tokens (in order):

[Geographical extent] [Data owner organization] [Language Code] [Scale] [Theme], where:

- **Geographic extent:** a rough approximation of the extent of the portrayal provided by the web service
  - Generally an ISO three letter country code (such as "USA") and, if necessary, a province, such as "USA-KY" for the state of Kentucky
  - See http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3 for ISO three-letter country codes
- **Data owner organization code**: an acronym or abbreviation for the entity that owns the data portrayed by the web service
  - The **data owner organization** *can be*, but is *not necessarily*, the name of the organization responsible for deploying the data as a web service
- **Language**: a two-letter ISO 639-1 code for the language in which the web service is portrayed
  - See http://en.wikipedia.org/wiki/List_of_ISO_639-1_codes for ISO language codes
- **Scale:** expresses the scale of the portrayal as a ratio (X:Y); large-scale portrayals should be abbreviated using SI symbols for million "M" and thousand "k"
  - An example of a valid scale in the service title is "1:1M"
- **Theme**: a short or abbreviated description of the data contained in the layer

Service Name: For version 1.3.0 services, the service name should be 'WMS'; for version 1.1.1, the service name should be 'OGC:WMS'

11.2.1 OneGeology Service Title Examples:

USA USGIN 1:3M Geology
USA-AZ AZGS ES 1:2k Faults

***Note: Language is omitted from the examples above, as they are in the default language English.

11.3 Layer Naming and Title Conventions

Each layer in a OneGeology-compliant web service should have a **Layer Title** and a **Layer Name**. The **Layer Name** should be identical to the **Layer Title**, using underscores instead of spaces and an abbreviated **Scale** token (see below)

Both **Layer Titles** and **Layer Names** should contain the following tokens (in order):

[Geographical Extent] [Data Owner Organization][Language Code][Scale][Theme], where

- **Geographic extent:** a rough approximation of the extent of the portrayal provided by the web service
  - Generally an ISO three letter country code (such as "USA") and, if necessary, a province, such as "USA-KY" for the state of Kentucky
  - See http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3 for ISO three-letter country codes
- **Data owner organization code**: an acronym or abbreviation for the entity that owns the data portrayed by the web service
  - The **data owner organization** *can be*, but is *not necessarily*, the name of the organization responsible for deploying the data as a web service
- **Language**: a two-letter ISO 639-1 code for the language in which the web service is portrayed
  - See http://en.wikipedia.org/wiki/List_of_ISO_639-1_codes for ISO language codes
- **Scale:** expresses the scale of the portrayal as a ratio (X:Y); large-scale portrayals should be abbreviated using SI symbols for million "M" and thousand "k"
  - An example of a valid scale in the service title is "1:1M"
  - **Layer Names** should abbreviate the **Scale** token by removing the **antecedent** and the **colon**. For example, a 1:23k scale would be abbreviated as 23k
- **Theme**: a short or abbreviated description of the data contained in the layer

11.3.1 Layer Title Examples

GBR BGS 1:625k Bedrock Age
USA USGIN 1:3M Geologic Age

11.3.2 Layer Name Examples

World_25M_Geol_Units
US-KY_KGS_24k_Faults
USA_USGIN_3M_Lithostratigraphy

11.3.3 A sample WMS GetCapabilities document demonstrating Layer Names and Layer Titles

Code Example 3: OneGeology-compliant Layer Name and Layer Title in a WMS GetCapabilities document

```
<Layer queryable="1" >
        <Name>USA_USGIN_3M-Geologic_Age</Name>
        <Title><![CDATA[USA USGIN 1:3M Geologic Age]]></Title>
        <Abstract><![CDATA[This layer provides a portrayal in which units are categorized according to the representative geologic age of the
        unit.]]></Abstract>
        <KeywordList>
                <Keyword>continent@North America</Keyword>
                <Keyword>geographicarea@USA</Keyword>
                <Keyword>dataprovider@USGIN</Keyword>
```

11.4 Additional Layer-level Metadata Requirements

Aside from **Layer Title** and **Layer Name** (discussed in Section 7.3), OneGeology standards require additional layer-level metadata, including the following:

Abstract: A description of your layer data; should be as comprehensive and interoperable as possible. You may wish to include other metadata in the abstract, such as information about your organization and other data you make available. You may also wish to include a statement on access constraints (source: http://www.onegeology.org/wmscookbook/2_6.html).

Keywords: Tags used to locate and categorize your layer. Required and conditional keywords are described in Table 5 (source: http://www.onegeology.org/wmscookbook/2_6.html):

Table 5: OneGeology Keywords

Keyword Description	Appearance in GetCapabilities Document	Example	Cardinality
Continent	continent@value	continent@North America	Required
Subcontinent	subcontinent@value	subcontinent@India	Conditional
Geographic area (country)	geographicarea@value	geographicarea@USA	Required
State (region or province)	subarea@value	subarea@Arizona	Conditional
Data Provider	dataprovider@value	dataprovider@USGIN	Required
Service Provider	serviceprovider@value	serviceprovider@AZGS	Required

Bounding Box: A rough indicator of the extent of your data; can be computed from data in the GeoServer 'Edit Layer' interface (Figure 17), or customized.

SRS: The spatial reference system used by your web service, indicated by an EPSG code.

Spatial reference systems specify a datum, a reference ellipsoid, and a coordinate system for cartographic representations. Most known spatial reference systems are associated with an EPSG code number; EPSG stands for European Petroleum Survey Group.

The default SRS is WGS 1984 (EPSG:4326). For the purposes of interoperability, your web service must be deployed using EPSG:4326.

MetadataURL: A link to an external ISO-valid metadata document. A **MetadataURL** is required for services deployed on OneGeology. To create an ISO Metadata record using ISO19115:2003, use the USGIN Metadata Wizard.

Readable Legend Graphic: URL(s) for the legend graphic. Readable Legend Graphic URLs should be present in the getCapabilities document by default when an SLD is assigned to a layer. For more information on SLDs, see Section 5, Styling.

12 Glossary

Attribute (GIS)

Within the context of geographic information systems, an attribute describes a feature. For example: attributes of a fault feature might include the latitude and longitude coordinates for each fault, as well as the fault's age, dip, and slip.

Database

A method of storing data. In a database, data is divided up into database records; in turn, database records are divided up into database fields. The advantage of a database is that it can be sorted and searched by field contents.

Though modern databases are usually digital, a physical example of a database is a card catalog in a public library. In a card catalog, data is divided up into individual cards, which are directly analogous to database records. Each card (record) in the catalog corresponds with and describes a book. The information about each book is divided up into fields: title; author; subject; publication date; etc. Digital databases can be in tabular format (that is, a table) in which rows represent individual records and columns constitute fields; or they can be viewed record-by-record.

Database Field

A subdivision of a database record in which a specific type of data is entered. Using the analogy of a card catalog in a public library: if the card catalog is directly analogous to a database, and if cards in the catalog are directly analogous to database records, then the different subdivisions of information found in each card in the catalog (title, author, publishing date, etc.) are all database fields.

Database fields are functionally similar to markup language elements.

Database Record

A subdivision of a database. Using the analogy of a card catalog in a public library, each record in a database is analogous to an index card in the card catalog; each card (record) corresponds with and describes an individual book. Database records are further subdivided into fields, which contain specific kinds of data (title, author, etc).

Element

Elements are logical document components found in markup languages such as XML and HTML. Elements simultaneously define the structure and content of a document; they are constituted by two markup language tags and the content between the two. For example:
<tag>content</tag>

The two tags and their content form an element. In a more concrete example, HTML uses the tag to demarcate text that should be Italicized. So, an Italicized element of an HTML document would appear as follows:

Italicized content

In a web browser, this element would produce the following result:

Italicized content

For a more detailed overview of elements, see the USGIN XML Tutorial.

Markup language elements are functionally similar to database fields.

Feature

The word *feature* may indicate a *geologic* feature, such as a fault, formation, or dike. Alternatively, a feature can be a GIS representation of a real-world resource. GIS features do not always correspond with geologic features because GIS software can be used to represent anthropogenic objects such as buildings, roads, or canals. The definition of the term feature therefore depends on its context: it can be used to mean either a geologic feature or a GIS feature.

GIS features are often described by attributes.

Interchange Format

An interchange format is an application-neutral format used to exchange data between systems. For example, if one needed to transfer data between an Oracle database and an Access database, an interchange format could be used to transfer this data from one database to the other.

Interoperability

The capability to communicate, execute programs, or transfer data in a manner that requires the user to have little or no specialized knowledge. Interoperability is based on a shared data schema and the use of standard vocabulary terms.

Markup Language

Markup languages, such as HTML and XML, use elements to annotate text in such a way that the structure of the document is distinguishable from the contents of the document. Consequently, markup language documents are very good for storing data, since the distinction between structure and content is transparent and immediately discernible.

Metadata

Literally, data about data. Metadata is used to organize, categorize, browse, and discover data resources. For more information, see the USGIN Metadata Tutorial.

Schema

A schema is the structural framework for a data model or collection of attributes that describe something. For practical purposes, schemas *structure* documents by dictating where and how data from a dataset should be entered into a document.

Uniform Resource Identifier (URI)

Uniquely identifies a resource. See the USGIN URI Tutorial for more information about URIs.

Web Service

Under the client-server relationship, a web service is a server-side protocol that specifies syntax used by client software applications to interact with web-accessible resources on the server. As long as a request for web-accessible resources from the server uses the appropriate syntax, web services can provide a response. Web service syntax defines required input parameters, operation output, and the results of any operations performed.

Web services decouple the client and the server, allowing the two systems to develop independently: as long as the web service is in place and specified syntax is used, changes may be made to either server or client without jeopardizing the server's ability to make web-accessible data available upon request or the client's ability to request web-accessible resources from the server. Consequently, web services can field requests from multiple different client platforms, including mobile devices.

Etymological Note: "hosting data as a service" is considered synonymous with "hosting data in accordance with a web service protocol." Likewise, data hosted in accordance with a web service protocol is often referred to as a web service.

Web services are well suited to the task of sharing digital geoscience information in accordance with **USGIN** specifications (for a number of reasons: owing to standardized data requests, they can be accessed from a variety of platforms; they are read-only and preserve data ownership; and they come in a number of different flavors for different purposes.

The Open Geospatial Consortium has provided specifications for several different flavors of web service that are relevant to geographic information systems, **USGIN**, and the State Geothermal Data project. These include, but are not limited to, **web map services (WMS)** and **web feature services (WFS).**

Web Map Services (WMS) provide a georeferenced map image within a geographic bounding box. Existing georeferenced map images are acquired by **getMap** requests sent by client software to the web map service; **getMap** requests specify the geographic area of the bounding box and the map that is desired. A variety of other parameters allow control over image size, map projection, and other details. The server responds by returning an image file (typically .tif, .jpg or .png) for the requested area.

For example, a **getMap** request could be used to request a georeferenced satellite photo of Arizona; if the satellite photo exists on the server as a web-accessible image file, the web service will respond to the **getMap** request by providing the client with the desired image.

Web map service image files are often based on a source shapefile or feature class. Map services support **FeatureInfo** requests at point locations; **FeatureInfo** request returns are not standardized and depend on server configuration. Map services are most useful for client applications designed to visually explore geographic relationships.

Web Feature Services (WFS) provide collections of georeferenced features described by attributes. Attributes are essentially a collection of database records that describe the location and characteristics of bound features.

For example a feature service containing data about river systems might provide features with linear geometry representing river segments; each feature might be described by attributes such as average flow rate, width, and depth for the segment.

Feature services are useful for geospatial analysis performed using data describing individual geographically located entities. For example, a client application used to calculate the most efficient route between points in a city utilizing real-time traffic information would require a representation of the streets as features described by attributes.

Owing to the large amount of data associated with WFS, these web services tend to require fast Internet connections.

XML

XML stands for Extensible Markup Language. XML documents are often used as interchange formats for database records, owing to the ability for XML elements to emulate database fields. To use XML documents as an interchange format, USGIN defines XML schemas in which each XML document corresponds with a specific database record and each element corresponds with data entered in a database field. In these documents, tags define the database fields.

For example: well_name is a field within the WellHeader content model. Any known oil and gas well names go in the well_name field. An XML schema for expressing database records within the WellHeader content model might appear as follows:

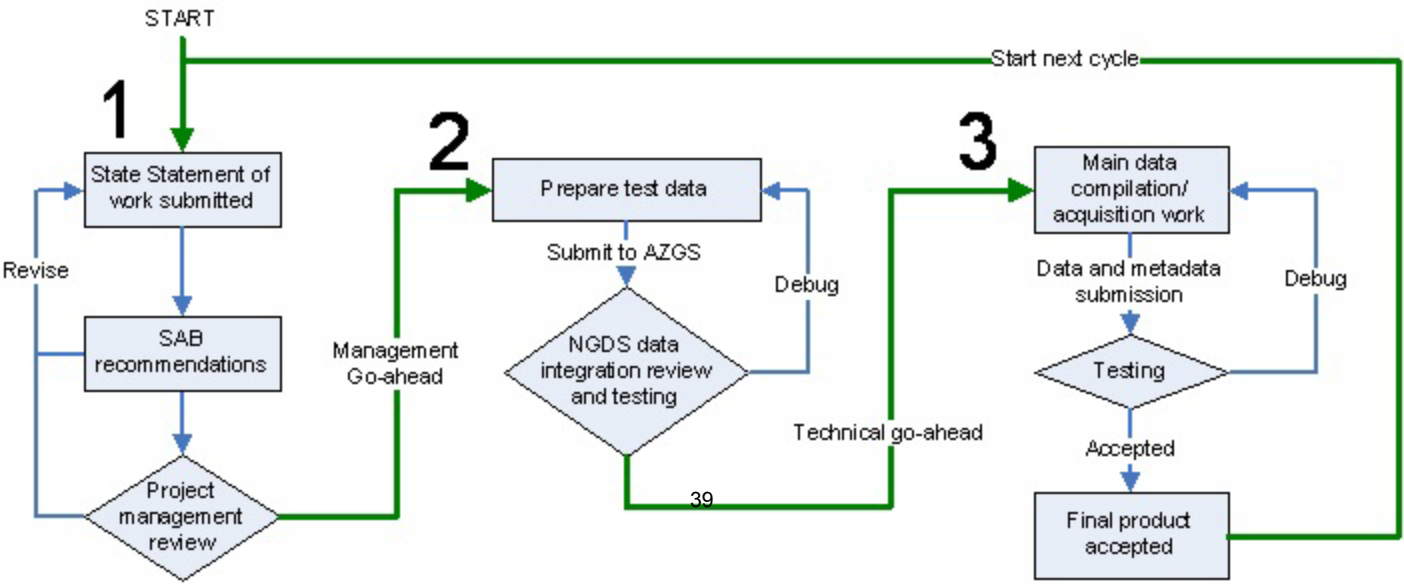
<content_model_name>WellHeader</content_model_name>

<database_record>

<well_name>Union Oil Company of California-South 7 Veysey</well_name>

In this example, the Union Oil Company of California-South 7 Veysey well name constitutes the content delineated by the <well_name> tag.

For a more detailed overview of XML, see the USGIN XML Tutorial.


START
Start next cycle
1
State Statement of work submitted
Revise
SAB recommendations
Project management review
Management Go-ahead
2
Prepare test data
Submit to AZGS
Debug
NGDS data integration review and testing
Technical go-ahead
3
Main data compilation/ acquisition work
Data and metadata submission
Debug
Testing
Accepted
Final product accepted

USGIN
U.S. Geoscience Information Network

USGIN Metadata Profile: Use of ISO metadata specifications to describe geoscience information resources

Version 1.2draft

Title:
USGIN Metadata Profile: Use of ISO metadata specifications to describe geoscience information resources

Latest released version:
http://lab.usgin.org/profiles/usgin-iso-19139-profile

Creator:
USGIN Standards and Protocols Drafting Team

Editors:
Stephen M. Richard and Wolfgang Grunberg

Creation date:
8/18/2009

Last revision date:
3/2/2014 12:11 PM

Document Status:
Live trunk version, post 1.2 edits

Publisher:
Arizona Geological Survey

Description:

This document describes recommended practices for using ISO19139 xml encoding of ISO 19115 and ISO 19119 metadata to describe a broad spectrum of geoscience resources. The document provides guidance for the population of ISO19139 encoded metadata documents to enable interoperability of catalog service clients with multiple servers conforming to this profile.

Contributor:

See acknowledgements

Document Identifier:

gin2010-009.1.2.x http://repository.usgin.org/uri_gin/usgin/dlio/337

Notices

Neither the USGIN project, nor any of the participating agencies, take any position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither do they represent that there has been any effort to identify any such rights.

This document and the information contained herein is provided on an "AS IS" basis and USGIN DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Revision History

Version	Date	Comments	By
0.1	2009-08-18	Initial draft document	Wolfgang Grunberg
0.7	2009-10-28	Revisions, addition of material, re-title, focus on use of ISO 19139. Revising adding material done, circulate for review, comment. No discussion of serviceIdentification section as yet; this will have to be added, at which point the file title will be revised.	Stephen Richard
0.7.1	2009-11-19	Revisions, crosschecked with ISO 19139 and NAP Profile, added dataset metadata example	Wolfgang Grunberg
0.8	2009-11-24	SMR review, editing based on W. Grunberg comments	Stephen Richard
0.8.1	2009-11-30	Codelist review, approved previous tracked changes	Wolfgang Grunberg
0.9	2009-12-03	Add section on distribution and distributor, review and accept changes	Stephen Richard
0.9.1	2009-12-04	First public release of draft dataset and dataset series profile, Word and PDF file preparation	Wolfgang Grunberg
0.9.3	2009-12-11	Review edits and comments. Split out data identification element into separate tables for MD_DataIdentification and SV_ServiceIdentification; add service metadata elements.	Stephen Richard
1.0.0	2010-01-11	Review dataset and service metadata sections; update metadata examples. Second public release of draft dataset and dataset series profile, Word and PDF file preparation	Wolfgang Grunberg
1.0.1	2010-01-20	Additional examples, codeList overhaul	Wolfgang Grunberg
1.1.0	2010-02-02	Change recommendation for codelist usage to use ISO identifiers; update codelist discussion. Update distribution format discussion. Review and formatting. Public release, Word and PDF file preparation	Stephen Richard, Wolfgang Grunberg
1.1.1	2010-03-04	Review, title page changes	Stephen Richard
1.1.2	2010-03-04	AZGS Contributed Reports release, Word and PDF file preparation	Wolfgang Grunberg
1.1.3	2010-11-16	Minor edits and bug fixes, modify title and title page text.	Wolfgang Grunberg, Stephen Richard
1.1.4	2010-11-18	Correct example document for distribution element conventions, add simple data quality statement, funding acknowledgement, update use case section	Stephen Richard
	2011-12-28	add ‘Recommended practices’ to title. Update language in distribution section.	Stephen Richard
	2013-06-13	Finish major review and edit of text to bring in alignment with practice that has evolved since 2009. Remove 'Recommended practices' in title, add 'USGIN Metadata Profile'. Remove 'service metadata example, it was too deeply flawed to fix. May add a service metadata example later.	Stephen Richard
	2014-01-23	Add hierarchyLevelName string values for use in metadata document in Table 1	Stephen Richard

Acknowledgement

Many individuals and organizations have contributed to or inspired the development of these Guidelines.

The USGIN Standards and Protocol Drafting Team include (alphabetically):

Ryan Clark - AZGS
Wolfgang Grunberg – AZGS
Dominic Owen - AZGS
Stephen M. Richard – AZGS

Funding Provided by (chronologically):

National Science Foundation under EAR-0753154 to the Arizona Geological Survey acting on behalf of the Association of American State Geologists; 2009.

U.S. Department of Energy under award DE-EE0001120 to Boise State University; 2009.

U.S. Department of Energy under award DE-EE1002850 to the Arizona Geological Survey acting on behalf of the Association of American State Geologists, 2010.

Special Thanks to (alphabetically):

ANZLIC
INSPIRE
MEDIN
NAP
NOAA

Contact Information

Arizona Geological Survey
416 W. Congress St., Suite 100.
Tucson, Arizona 85701-1381
Phone: 520.770.3500
Fax: 520.770.3505

Email

metadata@usgin.org

Online

http://usgin.org
http://lab.usgin.org

Table of Contents

1 Introduction 8
1.1 Normative References 8
1.2 Purpose 9
1.3 Terminology 9
1.4 ISO Schemas Location 10
2 Overview of the Profile 12
2.1 Quick Reference 12
2.1.1 Non-service resources 14
2.1.2 Service Resources 14
2.2 USGIN specification constraints and recommendations 15
2.3 USGIN specification extensions 15
2.4 General Objectives 15
2.5 Requirements 16
2.6 Use cases to be supported 16
2.7 Resources of interest 17
3 USGIN Usage of Metadata Elements 23
3.1 Core spatial dataset, dataset series, and service elements 23
3.2 Dataset Identification properties (MD_DataIdentification) 40
3.3 Service identification elements (SV_ServiceIdentification) 50
4 Usage notes for Metadata Elements 63
4.1 Metadata file identifier 63
4.2 Metadata hierarchy 63
4.3 Metadata Contact vs. Resource Citation vs. Resource Contact 63
4.4 Resource Title 64
4.5 Resource Abstract 64
4.6 Resource Type 64
4.7 Resource Locator 65
4.8 Unique Resource Identifier 65
4.9 Browse Graphics 65
4.10 Resolution and equivalentScale 66
4.11 Resource Language 66
4.12 Encoding of Vertical Extents 66
4.13 Content information: Entities and Attributes 66
4.14 Use of MD_Distribution and MD_Distributor 67
4.15 Distribution Format 69
4.15.1 Digital resources 69
4.15.2 Non digital resources 70
4.16 CI_OnlineResource 72
4.16.1 CI_OnlineResource protocol 73
4.16.2 CI_OnlineResource applicationProfile 73
4.16.3 CI_OnlineResource name and description 73
4.16.4 CI_OnlineResource function property 74
4.16.5 Scoped name in service distributions 75
4.17 Responsible parties and logos 75
4.18 Extensions to CharacterString 76
4.18.1 Web extensions 76
4.18.2 Language localization 76
4.18.3 Codelists 77

4.19 Geographic bounding box 84
4.20 Data Quality 84
4.20.1 Simple quality statement 84
4.21 Lineage 84
4.22 Temporal extents 86
4.23 Operation metadata 87
5 Abbreviations 88
6 References 89
6.1 Cited literature 89
7 Codelists 90
7.1 ServiceType 90
7.2 Linkage name conventions 92
8 Examples 94
8.1 USGIN ISO 19139 Minimum Dataset Metadata 94
8.2 USGIN ISO 19139 Dataset Metadata 102
8.3 ISO19110 Feature Catalog example 125

Tables and Figures

Table 1. Summary of resource types described by metadata for US Geoscience Information Network catalogs.. 14
Table 2. Description best practices for ISO19139 metadata elements in USGIN profile.. 18
Table 3. Dataset Identification properties (MD_DataIdentification) 34
Table 4. Service Identification properties (SV_ServiceIdentification) 43
Table 5. Example format strings for digital files.. 61
Table 6. USGIN Distribution formats for non digital resources. 62
Table 7. OnlineFunctionCode values from NAP (INCITS 453) and ESRI Geoportal toolkit v. 3.1.. 63
Table 8. Codelist crosswalk between ISO, NAP and USGIN. 69
Table 9. Usage of data quality scope description elements 75
Table 10. Inspire spatial data service type 82
Table 11. USGIN service type vocabulary.. 82
Table 12. USGIN Names to identify special linkage URL's for CI_Online Resource. 83
Figure 1. gmd:MD_Distribution_Type diagram 60

1 Introduction

A key component of a distributed information network is a catalog system, a collection of resources that allow data and service providers to register resources, and data consumers to locate and use those resources. Currently, many online catalogs are web pages with collections of URLs for services, or services are discovered accidently or by word of mouth. The vision is to enable a web client (portal) to search across one or more metadata registries without having to configure the client individually for each of the registries that will be searched. Thus, metadata providers can focus on data development, without having to also develop web clients to enable search of that metadata.

The Open Geospatial Consortium (OGC) Catalog Service for the Web (CSW) specification defines a collection of basic operations for searching catalogs of metadata via the web. Engineering the desired interoperability requires adding additional constraints on CSW operation; one of the major constraints is selection of the xml schema that will be used to encode metadata for the service. The core CSW specification requires use of a basic xml schema that includes content defined by the Dublin Core Metadata specification [Dublin Core, 2008-01-14]. This document concerns use of the ISO19115 and ISO19119 content models implemented using the ISO19139 xml schema for encoding of metadata content. Some more specific constraints on use of this implementation may be included in a separate document (USGIN, 2012) describing metadata constraints for different kinds of resources.

1.1 Normative References

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 19115 designates these two normative references:

- ISO 19115:2005, *Geographic information - Metadata*
- ISO 19115/Cor.1:2006, *Geographic information – Metadata, Technical Corrigendum*

ISO 19119 designates these normative references:

- ISO 19119:2005, *Geographic information - Services*
- ISO 19119:2005/Amd 1:2008, *Extensions of the service metadata model ISO 19108 designates*:
- ISO 19108:2005, *Geographic information – Temporal Schema*

ISO 639-2, Codes for the representation of names of languages - Part 2: Alpha-3 code control ISO 8601, Data elements and interchange formats - Information interchange - Representation of dates and times

ISO/TS 19139:2007, Geographic information - Metadata – XML Schema Implementation

OGC 07-006r1¸ OpenGIS Catalog Services Specification version 2.0.2, Corrigendum 2 release, 2007

OGC 07-045, OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile, Version 1.0.0, 2007

INCITS 453-2009, North American Profile of ISO 19115:2003 – Geographic Information – Metadata (NAP-Metadata), 2009, American National Standards Institute, Inc.

ISO 10646-1, Information technology — Universal Multiple-Octet Coded Character Set (UCS) — Part 1: Architecture and Basic Multilingual Plane

RFC 2119, Key words for use in RFCs to Indicate Requirement Levels, Network Working Group, 1997.

1.2 Purpose

The USGIN uses the ISO 19115/19119 specifications for metadata content, and the ISO 19139 XML schema for encoding this content search results provided by USGIN CSW services. This profile conforms to most of the provisions of the North American Profile of ISO metadata (INCITS 453-2009, referred to as NAP), except it allows multiple distributor-format-transferOptions bindings for resource distribution, and recommends use of ISO19115 codelist values. This USGIN document is meant to provide guidance on the use of the ISO19139 XML schema to encode metadata for geoscience resources, with sufficient detail that application developers can produce software clients that utilize the metadata for automated discovery, evaluation, and utilization workflows. The focus of the profile is to enable interoperable catalog services for discovery, evaluation, and access to information resource of interest to geoscientists.

1.3 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in Internet RFC 2119.

Application profile: a schema that consists of data elements drawn from one or more namespaces, combined together by implementers, and optimized for a particular local application. (Rachel Heery and Manjula Patel, 2000, http://www.ariadne.ac.uk/issue25/app-profiles/)

Catalog application: Software that implements a searchable metadata registry. The application must support the ability to register information resources, to search the registered metadata, to support the discovery and binding to registered information resources within an information community.

Codelist (also as Code list): a controlled vocabulary that is used to populate values for an XML element.

Data product specification: a definition of the data schema and value domains for a dataset. The data schema specifies entities (features), properties associated with each entity, the data type used to specify property values, cardinality for property values, and if applicable, other logical constraints that determine data validity. Value domains are specified for simple data types—strings or numbers, and may include controlled vocabularies for terminology required to specify some properties.

Dataset series: collection of datasets sharing the same product specification (ISO 19115). ISO 19115 does not define product specification. For the purposes of USGIN, a product specification defines a data schema, any required controlled vocabularies, and recommended practices for use of schema (see Data product specification).

Dataset: an identifiable collection of data (ISO19115). USGIN extends the concept of data items to include physical artifacts like books, printed maps and diagrams, photographs, and material samples--any identifiable resource of interest. DCMI definition is "Data encoded in a defined structure" with additional comment "Examples include lists, tables, and databases. A dataset may be useful for direct machine processing." A digital dataset is a collection of data items in which individual data items are identified and accessible. Metadata for the collection is a different type than metadata for individual items in the collection (dataset vs. features). Criteria for what unifies the collection are variable (topic, area, author...). Data items may represent intellectual content -- information content and organization (data schema) -- or may represent particular manifestations (formats) of an intellectual artifact.

Interoperability: "The capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units." ISO/IEC 2382-01 (SC36 Secretariat, 2003)

Metadata element: a discrete unit of metadata (ISO 19115), an attribute of a metadata entity. A metadata element contains some content specifying the value of the element; this content may be simple—a number or string, or may be another metadata entity.

Metadata entity: a named set of metadata elements describing some aspect of a resource.

Metadata register: an information store that contains a collection of registered metadata records, maintained by a metadata registry. (ISO 11179)

Metadata registry: an information system for assignment of unambiguous identifiers to administered metadata records. (ISO 11179)

Metadata section: Part of a metadata document consisting of a collection of related metadata entities and metadata elements (ISO 191115).

Metadata: data about a resource in some context. Generalize from ISO 11179 definition of metadata, which constrains the scope to data about data. For USGIN purposes, metadata may describe any resource—including electronic, intellectual, and physical artifacts. Metadata specify characteristics that can be queried and presented for evaluation and further processing by both humans and software.

Profile: set of one or more base standards and - where applicable - the identification of chosen clauses, classes, subsets, options and parameters of those base standards that are necessary for accomplishing a particular function [ISO 19101, ISO 19106]

Resource: An identifiable thing that fulfills a requirement. Usage here is closer to definition used in RDF (www.w3.org/TR/REC-rdf-syntax*),* generalized from ISO19115, which defines resource as an ‘asset or means that fulfills a requirement’ without defining asset or means. "An object or artifact that is described by a record in the information model of a catalogue" (OGC 07-006r1)
Service metadata: metadata describing the operations and information available from a server.

Source Specification: The specification or standard that is the subject of a profile.

User Community: A group of agents who decide to make a similar usage of a source specification in order to be able to interoperate.

Note that throughout this document, the names of XML elements are shown in `this typecase`. Long X-paths have been broken with non-breaking hyphen characters. Note that hyphens are not used in any XML attribute or element name, so if they appear in the text, they are strictly for better text wrapping. In Xpath expressions `/../` indicates that some elements have been omitted from the path.

1.4 ISO Schemas Location

ISO I9139 xml schemas are hosted by the International Organization for Standardization (ISO) at http://standards.iso.org/ittf/PubliclyAvailable-Standards/ISO_19139_Schemas/gmd/. The ISO Technical Committee 211 (TC211) also hosts the schemas at http://www.isotc211.org/2005/gmd/. These schema implement ISO Technical Specification 19139:2007 (dated 2007 Apr 17), which appears to include the changes from ISO 19115:2003 Cor 1;2006(E), but do not include any of the service metadata content.

The Open Geospatial Consortium (OGC) hosts a copy of the schemas in an online repository at http://schemas.opengis.net/iso/19139/. In the OGC repository, two versions are posted: 20060504 and 20070417. Unfortunately, these two directories both contain schema with the same target namespace, so there is no clear way to distinguish applications that are based on one or the other. The medatadaEntity.xsd in the two directories is identical; other schema have not been compared (but see discussion paper gin2009-005 at http://lab.usgin.org/node/269). The 20070417 directory contains schema implementing ISO Technical Specification 19139:2007 (dated 2007 Apr 17), apparently identical to that in the ISO repositories, but this is not declared in any included documentation (need metadata on the metadata schema!).

The 20070417 version of the ISO 19139 schemas references GML 3.2.1. However, there is no mention of the SRV namespace (http://www.isotc211.org/2005/srv) in this schema. The SRV namespace is required to specify information about dynamic, online services such as WFS and WMS, so the 20070417 version is not useful for metadata catalogs that register services.

In order to create metadata for both static datasets and dynamic, online services and for use with CSW, the OGC created an xml schema that merges the schema for ISO19115 (dataset metadata) and ISO19119 (service metadata) (see section D.1.5, page 105 in OGC 07-045). The way that was accomplished was by creating a schema located at http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd. This schema simply imports .. iso/19139/20060504/gmd/gmd.xsd and .. iso/19139/20060504/srv/srv.xsd. Thus for CSW 2.0.2 implementations, the 20060504 versions of the ISO19139 schema must be used.

2 Overview of the Profile

2.1 Quick Reference

This section provides a summary of the rules for an ISO 19139 XML document to be considered a valid USGIN metadata record

1. MD_Metadata/fileIdentifier/gco:CharacterString must be present, and may be interpreted by client applications as a string that is unique to each particular metadata record in the scope of the catalog from which it was obtained. This is typically a UUID autogenerated by the metadata editing tool or import process

2. MD_Metadata/language/gco:CharacterString will be assumed to be 'eng' (English) unless another value is specified. Any characters after the first three letters in this character string may be ignored by client applications.

3. Unless otherwise specified, the character set element will be assumed to be:

```
<gmd:characterSet>
      <gmd:MD_CharacterSetCode
codeListValue="utf8">UTF-8</gmd:MD_CharacterSetCode>    </gmd:characterSet>
```

Unless otherwise specified MD_Metadata/hierarchyLevel/MD_ScopeCode will be assumed to be:

```
<gmd:hierarchyLevel>
  <gmd:MD_ScopeCode  codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/
  ISO_19139_Schemas/resources/Codelist/gmxCodelists.xml#MD_ScopeCode"
    codeListValue="dataset">dataset</gmd:MD_ScopeCode>
</gmd:hierarchyLevel>
```

Metadata with scope code values that are not in the set {`collectionHardware`, `collectionSession`, `dataset`, `series`, `nonGeographicDataset`, `dimensionGroup`, `fieldSession`, `software`, `service`, `model`, `tile`} will be considered out of scope and may be ignored by a harvesting catalog.

4. MD_Metadata/hierarchyLevelName/gco:CharacterString is mandatory to identify the USGIN resource type from Table 1 (above). Default value is "Dataset". Encode hierarchy by including `hierarchyLevelName` elements for all broader resource categories. E.g. default should also include a `hierarchyLevelName`="Collection" element.

USGIN Hierarchy level name	**Broader Resource Type**
Collection	
Dataset	Collection
Catalog	Dataset
Physical artifact collection	Collection
Document	
Image	Document
StillImage	Image
Human-generated image	StillImage
Photograph	StillImage

Remote sensing Earth image	StillImage
Map	Human-generated image | StillImage
MovingImage	Image
Sound	Document
Text	Document
Hypertext document	Text
Model	
Physical artifact	
Service	
Software	
Stand-AloneApplication	Software
WebApplication	Software

5. USGIN requires at least one MD_Metadata/contact/CI_ResponsibleParty with role.CI_RoleCode@codeListValue = "originator" (CI_RoleCode element value = "originator") that identifies the original source of the metadata record. The CI_ResponsibleParty element must include a contact e-mail address (electronicMailAddress) or telephone number (voiceTelephone), and one of individualName, organisationName, or positionName must be present.

6. USGIN profile requires a MD_Metadata/dateStamp/gco:DateTime (Note this contrasts with INSPIRE mandate to use dateStamp/gco:Date) that specifies the date and time when the metadata record was created or last updated.

7. MD_Metadata/metadataStandardName/gco:CharacterString must be the string "ISO-USGIN" to identify a metadata record as conforming to this profile.

8. MD_Metadata/metadataStandardVersion/gco:CharacterString must be the string "1.2" to identify a metadata record as conforming to this profile.

9. MD_Metadata/identificationInfo/*MD_Identification*/citation/CI_Citation/title is mandatory. A meaningful title that is unique within the scope of the catalog should be provided for each resource that is described by a metadata record. This may be in a MD_DataIdentification or SV_ServiceIdentification element.

10. MD_Metadata/identificationInfo/*MD_Identification*/citation/CI_Citation/date is mandatory, and must be specified as an xml DataTime data type: `YYYY-MM-DDThh:mm:ssZ`. Some validators require time zone as well; recommend GMT indicated by 'Z', other time zones are indicated by "+" or "-" integer offset from GMT, e.g. "+7" for US MST.

11. MD_Metadata/identificationInfo/*MD_Identification*/citation/CI_Citation/responsibleParty is mandatory with role CI_RoleCode@codeListValue one of {originator, principalInvestigator, processor, author}. This element identifies the agent responsible for the intellectual content of the described resource. The CI_ResponsibleParty element must include a contact e-mail address (electronicMailAddress) or telephone number (voiceTelephone), and one of individualName, organisationName, or positionName must be present.

12. identificationInfo/*MD_Identification*/abstract is mandatory, but nilable. Should include a description of resource content and any other information that will help users evaluate and use the resource.

13. identificationInfo/*MD_Identification*/-extent/EX_Extent/-geographicElement/-EX_Ge

ographicBoundingBox is mandatory if the resource has a geographic footprint. Geographic extent must be represented by bounding box in WGS 84 decimal degrees. Non-geographic resources are given a location keyword 'non-geographic'.

2.1.1 Non-service resources

Any resource that is not a service is described by a metadata record with a MD_DataIdentification as the concrete implementation of the abstract *MD_Identification* element.

1. If the metadata describes a physical resource, identificationInfo/MD_DataIdentification/-pointOfContact/CI_ResponsibleParty must provide contact information for the agent who is the resource steward. Count of (individualName + organisationName + positionName) must be > 0. The CI_ResponsibleParty/role/CI_RoleCode is from CI_RoleCode codelist. ISO role codes for physical resource point of contact are {custodian, owner, pointOfContact}.

2. MD_Metadata/distributionInfo/MD_Distribution/distributor/MD_Distributor/distributorContact/CI_ResponsibleParty is mandatory with role.CI_RoleCode@codeListValue = "pointOfContact" (CI_RoleCode element value = "pointOfContact") that provides a contact point to report problems with accessing the resource through distributions associated with the distributor. The CI_ResponsibleParty element must include a contact e-mail address (electronicMailAddress) or telephone number (voiceTelephone), and one of individualName, organisationName, or positionName must be present.

3. MD_Metadata/distributionInfo/MD_Distribution/distributor/MD_Distributor/distributionOrderProcess/MD_StandardOrderProcess is mandatory if the described resource is not available online. This should be a free text explanation of how to access the described resource.

4. MD_Metadata/distributionInfo/MD_Distribution/distributor/MD_Distributor/distributorTransferOptions/-MD_DigitalTransferOptions/online/CI_OnlineResource/linkage/URL is mandatory if the described resource is available online. Note that if online access is through a service, additional CI_OnlineResource properties may be required to provide sufficient information to allow software agents to use the metadata to connect to the service.

2.1.2 Service Resources

Service metadata records in the USGIN profile are intended to describe a particular service instance from the point of view of a service implementation and enpoint. Each dataset exposed by the service should be described by a separate metadata record with MD_HierarachyLevel = 'Dataset'.

1. The srv:ServiceType element value must be populated from a string in the USGIN ServiceType codelist (Table 11).

2. Service status is required. Value is from MD_ProgressCode codelist. ISO Code names applicable to services include {`completed, obsolete, onGoing, planned, required, underDevelopment`}.

3. Service metadata records must provide a coupling type attribute specifying 'tight', 'mixed', or 'loose' coupling with particular data.
4. If coupling type is 'tight' or 'mixed' a at least one `coupledResource` element is mandatory.
5. At least one service operation MUST be described with the `operationDescription/-CharacterString` = 'serviceDescription'. The `identificationInfo/SV_ServiceIdentification/-containsOperations/SV_OperationMetadata/connectpoint` link for this operation MUST retrieve the service-specific self description document (e.g. WSDL, GetCapabilities, WADL). The `CI_OnlineResource` in this connectPoint elment must have `CI_OnlineResource/name` = "serviceDescription" (from the table in section 7.2 Linkage name conventions).

2.2 USGIN specification constraints and recommendations

Recommended practices

- DatasetURI if there is one. Nilable if not available.
- All resources should have a point of contact responsible party thaqt specifies the agent currently responsible for curation and maintenance of the resource.
- Allow multiple distributor-format-transfer option combinations for a single resource. This is critical information for enabling automated access to resources based on the metadata content. The ISO metadata model allows different approaches to binding between a distributor and resource distributions from that distributor when there are multiple distributors and distributions. This profile defines conventions for documenting those bindings.
- Representation of aggregated resources done using `identificationInfo/MD_Data-Identification/aggregationInfo/MD_AggregateInformation`, not `MD_Metadata/parentIdentifier`
- Vertical resource extent may use an elevation convention or a depth convention. CRS referenced to mean sea level, meters, measured positive up.
- Resolution expressed using `equivalentScale/MD_RepresentativeFraction/denominator` is mandatory.
- Recommended distribution format codes (Table 6) for `distributionFormat/name` introduced for categorization of physical resources, like a book, rock sample, paper document. USGIN recommends use of MIME types if they are registered for the format, and provides a recommended syntax for file formats that do not have corresponding MIME types.
- Table 12 is a list of recommended `CI_OnlineResource/name` strings used to identify special online resources link icons for branding.
- Use of a controlled vocabulary for a CharacterString data type value is indicated by including an xsi:type=" gml:CodeType" on the gco:characterString element, which then requires a codeSpace attribute (see 4.14.2-Non digital resources and 7.2-Linkage name conventions). This codeSpace should be the URI for the vocabulary used. The CharacterString element value MUST then be an identifier from that vocabulary. This convention turns the gco:CharacterString value into a GML scoped name or gco:LocalName element.

- If the metadata resource is not specific to a geographic extent, the *`MD_identification/`*`extent/-EX_Extent/geographicElement/` should have a

nilReason="inapplicable" and nil="true" attributes and the place keyword 'non-geographic' should be included in the descriptiveKeywords element.

2.3 USGIN specification extensions

Extensions to ISO19115, ISO119, ISO19139 introduced by USGIN profile:

Allow use of `identificationInfo/SV_ServiceIdentification/coupledResource/-SV_CoupledResource/ScopedName` defined by OGC 07-045 ISO profile for CSW 2.0.2, use to provide WMS layer names or WFS feature names for service requests.

2.4 General Objectives

The Profile defines:

- mandatory and conditional metadata sections, metadata entities, and metadata elements
- the minimum set of metadata elements for any resource in order to conform to the Profile
- the core metadata for geographic datasets
- optional metadata elements that allow for a more extensive standard description of resources
- some recommended practices to increase the utility and interoperability of metadata.

2.5 Requirements

M (mandatory). Metadata element must have a valid value.

C (conditional). Metadata element is mandatory based on values of other metadata elements in the metadata record.

O (optional). Metadata element may be null in a valid document.

X (not used). Metadata element is not used by a Profile. The element may be included where it is schema valid, but may be ignored by applications implementing the profile.

2.6 Use cases to be supported

This section includes a number of user scenarios for how we intend USGIN metadata to be used, and discussion of several basic approach requirements that guide metadata content recommendations. At its heart, the problem is to find resources of interest via the internet, based on criteria of topic, place, or time, evaluate resources for an intended purpose, and learn how to access those resources. Detailed metadata describing a resource data schema, describing service or application operation, or providing detailed descriptions of analytical techniques and parameter are outside the scope intended for USGIN metadata. Our contention is that this more domain/resource specific type information is better accounted for with linked documents utilizing schema appropriate to those specific resource. Some examples include OGC getCapabilities, WSDL, and ISO 19110 feature catalogs. For more in depth discussion of use cases, scenarios, and requirements, see Metadata Recommendations for Geoscience Resources (Richard and Grunberg, 2010).

- A user specifies a geographic bounding box or one or more text keywords to constrain the resources of interest, and searches a metadata catalog using these criteria. The user is presented with a web page containing a list of resources that meet the criteria, with links for each resource that provide additional detailed metadata, and direct access to the resource if an online version is accessible, e.g. as a web page, Adobe Acrobat document, or online application (see Accessing Resources, below).
- A client application provides user with a map window that contains some simple base map information (political boundaries, major roads and rivers). User wishes to assemble a variety of other data layers for a particular area for some analysis or data exploration, e.g. slope steepness, geologic units, bedding orientation, and vegetation type for a hazard assessment. User centers map view on area of interest, then using an 'add data' tab, accesses a catalog application that allows them to search for web services that provide the desired datasets. After obtaining the results and reviewing the metadata for the located services, user selects one or more to add to the table of contents for the client application. Response from catalog has sufficient information to enable the client application to load and use the resource (e.g. serviceType, OnlineResourceLinkage). More concrete instances of this case would be finding Web Map services to add as layers in an ESRI ArcMap project, borehole Web Feature Services to post borehole logs in a 3-D mapping application, or water chemistry data Web Feature Service to bring data into a spreadsheet or database.
- A catalog operator wishes to import and cache catalog records from a collaborating catalog that have been inserted or updated during the last month (harvest). This operation requires knowledge of the metadata standard and version used for the returned records.
- A user discovers an error in a metadata record for a resource that they have authored, and wishes to contact the metadata producer to request correction.
- A search returns several results that appear to contain the desired content, and user must select the most likely to meet their needs. Metadata should provide sufficient information to guide this decision.
- A project geologist at Company X is searching for data relevant to a new exploration target, and wishes to restrict the search to resources that are publicly available.

2.7 Resources of interest

Table 1 summarizes the geoscience information resources of interest to the community that can be registered and discovered using this metadata profile. Note that this collection of resource types includes several kinds of resources that are not typically associated with ISO19115/ISO19119, which were created specifically for geospatial resources. *Table 1. Summary of resource types described by metadata for US Geoscience Information Network catalogs. Resource type **names in bold** have been prioritized for implementation in version one catalogs. The Resource type names include the type hierarchy encoded with the broader (parent) resource type indicated in the Broader Resource Type column.*

esource ype erarchy	**Broader Resource Type**	**Source**	**Definition**	**string value for hierarchyLe elName**
ollection		DCMI resource Types http://dublinco	An aggregation of resources. A collection is described as a group; its parts may also be separately described. (from http://www.ukoln.ac.uk/metadata/-dcmi/collection-application-profile/): The term "collection" can be applied to any aggregation of physical or digital	collection

		re.org/documents/dcmi-type-vocabulary/	items. Those items may be of any type, so examples might include aggregations of natural objects, created objects, "born-digital" items, digital surrogates of physical items, and the catalogs of such collections (as aggregations of metadata records). The criteria for aggregation is arbitrary. Collections may contain any number of items and may have varying levels of permanence. A "collection-level description" provides a description of the collection as a unit: the resource described by a collection-level description is the collection, rather than the individual items within that collection. Collection-level descriptions are referred to in Heaney [2000] as "unitary finding-aids".	
ataset	Collection	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A collection of data items in which individual data items are identified and accessible. DCMI definition is "Data encoded in a defined structure." with additional comment "Examples include lists, tables, and databases. A dataset may be useful for direct machine processing." The container may be a stand-alone digital file (mdb, spreadsheet, table in a Word document), a web service, or an enterprise database; these are called different 'distributions'. Synonym: structured data collection. This resource type represents the intellectual artifact (FRBR 'work') -- the information content and organization; the dataset may have more than one manifestation (format) -- as a list, a table, databases, using different software implementations.	collection:dataset
atalog	Dataset	USGIN	A collection of data items that index resources, as in metadata records; a metadata registry. The resource represents the information content and organization. Catalogs are accessed using other resources, like an interactiveResource or Service, and may have different formats. This Resource type should be used to categorize web applications and services that are interfaces to metadata catalogs.	collection:dataset:catalog
hysical tifact ollection	Collection	USGIN	A collection of identifiable physical objects, unified based on some criteria. Criteria for defining a collection may be who collected, where curated, why collected, kind of material…	collection:physical artifact collection
ocument		USGIN	A packaged body of intellectual work; has an author, title, some status with respect to Review/authority/quality. USGS peer reviewed would be a 'status property'. A document may have a variety of physical manifestations (pdf file, hardbound book, tiff scan, Word processor document...), and versions may exist as the document is traced through some publication process. May be map, vector graphics, text. Sound, moving images are included as document types.	document
nage	Document	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A visual representation other than text. Examples include images and photographs of physical objects, paintings, prints, drawings, other images and graphics, animations and moving pictures, film, diagrams, maps, musical notation. Note that Image may include both electronic and physical representations.	document:image
tillImage	Image	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A static visual representation. Examples include paintings, drawings, graphic designs, plans and maps. Recommended best practice is to assign the type Text to images of textual materials if the intent of the image is to capture the textual content as opposed to the appearance of the medium containing the text. Instances of the type Still Image must also be describable as instances of the broader type Image. Subtype of Image.	document:image:stillimage
uman-generated image	StillImage	USGIN	Image produced by human drawing or painting, using any media. May be entirely product of human imagination, human perception of the world, or a human-modified photographic image.	document:image:stillimage:human-generate

				image
hotograph	StillImage	USGIN	Image produced by optical device with chemical or electronic image capture; represents things in the field of view directly as captured by the device. Photographs may be modified by human processing; there is a continuum between photographs and human-generated image. Distinction between the two is largely based on intention	document:ima e:stillimage:hu man-generate image:map
emote ensing arth image	StillImage	USGIN	Image of Earth surface acquired by an air born or earth-orbiting sensor. May be georeferenced such that location in the image directly corresponds to location on the earth.	document:ima e:stillimage:ph tograph
ap	Human-gen erated image	USGIN	Human-generated depiction of some part of the earth using a mathematical system of correspondence between geometry in the image and location on the earth.	document:ima e:stillimage:re mote sensing earth image
ovingImage	Image	DCMI resource Types http://dublinco re.org/docume nts/dcmi-type-vocabulary/	A series of visual representations imparting an impression of motion when shown in succession. Examples include animations, movies, television programs, videos, zoetropes, or visual output from a simulation. Instances of the type Moving Image must also be describable as instances of the broader type Image. Subtype of Image. Commonly include sound	document:ima e:moving imag
ound	Document	DCMI resource Types http://dublinco re.org/docume nts/dcmi-type-vocabulary/	A resource primarily intended to be heard. Various media and encoding schemes may be used, e.g. vinyl disks, digital files, magnetic tape recordings.	document:sou d
ext	Document	DCMI resource Types http://dublinco re.org/docume nts/dcmi-type-vocabulary/	A resource consisting primarily of words for reading. Examples include books, letters, dissertations, poems, newspapers, articles, archives of mailing lists. Note that facsimiles or images of texts are still of the genre Text.	document:tex
ypertext ocument	Text	USGIN	A body of text that contains navigable links allowing non-linear reading paths through the work. Links to documents or other resources outside of the document are possible; and the document itself may be packaged in discrete parts (e.g. pages, files). The criteria for determining document membership are somewhat arbitrary, but in general the document (analogous to a single web site) should contain related parts authored and managed by the same agent.	document:tex hypertext document
vent		DCMI resource Types http://dublinco re.org/docume nts/dcmi-type-vocabulary/	A non-persistent, time-based occurrence (perdurant). Metadata for an event provides descriptive information that is the basis for discovery of the purpose, location, duration, and responsible agents associated with an event. Examples include an exhibition, webcast, conference, workshop, open day, performance, battle, trial, wedding, tea party, and conflagration.	event
roject	Event	USGIN	A termporally extended event characterized by activity that has some stated purpose; projects typically have associated spatial extents, which represent the area of interest for the project. This extent serves as a mechanism to filter descriptions and concepts in the information system for those that may be related to the project based on spatial relationships. Projects in a large	event:project

			organization will likely have hierarchical (part-whole) relationships.	
odel		USGIN	Algorithm, workflow; an abstract representation of a collection of related processes, objects and relationships. A model resource may be related to various kinds of document that portray the model, or to software that implements the model, or with datasets as input or output.	model
hysical rtifact		DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	General category for physical resources that are indexed by metadata records; also root of an artifact type hierarchy. An identifiable physical object. Identification is always a function of some human intention, thus differentiating an artifact from other 'natural' things. Note that digital representations of these objects that are accessible on the WWW will be Images, Text or one of the other types.	physical artifa
ervice		DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A resource that provides one or more functions via a network interface designed for machine interaction. An implementation of an interface to some sort of digital resource, using either a 'pull' model in which client requests some content from the service, and receives that content in a single 'response' package, or a 'push' model in which client establishes connection and monitors for change events (update, new data…) from service. Difficult to draw line on when a service provides 'files' and when it provides 'data', because responses are always in a form that could be considered a file. Also includes interfaces to digital resources that provide a continuous (with some sampling interval and unit of packaging for distribution) feed of some sort of data.	service
oftware		USGIN	A computer program in source or compiled form. Examples include a C source file, MS-Windows .exe executable, or Perl script. Identity is associated with function, authorship, and programming environment.	software
tand-Alone-pplication	Software	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	Identifiable stand alone software application. Identity of resource is based on function performed, input and output requirements, and authorship. The same application may be packaged in different file formats to run in different software environments; thus an application might have one or more associated digital files. For this catalog scheme, stand alone applications are software that can be packaged in a single file and transferred between machines, unpackaged and compiled or installed on a computer meeting specified hardware and software environment conditions, to execute the described function on that computer, independent of any network connection.	software:stand alone-applicat on
teractive-esource	Software	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A resource requiring interaction from the user to be understood, executed, or experienced. Comment: Examples include forms on Web pages, applets, multimedia learning objects, chat services, or virtual reality environments. Interactive resources are software driven. From the point of view of the catalog, they are accessed by a URL to a web site that is the interface for operating the application. The application operates by interaction with one or more human participants. The application requires network connection to operate, is accessible via the internet, and requires human interaction.	software:inter ctive resource
tructured gital data em		USGIN	An individually identifiable item in a structured digital data collection. Characterized by a schema, and some particular values. In ISO11179 terms, this is an instance of a data element. Tagging, commenting, reviewing, rating community interaction with catalog will probably require metadata records about particular data items in cataloged datasets (including metadata items in catalogs.)	structured digital data item
ampling oint, site, ation	Structured digital data item	From ScienceBase item types, SMR redux	A resource that is a location-based container/base for observation data. These represent OGC O&M sampling features, and can be generalized to include other sampling geometry (borehole, image footprint)... Analogous in function to a keyword, but carries metadata on who located, when, why, how...	sampling poin

3 USGIN Usage of Metadata Elements

3.1 Core spatial dataset, dataset series, and service elements

Table 2 is a listing of ISO19115 metadata elements used to describe any resource. Tables 3 and 4 provide specifics for describing datasets and services. Note that in the USGIN context, dataset is construed quite broadly to include any kind of georeferenced information resource, including physical samples and hard copy documents. The service metadata elements are defined by ISO19119. The root element of ISO XML-encoded metadata is MD_Metadata. Elements are discussed in this table in the order that they appear in the metadata document. Not all elements are discussed in detail. In a number of places where USGIN makes no specific provisions, we defer to recommendations in the North American Profile for ISO metadata (INCITS 453, referred to as NAP). Note that throughout this and the subsequent tables, the names of XML elements are shown in `this typecase`. Long X-paths have been broken with non-breaking hyphen characters. Hyphens are not used in any XML attribute or element name, so if they appear in the text, they are strictly for text wrapping.

Table 2. Description best practices for ISO19139 metadata elements in USGIN profile. This table includes base elements.Elements are in the order that they appear in a metadata instance.

ISO 19115 (M/C/O) xPath from MD_Metadata	**NAP-USGIN M/C/O**	**Comments**
Metadata file identifier (O) `fileIdentifier`	M-M	Identifier for the metadata record, as opposed to `DatasetURI`, which resource. A unique metadata record identifier must be included to such as GetRecordById or harvest transactions. This identifier shou harvest operations. Ideally there is one metadata record describing that there should be a one-to-one mapping between metadata fileI DatasetURIs. However, not all described resources will have a Dat metadata record is a different resource from the resource it describ have the same identifier. The protocol used to generate the identifi long as it generates globally unique identifier strings. Services that (e.g. serviceURL and layerID) are expected to put the key values in there is technically no limit on the length of the identifier string, sug to keep the string length less than 255 so the string will fit in legacy fields. USGIN, ANZLIC, and the OGC CSW profiles for ISO metadata (OG the use of the UUID (Universally Unique Identifier) for the `fileIdent` `fileIdentifier` is used to identify duplicate copies of metadata reco metadata record from another (via `MD_DataIdentification/aggregatio` metadata from a described resource (e.g. `DS_Dataset/has/MD_Metadat` difference between the two metadata records then one can determ version by the content of other elements in the metadata record. Th metadata record should be the only one made publicly available in systems such as a catalog service.
Metadata language (M) `language`	M-M	The language string is composed of a three lower-case letter langu Most CSW client and server applications only support the three lett (Note: The recommended USGIN practice is different from the NAP including an alpha3 country code (ISO3166-1), given in uppercase,

		"<ISO639-2/T three letter language code><;><blank space><ISO3 country code>", e.g. fra; CAN. Language searches for the three le wildcard character will work with either convention.)
Metadata character set (C) `characterSet`	M-M	USGIN recommends use of ISO codelists: codeListValue="utf8" co **Hyperlink reference not valid.**. See 4.17.3 Codelists for discussio USGIN requires that a character set code is specified, but this will r value for all records in a catalog.
Parent metadata record (O) `parentIdentifier`	O-X	Not used in USGIN profile. USGIN CSW service implementations c be able to navigate parent links to obtain inherited metadata proper using parent links, so this element is not used. To represent relatio described resources use `MD_Identification/aggregationInfo`.
Resource type (C) `hierarchyLevel`	M-M	Cardinality is 1...*. NAP and ISO codelists are equivalent. See 4.1 discussion of encoding of codelist values. Due to interoperability pr mandates use of ISO codelists. At least one MD_ScopeCode codelist value is required; default is 'd catalog metadata, the codelist is {`collectionHardware, collect series, nonGeographicDataset, dimensionGroup, fieldSes service, model, tile`}; other values from the ISO 19115 Scope considered out of scope for USGIN catalog metadata and the meta ignored by harvesting catalogs. The European INSPIRE Implement (MD_IR_and_ISO_20090218) proscribes the code list for the first h element in an MD_Metadata document to be one of {`dataset, se` metadata set will be considered out of scope for the directive (see s Type); in order to maintain compatibility with INSPIRE catalogs, it is scope code from this subset. ISO Example – dataset metadata: <gmd:hierarchyLevel> <gmd:MD_ScopeCode codeList="http://standards.iso.org/ittf/PubliclyAvailableStanda ISO_19139_Schemas/resources/Codelist/gmxCodelists.xml#M codeListValue="dataset">dataset</gmd:MD_ScopeCode> </gmd:hierarchyLevel>
Resource hierarchy level name (C) `hierarchyLevelName`	O-M	USGIN makes this property mandatory to identify the USGIN resou (above). Default USGIN `hierarchyLevelName.CharacterString` is "Data by including `hierarchyLevelName` elements for all broader resource ca should also include a `hierarchyLevelName`="Collection" element. For `hierarchyLevelName.CharacterString` is "Service". As use cases deve rationale for definition of sub-categories of service, the resource ca expanded. (ISO 19115 assumes that the metadata hierarchy level "dataset" if it is not documented. NAP does not use it.) Example – dataset metadata: <gmd:hierarchyLevelName> <gco:CharacterString>Dataset</gco:CharacterString> </gmd:hierarchyLevelName> <gmd:hierarchyLevelName> <gco:CharacterString>Collection</gco:CharacterString> </gmd:hierarchyLevelName>
Metadata point of contact	M-M	Cardinality on contact is 1..*. USGIN requires at least one `CI_Resp`

(M) contact/CI_ResponsibleParty		role.CI_RoleCode@codeListValue = "originator" (CI_RoleCode elemen that identifies the original source of the metadata record. If the poin report errors, updates to metadata, etc. is different than the originat contact/CI_ResponsibleParty element may be included with role.CI_ "pointOfContact" (CI_RoleCode element value="pointOfContact" the metadata records wishes to identify itself in result records, this included in an additional MD_Metadata/contact/CI_ResponsibleParty el role.CI_RoleCode@codeListValue = "distributor". See 4.17.3 *Code* encoding of codelist values. ISO Role codes applicable in this cont {distributor, originator, pointOfContact}. The CI_ResponsibleParty element must include a contact e-mail add (electronicMailAddress) or telephone number phone/CI_Telephone/voi (individualName + organisationName + positionName) > 0. If the contactInfo/CI_Contact/onlineResource/CI_OnlineResource element fo ponsibleParty with role.CI_RoleCode@codeListValue = "originator" h name = "icon", the CI_OnlineResource/linkage/URL will be assumed to file (e.g. tif, png, jpg) for the metadata originator. This Icon will be d results to credit the metadata originator. Metadata harvesters must metadata originator information so that the origin of metadata recor should harvest the point of contact information if it is different. Othe applying to the metadata record (not the described resource) may b catalogs.
Metadata date stamp (M) dateStamp	M-M	USGIN profile requires use of dateStamp/gco:DateTime (Note this con mandate to use dateStamp/gco:Date). This is the date and time when was created or updated (following NAP). The dateStamp is assume reflect any change in the metadata record that the metadata publish through the USGIN catalog system. This is the time stamp that will to determine if a metadata needs to be updated in a harvesting cat
Metadata standard name (O) metadataStandardName	M-M	USGIN profile conformant metadata is indicated by using "ISO-USG "ISO-NAP-USGIN" has been used in the past and will also be recog should be discontinued. Use is mandatory to indicate that the metadata record conforms to
Metadata standard version (O) metadataStandardVersion	O-M	For this version of the USGIN profile, Must use the string "1.2". Use is mandatory to specify the version of the profile used
DataSet Identifier (O) dataSetURI	O-O	This is a string that uniquely identifies the described resource. If the resou should be included here; if the resource will be referenced from other met identifier here. Any kind of resource (not only datasets) may have an iden identifier is not specified, but some sort of documented scheme to assure u (UUID, URN…). USGIN is attempting to make the semantics of iden provision (see Unique resource identifier in Table 3, below) that the MD_DataIdentification/citation/CI_Citation/identifier/MD_Identifie resource. This may be identical with the resource described by the its value is identical to MD_Metadata/dataSetURI, or it may be a publica intellectual source of the described resource, in which case it is a d The MD_Distribution/transferOptions/MD_DigitalTransferOptions/onli used to specify URLs for access to the resource, even if the datasetURI is dereference to obtain a representation of the described resource. The dataS considered an opaque identifier. This will avoid ambiguity about where to

		access to a described resource. If the dataset is coupled to a service, the value of the MD_Metadata/d unique resource identifier used by srv:coupledResource to link the The *OpenGIS® Catalogue Services Specification 2.0.2 - ISO Meta* (OGC 07-045) Annex F recommends that MD_DataIdentification/ci identifier/MD_Identifier/code match the identifiers specified by SV_ServiceIdentification/operatesOn and SV_ServiceIdentification/c linking a described service to datasets that the service operates on fileIdentifier (above), this requires that a MD_DataIdentification/c CI_Citation/identifier that identifies the described resource is inclu record; its value must be the same as MD_Metadata/dataSetURI.
Other languages (C) locale	C-X	Other languages used in metadata free text description. If description in m provided, this property should indicate what those languages are. The prin metadata description is identified with MD_Metadata/language and charact languages are identified by MD_Metadata/locale/PT_locale elements, in provided according to ISO 639-2/T three-letter terminology codes in lowe country is provided according to ISO 3166-1 three-letter codes in upperca characterEncoding. See 4.17.3 *Codelists* for discussion of encoding USGIN Catalogs must harvest and process metadata content in the specified by MD_Metadata/language and characterSet, and may ig localized content.
[role] Resource spatial representation (O) spatialRepresentationInfo	O-O	Best practice is to include metadata for spatial representation if the georeferenced dataset. Metadata for Spatial data representation are der Metadata is instantiated as one or more of MD_GridSpatialRepresentation MD_VectorSpatialRepresentation, MD_Georectified, or MD_Georeferencea profile follows NAP for spatial representation metadata. Vector Spatial Re provided if point or vector objects exist in the dataset. If MD_VectorSpatia either spatialRepresentationInfo/MD_VectorSpatialRepresentation/to spatialRepresentationInfo/MD_VectorSpatialRepresentation/geometric or both." (NAP) MD_GridSpatialRepresentation or one of its subtypes (MD MD_Georeferenceable) should be provided if dataset objects are gridded. M used if the grid (image) is georeferenced, and MD_Georeferenceable is use georeferenced. Follow NAP optionality if these elements are used.
Resource spatial representation vector topology (O) spatialRepresentationInfo/-MD_VectorSpatialRepresentation/topologyLevel	C-C	Code that specifies the degree of complexity of spatial relationships betwe Value is from ISO codelist MD_TopologyLevelCode. (Code names in this lis {geometryOnly, topology1D, planarGraph, fullPlanarGrapl fullSurfaceGraph, topology3D, fullTopology3D, abstract for discussion of encoding of codelist values. It is unclear precisely w terms of the topology encoding. To be useful, assertion that topology is pr topological relationships that may be implicit in the encoded vector geom represented (e.g. by correlation tables—left poly, right poly for a polyline clarification should be included in the resource abstract. Mandatory if MD_VectorSpatialRepresentation is present.
Resource spatial representation vector geometric objects (O) spatialRepresentationInfo/-MD_VectorSpatialRepresentation/geometricObjects	C-C	"Identification of the objects used to represent features in the dataset." Pro count for the number of objects of each type. Use the ISO MD_GeometricOb Code names in this list are: {complex, composite, curve, point, 4.17.3 *Codelists* for discussion of encoding of codelist values. Mand MD_VectorSpatialRepresentation is present.

[role] Resource's spatial reference system (O) `referenceSystemInfo`	O-O?	Description of the spatial and/or temporal reference systems used in the da (`identificationInfo/spatialRepresentationType/MD_SpatialRepresenta` or (`../MD_SpatialRepresentationTypeCode` = "`grid`") or (`../MD_SpatialR "`tin`") implies count `referenceSystemInfo` >= 1) }. See 4.17.3 Codelis encoding of codelist values. Use ISO codelist.
Reference System identifier code (O) `referenceSystemInfo/-MD_Refe renceSystem/-referenceSystem Identifier/-RS_Identifier/co de`	C-C	If `referenceSystemInfo` is included, then the `RS_Identifier` element m code value. For USGIN the code should be a value from the EPSG Dataset register (http://www.epsg-registry.org/) in the form "EPSG:nnnn" code number for the CRS. If the CRS is not defined in the EPSG registry, specified in the NAP profile should be followed, e.g. the CRS shall be des 19111 and ISO/TS 19127, assigned an identifier, and registered with an au referenced here. The `RS_Identifier/codespace` in this case should identify where the CRS definition is registered, such that the definition can be loca USGIN purposes is to provide georeferenced data using one of the EPSG reference systems if this is possible.
Metadata extension information (O) `metadataExtensionInfo`	X-X	Not used in this profile.
Resource identification information (M) `identificationInfo`	M-M	Cardinality 1..*. The content of this element identifies the described resources that are not services, use `MD_DataIdentification` (**see Tal `SV_ServiceIdentification` is required (**see Table 4)**.
[role] Content information (O) `contentInfo`	O-O	Characteristics describing the feature catalog, coverage, or image `Information` is an abstract class. One or more of `MD_FeatureCatalogu` `MD_CoverageDescription` or `MD_ImageDescription` elements may be use content. `MD_FeatureCatalogueDescription` describes content in a featu like an ESRI geodatabase that may have more than one feature, e. polygons, fault line features, and point observation locations for stri Entity-attribute metadata for dataset should be documented using I are located at http://www.isotc211.org/2005/gfc/gfc.xsd. The ISO1 implementation of ISO19115 does not allow inclusion of the feature metadata record, but provides a citation element in the `gmd:content` `CatalogueDescription` element. The USGIN profile provision is to link using an xlink:href attribute on the `gmd:featureCatalogueCitation`, ide accompanying xlink:title attribute with the value "Link to ISO19110 Example 2 shows recommended encoding. An example feature cat ISO19110 XML is included in section 8.4, ISO19110 Feature Catal `MD_CoverageDescription` is for datasets that are one of the types liste MD_CoverageContentTypeCode: image, thematicClassification, ph coverage is a data structure that acts as a function to return values direct position within its spatiotemporal domain (OGC 07-067r5). In values for light intensity in a given wavelength range, thematicClas return codes corresponding to some domain concept, and physical coverages return values representing some physical quantity like m density, resistivity.
[role] Resource distribution information (O) `distributionInfo`	O-M	This element provides information to inform users how to obtain or resource, which is considered part of the minimum useful content th by a metadata record. USGIN profile specifies that at least one `MD_` and one `MD_Distribution/transferOptions` element are required, and

		provides the specified transfer options. See section 4.13 '*Use of MD_Distribution and MD_Distributor*' for instructions for more complicated combinations transfer options, and ordering instructions.
Resource distribution format (O) `distributionInfo/-MD_Distribution/distributionFormat`	O-O	Information on the format or physical manifestation of the resource • If the resource is a physical (not electronic) resource, like a boo document, the `distributionFormat/MD_Format/name` is mandatory, Table 6. USGIN Distribution formats for non digital resources. U http://resources.usgin.org/registry/distributionFormatNames201 • For digital resources, the format specifies the file type, either us formatted string. Pattern for digital resources that do not have a `[vendor:applicationName]/fileExtension`. The vendor an may not be applicable, and could be omitted, but the '/' and file always be present. • If the format consists of a single file, the file extension is a three abbreviation assigned by the vendor.
Resource distributor information (O) `distributionInfo/-MD_Distribution/-distributor/-MD_Distributor/`	O-C	*USGIN differs from NAP* in this case (but not with ISO19115) by all distributors, and binding between distributors, transfer options, and profile, each `distributor/MD_Distributor` is a binding between one o and the distributor formats that are available through that/those tra (`MD_DigitalTransferOptions/onLine/CI_OnlineResource` in particular). I available from the same distributor, or have different transfer option represented as different `distributor/MD_Distributor` instances. See *MD_Distribution and MD_Distributor*' for instructions on use of thes
Resource distributor responsible party (O) `distributionInfo/-MD_Distribution/-distributor/MD_Distributor/-distributorContact/-CI_ResponsibleParty`	C-M	`MD_Distributor` is required, which requires one `CI_ResponsibleParty`. role.CI_RoleCode@codeListValue = "`pointOfContact` " (CI_RoleC "`pointOfContact` ") is mandatory, to provide a contact point to rep accessing the resource through distributions associated with the di CI_ResponsibleParty element must include a contact e-mail addres (electronicMailAddress) or telephone number (electronicMailAddres individualName, organisationName, or positionName must be pres distributorContacts with different roles may be ignored by USGIN c applicable in this context include: {`resourceProvider, distribu` See section 4.17.3 '*Codelists*' for details on codelist encoding.
Resource distributor order process (O) `distributionInfo/-MD_Distribution/distributor/-MD_Distributor/-distributionOrderProcess/-MD_StandardOrderProcess`	O-C	Information on the availability of the resource including at least one and time, ordering instructions, or turnaround. For physical resourc are mandatory as these will typically indicate the method of access
Resource distributor format (O) `distributionInfo/-MD_Distribution/distributor/-MD_Distributor/-distributorFormat/MD_Format`	O-O	See section 4.14 'Distribution Format ' for instructions on use of the
Resource distribution transfer options (O)	C-C	`MD_DigitalTransferOptions` provides information on digital distributio section 4.13 '*Use of MD_Distribution and MD_Distributor*' for instru

distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransferOptions		element. If distribution information is included (MD_Distribution is n one one MD_Distribution/transferOptions element is required.
Resource distributor online distribution linkage (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransferOptions/-online/-CI_OnlineResource/linkage	M-M	Digital transfer options are "Technical means and media by which a from the distributor." The CI_OnlineResource/linkage/URL element mu URL to access the resource directly (see sections 4.13 to 4.15 for r
Resource distributor online distribution linkage (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransfer-Options/-online/-CI_OnlineResource/protocol	M-O	The CI_OnlineResource/protocol element is mandatory if access to tl additional protocol information that is stacked on the protocol speci CI_OnlineResource/linkage/URL prefix (e.g. http:, ftp:). Typically this w WMS-1.3.0, OpenSearch-1.1, or DAP-2.0. Ideally these strings sho dereferenced to learn something about the service specification.
Resource distributor online distribution linkage (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransfer-Options/online/-CI_OnlineResource/name	O-O	The CI_OnlineResource/name element may duplicate the file name if t but it is recommended to provide a user-friendly label for the file tha a user interface. For links to resource distribution end points (distributionInfo//CI_OnlineResource) that are not simple file access online order forms, mailto links to e-mail a request for the resource, special values defined in Table 13.
Resource distributor online distribution application profile (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransferOptions/-online/CI_OnlineResource/-applicationProfile	C-O	applicationProfile is mandatory if the CI_OnlineResource/linkage do HTML web page (or other standard format that will be recognized b another software application is needed to use a linked file resource is a service instance conforming to a profile. The applicationProfile specify the software using the following recommended syntax: "ver name/application version", e.g. "Microsoft:Word/2007", or "ESRI:Ar documents for which the service type and base protocol prefix on tl provide sufficient information to guide client software, the applicatio used to indicate a profile on the serviceType or some variation in de content conventions. See section 4.15 CI_OnlineResource for more
Resource distributor online distribution function (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransferOptions/-online/CI_OnlineResource/-function	O-O	CI_OnlineResource/function is mandatory if the linkage is not a simp downloadable file. USGIN recommended values for CI_OnlineFunc summarized in Table 7. If the resource is accessible as a web serv service should be separate metadata record with the dataset(s) exp service identified in the service metadata record as coupledResour
[role] Data quality information (O) dataQualityInfo	C-O	Either dataQualityInfo/DQ_DataQuality/report or dataQualityInfo/DQ_ mandatory if a dataQualityInfo element is present. dataQualityInfo/ required, with value from MD_ScopeCode: {collectionHardware dataset, series, nonGeographicDataset, dimensionGroup, software, service, model, tile}. See 4.17.3 *Codelists* for dis codelist values. See section 4.19 Data Quality for discussion of dat parts.

Data quality scope (O) dataQualityInfo/DQ_DataQuality/scope	C-C	Mandatory if DQ_DataQuality is not null. Specifies the extent of cha data quality information is reported. Value is from MD_ScopeCode: {collectionHardware, collectionSession, dataset, serie nonGeographicDataset, dimensionGroup, fieldSession, so model, tile}. See 4.17.3 *Codelists* for discussion of encoding of
Data quality scope level description (O) dataQualityInfo/-DQ_DataQuality/-scope/levelDescription	C-X	Not used by USGIN. See section 4.19 *Data Quality* for more dis levelDescription.
Data quality report (O) dataQualityInfo/-DQ_DataQuality/report	C-C	In this profile, data quality information is optional, but if included, th data quality information is a free text explanation of data quality co DQ_DataQuality[1]//report[1]//explanation/CharacterString element text discussion/description of data quality considerations for the ind of any specific data quality element to contain this explanation is ar be considered meaningful in this context. Other data quality information should be preserved by USGIN catal metadata that includes these elements, and presented when the us metadata record. Each DQ_DataQuality/report element requires that possible data quality elements must be present, and multiple repor within each DQ_DataQuality element. Each of these *AbstractDQ_eleme* nameOfMeasure, measureIdentification, measureDescription, evaluation evaluationMethodDescription, evaluationProcedure, and dateTime elem required result elements. The *AbstractDQ_element* /result is either a a DQ_QuantitativeResult, each of which has required and optional su of this report metadata should follow recommendations in NAP.
Data quality lineage (O) dataQualityInfo/-DQ_DataQuality/lineage	C-O	Not applicable to services. USGIN recommended practice is descri
Data quality lineage statement (O) dataQualityInfo/-DQ_DataQuality/lineage/-LI_Lineage/statement	C-C	If lineage information is included, a lineage/LI_Lineage/statement is specifies a "General explanation of the data producer's knowledg (NAP). USGIN recommended practice is described in section 4.19.
Data quality lineage source (O) dataQualityInfo/-DQ_DataQuality/lineage/-LI_Lineage/source	C-O	Each source/LI_Source element describes a source data resource th processStep. NAP provision is that LI_Source/description is manda LI_Source/sourceCitation and LI_Source/sourceExtent are not provi LI_Source/description includes the source medium name from the C MD_MediumNameCode, followed by <;><blank space> and a free "dvd; source satellite image." If the source is part of a processing chain, the LI_Source/processSte provides "Information about an event related to the creation proces (INCITS 453). This is interpreted to mean that the link from a sourc a process step for which the described source is an output. USGIN is described in section 4.19.
Data quality lineage process	C-O	An event in the development of the dataset. Each step requires a fr may have a free text rationale, dateTime stamp when process was

step (O) dataQualityInfo/-DQ_DataQuality/lineage/-LI_Lineage/processStep		CI_ResponsibleParty elements identifying parties involved in the proc many source/LI_Source associations to identify data that is input into practice recommended for USGIN is that source association from a inputs to a process, and processStep associations from a source ele resource to a process step that produced it. See USGIN recommen described in section 4.19.
[role] Portrayal catalog information (O) portrayalCatalogueInfo	O-O	portrayalCatalogueInfo/MD_PortrayalCatalogReference/portrayalCatal n element identifying a catalogue that contains symbols and rules t portrayal catalog is a collection of defined symbols used to depict, t a map. No documentation in ISO19115 about how this is supposed defines the structure of a Portrayal Catalogue. No USGIN recomme been adopted.
[role] Metadata constraint information (O) metadataConstraints	O-O	This element specifies use constraints for access to the metadata r for accessing the describe resource are in resourceConstraint/MD_Co MD_DatasetIdentification or MD_ServiceIdentification. Follow NAP f access constraints. NAP provision is that metadataConstraints/MD_Constraints/useLimitat MD_Constraints is used to specify metadataConstraints. When one of MD_LegalConstraints or MD_SecurityConstraints is used, useLimitation MD_LegalConstraints are specified by MD_RestrictionCode. ISO cod {copyright, patent, patentPending, trademark, license, intellectualPropertyRights, restricted, otherRestricti *Codelists* for discussion of encoding of codelist values. otherCons element required if accessConstraints or useConstraints is set to "oth example: "Data only to be used for the purposes for which they wer MD_SecurtyConstraints has various optional free text values, and a re MD_SecurityConstraints/classification from ISO MD_ClassificationC restricted, confidential, secret, topSecret}. See 4.17.3 of encoding of codelist values.
[role] Application schema information (O) applicationSchemaInfo	O-O	Information about the information schema of the resource applicationSchemaInfo/MD_ApplicationSchemaInformation element has name/CI_Citation, schemaLanguage free text, and constraintLanguage fr MD_ApplicationSchemaInformation element also allows inclusion of an document as ASCII, or a binary graphicsFile or softwareDevelopment applicationSchemaInfo elements may be used for different presenta schema, or for different kinds of schema (e.g. physical, logical, con
[role] Metadata maintenance information (O) metadataMaintenance	O-O	This element provides information about the maintenance schedule metadata record. Only one MD_MaintenanceInformation element may required MD_MaintenanceFrequencyCode. The ISO codelist is {co weekly, fortnightly, monthly, quarterly, biannually, a irregular, notPlanned, unknown}. See 4.17.3 *Codelists* for dis codelist values. Additional metadataMaintenance/MD_MaintenanceInformation/mai added to metadata records to document updates to the the content catalogs should use metadata maintenance notes to document any harvested metadata records; ideally a link to the original metadata provided.
[role] Series information (O) series	X-X	The MD_Metadata/series element that appears in the ISO19139 sche implement the metadata application model in ISO19115:2003 Figu class diagram defining the classes of geographic information to wh

		Not Used by USGIN.
[role] Described resource (O) `describes`	X-X	The `MD_Metadata/describes` element that appears in the ISO19139 s implement the metadata application model in ISO19115:2003 Figu class diagram defining the classes of geographic information to wh Not used by USGIN.
[role] Property type description (O) `propertyType`	X-X	The `MD_Metadata/propertyType` element that appears in the ISO1913 implement the metadata application model in ISO19115:2003 Figu class diagram defining the classes of geographic information to wh Not used by USGIN.
[role] Feature type description (O) `featureType`	X-X	Although an `MD_Metadata/featureType` element that appears in the IS appears to implement the metadata application model in ISO19115 a UML class diagram defining the classes of geographic informatio applies. Not used by USGIN.
[role] Feature attributes (O) `featureAttribute`	X-X	Although an `MD_Metadata/featureAttribute` element that appears in t appears to implement the metadata application model in ISO19115 a UML class diagram defining the classes of geographic informatio applies. Not used by USGIN.

3.2 Dataset Identification properties (MD_DataIdentification)

The difference between metadata for services, and metadata for other resources is in the `identificationInfo` part of the ISO19139 xml schema. This section documents use of `MD_DataIdentification` for metadata describing resources of interest in the geoscience information network that are not services. Service metadata utilizes the `SV_ServiceIdentification` element to provide a description and identification of a service (see 3.3 Service identification elements).

Note that the USGIN approach to metadata is predicated on the idea that the users first interest is finding information about some particular topic; how that information is accessed or acquired is a secondary consideration. Thus metadata records are about information resources (mostly datasets and other data products or publications) at the FRBR work level [FRBR, 2004]. Information on accessing data resources through service interfaces is considered a kind of distribution of the resource. The MD_DataIdentification part of a metadata record describes the information resource, and multiple `distributionInformation` elements specify access to that data through one or more services, or as various kinds of packaged, downloadable files. Given that a metadata record is about exactly one 'work', each metadata record has exactly one `MD_DataIdentification` element. Metadata records about services are targeted for the rather unusual situation in which a user is primarily interested in a specific kind of service and the details of its operation, typically this will be for web processing or brokering services. In practice, service protocols typically specify some kind of service specific self-description document (e.g. WSDL, WADL, OGC GetCapabilities, OpenSearch Description) that client software is more likely to understand than an ISO SV_ServiceIdentification element.

Table 3. Dataset Identification properties (MD_DataIdentification)

ISO 19115 (M/C/O) xPath from MD_Metadata	**NAP-USGIN M/C/O**	**Comments on MD_DataIdentification**
Resource citation (M) `identificationInfo/-MD_DataIdentification/-citation/CI_Citation`	M-M	CI_Citation cardinality exactly one required. The citation attribute pr identifies the intellectual origin of the content in the described resou a citation in a scientific journal. Required content for a `CI_Citation` el and `responsibleParty`. NOTE: if multiple identificationInfo sections a catalogs may ignore all but the first instance.
Resource title (M) `identificationInfo/-MD_DataIdentification/-citation/-CI_Citation/title`	M-M	A meaningful title that is unique within the scope of the catalog shou each resource that is described by a metadata record. USGIN recor that inform the human reader about the dataset's content as well as
Resource reference date (M) `identificationInfo/-MD_DataIdentification/-citation/CI_Citation/date/-CI_Date/date/`	M-M	Best practice is to include at least the date of publication or creation date of the resource reported in the citation corresponds to the most resource's content. `CI_Date` content includes a `date` and `dateType`. Da uses xs:dateTime data type, The dateTime data type is specified in "YYYY-MM-DDThh:mm:ss". `timezoneOffset` is optional (http://www.w3.org/TR/xmlschema11-2). Example date encoding: 2000-12-12T12:00:00+13:00, 2006-10-01T month or day is not known, encode as '01', for example '2006-01-01 the event used for the temporal aspect of the resource, and must be {creation, publication, revision}, which is a subset of the ISO DateTy date is distinct from the `dateStamp` for the metadata record (date and

		metadata update), or the EX_Extent/temporalElement that specifies the the resource content is applicable.
Unique resource identifier (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/-identifier/MD_Identifier	C-C	For USGIN, if the Citation has an identifier that is different from the described resource (MD_Metadata/dataSetURI), it must be included here value should be an identifier for the cited resource, without any assu http protocol. The identifier may be resolvable to a URL, if a protoco identifier scheme that is resolvable (e.g. http, doi…), but this is not n document, and should not be assumed when processing metadata doc The USGIN profile requires the use of MD_Identifier element to iden RS_Identifier must not be used. If additional codespace and version with the identifier, it should be encoded as MD_Identifier/authority/C alternateTitle and MD_Identifier/authority/CI_Citation/edition
Resource responsible party (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/-citedResponsibleParty	M-M	USGIN requires that the citation include at least one CI_ResponsibleP count of (individualName + organisationName + positionName) mus responsible party should identify the agent responsible for the intelle described resource. In this context, the CI_ResponsibleParty/role/CI_RoleCode@codeListValue is from {originat principalInvestigator, processor, author}, a subset of the ISO CI_Rol 4.17.3 *Codelists* for discussion of encoding of codelist values. For m content, the responsible party is what would normally be considered Best practice is to include point of contact information for the resour MD_DataIdentification/pointOfContact/CI_ResponsibleParty. Guidance would be helpful for consistency, but has not been developed as yet editor, contributor} View compiler and editor as equivalent roles, sin include compiler. Use the ISO19115-1 codelist.
Resource presentation form (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/-presentationForm	O-O	The form in which the original cited resource is available. Note that t original source of intellectual content in the described resource, and be different from the format for distribution described in the metadat be specified if there is a difference between the cited resource prese the distribution format(s) listed in the distributionInfo/MD_Distributi metadata record. This information is mostly for documentation; for access, the distribution formats are more relevant. presentationForm uses CodeList = CI_PresentationFormCode, with I {documentDigital, documentHardcopy, imageDigital, image mapDigital, mapHardcopy, modelDigital, modelHardcopy, p profileHardcopy, tableDigital, tableHardcopy, videoDigit audioDigital}. See section 4.17.3 *Codelists* for details on codelis
Resource series (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/series	O-O	Information about the (publication) series or collection of which the r NAP rule (name + issueIdentification) > 0 should be followed.
Resource other citation details (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/-otherCitationDetails	O-O	A complete recommended bibliographic citation should be specified
Resource collective title (O) identificationInfo/-MD_DataIde	O-C	Title of the combined resource that the cited resource is part of, for e resource may be a paper in an anthology, in which case the antholo

ntification/-citation/CI_Citation/-collectiveTitle		collective title. Required if the cited resource is part of such a collect
Resource abstract (M) identificationInfo/-MD_DataIdentification/abstract	M-M	A free text summary of the content, significance, purpose, scope, et Exactly one value. Content from metadata harvested in other format CSDGM…) that does not map into ISO metadata elements can be i the abstract.
Resource purpose (O) identificationInfo/-MD_DataIdentification/purpose	O-O	"Summary of the intentions for which the dataset was developed. Pu objectives for creating the dataset and what the dataset is to suppor
Resource status (O) identificationInfo/-MD_DataIdentification/status	O-O	Value is from MD_ProgressCode codelist. ISO values are {complet historicalArchive, deprecated, onGoing, planned, requir underdevelopment}. For USGIN discovery metadata, values shoul {completed, onGoing, deprecated}. See section 4.17.3 *Codelist* usage.
Resource point of contact (O) identificationInfo/-MD_DataIdentification/-pointOfContact	O-C	Contact information for the agent who is currently the steward for th information is mandatory for physical resources such as core, cuttin manuscripts. Count of (individualName + organisationName + positi The CI_ResponsibleParty/role/CI_RoleCode is from CI_RoleCode code for physical resource point of contact are {custodian, owner, poi point of contact role codes may apply for other resources. See sec for details on codelist usage.
Resource maintenance (O) identificationInfo/-MD_DataIdentification/-resourceMaintenance	O-O	This element provides information about the maintenance schedule resource (or some subset/part of the resource specified by the scop description) described by the metadata record. 0 to many MD_Mainter elements may be included. Different MD_MaintenanceInformation elem have different napMD_ScopeCode or MD_ScopeDescription. Usage of poorly described, and no actual examples of usage could be found; allow identification of a set of attribute or features (by name?), or fea attribute instances (identified how?), or a dataset, to which the main applies. Use MD_MaintenanceFrequencyCode codelist. ISO values daily, weekly, fortnightly, monthly, quarterly, biannua asNeeded, irregular, notPlanned, unknown}. See section 4.17 on codelist usage.
Graphic overview of resource (O) identificationInfo/-MD_DataIdentification/graphicOverview	O-O	Highly recommended to include a URL providing a web-accessible v of the resource if it is applicable to the described resource, particula datasets that may be represented by maps. If MD_BrowseGraphic is inc MD_BrowseGraphic/filename character string is mandatory. USGIN Rec is to provide a complete URL as a gco:characterString value for the fileType/CharacterString should be a registered MIME type or stand abbreviation (e.g. napMD_FileFormatCode). This is a repeatable ele may present different resolutions, or different parts of resource. Nam overview should provide sufficient information for user to distinguish
Resource format (O) identificationInfo/-MD_DataIdentification/resourceFormat	X-X	This element is not used by NAP or USGIN; this information is enco MD_Metadata/distributionInfo/MD_Distribution/ in USGIN metadata (*MD_Distribution and MD_Distributor*).
Resource keywords (O) identificationInfo/-MD_DataIdentification/-	O-O	Keywords should be provided to facilitate the discovery of metadata the user. Keywords should be grouped according to Keyword Type MD_KeywordTypeCode {discipline, place, stratum, tempor

descriptiveKeywords/MD_Keyword		thesaurus in which they are defined (if applicable). See section 4.1 details on codelist usage. USGIN requires that MD_Keyword/keyword contain a CharacterString (s USGIN best practice is to include keywords in English.
Condition applying to access and use of resource (O) identificationInfo/-MD_DataIdentification/-resourceConstraints/	O-O	Restrictions on the access and use of a resource or metadata. Follo specification of resourceConstraints. This attribute provides informati to the described resource itself. In some situations, the metadata ma learn of the existence of a resource that they may not actually be ab further clearance. Constraints may be represented by MD_Constraint, MD_SecurityConstraint.
Aggregation information (O) identificationInfo/-MD_DataIdentification/-aggregationInfo/-MD_AggregateInformation	O-O	This element includes either a citation for or identifier of an associat with the type of association between the datasets, and optionally the the dataset. MD_AggregateInformation requires either aggregateDataSetName/CI_Cita aggregateDataSetIdentifier/MD_Identifier. MD_AggregateInformation/ mandatory, from DS_AssociationTypeCode. ISO codelist includes { largerWorkCitation, partOfSeamlessDatabase, source, ster 4.17.3 *Codelists* for details on codelist usage. USGIN recommended a dereferenceable URI for that resource in aggregateDataSetIdentifi *[note this text updated 2013-10-25 to reflect the information we're ac* related resources that do not have a metadata record, aggregateData may be used; this element is optional if aggregateDataSetIdentifier h For USGIN profile, this property, rather than MD_Metadata/parentIdent to indicate relationships between described resources.
Spatial Representation Type (O) MD_DataIdentification/spatialRepresentationType/	O-O	value from MD_SpatialRepresentationTypeCode list. ISO codelist in grid, textTable, tin, stereoModel, video}. See section 4.17 on codelist usage.
Resource spatial resolution (O) MD_DataIdentification/-spatialResolution/-MD_resolution/equivalentScale/MD_RepresentativeFraction/-denominator	C-C	If the spatial representation type code is vector, grid, textTabl equivalentScale/../denominator MUST be provided to express spatia arbitrary convention proposed to facilitate interoperablity). If a equiva is available, that should be supplied as well. The resolution distance smallest length between two resolvable points in the dataset. To cal given a resolution distance, recommended practice is to divide the r meters by 0.0005. This assumes that the smallest distance resolvab human usage is 0.5 mm.
Resource language (O) identificationInfo/-MD_DataIdentification/language	M-O	Language for content of described resource. Default value is 'eng'. I applicable to the described resource, a nilReason='notApplicable' at included on an empty language element. Multiple instances of this e the linguistic content of the resource is available in multiple languag Use the three-letter language code from the ISO 639-2/T three lette lowercase. ISO 639 codelists are available at http://www.loc.gov/standards/iso639-2/php/code_list.php.
Topic category identificationInfo/-MD_DataIdentification/-topicCategory	C-C	A topicCategory code must be provided when hierarchyLevel is set "dataset series". Codes are from MD_TopicCategoryCode, the ISO {farming, biota, boundaries, climatologyMeterologyAtmos elevation, environment, geoscientificInformation, healt imageryBaseMapsEarthCover, intelligenceMilitary, inland oceans, planningCadastre, society, structure, transport utilitiesCommunication}. See section 4.17.3 *Codelists* for deta

		Most USGIN resources will have MD_TopicCategoryCode = "geoscientificInformation", which is the default value for this pr topic categorization should be done using keywords.
Resource content extent identificationInfo/-MD_DataIdentification/extent/-EX_Extent	C-C	Defines the spatial (horizontal and vertical) and temporal region to w the resource applies. For USGIN, a geographicElement/geographicBou provided for any resource with content related to some geographic l geoscience resources, temporal extent expressed using named time geologic time scale should be specified as thematic keywords. If the not related to a geographic area, the place keyword 'non-geographic in the descriptiveKeywords element. In some situations, geographic indicated using location keywords, but this is strongly discouraged b location-based searches using geographic coordinates will not find s cases, geocoding tools from services like Google can often be used approximate bounding box.
Resource content extent description identificationInfo/-MD_DataIdentification/extent/-EX_Extent/description	C-O	Free text that describes the spatial and temporal extent of the datas geographic place names are used to express the geographic extent specifies that these should be encoded using keyword with keyword Geographic names may be duplicated in the EX_Extent/description.
Resource content extent bounding box identificationInfo/-MD_DataIdentification/extent/-EX_Extent/geographicElement/-EX_GeographicBoundingBox	O-C	USGIN profile requires that if an EX_Extent/geographicElement is supp geographic bounding box with bounding latitude and longitude expre Geodesic System WGS 84 decimal degrees. The corner coordinates for the geographic bounding box must not c because this may result in fatal errors with some CSW implementati must thus be represented as tiny rectangles. USGIN recommended the actual point location in the lower left corner of the rectangle.
Resource content extent geographic description identificationInfo/-MD_DataIdentification/extent/-EX_Extent/geographicElement/-EX_GeographicDescription	C-X	Not used by USGIN profile. In USGIN metadata this information sh using keywords for which the MD_KeywordTypeCode = 'place'; the has the same content as EX_GeographicDescription/authority/CI_Cita is the same as the EX_GeographicDescription/code.
Resource content extent bounding polygon identificationInfo/-MD_DataIdentification/extent/-EX_Extent/geographicElement/-EX_BoundingPolygon	C-O	To improve interoperability, USGIN mandates the use of Geographi instead of bounding polygon, if a bounding polygon is included it ma harvesting catalogs. If only a bounding polygon is provided, a mini must be calculated and inserted in the EX_GeographicBoundingBox record is harvested into a USGIN catalog.
Resource temporal extent (O) identificationInfo/-MD_DataIdentification/extent/-EX_Extent/temporalElement/-EX_TemporalExtent/extent/-TimePeriod	O-O	Property contains information about temporal extent to which resour many geoscience resources, this would be the geologic time period resource applies. USGIN mandates use of TimePeriod for all temporal ex extent for beginPosition@frame and endPosition@frame attributes are #ISO time extents, USGIN requires the values for beginPosition@frame and endP populated using numeric time coordinates in Ma, measured positive increas at 1950 CE (see Temporal extents). The default frame attribute value for coordinates is "urn:cgi:trs:CGI:StandardGeologicTimeMa" ISO 8601 Default Example:

		<gml:TimePeriod gml:id="IdModern"> <gml:name>Y2kX</gml:name> <gml:beginPosition frame="#ISO-8601">2010-01-00T00:00:00</gml:beginPosition> <gml:endPosition frame="#ISO-8601">2010-12-31T24:00:00</gml:endPosition> </gml:TimePeriod> Geologic Time Example: <gml:TimePeriod gml:id="IdJurassic"> <gml:name>Jurassic</gml:name> <gml:beginPosition frame="urn:cgi:trs:CGI:StandardGeologicTime </gml:beginPosition> <gml:endPosition frame="urn:cgi:trs:CGI:StandardGeologicTimeMa ">135 </gml:endPosition> </gml:TimePeriod>
Resource spatio-temporal extent (O) identificationInfo/-MD_DataIdentification/extent/-EX_Extent/temporalElement/-EX_SpatialTemporalExtent/	O-X	Not used. Although use of EX_SpatialTemporalExtent is allowed by IS USGIN mandates encoding space time location with EX_TemporalExte EX_GeographicBoundingBox.
Resource vertical extent (O) identificationInfo/-MD_DataIdentification/extent/-EX_Extent/verticalElement/-EX_VerticalExtent	O-O	Vertical extent is used to describe the elevation location of resource Common geoscience examples are resources related to vertical loc depth in a water body. The borehole trace is the vertical coordinate (CRS) within which described resources are located using coordinat distance from the collar (or ground level, or Kelly bushing) of the bo EX_VerticalExtent has minimumValue, maximumValue that are real numbe For interoperability, USGIN mandates that if vertical extent is specifi the verticalCRS xlink:href attribute value must be "http://www.spatialreference.org/ref/epsg/5714/" (replaces URN "urn:ogc:def:crs:EPSG::5714", this value may appear in older metac VerticalCRS with origin at World mean sea level (MSL), with elevati positive in meters(http://www.epsg-registry.org/). If vertical extent is the verticalCRS xlink:href attribute value must be an identifier for a w that can be dereferenced to get information on the borehole or soun

3.3 Service identification elements (SV_ServiceIdentification)

Table 4. Service Identification properties (SV_ServiceIdentification). See discussion of the use of MD_DataIdentification and SV_ServiceIdentification preceding Table 3.

ISO 19115 and 19119 (M/C/O) xPath from MD_Metadata	NAP-USGIN M/C/O	Comments on SV_ServiceIdentification
Resource service citation (M) `identificationInfo[1]/-SV_ServiceIdentification/-citation/CI_Citation`	M-M	The citation attribute provides information for citing the described serv scientific citation purposes, a citation for the intellectual content of the presented by the service would be found in the `MD_DataIdentification/` for datasets identified in the `operatesOn` section of `SV_ServiceIdentifica` defined by Webster as "an act of quoting". For USGIN purposes, this information to identify the intellectual origin or authority for the content resource, along the lines of a citation in a scientific journal. The purpo the service is to identify a particular service instance as a unique entit for a `CI_Citation` element are `title`, `date`, and `responsibleParty`.
Resource title (M) `identificationInfo[1]/-SV_ServiceIdentification/-citation/CI_Citation/title`	M-M	USGIN recommends that the title in a service identification citation sh the particular service instance, and inform the human reader about th function, and context.
Resource reference date (M) `identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/date/-CI_Date/date/`	M-M	The citation date for a service may indicate the creation date, when th became operational, the publication date, when the service first becar revision date, which specifies the date of most recent update. If the s online, a notAvailable or superseded date may be specified. The `Date` to specify the signficance of a reported date. This date is distinct from metadata record, or the `EX_Extent/temporalElement` that specifies the ti the resource content is applicable. The data type for the `date` element is xs:date, defined thus "date uses date/timeSevenPropertyModel, with *·hour·*, *·minute·*, and *·second·* rec *·timezoneOffset·* remains ·optional" (http://www.w3.org/TR/xmlschem encoding: 2000-12-12+13:00, 2006-10-01. If the month or day is not k '01', for example '2006-01-01'. ISO CI_DateTypeCode names that ap `{creation, publication, revision}`. See section 4.17.3 *Codelists* usage.
Unique resource identifier (O) `identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-identifier/MD_Identifier`	C-O	For USGIN, because the Citation is for the service, this identifier should be i `MD_Metadata/dataSetURI`, and is therefore optional. For USGIN purposes, this element content value is only an identifier for the for accessing the service. The USGIN profile requires the use of MD_Ide identify resources. RS_Identifier may substitute for MD_Identifier in th but the USGIN profile requires use of MD_Identifer. If additional codes content is associated with the identifier, it should be encoded as `MD_Identifier/authority/CI_Citation/alternateTitle` and `MD_Identifier/authority/CI_Citation/edition`
Resource responsible party (O)	M-M	USGIN requires at least one `CI_ResponsibleParty` with a rule that count organisationName + positionName) > 0. For a service, the point of co questions or reporting problems should be in

identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-citedResponsibleParty		SV_ServiceIdentification/pointOfContact/CI_ResponsibleParty. The serv responsible party should identify the parties responsible for creating (i publishing the service. ISO Role code names applicable to a service c {originator, principalInvestigator, processor, author, pu codelist values {resourceProvider, custodian, owner} should b SV_ServiceIdentification/pointOfContact element. See section 4.17.3 on codelist usage.
Resource presentation form (O) identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-presentationForm	O-O	The form in which the service is available, which in the case of a servi service implementation described by the metadata record, so the infor generally very useful. This profile specifies no convention for its use. N to the original source of intellectual content in the datasets operated o should be in MD_DataIdentification/citation/CI_Citation elements that srv:operatesOn associations. presentationForm uses the CI_PresentationFormCode codelist; ISO co applicable to a service citation include {documentDigital, imageD mapDigital, modelDigital, profileDigital, tableDigital, audioDigital}. NAP adds {multimediaDigital, diagramDigital *Codelists* for details on codelist usage.
Resource series (O) identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/series	O-O	Information about the series or collection of which the cited service is (name + issueIdentification) > 0. At this point there is not much preced services into a formal series, so in general this element is probably no services. This profile specifies no convention for its use.
Resource other citation details (O) identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-otherCitationDetails	O-O	Free text information useful to identify and cite the described service i specifies no convention for its use.
Resource collective title (O) identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-collectiveTitle	O-O	Free text title of a "combined resource of which the service is a part." not much precedent for aggregating services into a collections, so in g probably not applicable to services. This profile specifies no conventio
Resource abstract (M) identificationInfo/-SV_ServiceIdentification/-abstract	M-M	A free text summary of the content, significance, purpose, scope, etc. described by this metadata. Exactly one value.
Resource purpose (O) identificationInfo/-SV_ServiceIdentification/-purpose	O-O	Text summary of the intentions for which the service was developed, i for creating the service and use cases it is designed to support. One v
Resource status (O) identificationInfo/-SV_ServiceIdentification/-status	M-M	Value is from MD_ProgressCode codelist. ISO Code names applicabl {completed, obsolete, onGoing, planned, required, underDev is synonymous with deprecated for this profile. See section 4.17.3 *Co* codelist usage.
Resource point of contact (O) identificationInfo/-SV_Servic	O-**O**^{R}	pointOfContact/CI_ResponsibleParty element for service metadata shou for a point of contact to report problems with the service. Element is o recommended! USGIN rule that count of (individualName + organisati

eIdentification/-pointOfContact		positionName) > 0. The CI_ResponsibleParty/role/CI_RoleCode@codeList CI_RoleCode; applicable name for the point of contact party are from {resourceProvider, custodian, owner}. Due to interoperability pr identifiers different from ISO identifiers for the same codelist elements use of ISO codelists. See section 4.17.3 *Codelists* for details on codel
Resource maintenance (O) identificationInfo/-SV_ServiceIdentification/-resourceMaintenance	O-O	This element provides information about the maintenance schedule or described by the metadata record. For a service, only one MD_Maintena elements may be included; for which the MD_ScopeDescription MD_Sc 'service'. If MD_MaintenanceInformation is present, then maintenanceAndUp mandatatory, populated by a MantenanceFrequencyCode; ISO name {continual, daily, weekly, fortnightly, monthly, quarterl annually, asNeeded, irregular, notPlanned, unknown}. See se for details on codelist usage. USGIN recommends following the NAP- practice that when SV_ServiceIdentification/status is set to "onGoing," e MD_MaintenanceInformation/dateOfNextUpdate or MD_MaintenanceInformati MaintenanceFrequency must be provided. Maintenance information specific to particular data the service presen in the dataset metadata for coupleResources associated with the serv
Graphic overview of resource (O) identificationInfo/-SV_ServiceIdentification/-graphicOverview	O-**O**[R]	Highly recommended to include a small image visual representation o provided by a map or image service. For geographic feature or data s overview might show the geographic distribution of available data. If M included, MD_BrowseGraphic/filename character string is mandatory. US practice is to provide a complete URL as a gco:characterString value property. Use napMD_FileFormatCode code values (http://www.fgdc.g register/codelists.html#IC_115) in fileType/CharacterString. See secti for details on encoding of the file format code, which is special becaus extension to the ISO base specification. Repeatable element; multiple values may present different resolutions resource. Names associated with overview should provide sufficient in distinguish these.
Resource format (O) identificationInfo/-SV_ServiceIdentification/-resourceFormat	O-X	The format of service response documents varies at the operation lev operation, different output formats may be requested. A listing of all po without bindings to the operations that respond with that format is not
Resource keywords (O) identificationInfo/-SV_ServiceIdentification/-descriptiveKeywords/MD_Keyword	O-O	Best Practice for USGIN profile metadata is to supply keywords to fac metadata records relevant to the user. **USGIN Keywords**: USGIN keyword vocabularies are in development this profile may include required keyword vocabularies. **Other Keywords**: Keyword Type - allowed ISO values from MD_Keyw {discipline, place, stratum, temporal, theme}. See section 4 details on codelist usage. MD_Keyword/keyword MUST contain a CharacterString element (see sect best practice is to include keywords in English.
Resource specific usage (O) identificationInfo/-SV_ServiceIdentification/-resourceSpec	O-X	Property not used by USGIN. Content may be ignored by clients.

ificUsage/		
Condition applying to access and use of resource (O) identificationInfo/-SV_ServiceIdentification/-resourceConstraints/	O-O	Information specifying restrictions on the access to and use of the des Constraints may be represented by MD_Constraint, MD_LegalConstraint, MD_SecurityConstraint. The attribute MD_Constraint/useLimitation is ma MD_LegalConstraint or MD_SecurityConstraint is provided. ISO19119 d as SV_ServiceIdentification/restrictions. Follow NAP specification th tification/resourceConstraints is to be used, and SV_ServiceIdentific not to be used. The metadataConstraints may allow a user to learn of th resource that they may not actually be able to access without further c
Aggregation information (O) identificationInfo/-SV_ServiceIdentification/-aggregationInfo/-MD_AggregateInformation	O-O	This element includes either a citation for or identifier of an associated along with the type of association, and optionally the activity that prod MD_AggregateInformation requires either aggregateDataSetName/CI_Citati SetIdentifier/MD_Identifier. associationType is mandatory, from DS_A ISO code names deemed applicable to service aggregation by this pr {crossReference, largerWorkCitation, source, }. See section details on codelist usage. Aggregation might be used to associate me layers with metadata for a service that provides a collection of layers (to cross reference services that can be chained, or to identify a resou to a service. If the related resource has an associated metadata record, USGIN rec is to include the identifier for that metadata record in aggregateDataSetIdentifier/MD_Identifier. For related resources that d metadata record, aggregateDataSetName/CI_Citation may be used; this aggregateDataSetIdentifier has a value. For USGIN profile, this property, rather than MD_Metadata/parentIdentif to indicate relationships between described resources. No testing has this element, or examples found for actual usage.
Resource service type (M) identificationInfo/-SV_ServiceIdentification/-serviceType	M-M	Exactly one value required. USGIN mandates use of a LocalName value (http://schemas.opengis.net/iso/19139/20060504/srv/serviceMetadata localName or ScopedName). There is not as yet a standard registry o identifiers that can serve as an authority for serviceTypes. An interim and identifiers is included in section 7.1 ServiceType (with the ad hoc 'http://resources.usgin.org/registry/serviceType201001'). Valid value are {WMS, WFS, WCS, CSW, …} Example: <srv:serviceType> <gco:LocalName codeSpace= "http://resources.usgin.org/registry/serviceType201001">WMS</gc </srv:serviceType>
Resource service type version (O) identificationInfo/-SV_ServiceIdentification/-serviceTypeVersion	O-C	Multiple serviceTypeVersion tags may not be implemented in some ha applications - USGIN recommends a reverse chronological order for s Constraint: if various versions are available, it is mandatory to list vers supported. Default is oldest version of service.
Resource service access properties (O) identificationInfo/-SV_ServiceIdentification/-accessProper	O-O	Optional MD_StandardOrderProcess element to provide information on the a service which include: fees, available date and time, ordering instructions, tu instructions and turnaround are not applicable to web services.

ties		
Resource service restrictions (O) identificationInfo/-SV_ServiceIdentification/-restrictions	O-X	Not used by USGIN; use resourceConstraints as per NAP.
Keywords (O) identificationInfo/-SV_ServiceIdentification/-keywords	O-X	Not used by USGIN; use descriptiveKeywords as per NAP
Resource service content extent (O) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent	C-C	Defines the spatial (horizontal and vertical) and temporal region to wh by the service apply. USGIN requires. Best Practice for USGIN is to in bounding box that specifies the extent to which resource content appl related to some geographic location. For geoscience resources, the te be expressed using time ordinal eras from a geologic time scale if the some particular geologic time. USGIN specifies count(description + geographicElement + temporalElem
Resource service content extent description ( ) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/description	C-C	Free text that describes the spatial and temporal extent of the dataset that description is mandatory if a geographicElement or temporalElement geographic place names are used to express the geographic extent th encoded using keywords with keyword type code = 'place'. Geograph duplicated in the EX_Extent/description.
Resource service content extent bounding box ( ) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/-geographicElement/-EX_GeographicBoundingBox	O-C	USGIN profile requires that if an EX_Extent/geographicElement is supplie geographic bounding box with bounding latitude and longitude expres decimal degrees. The corner coordinates for the geographic bounding box must not coi because this may result in fatal errors with some CSW implementation must thus be represented as tiny rectangles. USGIN recommended p actual point location in the lower left corner of the rectangle.
Resource service content extent geographic description ( ) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/geographic-Element/EX_Geographic-Description	C-X	Not used by USGIN profile, use keyword with type code = 'place'. Fo this information should be encoded using keywords, for which the MD = 'place'; the thesaurus/CI_Citation has the same content as EX_Geogr authority/CI_Citation, and the keyword is the same as the EX_Geograph
Resource service content extent bounding polygon ( ) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/-geographicElement/-EX_BoundingPolygon	C-X	To improve interoperability, USGIN mandates use of Geographic Bou polygons may be present, but may be ignored by harvesters.
Resource service temporal extent (O)	O-O	Information specifying the temporal extent to which resource is applica geoscience resources, this would be the geologic time period(s) to wh applies. Although the ISO19139 xml schema allows temporal extents to be

identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/temporalElement/EX_TemporalExtent/-extent/TimePeriod		ordered eras, TimePeriod MUST be used for temporal extent in USGIN p to make metadata interoperable. Values for beginPosition@frame and endPos populated if a temporal extent is provided. The default frame property value standard calendar date and time. For geologic time extents, USGIN requires beginPosition@frame and endPosition@frame to be populated using numeric measured positive increasing older with an origin at 1950 CE (see *Tempora* frame attribute value for geologic time coordinates is "urn:cgi:trs:CGI:Standa See section 4.21, below.
Resource service spatio-temporal extent (O) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/-temporalElement/-EX_SpatialTemporalExtent/	O-X	Although use of EX_SpatialTemporalExtent is allowed by ISO19139 and practice is to encode space time location with EX_TemporalExtent and E Box. Other optional extent elements may be included, but they may be implementations processing the metadata document.
Resource service vertical extent (O) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/-verticalElement/-EX_VerticalExtent	O-O	Vertical extent is used to describe the elevation location of resources described service instance, if applicable. Common geoscience examp related to vertical location in a borehole or depth in a water body. The vertical coordinate reference system (CRS) within which described re using coordinates measured in linear distance from the collar (or grou bushing) of the borehole. EX_VerticalExtent has minimumValue, maximumValue that are real numbers For interoperability, USGIN mandates that if vertical extent is specifie verticalCRS xlink:href attribute value must be "http://www.spatialreference.org/ref/epsg/5714/"; this is the VerticalCR mean sea level (MSL), with elevations measured up positive in meters (http://www.epsg-registry.org/). If vertical extent is specified using dep xlink:href attribute value must be an identifier for a well or depth sound dereferenced to get information on the borehole or sounding geometry other information, reported depths should be assumed to be true verti the mean elevation of the geographic extent reported for the resource other information, reported depths should be assumed to be true verti the mean elevation of the geographic extent reported for the resource
Coupled Resource ( ) identificationInfo/-SV_ServiceIdentification/-coupledResource	O-O	This element correlates operations (identified by operationName) with d identifier). For logical consistenty, and SV_coupledResource/identifier equal to MD_DataIdentification/citation/CI_Citation/identifier/MD_Ide dataset that is the target of a SV_ServiceIdentification/operatesOn elem inline MD_DataIdentification/citation../code element, or a @uuidref att is necessary to implement the many-to-many relationship between da operations in a single service.
Coupled Resource operation name (M) identificationInfo/-SV_ServiceIdentification/-coupledResource/-SV_CoupledResource/-operationName	M-C	Mandatory if operations are specific to particular coupled resources. S service operation: GetMap, GetFeature, etc. There is no internal chec record that the given operation name is valid.
Coupled Resource identifier (M)	M-C	Mandatory in service metadata record if coupling type is tight or mixed tightly coupled dataset. Equal to MD_DataIdentification/citation/CI_Ci

identificationInfo/-SV_ServiceIdentification/-coupledResource/-SV_CoupledResource/identifier		identifier/MD_Identifier/code for a dataset that is the target of a SV_Se operatesOn element (either in an inline MD_DataIdentification/citation. @uuidref attribute).
Coupled Resource scoped name (X) identificationInfo/-SV_ServiceIdentification/-coupledResource/-SV_CoupledResource/ScopedName	X-O	OGC 07-045 application profile for ISO metadata using CSW 2.0.2 ex SV_CoupledResource with a ScopedName, defined as a scoped iden in the context of the given service instance (e.g. layer name or feature necessary for users to generate service requests (like GetMap or Get ISO service metadata.
Service coupling type (M) identificationInfo/-SV_ServiceIdentification/-couplingType	M-M	Type of coupling between service and associated data, if applicable. C by USGIN: • loose - Service instance is not associated with a specific datas datasetcollection. Loosely coupled services may have an asso types through the service type definition; no MD_DataIdentifica described. • mixed - Service instance is associated with a specific dataset but this can also be used with external data (i.e. data that is no operatesOn association). MD_DataIdentification describes the instance. • tight - Service instance only operates with specific data; MD_D element MUST be provided for the coupled data.
Service operations (M) identificationInfo/-SV_ServiceIdentification/-containsOperations	M-M	This element describes operations performed by the service, but the I includes insufficient detail to completely describe all parameters neces connection to a service. Widely used xml formats exist to describe ser including OGC getCapabilities.xml and W3C Web Service Description srv:containsOperations is a required element in a valid SV_ServiceIdent instance. At least one service operation MUST be described with the operationDescription/CharacterString = 'serviceDescription'. The identificationInfo/SV_ServiceIdentification/containsOperations/SV_Op connectpoint link for this operation MUST retrieve the service-specific document (e.g. WSDL, GetCapabilities, WADL).
Service operation name (M) identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/-operationName	M-M	Operation name following convention of the described service; manda associated with SV_OperationMetadata element with the operationDes String = 'serviceDescription'.
Service operation distributed computing platforms (M) identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/DCP	M-M	mandatory instance is associated with SV_OperationMetadata eleme operationDescription/CharacterString = 'serviceDescription'. USGIN p provision for use of this element.
Service operation description (O)	O-M	mandatory instance is associated with SV_OperationMetadata eleme operationDescription/CharacterString = 'serviceDescription'.

identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/-operationDescription		
Service operation invocation name (O) identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/-invocationName	O-O	service instance specific string to use in requests to invoke an operati
Service operation online resource (M) identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/-connectpoint	M-C	For loosely coupled services, operations and endpoints need to be de information endpoint (CI_OnlineResource) that returns a service self des MUST always be provded. The CI_OnlineResource/linkage URL should request if a single http GET request is sufficient. If the request require message package, or does not use standard web protocols, the CI_OnlineResource/description MUST provide instructions for executing
Service operates on (O) identificationInfo/-SV_ServiceIdentification/-operatesOn	O-C	"Provides information on the datasets that the service operates on" (IS If SV_ServiceIdentification/-couplingType is 'tight' or 'mixed', then ope must include a valid MD_DataIdentification element inline or by @uuidre explicitly links to an existing dataset metadata record that describes th The value of SV_ServiceIdentification/operatesOn@uuidref or SV_ServiceIdentification/operatesOn/MD_DataIden-tification/citation/CI_Citation/identifier/MD_Identifier/code must c the SV_ServiceIdentification/coupledResource/MD_CoupledResource/ident metadata record for the coupled dataset is a separate gmd:MD_Meta service described in the service metadata record should be identified the dataset. Explicitly linked reference example: <srv:operatesOn uuidref="13ce1e84-c887-4fd8-b888-8d021b1fa4c2" xlink:href=http://resources.azgs.org/geonetwork/srv/en/metada xlink:title="azgs:azgeochron"/>

4 Usage notes for Metadata Elements

This section presents additional information and discussion to supplement that in Table 2, Table 3, and Table 4.

4.1 Metadata file identifier

`MD_Metadata/fileIdentifier` is unique identifier for the metadata file. Some metadata profiles suggest that the metadata file identifier ID should be the same as the ID for the described resource. In the USGIN scheme, the metadata record is considered a separate resource, with a distinct identifier, from the resource it describes. The described resource identifier is the Unique resource identifier (`DatasetURI`, 4.8, below).

4.2 Metadata hierarchy

The ISO19115 specification (especially Annex H) discusses the use of metadata hierarchy, in which a resource may inherit metadata properties from parent metadata records in the hierarchy. For example a `dataset` in a dataset series might inherit all of the metadata content from the parent dataset series metadata record, except for dataset-specific data quality metadata. The linkage would be made through `MD_Metadata/parentIdentifier`. This kind of nesting, effectively database normalization, makes sense in a data management environment, but in a discovery environment, it present problems. Many queries would require joins with the parent metadata record, and software clients would be required to navigate the parent links to acquire 'inherited' properties from 'parent' records. For catalog service purposes, USGIN mandates that in metadata records returned by services, all inherited properties in such a hierarchy should be included explicitly in the metadata document, as opposed to implicitly through the `parentIdentifier` link. Internal document links may be used where allowed by the xml schema for identified elements repeated in a single response document. Resolvable xlink:href links to external resource are technically possible, but require that client software is xlink-aware, which is not always the case.

4.3 Metadata Contact vs. Resource Citation vs. Resource Contact

There are various locations to store contact information within an ISO 19139 metadata record. Here is a summary of the required contact properties and their significance as it pertains to the USGIN Profile.

- `MD_Metadata/contact/CI_ResponsibleParty` or "metadata point of contact" describes how to contact the party responsible for the **metadata** record to allow users to report errors, updates to metadata etc. The mandatory `CI_RoleCode` is set to "`pointOfContact`".
- `MD_Metadata/identificationInfo/[MD_DataIdentification || SV_ServiceIdentification]/-citation/CI_Citation/citedResponsibleParty/CI_ResponsiblerParty` provides information to identify the **intellectual origin** of the content in the described resource. This is straight forward when citing library resources (books, journals, etc.) but less clear when defining the intellectual origin of, for example, physical samples. The mandatory `CI_RoleCode` is set to one of the ISO codelist values {`custodian, owner, distributor, originator, pointOfContact, principalInvestigator, publisher, author`}.

- `MD_Metadata/identificationInfo/[MD_DataIdentification || SV_ServiceIdentification]/-pointOfContact/CI_ResponsibleParty` or "resource point of contact" contains information on who to contact to **access** the described resource. The mandatory `CI_RoleCode` is set to one of the ISO codelist values {`resourceProvider, custodian, owner, user, distributor, originator, pointOfContact, principalInvestigator, processor, publisher, author`}.

Additional contact information in the distribution section of the metadata provides point of contact for individual distribution processes to report problems with access to resources via different distributions.

4.4 Resource Title

Resource titles should provide sufficient information to distinguish the resource from other similar resources. They are not required to be globally unique, but bear in mind that users will be presented only with the resource title in typical brief search response documents. It is thus a disservice to have significant duplication of title strings or uninformative titles.

4.5 Resource Abstract

Ideally the resource abstract provides a succinct summary of the content of the resource, the purpose for which it was originally created, some indication of important quality parameters to help evaluate fitness for other purposes, any significant constraints on use of the resource, and a list of distribution options. After reading the abstract a user should have a good idea of whether the described resource contains the information they need and is accessible in a fashion that they can use.

4.6 Resource Type

The ISO 19115 `MD_Metadata/hierarchyLevel` property provides a high level categorization of resource types. The European INSPIRE Implementing Rules (MD_IR_and_ISO_20090218) proscribes the code list for the first hierarchyLevel xml element in an MD_Metadata document to be one of {`dataset, service, series`}, or the metadata set will be considered out of scope for the directive. Thus, metadata meant to be utilized by INSPIRE catalogs must follow this rule. The full ISO MD_ScopeCode list has a wider (and more useful) variety of resource categories; one or more hierarchyLevel elements using these codes could follow the first one with an INSPIRE-valid code in the first element to maintain INSPIRE compliance.

Table 1 in this document includes a more geoscience-domain-specific list of resource types, and values from this list should be used in one or more `hierarchyLevelName` elements. To enable resource-category-type searches to find narrower subcategories without complex query processing, `hierarchyLevelName` elements for the resource type and all broader/more general resource type categories should be included. The hierarchical categorization of the resources is encoded with the most specific category first, and progressively broader categories listed subsequently. Thus, harvesters that only take the first `hierarchyLevelName` element will get the most specific value. For example, if the resource is a photograph:

```
<gmd:hierarchyLevelName>
  <gco:CharacterString>Photograph</gco:CharacterString>
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
  <gco:CharacterString>StillImage</gco:CharacterString>
```

```
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
  <gco:CharacterString>Image</gco:CharacterString>
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
  <gco:CharacterString>Document</gco:CharacterString>
</gmd:hierarchyLevelName>
```

Note that the distinction of resource type and format is not always clear. Table 1 attempts to define resource types that are not specifically bound to a particular format, but are defined based on the kind of content. Format is interpreted as relating to specific approaches to encoding content and committing it to some sort of media.

4.7 Resource Locator

URL's for online access to resources are encoded in USGIN ISO 19139 metadata documents in the element `MD_Distribution/transferOptions/MD_DigitalTransferOptions/online/CI_Online-Resource`. Consistent use of this rule eliminates ambiguity on where to locate the URL to access a resource. Conventions for use of the `CI_OnlineResource` subelements `protocol`, `applicationProfile`, `name`, `description`, and `function` to enable metadata clients to reliably access referenced resources are discussed in section 4.15 CI_OnlineResource.

4.8 Unique Resource Identifier

The `MD_Metadata/DataSetURI` property should be a globally unique identifier for the described resource. The protocol used for this identifier is not specified by the USGIN Profile, but if it does not have a know resolution service, the capabilities document for a CSW service providing the metadata should have at least a text explanation of how to resolve URI's used by the service. Protocols with available resolvers include http (use the WWW DNS system) and doi (http://dx.doi.org/). Some authorities using urn: protocols are also implementing or have resolver services in place.

4.9 Browse Graphics

Links to browse graphics are useful to assist users evaluate resources that have a graphical aspect. The `MD_BrowseGraphic` element includes a fileType attribute that is important to inform client software accessing and displaying the graphic. For the USGIN profile, use standard MIME types in the `gmd:fileType/gco:CharacterString`. See the official Internet Assigned Numbers Authority (IANA) Mime Media Type registry at http://www.iana.org/assignments/media-types for a list of registered MIME types and their scope.

```
<gmd:MD_BrowseGraphic>
  <gmd:fileName>
        <gco:CharacterString>http://publicdocs.mnr.gov.on.ca/View.asp?-
        Document_ID=9632&Attachment_ID=18204</gco:CharacterString>
  </gmd:fileName>
  <gmd:fileDescription>
      <gco:CharacterString>Base Map from OMNR</gco:CharacterString>
  </gmd:fileDescription>
```

```
  <gmd:fileType>                <!-- This is a MIME Type -->
      <gco:CharacterString>image/jpg</gco:CharacterString>
  </gmd:fileType>
</gmd:MD_BrowseGraphic>
```

Code example 1. Encoding url, display name and file type for browse graphic.

4.10 Resolution and equivalentScale

For spatial datasets, some indication of the resolution of the data is very useful for evaluating fitness for use. From a data perspective, resolution in a spatial dataset specifies the the smallest resolvable distance between two points in the Earth for that dataset. For a grid or coverage, this would be the average distance between sample points. From data portrayal perspective, an `equivalentScale` is reported, representing the scale at which the portrayal was intended to be viewed when it was constructed. If the spatial representation type code is `vector, grid, textTable` or `tin`, an `equivalentScale/../denominator` MUST be provided to express spatial resolution. This is an arbitrary convention adopted to facilitate interoperablity. If a `equivalentScale/../distance` is available, that should be supplied as well. To calculate `equivalentScale` given a resolution distance, recommended practice is to divide the resolution distance in meters by 0.0005. This assumes that the smallest visually resolvable distance in a map display for human usage is 0.5 mm.

4.11 Resource Language

USGIN metadata is assumed to use American English and metadata documents should be returned in that language. The ISO 19139 XML includes elements and provisions to implement multiple languages in a single metadata package using `PT_Text` and `LocalizedCharacterString`, but in order to avoid complexity with USGIN recommended practice is to implement metadata search for different languages as different services. Each service would provide `CharacterStrings` in the same, single language, specified by the `MD_Metadata/language` element in returned metadata records.

4.12 Encoding of Vertical Extents

`EX_VerticalExtent` has `minimumValue`, `maximumValue` that are real numbers, and a `verticalCRS`. The vertical coordinate reference system (CRS) specifies how the bounds of the vertical extent are georeferenced. The USGIN profile accepts two approaches. Elevation is reported with reference to Earth mean sea level, measured positive upward. On the other hand, subsurface or subaqueous data are typically referenced relative to a borehole trace or sounding, measured positive downward from a local datum at the borehole or sounding orgin. The origin is typically a borehole collar at the ground surface, the Kelly bushing or drill floor on a drilling rig, or a ship deck., For to facilitate interoperability.

For interoperability, USGIN mandates that both elevation and depth vertical extent bounds MUST be specified in meters. In addition, if vertical extent is specified as an elevation, the `verticalCRS` xlink:href attribute value MUST be "http://www.spatialreference.org/ref/epsg/5714/"; this is the VerticalCRS with origin at World mean sea level (MSL), with elevations measured up positive in meters (http://www.epsg-registry.org/). If vertical extent is specified using depth, the `verticalCRS` xlink:href attribute value must be an identifier for a well or depth sounding that can be dereferenced to get information on the borehole or sounding geometry.

4.13 Content information: Entities and Attributes

One of the major deficiencies of the ISO 19115/19139 specifications is the absence of model elements to describe the entities and attributes in standard tabular datasets. Description of dataset schema using the ISO Technical Committee 211 (TC211) framework requires use of ISO19110. A standard XML schema for implementing this specification has recently become available, and entity-attribute metadata for dataset should be documented using the schemas located at http://www.isotc211.org/2005/gfc/gfc.xsd. The ISO19139 XML implementation of ISO19115 does not allow inclusion of the feature catalog inside the metadata record, but provides a citation element in the `gmd:contentInfo/gmd:MD_FeatureCatalogueDescription` element. The USGIN profile provision is to link to the feature catalog using an xlink:href attribute on the `gmd:featureCatalogueCitation`, identified as such by an accompanying xlink:title attribute with the value "Link to ISO19110 feature catalog". Code Example 2 shows recommended encoding. Section 8.4, ISO19110 Feature Catalog example, provides an example encoding of entity-attribute information for a feature dataset.

```
<gmd:contentInfo>
  <gmd:MD_FeatureCatalogueDescription>
    <gmd:includedWithDataset>
     <!-- true if a FeatureCatalog document is packaged with the dataset -->
       <gco:Boolean>false</gco:Boolean>
    </gmd:includedWithDataset>
   <!-- a dataset may contain multiple feature types. For an MS Access type database, each
   table    could be considered a featue type -->
        <gmd:featureTypes>
   <!-- the codespace is the namespace URI, should dereference to obtain XML schema (if
   applicable)-->
          <gco:LocalName codeSpace="
   http://xmlns.geosciml.org/geosciml-portrayal/2.0">GeologicContact_2</gco:LocalName>
        </gmd:featureTypes>
   <!-- follow pattern used by Stanford library http://opengeoportal.org/archives/category/-
   uncategorized); href is to an ISO19110 feature catalog,
   with featuretype identified by fragment ID -->
        <gmd:featureCatalogueCitation xlink:href="
   http://schemas.usgin.org/featurecatalog/12.xml#geologiccontact"
          xlink:title="Link to ISO19110 feature catalog"/>
   </gmd:MD_FeatureCatalogueDescription>
  </gmd:contentInfo>
```

Code Example 2. Encoding of feature catalog link for entity attribute description of datasets. See 8.4 ISO19110 Feature Catalog example for and example instance document.

4.14 Use of MD_Distribution and MD_Distributor

The ISO19115 model provides two possible paths for specifying information about how a user can access the resource. The `MD_Distribution` element may have 0 to many `distributionFormat`, `distributor`, and `transferOptions` child elements (see Figure 1). On the other hand, each of the `distributor` child elements may have 0 to many `distributorFormat` and `distributorTransferOption` elements. Several major existing applications that consume ISO19139 xml metadata files are configured to expect format and transfer option information to be at the `MD_Distribution/-distributionFormat` and `MD_Distribution/transferOptions` path. This works fine as long as there are not different format or transfer options from different distributors, or different transfer options for different formats. In these cases, a binding between distributor, format, and transfer options

necessitates use of the MD_Distribution/distributor/MD_Distributor path to distributorFormat and distributorTransferOptions (and distributionOrderProcess) information that works together. In order to accommodate both existing applications that utilize content in the MD_Distribution/-distributionFormat and MD_Distribution/transferOptions elements, and situations that require binding between distributor, order process, format, and transfer options, the USGIN profile mandates that if multiple MD_Distribution/distributionFormat or MD_Distribution/transferOptions elements are included in a document, all formats must be available via all the specified transfer options, and the content of these elements should be included in line. If multiple MD_Distribution/distributor elements are present, without child MD_Distributor/distributorFormat or MD_Distributor/distributorTransferOptions elements, then all formats and transfer options are available from all distributors.
To specify different bindings between distributor, order process, format, and transfer options, a separate MD_Distribution/distributor/MD_Distributor instance is included for each binding. At least one MD_Distribution/distributionFormat and one MD_Distribution/transferOptions element MUST be included for applications that expect content in these elements, and the format and transfer options specified by these elements should apply to the first distributor/MD_Distributor element. Repeated CI_ResponsibleParty, MD_StandardOrderProcess, MD_Format or MD_DigitalTransferOption elements in the distributor/MD_Distributor elements should be specified by reference (xlink:href to gml:id of first occurrence of the element within the document). The implication is that the distributionOrderProcess/MD_StandardOrderProcess, distributorFormat/MD_Format, and distributorTransferOptions/MD_DigitalTransferOptions child elements of a single MD_Distributor are all compatible with each other. Following this convention will produce schema valid metadata, and the distributor elements after the first can be ignored by applications expecting only a single distributor element.

In summary, for the USGIN profile, each distributor/MD_Distributor is a binding between one or more transfer options and the distributor formats that are available through that/those transfer options (MD_DigitalTransferOptions/onLine/CI_OnlineResource in particular). If different formats are available from the same distributor, but have different transfer options, these should be represented as different distributor/MD_Distributor instances.

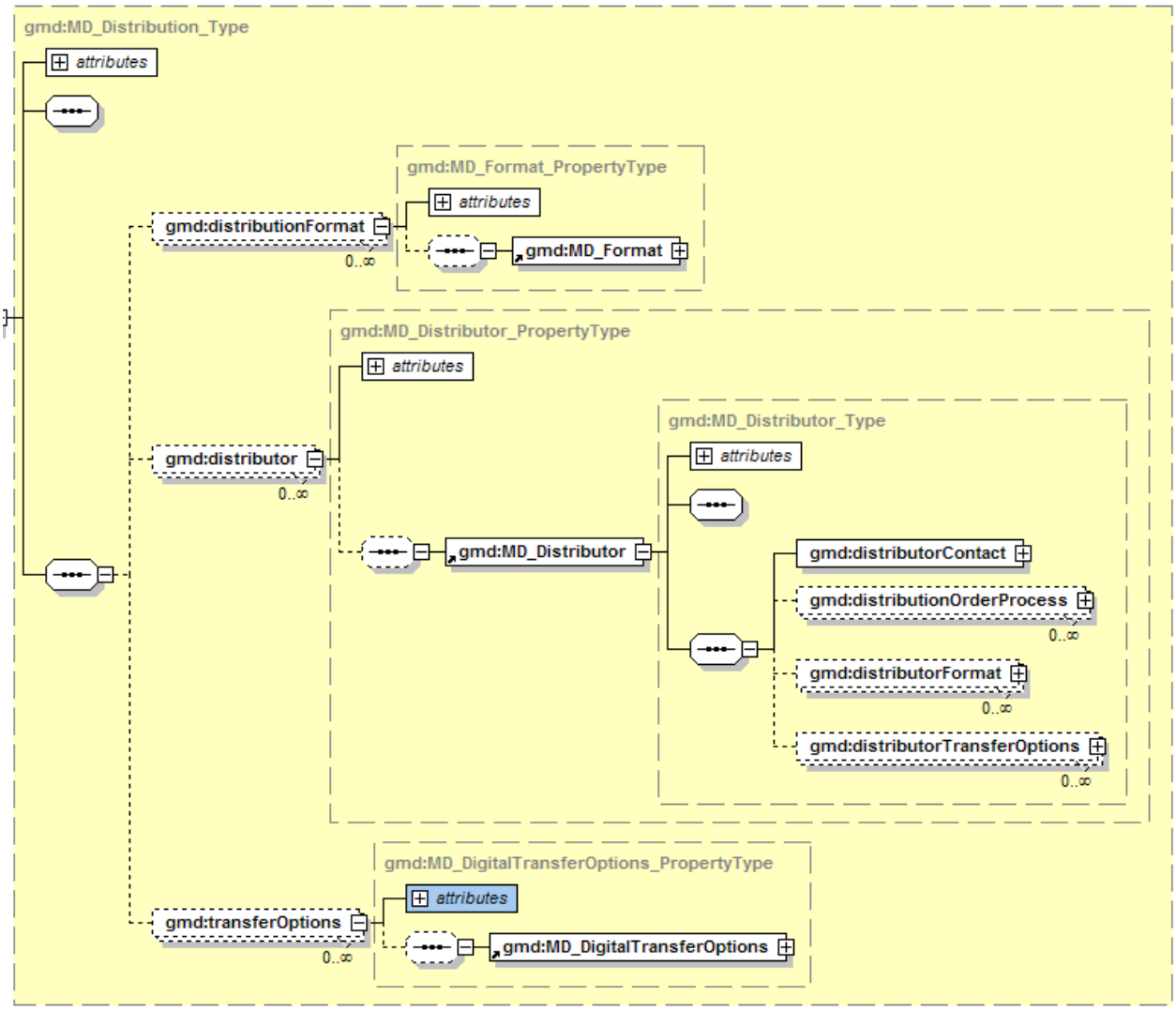


Figure 1. gmd:MD_Distribution_Type diagram

4.15 Distribution Format

If the resource distribution is a physical object, like a book, rock sample, paper document, the `distribution/../MD_Format/name` MUST be a term from the USGIN distribution format codelist (see Table 6). Note that format is partially orthogonal from resource type (Table 1). A document may be available in various digital (pdf, tiff, doc, txt) or non-digital (book, loose sheets) formats.

4.15.1 Digital resources

The format vocabulary needs to be designed to work in the framework of the `distribution/../-MD_DigitalTransferOptions`, which provides protocol, applicationProfile, name, and function subelements for online resources (CI_OnlineResource), and `MD_MediumNameCode` and `MD_Medium-FormatCode` for offline electronic resources (`MD_Medium`). For digital resources it provides terms to specify file-format information that does not have any other obvious home. Examples in INCITS 453, INSPIRE 19115/19, and ANZLIC 2007 populate `MD_Format/name` with values like 'ESRI ARC/INFO Coverage', 'ESRI shapefile', 'ESRI ARC/INFO Export e00', and 'MapInfo MID/MIF' all pertain to digital resources. If a MIME format (http://www.iana.org/assignments/media-types/)

is defined for a digital file format, the MIME media-type code should be used. If no appropriate MIME type is registered with IANA, USGIN mandates that the distribution format for digital resources should specify the file format using a pattern that includes vendor, application name, and file extension.
Pattern for digital resources: `[vendor:applicationName]/fileExtension`. The vendor and application names may not be applicable, and could be omitted, but the '/' and file extension should always be present. If the format consists of a single file, the file extension is a three letter file-type abbreviation assigned by the vendor. If the format consists of a package of files (e.g. an ArcGIS file geodatabase), the file extension is a name that in most cases should be obvious from vendor usage. The accompanying `MD_Format/version` value should indicate the version of application software if the format is specific to some version.
Service metadata describes a service as a resource, in which case distribution is effectively equivalent to the set of operations accessible at the service instance endpoint. The ISO 19119 service metadata model does not provide a clear way to associate individual operation responses with possible output formats for those responses. If all output formats are applicable to all service requests, they can be listed in `distributioninfo` as a collection of `distributionFormat/MD_Format` elements, but this is not a general solution. Another approach (e.g. GeoNetwork OpenSource, v. 2.4) to documenting output formats is to put format information in a collection of `srv:connectPoint/CI_OnlineResource/protocol` elements, with `connectPoint` elements for each format available on each operation. The USGIN profile convention is that operation metadata is best conveyed to metadata consumers by providing a link to a service-specific description file (getCapabilities or WSDL, see 4.22 Operation metadata), and this MUST be provided in a service metadata record by a `SV_OperationMetadata` element with an `operationDescription` value of 'serviceDescription'. The output formats offered by the service should be listed in a `distributionInfo` element as a collection of `distributionFormat/MD_Format` elements if all formats are applicable to all service requests, or if the mapping between requests and formats is obvious. Encoding of the format name should use whatever convention is used by the service to specify that output format in requests made to the service.
Table 5. Example format strings for digital files. These are to be used only if an appropriate MIME type is not defined.

ESRI:ARCINFO/Coverage
/shapefile
ESRI:ARCINFO/e00
PitneyBowes:MapInfo/mid
ESRI:ArcGIS/mdb
ESRI:ArcGIS/fileGeodatabase
Microsoft:Access/mdb

4.15.2 Non digital resources

The `MD_Format` element is the only format information for resources that do not have digital transfer options, and USGIN proposes Table 6 as a vocabulary for use to specify format of non-digital resources. Although this codelist could be implemented as a schema extension, for the time being we propose to use it as a controlled vocabulary specified by profile and practice, rather than schema. Use of such controlled vocabulary can be indicated by using xsi:type on the gco:characterString element to make the type gml:CodeType, which then requires a codeSpace

attribute. The distribution format Identifier from Table 6 should be used as the element value. Example encoding:

```
<gmd:MD_Format>
  <gmd:name>
      <gco:CharacterString xsi:type=”gml:CodeType”
    codeSpace="http://resources.usgin.org/registry/distributionFormatNames201001">
    sample:core</gco:CharacterString>
  </gmd:name>
  <gmd:version nilReason="notapplicable"/>
<gmd:MD_Format>
```

Table 6. USGIN Distribution formats for non digital resources. URI for this codelist is http://resources.usgin.org/registry/distributionFormatNames201001

Identifier	Name	Parent format	Scope
physicalArtifact	Physical artifact		described resource is a physical object
sample	Sample	physicalArtifact	Use for uncategorized sample.
sample:core	Core	sample	Cylindrical rock sample extracted from Earth with a coring drill
sample:cuttings	Cuttings	sample	Small rock fragments recovered from drilling process as sample of material being drilled
sample:fluid	Fluid	sample	Sample of a fluid
sample:handSample	Hand sample	sample	Single piece or pieces of material.
hardCopy	page-size document	physicalArtifact	A physical copy of a document on paper, film, or other similar material.
hardCopy:book	Book	hardcopy	Manuscript printed on paper, bound into a single volume
hardCopy:manuscript	Manuscript	hardCopy	Other printed or written representation on physical media, usually paper or mylar, includes unbound books, index cards, loose notes, file folders of papers
hardCopy:printedImage	Printed image	hardCopy	Image on paper or other opaque or semi-opaque media.

printedImage:paperMap	oversize sheet	printedImage	Map image on a single sheet
hardCopy:filmImage	photographic film	hardCopy	Image on film, viewed by passing light through the film. Includes single still images and collections of connected images for a movie.
audioRecording:tape	magnetic tape	physicalArti-fact	use for sound resources that are recorded on magnetic tape.
audioRecording:otherMedia	vinyl disk, wax cylinders	physicalArtifact	

4.16 CI_OnlineResource

The CI_OnlineResource element is used in a variety of contexts in ISO19115 content (see Table 7). This complex-content-element provides a linkage to some online resource related to the metadata or the described context resource. It is thus analogous to atom:link (IETF RFC-4287) or the link element described in IETF RFC-5988. For links that locate online documents accessible using standard browser and file type resolution technology, the link can be as simple as a single URL element that retrieves a representation of the resource. There are many other kinds of related resources that the onlineResource element may point to, including web services that provide access to a dataset resource, metadata for related or source resources, specifications for standards or extensions to standards.

Table 7. Contexts for CI_OnlineResource elements

Context	usage
contactInfo/onlineResource	Link to online resource to assist contacting a referenced individual or organization.
distributionInfo/transferOptions	link to a representation of the resource that is the subject of the metadata record
metadataExtension	link to document describing specifications for elements and attributes extending the base standard.

For the USGIN, it is desired that such links are accompanied by sufficient description that the linked resource can be accessed and provided to the metadata user automatically, with little or no operator intervention other than clicking on a link in a presentation of the metadata. A content specification for such machine-actionable links is discussed in a separate USGIN discussion paper. This approach essentially considers the metadata as a sort of hypermedia in which the CI_OnlineResource elements define the 'affordances', or actions that are available to a client.

The CI_OnlineResource element contains 6 child elements. The only required content is a URL that will access the online resource. All other properties are optional. The `protocol`, `applicationProfile`, `name` and `description` properties are all free text with loosely defined semantics as indicated by the element names.

4.16.1 CI_OnlineResource protocol

Since the base level web locator scheme (ftp, http) is indicated by the prefix of the URL (IETF RFC 1738), USGIN metadata uses the `protocol` attribute to specify a higher-level protocol in the network stack, e.g. a serviceType for data accessed through remote-procedure-call-on-HTTP type services.

4.16.2 CI_OnlineResource applicationProfile

The use of `applicationProfile` for file-based resources accessed via URL is discussed in Table 2. For links to documents for which the service type (encoded in the `CI_OnlineResource/protocol` property) and base protocol prefix in the `CI_OnlineResource/link` URL (e.g. 'http:') do not provide sufficient information to guide client software, the `applicationProfile` property is used to indicate a profile on the serviceType or some variation in document encoding or content conventions. For example WFS services may offer different features in different namespaces or encoding schemes, a catalog may offer different metadata encoding, or a resource-oriented service may offer representations using different encoding schemes. The same scheme may be used with different conventions, for instance different profiles for the use of ISO19139 or csw:record XML metadata encoding. RDF representations may be offered in XML, Turtle, or N3 encoding. Although these variations may be deducible using content negotiation or by accessing and parsing a service description document, much simple client logic is possible if the information is provided up front with the link. The actual string value that should be provided to specify an application profile should be defined by the agent originating the profile.

4.16.3 CI_OnlineResource name and description

In order to provide further clarity and guidance for the utilization of linked resources USGIN recommends use of the `CI_OnlineResource/name` property as indicated in Table 13. In order to identify the linkage element that locates the service description document (see 4.22 Operation metadata), USGIN mandates using `CI_OnlineResource/name` = "serviceDescription" in the `CI_OnlineResource` element with the linkage to the service description. The assumption is that any client that can connect to the service will know what to expect as the service description provided by that service.

The `description` property may be used to provide information about the online resource, and more usefully, to provide an explanation of how the other content of the `CI_OnlineResource` element is to be used to access the resource. The `linkage/description` should provide any necessary additional information in a text narrative as well. In some cases the protocol, applicationProfile and name properties may be insufficient to enable machine access to the resource through the provided link. The `CI_OnlineResource/description` element may include key values pairs to provide additional necessary information. USGIN recommendataion is to encode these in a 'parameter' object using JSON syntax. The parameters value is a list of key:value pairs enclosed in curly brackets ('{"key":"value", "key1":"value1"…}'). The keys should be the exact string that is required for the data access request parameter. For example, a dataset distributed through a particular layer in a multi-layer WMS:

```
<gmd:description>
  <gco:CharacterString>
    Whatever descriptive text you want.
    parameters:{"layers":"gtp_datagap_well_data_collection",
                "styles":"wellsDataSLD"}
  </gco:CharacterString>
```

```
</gmd:description>
```

The above example also shows how to encode multiple parameters.

In the case of a dataset distributed through a particular feature type in a multi-feature WFS:

```
<gmd:description>
  <gco:CharacterString>
    Whatever descriptive text you want.
    parameters:{"typeName":"BoreholeLithInterval2.0"}
  </gco:CharacterString>
</gmd:description>
```

In some cases, a WFS may offer multiple feature types based on the same base dataset, for example a geologic map service that offers a separate contacts and faults feature types from the same geologic lines dataset. In this case, the typeName values will be a comma-delimited list:

```
<gmd:description>
  <gco:CharacterString>
    Whatever descriptive text you want.
    parameters:{"typeName":"contacts,faults"}
  </gco:CharacterString>
</gmd:description>
```

4.16.4 CI_OnlineResource function property

The `function` property is populated from a codelist (see Table 8) and is used to indicate the expected action that actuating the link will trigger. (see section 4.17.3 *Codelists*).

Table 8. OnlineFunctionCode values from NAP (INCITS 453) and ISO 19115. '(ISO)' after the text in column 2 indicates function codes from ISO19115 (2006).

OnLineFunctionCode	USGIN profile usage
browsing	`CI_OnlineResource/linkage` is a valid URL for a web application that enables user to explore and seek information about the resource from a Web browser
download	`CI_OnlineResource/linkage` is a valid URL that will initiate transfer of a file containing the described resource to the local system. (ISO)
emailService	`CI_OnlineResource/linkage` is a valid mailto: link that will initiate an e-mail message to the correct party to request access to the described resource. (NAP)
information	`CI_OnlineResource/linkage` is a valid URL that will access a resource that provides information about the resource content, for example an explanatory web page, a downloadable document with narrative text describing the resource, or an XML encoded metadata document. (ISO)
offlineAccess	`CI_OnlineResource/linkage` is a valid URL that will access a web page providing instructions for requesting the resource from the provider. (ISO)

order	CI_OnlineResource/linkage is a valid URL that will access a web page to initiate an ordering process for obtaining the resource. (ISO)
search	CI_OnlineResource/linkage is a valid URL that will access a search interface for seeking out specific information content contained by resource, e.g. the metadata describes a database, and this linkage accesses a search interface to search the database (ISO)
upload	CI_OnlineResource/linkage is a valid URL for a web interface to transfer data from a local storage device or system to be included in the described resource. (NAP)
webService	CI_OnlineResource/linkage is a URL that accesses a standard web service. If the CI_OnlineResource/name is 'ServiceDescription' then the link is to a machine-readable standard service description document as defined by the service specification. Example description document may be a Web Services Description Language (WSDL) file or OGC getCapabilities file. (NAP)

4.16.5 Scoped name in service distributions

A distribution option for a dataset might be through a Web Map Service that contains one or more layers portraying the described dataset, along with layers portraying other datesets. Likewise, a dataset may be distributed as a particular feature type in a Web Feature Service that offers multiple features. Analogous internal knowledge may be required for various other service protocols. To completely specify access to the correct data representation requires knowing the layer name, feature type name, or some other information in addition to the base service connection information provided by the existing properties. USGIN recommended practice for this common situation is to include a parameters section in the linkage/description (see 4.16.3 CI_OnlineResource name and description, above).

4.17 Responsible parties and logos

Metadata should include a URL that locates a thumbnail logo for organizations related to the metadata origination, the organization hosting the catalog that returned the metadata, the organization that originated the data, and the organization hosting online services that provide access to the data. The standard place to put URL's in ISO19139 metadata is in the CI_Contact/onlineResource/CI_OnlineResource/linkage attribute. For URL's that indicate icon thumbnails, the CI_OnlineResource/name should be 'icon'.

The metadata originator information should be in a MD_Metadata/contact/CI_ResponsibleParty element with role code 'originator' to identify the original source of the metadata record, for which the CI_Contact/../CI_OnlineResource/linkage is a URL that points to an Icon for the metadata originator. This Icon will be displayed in search results to credit the metadata originator. Metadata harvesters should harvest and maintain this information so that the origin of metadata records can be credited.

The organization hosting the catalog that returned the metadata record should specified in a MD_Metadata/contact/CI_ResponsibleParty element with role code 'distributor', for which the CI_Contact/../CI_OnlineResource/linkage is a URL that points to an Icon for the metadata server hosting organization. This information need not be harvested, because it will be replaced by information describing the harvesting catalog service.

The organization that originated the data is specified by MD_Metadata/identificationInfo/MD_Data-

Identification/citation/../CI_ResponsibleParty with RoleCode ='originator', and /CI_OnlineResource/name='icon'. This will distinguish the citation responsible party element containing the icon linkage from CI_ResponsibleParty elements with RoleCode='author' or 'editor', which would provide an online linkage directly to the responsible party as specified by CI_OnlineResource protocol, applicationProfile, name, function, and description elements.

The organization hosting a service providing online access to described data is specified by MD_Metadata/distributionInfo/MD_Distribution/distributor/MD_Distributor/distributorContact/-CI_ResponsibleParty with RoleCode ='resourceProvider' or 'distributor', and ../CI_Online-Resource/name='icon'. Because the cardinality of distributorContact responsible party and online resources is 1, only one linkage can be provided for a distributor, and the metadata author must decide whether that will be a link to an icon, or a link to a web site or other resource related to the distributor.

```
<gmd:contact>
  <gmd:CI_ResponsibleParty>
    <gmd:organisationName>
      <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
    </gmd:organisationName>
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:onlineResource>
          <gmd:CI_OnlineResource>
            <gmd:linkage>
              <gmd:URL>http://www.azgs.az.gov/logo/metadata/azgs.png</gmd:URL>
            </gmd:linkage>
            <gmd:name>
              <gco:CharacterString>icon</gco:CharacterString>
            </gmd:name>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
      </gmd:CI_Contact>
    </gmd:contactInfo>
    <gmd:role>
      <gmd:CI_RoleCode
     codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_RoleCod
     e"
        codeListValue="originator">originator</gmd:CI_RoleCode>
    </gmd:role>
  </gmd:CI_ResponsibleParty>
</gmd:contact>
```

4.18 Extensions to CharacterString

4.18.1 Web extensions

ISO 19139 defines several extensions to gco:CharacterString in the gmx namespace. These are defined as members of an xml substitution group for gco:CharacterString, and include gmx:Anchor, gmx:FileName, and gmx:MimeFileType. gmx:Anchor is used for URL's linking to online web resources, and include a URI attribute associated with the character string that is the human-readable label for the link. gmx:FileName adds a filename URI attribute that specifies a

machine-readable absolute path to the location of the file, the human readable file name specified by the character string. `gmx:MimeFileType` adds a MIME type/subtype attribute to a character string that specifies a human readable file type. The gmx namespace is not imported into other ISO19139 schema in the normative schema. In order to create schema-valid documents that use these extensions, explicit namespace-declaration must be made to the gmx schema in instance documents. For the USGIN profile, these extension classes are not recommended, and may be ignored by client software.

4.18.2 Language localization

ISO 19139 defines an extension to `gco:CharacterString` that allows substitution by `PT_FreeText` or `LocalisedCharacterString`. `LocalisedCharacterString` adds a `locale/PT_Locale` property to the `CharacterSTring` element that can specify the language, country, and character encoding for the string. `PT_FreeText` allow substitution of a collection of `LocalisedCharacterString` elements for any `CharacterString`, each localized to a different language/country.

These various possibilities create potential to break interoperability. To avoid this problem, and to avoid complexity with `PT_Text` and `LocalizedCharacterString`, USGIN profile practice is to implement services for different languages as different catalog service instances, each of which serves CharacterStrings in a single language specified by the `MD_Metadata/language` element.

4.18.3 Codelists

ISO 19139 defines a "CodeListValue_Type" XML Class Type with three attributes:

`codeList`: contains a URL that references a codeList definition within a registry or a codelist catalogue. As currently used in the metadata services we have studied, the codeList is not used to identify a vocabulary; rather it provides a locator (functionally equivalent to xlink:href) for an online resource, typically a web page or xml file, that contains a listing of the codelist with the code values and scope notes. Different services provide different URL's, possibly linking to different kinds of resources (e.g. web page or xml file), for the same codelist. Thus, in practice, the values in this attribute can not be used for automated determination of the code list in use in a metadata document.

`codeListValue`: carries the value from the 'name' column of the codelist tables in ISO 19115, Annex B. This value serves to identify the codelist meaning in the context of the codelist catalogue (or registry) located by the `codeList` attribute. The codelist catalogue is expected to contain an explicit name and definition of the value in the default language of the metadata, as well as alternate expressions in different code spaces, some of them corresponding to the different locales supported by the metadata.

The `codeSpace` attribute is an optional identifier (URI); when present it refers to an alternative expression of the codelist value definition. In the example in ISO19139, section 8.5.5.1 (p. 30), the codeSpace URI for the domain code is the string "domainCode", and the value from the domainCode column in a codelist definition table in ISO 19115, Annex B is included as the value of the xml CodeList element in this case. This appears to be an attempt to allow alternate codeListValue strings to be associated with the codelist value concept for language localization or use of alternate identifiers.

Codelist elements in the ISO19139 XML schema are assigned to type CodeListValue_Type, and also included in a substitution group for `gco:CharacterString`. These codeList elements are thus substitutable for elements typed `gco:CharacterString`. Consequently, any CodeList instance

is an XML element that takes a string value (corresponding to the element value for a gco:CharacterString), and has three XML attributes defined by the CodeListValue_Type XML Class Type. A corresponding XML Class Property Type is defined for each of these CodeList elements, and this property type is used to restrict the values in XML `CharacterString` attributes to the code list.

The ISO specification uses an unfortunate choice of name for the 'codeListValue' attribute that is defined to be a identifier, apparently with the intention that it is a language-neutral concept identifier that might be associated with various language-localized labels for the concept. NAP CodeList registries (http://www.fgdc.gov/nap/metadata/register) contrast with the codelists defined in the tables in ISO 19115 Annex B in that the identifier (the 'name' column the ISO19115 Annex B tables) is an integer identifier with the prefix 'RI_'. This would appear to correspond functionally to the 'domainCode' values in the ISO19115 Annex B tables, which ISO19139 indicates should be the codeListValue when the `codeSpace`="domainCode".
NAP and INSPIRE usage is consistent with the ISO19139 definition of codeListValue as an identifier, with the name or label for the codeList concept included as the value of the CodeListValue attribute. The 'name' column in ISO 19115, Annex B tables, which is described as the content for the codeListValue by ISO19139, contains English words that are the same as the labels one would use in English. In the CT_CodeListCatalogues in the ISO publicly available standards registry for ISO 19139 (http://standards.iso.org/ittf/PubliclyAvailableStandards/-ISO_19139_Schemas/resources/), the CodeListDictionary/codeEntry/CodeDefinition elements only include gml:description and gml:identifier elements, but no gml:name elements. Based on this ISO guidance, USGIN profile mandates that controlled vocabulary terms from codelists are encoded thus:

Example 1:

```
<gmd:CI_DateTypeCode codeList=" http://standards.iso.org/ittf/PubliclyAvailableStandards/-
    ISO_19139_Schemas/resources/#CI_DateTypeCode"
    codeListValue="creation"/>
```

Example 2

```
<gmd:MD_CharacterSetCode
    codeList="
    http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_DateTypeCode"
    codeListValue="utf8"/>
```

Or

```
<gmd:MD_CharacterSetCode
    codeList="
    http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_DateTypeCode"
    codeListValue="utf8">UTF-8</gmd:MD_CharacterSetCode>
```

The alternate version of example 2 includes label string for the codelist concept as the element value string. Most applications appear to use only the codeListValue string for querying and content inspection. In order to indicate that standard ISO codelists are being used, the `codeList` attribute value MUST be " http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml ", which is the web location for the authoritative codelist dictionary XML document. Including the fragment identifier name for the particular codelist is informative only; this is the terminal token of the URL after the '#' character, and will be ignored by browsers.

The use of other codelists is implemented by providing a `codeList` attribute that points to the codelist actually used. The following example shows use of a DateTypeCode from the code list

in the North American Profile:

```
<gmd:CI_DateTypeCode
        codeList="http://www.fgdc.gov/nap/metadata/register/registerItemClasses.html#IC_87"
        codeListValue="superseded">superseded</gmd:CI_DateTypeCode>
```

Note that the ISO codelists use the name value for the `codeListValue` attribute. This is consistent with ISO codelist practice but is in contrast to the NAP codeList example encoding in Appendix E of NAP profile document (INCITS 453, 2009). In these examples the codeListValue is the 'Item identifier' from the NAP registry specified by the codeList, with the prefix 'RI_' added.

The ISO codelists are in much wider use at this time than the NAP codelists (as far as we can tell from surveying existing services), but we recognize that some of the terms added in the NAP codelists may be required for metadata describing some of the resources in the USGIN scope (Table 1). Table 8 summarizes differences between the ISO and NAP codelists. The recommended practice is to use ISO codelists wherever possible, encoded as in the examples above. NAP codes may be used where necessary, with the appropriate codeList URL: " http://www.fgdc.gov/nap/metadata/register/registerItemClasses.html" and optional codelist-specific document fragment token (e.g. #IC_87)
If a new codelist is created to restrict text in an ISO element whose type is simply CharacterString (e.g. HierarchyLevelName), then `characterString` element values are encoded by soft-typing the element that takes the character string using the xsi:type attribute. The following example uses the FileFormatCodeList, which is the only code list vocabulary added to the collection of codelists defined by ISO 19115 by the North American Profile.

```
<gmd:fileType xsi:type="gml:CodeType"
        codeList="http://www.fgdc.gov/nap/metadata/register/registerItemClasses.html#IC_115"
        codeListValue="jpg">
  <gco:CharacterString>jpg</gco:CharacterString>
</gmd:fileType>
```

As a convention for using controlled vocabularies on `gco:characterString` elements without the overhead of a defining a new namespace and xml schema, USGIN proposes that use of a controlled vocabulary be indicated by adding the attribute `xsi:type="gml:CodeType"` on the `gco:CharacterString` element. The type gml:CodeType requires a codeList attribute (see 4.14.2 Non digital resources and 7.2 Linkage name conventions), which MUST be the URI for the vocabulary used, with the implication that the `gco:CharacterString` element value and the codeListValue attribute will then be an identifier from that vocabulary. This essentially turns the CharacterString into a GML scoped name or gco:LocalName element.
Table 9. Codelist crosswalk between ISO, NAP and USGIN.

Codelist (ISO / NAP)	Coded Values/Names	Comments
`CI_DateTypeCode`	`creation, publication, revision`	ISO 19115 (B.5.2)
`napCI_DateTypeCode`	`…, notAvailable, inForce, adopted, deprecated, superseded`	NAP expansion
`CI_OnLineFunction-Code`	`download, information, offlineAccess, order, search`	ISO 19115 (B.5.3)
`napCI_OnLineFunction-Code`	`…, upload, webService, emailService, browsing, fileAccess, webMapService`	NAP expansion
`CI_PresentationForm-Code` `napCI_PresentationFo`	`documentDigital, documentHardcopy, imageDigital, imageHardcopy, mapDigital, mapHardcopy, modelDigital, modelHardcopy, profileDigital,`	ISO 19115 (B.5.4)

rm-Code	profileHardcopy, tableDigital, tableHardcopy, videoDigital, videoHardcopy, audioDigital	
	…, audioHardcopy, multimediaDigital, multimediaHardcopy, diagramDigital, diagramHardcopy	NAP expansion
CI_RoleCode napCI_RoleCode	resourceProvider, custodian, owner, user, distributor, originator, pointOfContact, principalInvestigator, processor, publisher, author	ISO 19115 (B.5.5)
	…, collaborator, editor, mediator, rightsHolder	NAP expansion
DQ_EvaluationMethod-TypeCode napDQ_Evaluation-MethodTypeCode	directInternal, directExternal, indirect	ISO 19115 (B.5.6)
DS_AssociationType-Code napDS_Association-TypeCode	crossReference, largerWorkCitation, partOfSeamlessDatabase, source, stereoMate	ISO 19115 (B.5.7)
	…, isComposedOf	NAP expansion
DS_InitiativeType-Code napDS_Initiative-TypeCode	campaign, collection, exercise, experiment, investigation, mission, sensor, operation, platform, process, program, project, study, task, trial	ISO 19115 (B.5.8)
MD_CellGeometryCode	point, area	ISO 19115 (B.5.9)
napMD_CellGeometry-Code	…, voxel	NAP expansion
MD_CharacterSetCode napMD_CharacterSet-Code	ucs2, ucs4, utf7, utf8, utf16, 8859part1, 8859part2, 8859part3, 8859part4, 8859part5, 8859part6, 8859part7, 8859part8, 8859part9, 8859part10, 8859part11, 8859part13, 8859part14, 8859part15, 8859part16, jis, shiftJIS, eucJP, usAscii, ebcdic, eucKR, big5, GB2312	ISO 19115 (B.5.10)
MD_Classification-Code	unclassified, restricted, confidential, secret, topSecret	ISO 19115 (B.5.11)
napMD_ClassificationCode	…, sensitive, forOfficialUseOnly	NAP expansion
MD_CoverageContent-TypeCode napMD_Coverage-ContentTypeCode	image, thematicClassification, physicalMeasurement	ISO 19115 (B.5.12)
MD_DataTypeCode not used by NAP	class, codelist, enumeration, codelistElement, abstractClass,	ISO 19115 (B.5.13) – The

and USGIN	aggregateClass, specifiedClass, datatypeClass, interfaceClass, unionClass, metaClass, typeClass, characterString, integer, association	MD_MetadataExtension Information element and its codelists are not used by NAP and USGIN.
MD_DimensionName-TypeCode napMD_DimensionName-TypeCode	row, column, vertical, track, crossTrack, line, sample, time	ISO 19115 (B.5.14)
MD_GeometricObject-TypeCode napMD_Geometric-ObjectTypeCode	complex, composite, curve, point, solid, surface	ISO 19115 (B.5.15)
MD_ImagingCondition-Code napMD_Imaging-ConditionCode	blurredImage, cloud, degradingObliquitty, fog, heavySmokeOrDust, night, rain, semiDarkness, shadow, snow, terrainMasking	ISO 19115 (B.5.16)
MD_KeywordTypeCode napMD_KeywordType-Code	discipline, place, stratum, temporal, theme	ISO 19115 (B.5.17)
	…, product, subTopicCategory	NAP expansion
MD_Maintenance-FrequencyCode napMD_Maintenance-FrequencyCode	continual, daily, weekly, fortnightly, monthly, quarterly, biannually, annually, asNeeded, irregular, notPlanned, unknown	ISO 19115 (B.5.18)
	…, semimonthly	NAP expansion
MD_MediumFormatCode napMD_MediumFormat-Code	cpio, tar, highSierra, iso9660, iso9660RockRidge, iso9660AppleHFS	ISO 19115 (B.5.19)
	…, UDF	NAP expansion
MD_MediumNameCode napMD_MediumNameCode	cdRom, dvd, dvdRom, 3halfinchFloppy, 5quarterInchFloppy, 7trackTape, 9trackTape, 3480Cartridge, 3490Cartridge, 3580Cartridge, 4mmCartridgeTape, 8mmCartridgeTape, digitalLinearTape, onLine, satellite, telephoneLink, hardcopy, hardcopyDiazoPolyester08, hardcopyCardMicrofilm, hardcopyMicrofilm240, hardcopyMicrofilm35, hardcopyMicrofilm70, hardcopyMicrofilmGeneral, hardcopyMicrofilmMicrofiche, hardcopyNegativePhoto, hardcopyPaper	ISO 19115 (B.5.20)
	…, hardcopyDiazo, hardcopyPhoto, hardcopyTracedPaper, hardDisk, USBFlashDrive,	NAP expansion

	1quarterInchCartridgeTape	
MD_ObligationCode not used by NAP and USGIN	mandatory, optional, conditional	ISO 19115 (B.5.21) - The MD_MetadataExtensionInformation element and its codelists are not used by NAP and USGIN.
MD_PixelOrientation-Code napMD_Pixel-OrientationCode	center, lowerLeft, lowerRight, upperRight, upperLeft	ISO 19115 (B.5.22)
MD_ProgressCode napMD_ProgressCode	completed, historicalArchive, obsolete, onGoing, planned, required, underDevelopment	ISO 19115 (B.5.23)
	…, proposed	NAP expansion
MD_RestrictionCode napMD_Restriction-Code	copyright, patent, patentPending, trademark, license, intellectualPropertyRights, restricted, otherRestrictions	ISO 19115 (B.5.24)
	…, licenseUnrestricted, licenseEndUser, licenseDistributor, privacy, statutory, confidential, sensitivity	NAP expansion
MD_ScopeCode napMD_ScopeCode	attribute, attributeType, collectionHardware, collectionSession, dataset, series, nonGeographicDataset, dimensionGroup, feature, featureType, propertyType, fieldSession, software, service, model, tile	ISO 19115 (B.5.25)
MD_Spatial-RepresentationTypeCode napMD_Spatial-RepresentationTypeCode	vector, grid, textTable, tin, stereoModel, video	ISO 19115 (B.5.26)
MD_TopicCategoryCode napMD_TopicCategory-Code	farming, biota, boundaries, climatologyMeterologyAtmosphere, economy, elevation, environment, geoscientificInformation, health, imageryBaseMapsEarthCover, intelligenceMilitary, inlandWater, location, oceans, planningCadastre, society, structure, transportation, utilitiesCommunication	ISO 19115 (B.5.27)
MD_TopologyLevelCode napMD_TopologyLevel-Code	geometryOnly, topology1D, planarGraph, fullPlanarGraph, surfaceGraph, fullSurfaceGraph, topology3D, fullTopology3D, abstract	ISO 19115 (B.5.28)

SV_CouplingType napSV_CouplingType	loose, mixed, tight	ISO 19119 (Amendment 1; C.2.8)
SV_Parameter-Direction napSV_Parameter-Direction	in, out, in/out	ISO 19119 (Amendment 1; C.2.9
LanguageCode	see http://www.loc.gov/standards/iso639-2/php/code_list.php	no complete NAP or ISO registry found
not used by ISO nap_DCPList	XML, CORBA, JAVA, COM, SQL, WebServices	NAP specific codelist – not used by USGIN due to poorly defined semantics and use.
not used by ISO napMD_FileFormatCode	bil, bmp, bsq, bzip2, cdr, cgm, cover, csv, dbf, dgn, doc, dwg, dxf, e00, ecw, eps, ers, gdb, geotiff, gif, gml, grid, gzip, html, jpg, mdb, mif, pbm, pdf, png, ps, rtf, sdc, shp, sid, svg, tab, tar, tiff, txt, xhtml, xls, xml, xwd, zip, wpd	NAP specific codelist – not formally used by USGIN, but these character strings should are to be used to populate fileType elements.

4.19 Geographic bounding box

USGIN profile requires that if an `EX_Extent/geographicElement` is supplied, at least one geographic bounding box with bounding latitude and longitude expressed using WGS 84 decimal degrees MUST be provided.

The corner coordinates for the geographic bounding box must not coincide in one point, because this may result in fatal errors with some CSW implementations. Point locations must thus be represented as tiny rectangles. USGIN recommended practice is to place the actual point location in the lower left corner of the rectangle.

4.20 Data Quality

4.20.1 Simple quality statement

For resource evaluation purposes, to answer a user's question 'is the described resource good enough for my purposes?', a free text statement is generally sufficient in place of a detailed xml document. For example, this might be a statement something like "*These data were compiled by the authors from field sheets and notes by scanning paper copies, georeferencing and digitizing on screen. Station locations are based on Garmin 12 GPS readings, except locally where they have been adjusted for consistency with the base map. Original GPS coordinates are reported in the station table. These data have been reviewed for completeness of description and internal consistency, but have not been independently field checked.*" This is the

only kind of quality information available for many resources; it is neither a quantitative measure nor technically a conformance result. A simple qualityStatement would suffice and provide significant value. There is no ISO19115 element for such a quality statement.

To implement this simple case, the proposed solution is that the first report conformance explanation in each DQ_DataQuality element instance would contain a free text discussion of data quality considerations for the indicated scope for that element. Thus, text at the XPath `DQ_DataQuality//report[1]//result[1]//explanation/CharacterString` will be interpreted as a self-explanatory quality statement. The use of any specific data quality element (e.g. DQ_QuantitativeAttributeAccuracy, DQ_RelativeInternalPositionalAccuracy..) to contain this explanation is arbitrary and may be ignored by client applications.

4.21 Lineage

Lineage in data quality section is a collection of processing steps that describe the provenance of a resource. Each step has some input resources, identified by source citations associated with the process step. The `LI_ProcessStep` element does not directly identify its output resource, so in a lineage that involves a chain of steps with intermediate resources, the sourceStep association from `LI_Source` links a resource to a processing step that it is output from.

A `LI_Lineage` element that contains a `source` but no `processStep` elements is redundant; this source information should be represented with the `MD_DataIdentification/citation/CI_Citation` element.

A GIS dataset originally digitized from a published geologic map, put online, obtained by an online download, and reprojected would report one `processStep` (reprojection) with `source/LI_Source` that has a `CI_Citation` for the downloaded data. This `LI_Source` would have a `sourceStep` pointing to an `LI_ProcessStep` for the original digital conversion from the paper map, and the `LI_ProcessStep/source/LI_Source` would contain the citation for the original paper map.

In order to enable XPath queries for any of the sources or processSteps in a processing chain, all related `LI_Source` and `LI_ProcessStep` elements should be directly nested within the `LI_Lineage` element, and the `processStep/source` and `LI_Source/sourceStep` associations should be by internal document references (see Code Example 3).

Code Example 3: A complex processing and source history using LI_Lineage.

```
<?xml version="1.0" encoding="UTF-8"?>
<LI_Lineage
  xmlns="http://www.isotc211.org/2005/gmd"
  xmlns:gco="http://www.isotc211.org/2005/gco"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.isotc211.org/2005/gmd
 http://schemas.opengis.net/iso/19139/20070417/gmd/dataQuality.xsd">
  <statement>
    <CharacterString>The digital data described by this metadata was originally compiled
        digitally from two published maps; this digital dataset was then reprojected to produce
        the described resource.</CharacterString>
 </statement>
 <processStep>
```

```
      <LI_ProcessStep id="1">
        <description>
          <CharacterString>digital compilation of 2 Maps</CharacterString>
        </description>
        <source xlink:href="#10"/>
        <source xlink:href="#20"/>
      </LI_ProcessStep>
    </processStep>
    <processStep>
      <LI_ProcessStep id="2">
        <description>
          <CharacterString>digital map compilation reprojected, should have some way to specify
            projection parameters?, output is LI_Source id=70 </CharacterString>
        </description>
        <source xlink:href="#40"/>
      </LI_ProcessStep>
    </processStep>
    <source>
      <LI_Source id="40">
        <description>
          <CharacterString>a digital compilation of 2 maps, output of processStep ID=1, input into
            reprojection process</CharacterString>
        </description>
        <sourceStep xlink:href="1"/>
      </LI_Source>
    </source>
    <source>
      <LI_Source id="10">
        <description>
          <LocalisedCharacterString>ultimate
            source--somepublishedmap</LocalisedCharacterString>
        </description>
  <!--no source processing recorded for production of paper map so no sourceStep-->
      </LI_Source>
    </source>
    <source>
      <LI_Source id="20">
        <description>
          <LocalisedCharacterString>another published map</LocalisedCharacterString>
        </description>
      </LI_Source>
    </source>
    <source>
      <LI_Source id="70">
        <description>
          <LocalisedCharacterString>a reprojected version of the digital
              compilation</LocalisedCharacterString>
        </description>
        <sourceStep xlink:href="2"/>
      </LI_Source>
    </source>
```

```
</LI_Lineage>
```

An `LI_Lineage` may be constructed that involves a number of resources and processing steps, and this lineage may be referenced by metadata for all the resources involved in the processing. The `LI_Lineage/source/LI_Source/sourceCitation/CI_Citation/identifier/MD_Identifier` is a reference to the `MD_Metadata/fileIdentifier` for the metadata for each resource in the chain. This approach allows the metadata record to record relationships through process steps between resources.

4.22 Temporal extents

Resource temporal extent (`identificationInfo/MD_DataIdentification/extent/EX_Extent/-temporalElement/EX_TemporalExtent/extent/TimePeriod`) is used to specify the temporal interval to which the content of a resource applies. Default reference frame for time is calendar date and time encoded using ISO-8601:

```
<gml:TimePeriod gml:id="Id2010">
    <!-- Note that gml:TimePeriod requires a gml:id string that must be unique
      in the scope of the instance document, and must have an alphabet
      character ('a-z', 'A-Z') as its first character -->
  <!-- USGIN requires the beginPosition and endPosition's frame property to be defined.
  The default value is #ISO-8601. -->
    <gml:beginPosition frame="#ISO-8601">2010-01-00T00:00:00Z </gml:beginPosition>
    <gml:endPosition frame="#ISO-8601">2010-12-31T24:00:00Z</gml:endPosition>
</gml:TimePeriod>
```

<gml:endPosition indeterminatePosition="now"/> is the correct way to represent "Present" in ISO or GML as one of the boundaries of a timePeriod.

The ISO 19139 xml schema allows temporal location to be quantified by a gml:TimeInstant or gml:TimePeriod element. In order to promote interoperability, the USGIN profile mandates use of gml:TimePeriod for specifying temporal extent for a resource.

For geologic time extents, the time coordinates for the beginPosition and endPosition should be expressed numerically in Ma. This convention allows search for resources pertinent to intervals of geologic time using simple numeric comparisons instead of the complex hierarchical concept expansions that would be necessary to use named eras from a stratigraphic time scale. Encoding example:

```
<EX_TemporalExtent>
  <extent>
    <gml:TimePeriod gml:id="y34096">
      <gml:beginPosition
        frame="urn:CGI:TemporalCRS:cgi:standardGeologyMa">220</gml:beginPosition>
      <gml:endPosition
     frame="urn:CGI:TemporalCRS:cgi:standardGeologyMa">140 </gml:endPosition>
    </gml:TimePeriod>
  </extent>
</EX_TemporalExtent>
```

The frame for the beginPosition and endPosition is a URI for standard geologic time, measured positive getting older, with an origin at 1950 CE, in units of millions of years.

4.23 Operation metadata

See the discussion on the use of service metadata in the USGIN profile above, preceding Table 3. The srv namespace elements based on ISO 19119 provide insufficient information and guidelines to access data through services without apriori knowledge of the service protocol. This is due in part to weakly defined semantics and use cases for the elements that are there (DCP, applicationProfile, protocol, MD_Format, serviceType, operationName vs. invocationName, connectPoint), and partly due to incomplete content model (where to put allowed outputFormat parameter values or supported query operations for CSW or WMS, what is syntax for delimiting parameters, are http request headers used). The ISO 19119 model for service metadata does not include a mechanism to specify valid values for operation parameters. For instance, OGC WMS and CSW services both support an output format parameter, and OGC capabilities documents provide a listing of the supported output formats, but where do these go in ISO19139 xml documents? Does the described service support http POST or GET method? This information is necessary in order to compose valid service requests.

Based on these considerations, the USGIN profile mandates that a service metadata record MUST include at least one service operation with the `operationDescription/CharacterString` = 'serviceDescription'. The `identificationInfo/SV_ServiceIdentification/containsOperations/-SV_OperationMetadata/connectpoint` link for this operation MUST retrieve the service-specific self description document (e.g. WSDL, GetCapabilities, WADL). The `CI_OnlineResource` in this connectPoint elment must have `CI_OnlineResource/name` = "serviceDescription" (from the table in section 7.2 Linkage name conventions).

WSDL and getCapabilities were designed to describe service operation, and it seems counterproductive to invent another scheme to do the same thing. Service metadata for processing services that are not coupled to particular input datasets should provide `srv:SV_OperationMetadata` to describe the operations offered by the service to inform human users. Because of the difficulty in creating usable abstract model that accounts for any and all possible services, it makes more sense to allow service description documents specific to different service frameworks to describe all the details in a machine-actionable format.

5 Abbreviations

CSW	Metadata Catalog for the Web. Also abbreviated as CS-W and CS/W
GeoSciML	Geoscience Markup Language
GML	Geographic Markup Language
GUID	Global Unique Identifier
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
UML	Unified Modeling Language
URI	Universal Resource Identifier
USGIN	U.S. Geoscience Information Network
WCS	Web coverage Service
WFS	Web Feature Service
XML	eXtensible Markup Language
XSD	XML Schema Definition
XSL	eXtensible Stylesheet Language
XSLT	XSL Transformations
XLink	XML Linking Language

6 References

6.1 Cited literature

[ANZLIC, 2007] ANZLIC Metadata Profile Guidelines, Version 1.0: Turner, ACT, ANZLIC - the Spatial Information Council, ISBN: 978-0-646-46940-9, 372 p., available at http://www.osdm.gov.au/Metadata/ANZLIC+metadata+resources/default.aspx (accessed 2010-11-18).

[Dublin Core] 2008-01-14 Dublin core Metadata Element Set, Version 1.1: Dublin Core Metadata Initiave, accessed at http://dublincore.org/documents/dces/.

[INSPIRE ISO19115/119] Drafting Team Metadata and European Commision Joint Research Centre, 2009-02-18, INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119,v. 1.1: European Commission Joint Research Centre, MD_IR_and_ISO_20090218, available at http://inspire.jrc.ec.europa.eu/index.cfm/pageid/101 (accessed 2010-11-18).

Franklin, Michael, Halevy, Alon, and Maier, David, 2005, From databases to dataspaces: a new abstraction for information management: ACM SIGMOD Record, V. 34, No. 4, ISSN:0163-5808.

[FRBR] Tillett, Barbara, 2004-03-02, What is FRBR?: Washington, DC, Library of Congress Cataloging Distribution Service, 8 pages. (originally published in Technicalities v. 25, no. 5, 2003), available at http://www.loc.gov/cds/downloads/FRBR.PDF (accessed 2013-06-12).

Heaney, Michael, 2000-01-14, An analytical model of Collections and their Catalogues: UK Office for Library and Information Networking, Third Issue, available at http://www.ukoln.ac.uk/metadata/-rslp/model/amcc-v31.pdf (accessed 2013-03-18).

Richard, S. M., and Grunberg, Wolfgang, editors, 2010, Metadata Recommendataions for Geoscience Resources: U.S. Geoscience Information Network Best Practices Document, Doc ID gin2010-11, v. 1.0.3, available at http://lab.usgin.org/profiles/doc/metadata-content-recommendations (accessed 2010-11-18).

7 Codelists

7.1 ServiceType

INSPIRE metadata Implementing Rules (*OJ L 326, 4.12.2008)* section D3 mandate the use of the value domain listed in Table 11 to categorize spatial data service types. These values are better suited for CI_OnlineFunctionCode used to specify CI_OnlineResource/online/Function. The USGIN team interprets the ISO scope notes to allow more useful content for service type, specifying an actual service specification like OGC WMS. USGIN draft ServiceType vocabulary is reported in Table 12.

Table 11. INSPIRE SPATIAL DATA SERVICE TYPE (for information only, not used by USGIN)

Type	Description
discovery	Discovery Service
view	View Service
download	Download Service
transformation	Transformation Service
invoke	Invoke Spatial Data Service
other	Other Services

Table 12. USGIN service type vocabulary. This is an interim listing of serviceTypes. The code list URI for this registry is http://resources.usgin.org/registry/serviceType201001. See also http://dclite4g.xmlns.com/ws.rdf for another rdf-based vocabulary. These should probably be merged.[NOTE 2014-03-03. Actual USGIN metadata is not using these precise identifiers. An activity is underway at https://github.com/OSGeo/Cat-Interop to develop a community vocabulary of service types.

Identifier	Name	Description
WMS	OGC Web Map service	provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases. A WMS request defines the geographic layer(s) and area of interest to be processed. The response to the request is one or more geo-registered map images (returned as JPEG, PNG, etc) that can be displayed in a browser application. The interface also supports the ability to specify whether the returned images should be transparent so that layers from multiple servers can be combined or not. (http://www.opengeospatial.org/standards/wms)
WFS	OGC Web Feature service	WFS offers direct fine-grained access to geographic information at the feature and feature property level; allows clients to retrieve or modify specific data items. That data can then be used for a wide variety of purposes, including purposes other than their producers' intended ones. http://www.opengeospatial.org/standards/wfs
WCS	OGC Web coverage	interface and operations that enable interoperable access to geospatial "coverages" [http://www.opengeospatial.org/ogc/-

	service	glossary/c]. The term "grid coverages" typically refers to content such as satellite images, digital aerial photos, digital elevation data, and other phenomena represented by values at each measurement point.
CSW	OGC Web catalog service	publish and search collections of descriptive information (metadata) about geospatial data, services and related resources. Providers of resources use catalogues to register metadata that conform to the provider's choice of an information model; such models include descriptions of spatial references and thematic information. (http://www.opengeospatial.org/standards/cat)
SOS	OGC Sensor observation service	manage deployed sensors and retrieve sensor data, specifically "observation" data. E.g. in-situ sensors (e.g., water monitoring) or dynamic sensors (e.g., satellite imaging). (http://www.opengeospatial.org/standards/sos)
WPS	OGC Web Processing service	rules for standardizing how inputs and outputs (requests and responses) for geospatial processing services, such as polygon overlay, are exchanged on the web to link services; defines how a client can request the execution of a process, and how the output from the process is handled, and an interface to register geospatial processes and for client to discovery and bind to those processes. (http://www.opengeospatial.org/standards/wps)
SPS	OGC Sensor planning service	interfaces for exchanging information about the capabilities of a sensor and how to task the sensor; designed to support queries to determine the feasibility of a sensor planning request, to submit such a request, to inquire about the status of such a request, to update or cancel such a request, and to request information about other OGC Web services that provide access to the data collected by the requested task.
OpenDAP	Open source data access protocol	(http://opendap.org/) a protocol for requesting and transporting data across the web. DAP 2.0 uses HTTP to frame the requests and responses. For details on DAP, see Data Access Protocol (DAP), version 2 which is a technical description of the Data Access Protocol.
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting	provides an application-independent interoperability framework based on metadata harvesting.

Example usage:

```
<srv:serviceType>
  <gco:LocalName
codeSpace="http://resources.usgin.org/registry/serviceType201001">WMS</gco:Local
Name>
</srv:serviceType>
```

7.2 Linkage name conventions

The cardinality of the `online` element in DigitalTransferOptions is 0..*. In order to distinguish the nature of various linkages that might be provided, above and beyond function, protocol, and

applicationProfile, USGIN profile mandates use of the following names to associate with links to identify important linkages.

Table 13. USGIN Names to identify special linkage URL's for CI_Online Resource. CodeList URI = http://resources.usgin.org/registry/linkageName201001

Identifier	Name (eng)	Usage
icon	Icon	linkage url is link to a thumbnail icon. Icon pixel height and width range?
serviceDescription	web service description	linkage url is link to getCapabilities or WSDL that describes a service using a formal syntax such that computer programs can automate connection to the service. OnlineFunctionCode for CI_OnlineResource should be 'Information'
baseURL	web service endpoint	Base url for service. Assumes that ServiceType specifies a well know service type such that requests can be constructed without significant additional information. OnlineFunctionCode for CI_Online-Resource should be 'webService'
serviceClient	web application	URL is linkage to a web application that allows the user to access the service. OnlineFunctionCode for CI_OnlineResource should be 'Browsing'
webpage	access information	URL locates a web page with instructions for accessing the service. This provides the user with information to implement a connection to the service, but does not enable automated service access. OnlineFunctionCode for CI_Online-Resource should be 'Information'
~~serviceMetadata~~	~~Service Metadata~~	~~linkage URL is link to a complete service metadata record. Use for distribution linkage to a service for which more information that can be provided by the CI_OnlineResource element is necessary to successfully automate access to the context resource through the service. OnlineFunctionCode for CI_Online-Resource should be 'Information'~~
download		

email request		
offline access		

Example usage:

```
<gmd:CI_OnlineResource>
  <gmd:linkage>
    <gmd:URL>http://75.101.143.247:8080/gsvr/wms?SERVICE=WMS&amp;REQUEST=getCa
      pabilities</gmd:URL>
  </gmd:linkage>
  <gmd:protocol>
    <gco:CharacterString>OGC:WMS</gco:CharacterString>
  </gmd:protocol>
  <gmd:name>
    <gco:CharacterString xsi:type=”gml:CodeType”
      codeSpace="http://resources.usgin.org/registry/linkageName201001">
      serviceDescription</gco:CharacterString>
  </gmd:name>
</gmd:CI_OnlineResource>
```

Use of such controlled vocabulary can be indicated by using xsi:type on the gco:characterString element to make the type gml:CodeType, which then requires a codeSpace attribute. The distribution format Identifier from Table 6 should be used as the element value. For compatibility with systems that can not process this encoding, the code identifier should be included as the element value as well as the codeListValue.

8 Examples

8.1 USGIN ISO 19139 Minimum Dataset Metadata

In the following listing, text in green is comments; XML elements are in blue, XML attributes are in black, and attribute values are in purple.

This example metadata record describes itself.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--*********************************************************************************************
*** Minimum example of a ISO 19139 Geospatial Dataset Metadata
*** based on the USGIN v1.2 Profile
*** by USGIN Standards and Protocols Drafting Team
*** U.S. Geoscience Information System (USGIN) - http://lab.usgin.org
*** Contributors: Wolfgang Grunberg, Stephen M Richard
*** 2013-06-13
***
*** DISCLAIMER: this is not an authoritative metadata example but an aide to get started.
*** Refer to the USGIN profile document for more specific and complete guidelines.
***
*** Validated against http://www.isotc211.org/2005/gmd (ISO 19115, CSW 2.0.2 AP ISO 1.0).
*** Follows the USGIN ISO 19139 Dataset Metadata Profile v1.2.
***
*** NOTES:
*** - Codelists:
*** Most ISO metadata profiles and applications use ISO codelists or codelists that use ISO's
                codelist names.
***
*** - Language code:
*** Use ISO <ISO639-2/T three letter language code - lower case> formatting.
***
*** KEY: (NAP-USGIN) - M/C/O/X (Mandatory, Conditional, Optional, Not Used)
***
*********************************************************************************************-->
        <!-- USGIN ISO 19139 geospatial dataset metadata record -->
<gmd:MD_Metadata
                    xmlns:gmd="http://www.isotc211.org/2005/gmd"
                     xmlns:gco="http://www.isotc211.org/2005/gco"
                     xmlns:gml="http://www.opengis.net/gml"
                     xmlns:xlink="http://www.w3.org/1999/xlink"
                     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                     xsi:schemaLocation="http://www.isotc211.org/2005/gmd
            http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd">
        <!-- (M-M) Metadata file identifier - A unique File Identifier (GUID) - USGIN recommends
                using a valid Universally Unique Identifier (UUID) -->
  <gmd:fileIdentifier>
    <gco:CharacterString>08fb00c8-0882-4bf7-b07f-fd37050c5efc</gco:CharacterString>
```

```
</gmd:fileIdentifier>
          <!-- (M-M) Metadata language - ISO <ISO639-2/T three letter language code - lower
      case formatting. -->
<gmd:language>
   <gco:CharacterString>eng</gco:CharacterString>
</gmd:language>

        <!-- (M-M) Metadata character set   - USGIN requires that a character set code is
              defined to facilitate CSW servers (deegree, GeoNetwork, etc.).   -->
<gmd:characterSet>
      <!-- MD_CharacterSetCode names: {ucs2, ucs4, utf7, utf8, utf16, 8859part1, 8859part2,
   8859part3, 8859part4, 8859part5, 8859part6, 8859part7, 8859part8, 8859part9, 8859part10,
   8859part11, 8859part13, 8859part14, 8859part15, 8859part16, jis, shiftJIS, eucJP, usAscii,
   ebcdic, eucKR, big5, GB2312} -->
    <gmd:MD_CharacterSetCode
   codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Charact
   erSetCode" codeListValue="utf8">UTF-8</gmd:MD_CharacterSetCode>
</gmd:characterSet>
         <!-- (M-M) Resource type - Define if this record is a: dataset (default), service, feature,
      software, etc. -->
<gmd:hierarchyLevel>
         <!-- MD_ScopeCode code names: {attribute, attributeType,
   collectionHardware,collectionSession, dataset, series, nonGeographicDataset,
   dimensionGroup, feature, featureType, propertyType, fieldSession, software, service, model,
   tile}. -->
   <gmd:MD_ScopeCode
     codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Scope
     Code" codeListValue="dataset">dataset</gmd:MD_ScopeCode>
</gmd:hierarchyLevel>
         <!-- (O-M) Resource hierarchy level name - USGIN makes this property mandatory to
              identify the USGIN resource type (see USGIN Profile, "Resources of Interest").
              Default CharacterString is “Dataset” Encode hierarchy by including
              hierarchyLevelName elements for all broader resource categories. E.g. default
              should also include a hierarchyLevelName=”Collection” element. For services
              USGIN hierarchyLevelName/CharacterString is “Service”. -->
<gmd:hierarchyLevelName>
   <gco:CharacterString>Dataset</gco:CharacterString>
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
   <gco:CharacterString>Collection</gco:CharacterString>
</gmd:hierarchyLevelName>
         <!-- (M-M) Metadata point of contact - Point of contact for the metadata record, e.g. for
              users to report errors, updates to metadata, etc. -->
<gmd:contact>
   <gmd:CI_ResponsibleParty>
    <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
    <!-- only one is mandatory, this example has organisationName -->
      <gmd:organisationName>
        <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
      </gmd:organisationName>
```

```
      <!-- Metadata contact information, either a phone number or e-mail address is required.
      This example has an e-mail address. ISO19115 makes contact role mandatory as well.
                -->
     <gmd:contactInfo>
       <gmd:CI_Contact>
         <gmd:address>
           <gmd:CI_Address>
  <!-- (O-M) Metadata point of contact e-mail address to meet profile requirement -->
             <gmd:electronicMailAddress>
               <gco:CharacterString>metadata@azgs.az.gov</gco:CharacterString>
             </gmd:electronicMailAddress>
           </gmd:CI_Address>
         </gmd:address>
       </gmd:CI_Contact>
     </gmd:contactInfo>
  <!-- (M-M) ISO 19139 Mandatory: contact role -->
     <gmd:role>
  <!-- CI_RoleCode names: {resourceProvider, custodian, owner, user, distributor, originator,
   pointOfContact, principalInvestigator, processor, publisher, author}   -->
     <gmd:CI_RoleCode
       codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_RoleC
       ode" codeListValue="pointOfContact">point of contact</gmd:CI_RoleCode>
     </gmd:role>
   </gmd:CI_ResponsibleParty>
</gmd:contact>
       <!-- (M-M) Metadata date stamp - USGIN profile requires use of
    dateStamp/gco:DateTime (Note this contrasts with INSPIRE mandate to use
    dateStamp/gco:Date). This is the date and time when the metadata record was created or
    most recently updated (following NAP). -->
<gmd:dateStamp>
  <!-- Requires an extended ISO 8601 formatted combined UTC date and time string
   (YYYY-MM-DDTHH:MM:SS) Some validators require time zone as well; recommend using
   GMT indicated by 'Z' -->
   <gco:DateTime>2010-01-14T10:00:00Z</gco:DateTime>
</gmd:dateStamp>
     <!-- (M-M) metadata standard - USGIN profile conformant metadata is indicated by the
               string “ISO-USGIN" -->
<gmd:metadataStandardName>
   <gco:CharacterString>ISO-USGIN</gco:CharacterString>
</gmd:metadataStandardName>
     <!-- (O-M) USGIN profile version -->
<gmd:metadataStandardVersion>
   <gco:CharacterString>1.2</gco:CharacterString>
</gmd:metadataStandardVersion>
      <!-- (M-M) Resource identification information - At least one of MD_DataIdentification
               (dataset, dataset series) or SV_ServiceIdentification (service) is required.   -->
<gmd:identificationInfo>
    <!-- Resource Dataset or Dataset Series Identification -->
   <gmd:MD_DataIdentification>
      <!-- (M-M) Resource citation - Information to identify the intellectual origin of the content in
               the described resource, along the lines of a citation in a scientific journal. Required
```

```
            content for a CI_Citation element are title, date, and responsibleParty. If the
            subject of the metadata record is derived from other sources, such that it is not
            considered identical with those sources, the provenance should be documented in
            LI_Lineage//LI_Source elements. -->
    <gmd:citation>
      <gmd:CI_Citation>
     <!-- (M-M) Resource title - Use titles that inform the human reader about the dataset's
            content; titles should be unique in the context of the metadata catalog that
            contains the metadata record. -->
        <gmd:title>
          <gco:CharacterString>USGIN minimum metadata example XML
           file.</gco:CharacterString>
        </gmd:title>
     <!-- (M-M) Resource reference date - Best practice is to include at least the date of
            publication or creation of the resource. The date of the resource reported in the
            citation corresponds to the resource's last update version according to its update
            frequency. CI_Date content includes a date and dateType. Date for USGIN profile
            uses gco:DateTime. -->
        <gmd:date>
          <gmd:CI_Date>
            <gmd:date>
<!-- Requires an extended ISO 8601 UTC date and time string (YYYY-MM-DDTHH:MM:SS)
 Some validators require time zone as well; recommend GMT indicated by 'Z' -->
              <gco:DateTime>2010-01-14T09:30:47Z</gco:DateTime>
            </gmd:date>
<!-- dateType is mandatory per ISO19115 -->
            <gmd:dateType>
<!-- CI_DateTypeCode names: {creation, publication, revision}. -->
              <gmd:CI_DateTypeCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
                CI_DateTypeCode"codeListValue="publication">publication</gmd:CI_DateTyp
                eCode>
            </gmd:dateType>
          </gmd:CI_Date>
        </gmd:date>
      <!-- (M-M) Resource responsible party - The citation attribute provides information for
            citing the intellectual origin of the described resource, analogous to a citation in a
            scientific journal. The citedResponsibleParty should specify the individual(s) and
            organizations responsible for the creation or origination of the resource. For most
            intellectual content, the responsible party is what would normally be considered
            the author of a work.    -->
        <gmd:citedResponsibleParty>
          <gmd:CI_ResponsibleParty>
 <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
 <!-- only one is mandatory, this example has organisationName -->
            <gmd:organisationName>
              <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
            </gmd:organisationName>
        <!-- cited responsible party contact information, either a phone number or e-mail
            address is required. This example has a voice telephone number. ISO19115
            makes contact role mandatory as well. -->
```

```
            <gmd:contactInfo>
              <gmd:CI_Contact>
                <gmd:phone>
                  <gmd:CI_Telephone>
                    <gmd:voice>
                      <gco:CharacterString>520-770-3500</gco:CharacterString>
                    </gmd:voice>
                  </gmd:CI_Telephone>
                </gmd:phone>
              </gmd:CI_Contact>
            </gmd:contactInfo>
  <!-- (M-M) ISO 19139 Mandatory: contact role   -->
            <gmd:role>
 <!-- CI_RoleCode names: {resourceProvider, custodian, owner, user, distributor, originator,
  pointOfContact, principalInvestigator, processor, publisher, author} -->
                                                        <gmd:CI_RoleCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
                CI_RoleCode"codeListValue="pointOfContact">point of
                contact</gmd:CI_RoleCode>
            </gmd:role>
          </gmd:CI_ResponsibleParty>
        </gmd:citedResponsibleParty>
      </gmd:CI_Citation>
    </gmd:citation>

    <!-- (M-M) Resource Abstract - A free text summary of the content, significance, purpose,
            scope, etc. of the resource. Exactly one value. -->
    <gmd:abstract>
      <gco:CharacterString>Example for the minimum required elements in a USGIN dataset
        metadata record. Note that this example includes conditional minimum elements that
        may or may not apply to a specific resource and its metadata. Although the resource is
        non-geographic, and extent element is included for example
        purposes.</gco:CharacterString>
    </gmd:abstract>
    <!-- (O-C) Resource point of contact (access contact) - CI_ResponsibleParty element here
            would contain information for point of contact to access the resource. This
            information is mandatory for physical resources such as core, cuttings, samples,
            manuscripts. This example metadata record is not about a physical resource, so
            the element is empty (included here only to anchor this comment... -->
<gmd:pointOfContact/>
    <!-- (O-C) Descriptive Keywords - Only required if the described resource is
            non-geographic, in which case the single keyword 'non-geographic' is mandatory.
            This example instance is about itself, and example metadata record, and is thus
            non-geographic so the keyword is mandatory. Note that and example bounding
            box is included for information purposes as well. -->
    <gmd:descriptiveKeywords>
      <gmd:MD_Keywords>
        <gmd:keyword>
          <gco:CharacterString>non-geographic</gco:CharacterString>
        </gmd:keyword>
```

```
            <gmd:type>
          <!-- Keyword Type - allowed values from MD_KeywordTypeCode names: {discipline,
        place, stratum, temporal, theme} -->
                <gmd:MD_KeywordTypeCode
              codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Ke
              ywordTypeCode" codeListValue="place">place</gmd:MD_KeywordTypeCode>
               </gmd:type>
            </gmd:MD_Keywords>
          </gmd:descriptiveKeywords>
          <!-- (M-M) Resource language - Multiple instances of this element indicate that the
         linguistic content of the resource is available in multiple languages -->
          <gmd:language>
        <!-- ISO 639-2/T three-letter language code (http://www.loc.gov/standards/iso639-2/). -->
             <gco:CharacterString>eng</gco:CharacterString>
          </gmd:language>
          <!-- (C-C) Topic category - A topicCategory code must be provided when hierarchyLevel
                  is set to "dataset" (this example document) or "dataset series". Most USGIN
                  resources will have MD_TopicCategoryCode = "geoscientificInformation", which is
                  the default value for this profile.   More specific topic categorization should be
                  done using keywords. -->
          <gmd:topicCategory>
        <!-- MD_TopicCategoryCode names: {farming, biota, boundaries,
         climatologyMeterologyAtmosphere, economy, elevation, environment,
         geoscientificInformation, health, imageryBaseMapsEarthCover, intelligenceMilitary,
         inlandWater, location, oceans, planningCadastre, society, structure, transportation,
         utilitiesCommunication} -->

    <gmd:MD_TopicCategoryCode>geoscientificInformation</gmd:MD_TopicCategoryCode>
          </gmd:topicCategory>
          <!-- (C-C) Resource content extent - Defines the spatial (horizontal and vertical) and
                  temporal region to which the content of the resource applies. For USGIN, a
                  geographicElement/geographicBoundingBox must be provided for any resource
                  with content related to some geographic location. If the described resource is not
                  related to a geographic area, the place keyword 'non-geographic' should be
                  included in the descriptiveKeywords element. -->
          <!-- note that although this example record is about a non-geographic resource, bounding
                  boxes are anticipated for most resources so an example element is included, but it
                  is not required. -->
          <gmd:extent>
            <gmd:EX_Extent>
          <!-- (O-C) Resource content extent bounding box -USGIN profile requires that if an
                  EX_Extent/geographicElement is supplied, it include a geographic bounding box
                  with bounding latitude and longitude expressed using WGS 84 decimal degrees.
                  The corner coordinates for the geographic bounding box must not coincide in one
                  point, because this may result in fatal errors with some CSW implementations.
                  Point locations must thus be represented as tiny rectangles. USGIN recommended
                  practice is to place the actual point location in the lower left corner of the rectangle.
                  -->
              <gmd:geographicElement>
                <gmd:EX_GeographicBoundingBox>
                  <gmd:westBoundLongitude>
```

```
                <gco:Decimal>-109.911001</gco:Decimal>
              </gmd:westBoundLongitude>
              <gmd:eastBoundLongitude>
                <gco:Decimal>-109.910999</gco:Decimal>
              </gmd:eastBoundLongitude>
              <gmd:southBoundLatitude>
                <gco:Decimal>34.772899</gco:Decimal>
              </gmd:southBoundLatitude>
              <gmd:northBoundLatitude>
                <gco:Decimal>34.772901</gco:Decimal>
              </gmd:northBoundLatitude>
            </gmd:EX_GeographicBoundingBox>
          </gmd:geographicElement>
        </gmd:EX_Extent>
      </gmd:extent>
    </gmd:MD_DataIdentification>
  </gmd:identificationInfo>
      <!-- Distribution (O-M) This element provides information to inform users how to obtain or
           access the described resource, which is considered part of the minimum useful
           content that should be provided by a metadata record. USGIN profile specifies that
           at least one MD_Distribution/distributor and one MD_Distribution/transferOptions
           element are required, and that the specified distributor provides the specified
           transfer options. -->
  <gmd:distributionInfo>
      <gmd:MD_Distribution>
          <gmd:distributor>
              <gmd:MD_Distributor>
       <!-- (O-C) distributor point of contact - CI_ResponsibleParty element here would contain
           information on who to contact if there is a problem with accessing the resource
           following the instructions in this distribution element. This information is
           mandatory, requiring at least a name (one of individual, organization, or position)
           and a voice telephone number or e-mail address. This example includes a position
           name and e-mail address. -->
                  <gmd:distributorContact>
                      <gmd:CI_ResponsibleParty>
                          <gmd:positionName>
                            <gco:CharacterString>Web Master</gco:CharacterString>
                          </gmd:positionName>
                          <gmd:contactInfo>
                              <gmd:CI_Contact>
                                  <gmd:address>
                                      <gmd:CI_Address>
                                          <gmd:electronicMailAddress>

<gco:CharacterString>webmaster@usgin.org</gco:CharacterString>
                                          </gmd:electronicMailAddress>
                                      </gmd:CI_Address>
                                  </gmd:address>
                              </gmd:CI_Contact>
                          </gmd:contactInfo>
                          <gmd:role>
```

```
    <!-- CI_RoleCode names: pointOfContact is mandatory for at least one distributor
                responsible party -->
                        <gmd:CI_RoleCode
                         codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.x
                         ml#CI_RoleCode"codeListValue="pointOfContact">point of
                         contact</gmd:CI_RoleCode>
                                </gmd:role>
                            </gmd:CI_ResponsibleParty>
                        </gmd:distributorContact>
                    </gmd:MD_Distributor>
                </gmd:distributor>
                <gmd:transferOptions>
                    <gmd:MD_DigitalTransferOptions>
                        <gmd:onLine>
                            <gmd:CI_OnlineResource>
        <!-- The CI_OnlineResource/linkage/URL element must contain complete URL to access
                the resource -->
                                <gmd:linkage>
                         <gmd:URL>http://repository.usgin.org/uri_gin/usgin/dlio/337</gmd:URL>
                                </gmd:linkage>
                            </gmd:CI_OnlineResource>
                        </gmd:onLine>
                    </gmd:MD_DigitalTransferOptions>
                </gmd:transferOptions>
            </gmd:MD_Distribution>
        </gmd:distributionInfo>
    </gmd:MD_Metadata>
```

8.2 USGIN ISO 19139 Dataset Metadata

In the following listing, text in green is comments; XML elements are in blue, XML attributes are in red, and attribute values are in purple.

This example metadata record describes a scanned well log document.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--*****************************************************************************************
*** Example ISO 19139 Geospatial Dataset Metadata based on the USGIN v1.2 Profile
*** by USGIN Standards and Protocols Drafting Team
*** U.S. Geoscience Information System (USGIN) - http://lab.usgin.org
*** Contributors: Wolfgang Grunberg, Stephen M Richard
*** 2013-06-14
***
*** DISCLAIMER: this is not an authoritative metadata example but an aide to get started; refer
                to
*** the USGIN profile document for more specific and complete guidelines.
***
*** Validated against http://www.isotc211.org/2005/gmd (ISO 19115, CSW 2.0.2 AP ISO 1.0).
*** Follows the USGIN ISO 19139 Dataset Metadata Profile v1.2.
***
*** NOTES:
*** - Codelists:
*** Most ISO metadata profiles and applications use ISO codelists. The codeList attribute points
                to a Uniform Resource Locator (URL) that locates a resource describing the code
                list values. It can be a URN or a URL.
***
*** KEY: - M/C/O/X (Mandatory, Conditional, Optional, Not Used)
***
*****************************************************************************************-->
     <!-- USGIN ISO 19139 geospatial dataset metadata record -->
<gmd:MD_Metadata   xmlns="http://www.isotc211.org/2005/gmd"
                xmlns:gco="http://www.isotc211.org/2005/gco"
                xmlns:gmd="http://www.isotc211.org/2005/gmd"
                xmlns:gmi="http://www.isotc211.org/2005/gmi"
                xmlns:gml="http://www.opengis.net/gml"
                xmlns:gmx="http://www.isotc211.org/2005/gmx"
                xmlns:gts="http://www.isotc211.org/2005/gts"
                xmlns:srv="http://www.isotc211.org/2005/srv"
                xmlns:xlink="http://www.w3.org/1999/xlink"
                xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                xsi:schemaLocation="http://www.isotc211.org/2005/gmd
                http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd">
      <!-- (M) Unique Identifier for the metadata record - USGIN recommends using a valid
                Universally Unique Identifier (UUID) -->
  <gmd:fileIdentifier>
    <gco:CharacterString>00C02E67-F1ED-473DA240068CCB041A73</gco:CharacterString>
  </gmd:fileIdentifier>
    <!-- (M) Metadata language ISO639-2/T three letter language code - lower case. -->
```

```
<gmd:language>
  <gco:CharacterString>eng</gco:CharacterString>
</gmd:language>
                    <!-- (M) Metadata character set   - default is "utf8", USGIN requires that a
              character set code is defined to facilitate CSW servers (deegree, GeoNetwork,
              etc.). -->
<gmd:characterSet>
                    <!-- MD_CharacterSetCode names: {ucs2, ucs4, utf7, utf8, utf16, 8859part1,
              8859part2, 8859part3, 8859part4, 8859part5, 8859part6, 8859part7, 8859part8,
              8859part9, 8859part10, 8859part11, 8859part13, 8859part14, 8859part15,
              8859part16, jis, shiftJIS, eucJP, usAscii, ebcdic, eucKR, big5, GB2312}.   -->
  <gmd:MD_CharacterSetCode
  codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Charact
  erSetCode" codeListValue="utf8">UTF-8</gmd:MD_CharacterSetCode>
</gmd:characterSet>

                    <!-- (M-M) Resource type - Define if this record is a: dataset (default),
              service, feature, software, etc. -->
<gmd:hierarchyLevel>
                    <!-- MD_ScopeCode code names: {attribute, attributeType,
              collectionHardware, col-lectionSession, dataset, series, nonGeographicDataset,
              dimensionGroup, feature, fea-tureType, propertyType, fieldSession, software,
              service, model, tile}. -->
  <gmd:MD_ScopeCode
    codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Scope
    Code" codeListValue="dataset">dataset</gmd:MD_ScopeCode>
</gmd:hierarchyLevel>
                    <!-- (O-M) Resource hierarchy level name - USGIN makes this property
              mandatory to identify the USGIN resource type (see USGIN Profile, "Resources of
              Interest"). Default USGIN hierarchyLevelName CharacterString is “Dataset”.
              Encode hierarchy by including hierarchyLevelName elements for all broader
              resource categories. E.g. default should also include a
              hierarchyLevelName=”Collection” element. For services USGIN
              hierarchyLevelName.CharacterString is “Service” -->
<gmd:hierarchyLevelName>
  <gco:CharacterString>Dataset</gco:CharacterString>
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
  <gco:CharacterString>Collection</gco:CharacterString>
</gmd:hierarchyLevelName>
                    <!-- (M-M) Metadata point of contact - Point of contact use- to report errors,
              updates to metadata -->
<gmd:contact>
  <gmd:CI_ResponsibleParty>
                  <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
    <gmd:individualName>
      <gco:CharacterString>Stephen Richard</gco:CharacterString>
    </gmd:individualName>
    <gmd:organisationName>
      <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
    </gmd:organisationName>
```

```
<gmd:positionName>
  <gco:CharacterString>Metadata Czar</gco:CharacterString>
</gmd:positionName>
<gmd:contactInfo>
  <gmd:CI_Contact>
    <gmd:phone>
      <gmd:CI_Telephone>
        <gmd:voice>
          <gco:CharacterString>520.770.3500</gco:CharacterString>
        </gmd:voice>
        <gmd:facsimile>
          <gco:CharacterString>520.770.3505</gco:CharacterString>
        </gmd:facsimile>
      </gmd:CI_Telephone>
    </gmd:phone>
    <gmd:address>
      <gmd:CI_Address>
        <gmd:deliveryPoint>
          <gco:CharacterString>416 W. Congress St., Suite 100</gco:CharacterString>
        </gmd:deliveryPoint>
        <gmd:city>
          <gco:CharacterString>Tucson</gco:CharacterString>
        </gmd:city>
        <gmd:administrativeArea>
          <gco:CharacterString>Arizona</gco:CharacterString>
        </gmd:administrativeArea>
        <gmd:postalCode>
          <gco:CharacterString>85701-1381</gco:CharacterString>
        </gmd:postalCode>
        <gmd:country>
          <gco:CharacterString>USA</gco:CharacterString>
        </gmd:country>
          <!-- (O-M) Metadata point of contact e-mail address - mandatory if voice phone
        number not supplied -->
        <gmd:electronicMailAddress>
          <gco:CharacterString>metadata@azgs.az.gov</gco:CharacterString>
        </gmd:electronicMailAddress>
      </gmd:CI_Address>
    </gmd:address>
          <!-- (O-O) online resources - this is the online resource to contact the metadata
       person-->
    <gmd:onlineResource>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
          <gmd:URL>http://www.azgs.az.gov</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
          <gco:CharacterString>http</gco:CharacterString>
        </gmd:protocol>
        <gmd:description>
          <gco:CharacterString>Arizona Geological Survey Web
```

```
                Site</gco:CharacterString>
            </gmd:description>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
        <gmd:hoursOfService>
            <gco:CharacterString>8 AM to 5 PM Mountain Standard time (no daylight
            savings)</gco:CharacterString>
        </gmd:hoursOfService>
        <gmd:contactInstructions>
          <gco:CharacterString>Contact Steve Rauzi [Ste-ve.Rauzi@azgs.az.gov] or call Oil
            and Gas Commission Staff at Arizona Geological Survey,
            520-770-3500.</gco:CharacterString>
        </gmd:contactInstructions>
      </gmd:CI_Contact>
    </gmd:contactInfo>
              <!-- (M-M) ISO 19139 Mandatory: contact role -->
    <gmd:role>
            <!-- a responsible party with role=pointOfContact is manda-tory for metadata
          contact -->
      <gmd:CI_RoleCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_Rol
        eCode" codeListValue="pointOfContact">pointOfContact</gmd:CI_RoleCode>
    </gmd:role>
  </gmd:CI_ResponsibleParty>
</gmd:contact>
                  <!-- (X-O) Metadata should include a URL that locates a thumbnail logo for
          organizations related to the metadata origination, the organization hosting the
          catalog that returned the metadata, the organization that originated the data, and
          the organization hosting online services that provide access to the data. -->
<gmd:contact>
  <gmd:CI_ResponsibleParty>
    <gmd:organisationName>
      <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
    </gmd:organisationName>
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:address>
          <gmd:CI_Address>
            <gmd:electronicMailAddress gco:nilReason="missing">
              <gco:CharacterString>metadata@usgin.org</gco:CharacterString>
            </gmd:electronicMailAddress>
          </gmd:CI_Address>
        </gmd:address>
        <gmd:onlineResource>
          <gmd:CI_OnlineResource>
              <!-- Icon image file (e.g. tif, png, jpg, gif) for the metadata originator. This Icon
          will be displayed in search results to credit the metadata originator. -->
            <gmd:linkage>
              <gmd:URL>http://www.azgs.az.gov/logo/metadata/azgs.png</gmd:URL>
            </gmd:linkage>
              <!-- URL that locates an icon thumbnails, the CI_OnlineResource/name MUST
```

```
            be 'icon' -->
            <gmd:name>
              <gco:CharacterString>icon</gco:CharacterString>
            </gmd:name>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
      </gmd:CI_Contact>
    </gmd:contactInfo>
                <!-- none of the codelist roles are applicable -->
    <gmd:role gco:nilReason="inapplicable"/>
  </gmd:CI_ResponsibleParty>
</gmd:contact>
                <!-- (M-M) Metadata date stamp - USGIN profile requires use of
            dateS-tamp/gco:DateTime (Note this contrasts with INSPIRE mandate to use
            dateS-tamp/gco:Date). This is the date and time when the metadata record was
            created or up-dated (following NAP). -->
<gmd:dateStamp>
  <gco:DateTime>2009-11-17T10:00:00</gco:DateTime>
</gmd:dateStamp>
                <!-- (M-M) USGIN profile conformant metadata is indicated by using
            "ISO-USGIN" -->
<gmd:metadataStandardName>
  <gco:CharacterString>ISO-USGIN</gco:CharacterString>
</gmd:metadataStandardName>
                <!-- (O-M) USGIN profile version, mandatory in this profile -->
<gmd:metadataStandardVersion>
  <gco:CharacterString>1.2</gco:CharacterString>
</gmd:metadataStandardVersion>
                <!-- (O-C) Dataset Identifier - For USGIN, this is a string that uniquely
            identifies the de-scribed resource. If the resource has an identifier, it should be
            included here; if the re-source will be referenced from other metadata, it must have
            an identifier here. If the da-taset is coupled to a service, the value of the
            MD_Metadata/dataSetURI attribute is the unique resource identifier used by
            srv:coupledResource to link the service with the da-taset. For the USGIN profile,
            the MD_Distribution/transferOptions/MD_DigitalTransferOptions/
            online/CI_OnlineResource is used to specify URLs for access to the resource. -->
<gmd:dataSetURI>

    <gco:CharacterString>http://azgs.az.gov/resource/00C02E67-F1ED-473D-A240-068CC
    B041A73</gco:CharacterString>
</gmd:dataSetURI>
                <!-- gmd:local would be here, but USGIN profile convention is that each
            catalog service is specific to one language, rather that expecting clients to be able
            to process gmd:PT_Locale and gmd:LocalisedCharacterString. gmd:local
            elements may be ignored by USGIN clients -->
                <!-- (O-O) Resource spatial representation - Spatial representation
            information for the da-taset (resource). Best practice is to include metadata for
            spatial representation if the de-scribed resource is a georeferenced dataset. -->
                <!-- in this example, the described resource is a document that does not
            contain geolo-cated data items, so spatial representation is not applicable -->
<gmd:spatialRepresentationInfo gco:nilReason="inapplicable"/>
```

```
                <!-- (O-O) Resource's spatial reference system - Description of the spatial
            and/or tem-poral reference systems used in the dataset. If
            {/MD_SpatialRepresentationTypeCode = "vector") or
            (../MD_SpatialRepresentationTypeCode = "grid") or
            (../MD_SpatialRepresentationTypeCode = "tin") then count referenceSystemInfo
            >= 1) }not applicable to the document described by this record -->
<gmd:referenceSystemInfo gco:nilReason="inapplicable"/>
                <!-- ********************************************************************************* -->
                <!-- (M-M) Resource identification information - At least one of
            MD_DataIdentification (da-taset, dataset series) or SV_ServiceIdentification
            (service) is required.   -->
<gmd:identificationInfo>
                <!-- Resource Dataset or Dataset Series Identification. In practice, anything
            that is not a service is described using MD_DataIdentification -->
  <gmd:MD_DataIdentification>
                <!-- (M-M) Resource citation - Information to identify the intellectual origin of the
            content in the described resource, along the lines of a citation in a scientific
            journal. Required content for a CI_Citation element are title, date, and
            responsibleParty. If the subject of the metadata record is derived from other
            sources, such that it is not considered identical with those sources, the
            provenance should be documented in LI_Lineage//LI_Source elements. -->
    <gmd:citation>
      <gmd:CI_Citation>
                <!-- (M-M) Resource title - Use titles that inform the human reader about the
            da-taset's content; titles should be unique in the context of the metadata catalog
            that contains the metadata record. -->
        <gmd:title>
          <gco:CharacterString>
            Scanned Borehole Compensated Sonic Log for 0391, Kerr-McGee08 Navajo
          </gco:CharacterString>
        </gmd:title>
                <!-- (M-M) Resource reference date - Best practice is to include at least the
            date of publi-cation or creation of the resource. The date of the resource reported
            in the citation corre-sponds to the resource's last update version according to its
            update frequency. CI_Date content includes a date and dateType. Date for USGIN
            profile uses gco:DateTime. -->
        <gmd:date>
          <gmd:CI_Date>
            <gmd:date>
                <!-- Requires an extended ISO 8601 UTC date and time string (YYYY-MM
            DDTHH:MM:SS) Some validators require time zone as well; recommend GMT
            indicated by 'Z' -->
              <gco:DateTime>2001-12-17T09:30:47</gco:DateTime>
            </gmd:date>
            <gmd:dateType>
                <!-- dateType is mandatory per ISO19115 -->
              <!-- CI_DateTypeCode values in citation context: {creation, publication,
            revision}. -->
              <gmd:CI_DateTypeCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
                CI_DateTypeCode"
```

```
            codeListValue="publication">publication</gmd:CI_DateTypeCode>
      </gmd:dateType>
    </gmd:CI_Date>
  </gmd:date>
          <!-- (C-C) Unique resource identifier   - For USGIN purposes, this element
       content value should be considered an identifier for the citation, without any
       assumption that it will use http protocol. The identifier may be resolvable to a URL,
       if a protocol prefix specifies an identifier scheme that is resolvable (e.g. http,
       doi…), but this is not necessary for a valid document, and should not be assumed
       when processing metadata documents. IF the Ci-tation has an identifier that is
       different from the identifier for the described resource (MD_Metadata/dataSetURI),
       it must be included here. The USGIN profile requires use of MD_Identifer. If
       additional codespace and version content is associated with the identifier, it should
       be encoded as MD_Identifier/authority/ CI_Citation/ alternateTitle and
       MD_Identifier/ authority/ CI_Citation/ edition   -->
  <gmd:identifier>
    <gmd:MD_Identifier>
      <gmd:code>
          <!-- 13 digit ISBN example -->
        <gco:CharacterString>isbn:000-0-000-00000-0</gco:CharacterString>
      </gmd:code>
    </gmd:MD_Identifier>
  </gmd:identifier>
          <!-- (M-M) Resource responsible party - The citation attribute provides
       information for cit-ing the intellectual origin of the described resource, analogous to
       a citation in a scientific journal. The citedResponsibleParty should specify the
       individual(s) and organizations re-sponsible for the creation or origination of the
       resource. For most intellectual content, the responsible party is what would
       normally be considered the author of a work.   -->
          <!-- in the context of the well log described by this example, the originator
       should be considered the operator who ran the log (individualName), or the
       logging company that ran the log; the drilling operator migh also be considered as
       originator. Unfortuantely, in this case that information is missing, so the provided
       contact information is for the custodian, not the author or originator. -->
  <gmd:citedResponsibleParty>
    <gmd:CI_ResponsibleParty>
          <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
      <gmd:individualName>
        <gco:CharacterString>Steve Rauzi</gco:CharacterString>
      </gmd:individualName>
      <gmd:organisationName>
        <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
      </gmd:organisationName>
      <gmd:positionName>
        <gco:CharacterString>Oil and Gas Administrator</gco:CharacterString>
      </gmd:positionName>
          <!-- cited responsible party contact information, either a voice phone number or
       e-mail address is required. Everything else is optional -->
      <gmd:contactInfo>
        <gmd:CI_Contact>
          <gmd:phone>
```

```
            <gmd:CI_Telephone>
              <gmd:voice>
                <gco:CharacterString>520-770-3500</gco:CharacterString>
              </gmd:voice>
              <gmd:facsimile>
                <gco:CharacterString>520-770-3505</gco:CharacterString>
              </gmd:facsimile>
            </gmd:CI_Telephone>
          </gmd:phone>
          <gmd:address>
            <gmd:CI_Address>
              <gmd:deliveryPoint>
                <gco:CharacterString>416 W. Congress St., Suite
                  100</gco:CharacterString>
              </gmd:deliveryPoint>
              <gmd:city>
                <gco:CharacterString>Tucson</gco:CharacterString>
              </gmd:city>
              <gmd:administrativeArea>
                <gco:CharacterString>Arizona</gco:CharacterString>
              </gmd:administrativeArea>
              <gmd:postalCode>
                <gco:CharacterString>85701</gco:CharacterString>
              </gmd:postalCode>
              <gmd:country>
                <gco:CharacterString>USA</gco:CharacterString>
              </gmd:country>
              <gmd:electronicMailAddress>
                <gco:CharacterString>Steve.rauzi@azgs.az.gov</gco:CharacterString>
              </gmd:electronicMailAddress>
            </gmd:CI_Address>
          </gmd:address>
        </gmd:CI_Contact>
      </gmd:contactInfo>
        <!-- (M-M) ISO 19139 Mandatory: contact role   -->
      <gmd:role>
        <!-- CI_RoleCode values in this context:   {originator, principalInvestigator,
       processor, author, custodian} -->
        <gmd:CI_RoleCode
          codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
          CI_RoleCode" codeListValue="custodian">custodian</gmd:CI_RoleCode>
      </gmd:role>
    </gmd:CI_ResponsibleParty>
  </gmd:citedResponsibleParty>
        <!-- (O-C) Dataset Presentation Form - The form in which the original cited
       resource is available. Note that the citation is to the original source of intellectual
       content in the de-scribed resource, and its presentation may be different from the
       format for distribution de-scribed in the metadata. This element should be specified
       if there is a difference between the cited resource presentation format and the
       distribution format(s) listed in the distribu-tionInfo/MD_Distribution section of the
       metadata record. This information is mostly for documentation; for discovery and
```

```
        data access, the distribution formats are more relevant. -->
          <!-- in this example, the original log was acquired by a pen on a strip chart
        recorder... -->
  <gmd:presentationForm>
            <!-- CI_PresentationFormCode names: {documentDigital, documentHardcopy,
        im-ageDigital, image-Hardcopy, mapDigital, mapHardcopy, modelDigital,
        model-Hardcopy, profileDigital, profileHardcopy, tableDigital, tableHardcopy,
        videoDigital, videoHardcopy, audioDigital} -->
      <gmd:CI_PresentationFormCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_
        PresentationFormCode" codeListValue="documentHardcopy">hardcopy on
        pa-per</gmd:CI_PresentationFormCode>
  </gmd:presentationForm>
          <!-- (O-O) Resource series - Information about the series or collection of which
        the cited resource is a part. -->
    <gmd:series>
      <gmd:CI_Series>
        <gmd:name>
        <!-- Name of the publication series or aggregate dataset of which the referenced
        dataset is a part. -->
          <gco:CharacterString>AZGS well log library</gco:CharacterString>
        </gmd:name>
      </gmd:CI_Series>
    </gmd:series>
          <!-- (O-O) Resource other citation details; Recommended practice: a complete
        recommended bibliographic citation should be specified in this element. -->
    <gmd:otherCitationDetails>
      <gco:CharacterString>
        Scanned Borehole Compensated Sonic Log for 0391, Kerr-McGee08 Navajo,1967:
        Tucson, AZ, Arizona Geological Survey Well Log Li-brary, scanned log file, 1 pdf
        document.
      </gco:CharacterString>
    </gmd:otherCitationDetails>
          <!-- (O-C) Resource collective title - Title of the combined resource that the cited
        resource is part of, for example the cited resource may be a paper in an anthology,
        in which case the anthology title would be the collective title. Required if the cited
        resource is part of such a collective work. -->
    <gmd:collectiveTitle gco:nilReason="inapplicable"/>
  </gmd:CI_Citation>
</gmd:citation>
            <!-- (M-M) Resource Abstract - A free text summary of the content,
        significance, purpose, scope, etc. of the resource. Exactly one value. -->
<gmd:abstract>
  <gco:CharacterString>Digital files containing Tiff images of scanned logs. Scanned
    using Neutra scanner hardware. Well penetrates Co-conino, Hermosa, Mississippian,
    Devonian (McCraken, Aneth) and bottoms in Precambrian. Total depth 3647 feet. No
    other logs were run. Well abandoned in 1973. </gco:CharacterString>
</gmd:abstract>
            <!-- (O-O) Resource purpose - Summary of the intentions for which the
        dataset was developed. Purpose includes objectives for creating the dataset and
        what the dataset is to support. No USGIN conventions for use of this element. -->
```

```
<gmd:purpose/>
                <!-- (O-O) Resource Status -   -->
<gmd:status>
              <!-- MD_ProgressCode: For USGIN discovery metadata, values should be
          re-stricted to {completed, onGoing, deprecated}. Obsolete is synonymous with
          deprecated.   -->
  <gmd:MD_ProgressCode
     codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Pr
     ogressCode" codeListValue="completed">completed</gmd:MD_ProgressCode>
</gmd:status>
              <!-- (O-C) Resource point of contact - CI_ResponsibleParty element contains
          information for point of contact to access the resource. This information is
          mandatory for physical resources such as core, cuttings, samples, manuscripts -->
<gmd:pointOfContact>
  <gmd:CI_ResponsibleParty>
    <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
    <gmd:individualName>
      <gco:CharacterString>Steve Rauzi</gco:CharacterString>
    </gmd:individualName>
    <gmd:organisationName>
      <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
    </gmd:organisationName>
    <gmd:positionName>
      <gco:CharacterString>Oil and Gas Administrator</gco:CharacterString>
    </gmd:positionName>
          <!-- (O-C) Contact Information - If a resource point of contact is required then
          (phone + deliveryPoint + electronicMailAddress) > 0. -->
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:phone>
          <gmd:CI_Telephone>
            <gmd:voice>
              <gco:CharacterString>520-770-3500</gco:CharacterString>
            </gmd:voice>
            <gmd:facsimile>
              <gco:CharacterString>520-770-3505</gco:CharacterString>
            </gmd:facsimile>
          </gmd:CI_Telephone>
        </gmd:phone>
        <gmd:address>
          <gmd:CI_Address>
            <gmd:deliveryPoint>
              <gco:CharacterString>416 W. Congress St., Suite
                  100</gco:CharacterString>
            </gmd:deliveryPoint>
            <gmd:city>
              <gco:CharacterString>Tucson</gco:CharacterString>
            </gmd:city>
            <gmd:administrativeArea>
              <gco:CharacterString>Arizona</gco:CharacterString>
            </gmd:administrativeArea>
```

```
            <gmd:postalCode>
              <gco:CharacterString>85701</gco:CharacterString>
            </gmd:postalCode>
            <gmd:country>
              <gco:CharacterString>USA</gco:CharacterString>
            </gmd:country>
            <gmd:electronicMailAddress>
              <gco:CharacterString>Steve.rauzi@azgs.az.gov</gco:CharacterString>
            </gmd:electronicMailAddress>
          </gmd:CI_Address>
        </gmd:address>
      </gmd:CI_Contact>
    </gmd:contactInfo>
          <!-- (M-M) ISO 19139 Mandatory: contact role -->
    <gmd:role>
          <!-- CI_RoleCode names: ISO role codes for physical resource point of contact
        are {custodian, owner, pointOfContact}; other point of contact role codes may
        apply for other resources.     -->
                                    <gmd:CI_RoleCode
        codeList=http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_
        RoleCode codeListValue="pointOfContact">point of contact</gmd:CI_RoleCode>
    </gmd:role>
  </gmd:CI_ResponsibleParty>
</gmd:pointOfContact>

          <!-- (O-O) Resource Maintenance - This element provides information about
        the mainte-nance schedule or history of the resource (or some subset/part of the
        resource specified by the scope and scope description) described by the metadata
        record. 0 to many MD_MaintenanceInformation elements may be included.     -->
<gmd:resourceMaintenance>
  <gmd:MD_MaintenanceInformation>
    <gmd:maintenanceAndUpdateFrequency>
          <!-- MD_MaintenanceFrequencyCode names: {continual, daily, weekly,
        fort-nightly, monthly, quarterly, biannually, annually, asNeeded, irregular,
        not-Planned, un-known} -->
      <gmd:MD_MaintenanceFrequencyCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
        _MaintenanceFrequencyCode" codeListValue="asNeeded">as
        need-ed</gmd:MD_MaintenanceFrequencyCode>
    </gmd:maintenanceAndUpdateFrequency>
  </gmd:MD_MaintenanceInformation>
</gmd:resourceMaintenance>
          <!-- (O-O) Graphic overview of resource - USGIN best practice is to provide
        xlink:href URL to file if it is available online, as an attribute of the
        MD_BrowseGraphic element. If   MD_BrowseGraphic is included,
        MD_BrowseGraphic/filename character string is manda-tory. Recommended
        practice is to use the Anchor extension of CharacterString xml ele-ment from
        ISO19139, which provides a url as an attribute and a text string as a label for the
        link.   -->
<gmd:graphicOverview>
  <gmd:MD_BrowseGraphic>
```

```
            <!-- file name is required. USGIN Recommended practice is to provide a
          com-plete URL as the CharacterString value -->
      <gmd:fileName>
        <gco:CharacterString>

          http://azgs.az.gov/resource/00C02E67-F1ED-473D-A240-068CCB041A73/previe
          w.jpg
        </gco:CharacterString>
      </gmd:fileName>
      <gmd:fileDescription>
        <gco:CharacterString>preview of scanned well log</gco:CharacterString>
      </gmd:fileDescription>
            <!-- Recommended practice is to provide a registered MIME type -->
      <gmd:fileType>
        <gco:CharacterString>image/jpg</gco:CharacterString>
      </gmd:fileType>
    </gmd:MD_BrowseGraphic>
  </gmd:graphicOverview>
                  <!-- (X-X) Resource Format - This element is not used by NAP or USGIN;
          this information is encoded in MD_Metadata/distributionInfo/MD_Distribution/ in
          USGIN metadata. -->
  <gmd:resourceFormat/>
                <!-- (O-O) Resource keywords - Best Practice for USGIN profile metadata is
          to supply keywords to facilitate the discovery of metadata records relevant to the
          user. USGIN re-quires that MD_Keyword/keyword contain a CharacterString.
          USGIN best practice is to include keywords in English -->
  <!-- Theme keywords -->
  <gmd:descriptiveKeywords>
    <gmd:MD_Keywords>
      <gmd:keyword>
        <gco:CharacterString>Scanned Gamma Ray Neutron</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>NMAL</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>borehole</gco:CharacterString>
      </gmd:keyword>
            <!-- Keyword Type: allowed values from MD_KeywordTypeCode
          names:{discipline, place, stratum, temporal, theme}   -->
      <gmd:type>
                                              <gmd:MD_KeywordTypeCode
          codeList=”http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
          KeywordTypeCode”
          codeListValue="theme">theme</gmd:MD_KeywordTypeCode>
      </gmd:type>
    </gmd:MD_Keywords>
  </gmd:descriptiveKeywords>
  <!-- Temporal keywords -->
  <gmd:descriptiveKeywords>
    <gmd:MD_Keywords>
```

```
      <gmd:keyword>
        <gco:CharacterString>Frasian</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>Upper Devonian</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>Devonian</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>Paleozoic</gco:CharacterString>
      </gmd:keyword>
            <!-- Keyword Type - allowed values from MD_KeywordTypeCode names:
          {discipline, place, stratum, temporal, theme} -->
      <gmd:type>
        <gmd:MD_KeywordTypeCode
          codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
          _KeywordTypeCode"
          codeListValue="temporal">temporal</gmd:MD_KeywordTypeCode>
      </gmd:type>
    </gmd:MD_Keywords>
  </gmd:descriptiveKeywords>
                <!-- Place keywords; if a resource is not associated with a spatial location, the
          key-word 'non-geographic' should be included as a place keyoword -->
  <gmd:descriptiveKeywords>
    <gmd:MD_Keywords>
      <gmd:keyword>
        <gco:CharacterString>Arizona</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>T41N R27E S22 NE NE</gco:CharacterString>
      </gmd:keyword>
      <gmd:type>
            <!-- Keyword Type - allowed values from MD_KeywordTypeCode names:
          {discipline, place, stratum, temporal, theme} -->
        <gmd:MD_KeywordTypeCode
          codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
          _KeywordTypeCode"
          codeListValue="place">place</gmd:MD_KeywordTypeCode>
      </gmd:type>
    </gmd:MD_Keywords>
  </gmd:descriptiveKeywords>
                <!-- (O-O) Condition applying to access and use of resource - Follow NAP for
          specifica-tion of resourceConstraints. This attribute provides information for access
          control to the described resource itself. In some situations, the metadata may allow
          a user to learn of the existence of a resource that they may not actually be able to
          access without further clearance. Constraints may be represented by
          MD_Constraint, MD_LegalConstraint, or MD_SecurityConstraint.
          Example here of license statement. -->

  <gmd:resourceConstraints>
```

```
        <gmd:MD_LegalConstraints>
          <gmd:useLimitation>
            <gco:CharacterString>License: CC-BY 4.0
(http://creativecommons.org/licenses/by/4.0/):
    Attribution — You must give appropriate credit, provide a link to the license, and indicate if
changes were made. You may do so in any reasonable manner, but not in any way that
suggests the licensor endorses you or your use.
    No additional restrictions — You may not apply legal terms or technological measures that
legally restrict others from doing anything the license permits.
    Notices:  You do not have to comply with the license for elements of the material in the
public domain or where your use is permitted by an applicable exception or limitation.
     No warranties are given. The license may not give you all of the permissions necessary for
your intended use. For example, other rights such as publicity, privacy, or moral rights may limit
how you use the material.
                    </gco:CharacterString>
          </gmd:useLimitation>
          <gmd:useConstraints>
            <gmd:MD_RestrictionCode
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resourc
es/Codelist/gmxCodelists.xml#CI_RestrictionCode" codeListValue="license"/>
          </gmd:useConstraints>
        </gmd:MD_LegalConstraints>
      </gmd:resourceConstraints>

                <!-- (O-O) Aggregation information - The citation for or name of an aggregate
            dataset, the type of aggregate dataset, and optionally the activity which produced
            the dataset. -->
      <gmd:aggregationInfo>
              <!-- MD_AggregateInformation requires either
            aggregateDataSetName/CI_Citation or aggregateDataSetIdentifier/MD_Identifier.
            -->
        <gmd:MD_AggregateInformation>
            <!-- Related dataset name -->
          <gmd:aggregateDataSetName>
            <gmd:CI_Citation>
              <gmd:title>
                <gco:CharacterString>Related Resource's Title</gco:CharacterString>
              </gmd:title>
              <gmd:date>
                <gmd:CI_Date>
                  <gmd:date>
                    <gco:DateTime>2001-12-17T09:30:47</gco:DateTime>
                  </gmd:date>
                  <gmd:dateType>
                    <gmd:CI_DateTypeCode
                        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.
                        xml#CI_DateTypeCode"
                        codeListValue="publication">publication</gmd:CI_DateTypeCode>
                  </gmd:dateType>
                </gmd:CI_Date>
              </gmd:date>
```

```
      </gmd:CI_Citation>
    </gmd:aggregateDataSetName>
          <!-- Data Set Identifier -->
    <gmd:aggregateDataSetIdentifier>
      <gmd:MD_Identifier>
        <gmd:code>

              <gco:CharacterString>00000000-0000-0000-0000-000000000000</gco:Char
            acterString>
        </gmd:code>
      </gmd:MD_Identifier>
    </gmd:aggregateDataSetIdentifier>
          <!-- (M-M) Association Type is mandatory.. -->
    <gmd:associationType>
          <!-- Use DS_AssociationTypeCode names: {crossReference, largerWorkCitation,
        partOfSeamlessDatabase, source, stereoMate}   -->
      <gmd:DS_AssociationTypeCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#DS
        _AssociationTypeCode" codeListValue="crossReference">cross
        reference</gmd:DS_AssociationTypeCode>
    </gmd:associationType>
  </gmd:MD_AggregateInformation>
</gmd:aggregationInfo>
            <!-- (M-M) Resource language - Multiple instances of this element indicate
        that the lin-guistic content of the resource is available in multiple languages -->
<gmd:language>
          <!-- use ISO639-2/T three letter language code in lower case. -->
  <gco:CharacterString>eng</gco:CharacterString>
</gmd:language>
            <!-- character encoding of resource content. Not applicable to scanned image
        -->
<gmd:characterSet gco:nilReason="inapplicable"/>
            <!-- (C-C) Topic category - A topicCategory code must be provided when
        hierarchyLevel is set to "dataset" (this example document) or "dataset series".
        Most USGIN resources will have MD_TopicCategoryCode =
        "geoscientificInformation", which is the default value for this profile.   More specific
        topic categorization should be done using keywords. -->
<gmd:topicCategory>
          <!-- MD_TopicCategoryCode names: {farming, biota, boundaries,
        climatologyMeterol-ogyAtmosphere, economy, elevation, environment,
        geoscientificInformation, health, im-ageryBaseMapsEarthCover,
        intelligenceMilitary, inlandWater, location, oceans, planning-Cadastre, society,
        structure, transportation, utilitiesCommunication} -->
<gmd:MD_TopicCategoryCode>geoscientificInformation</gmd:MD_TopicCategoryCode>
</gmd:topicCategory>
            <!-- (C-C) Resource content extent - Defines the spatial (horizontal and
        vertical) and tem-poral region to which the content of the resource applies. For
        USGIN, a ge-ographicElement/¬geographic-BoundingBox must be provided for
        any resource with con-tent related to some geographic location. If the described
        resource is not related to a geo-graphic area, the place keyword 'non-geographic'
        should be included in the descriptive-Keywords element. -->
```

```
<gmd:extent>
  <gmd:EX_Extent>
          <!-- (C-C) Resource Content extent description - Free text that describes the
        spatial and temporal extent of the dataset. Note that if geographic place names are
        used to express the geographic extent, USGIN profile specifies that these should
        be encoded using key-word with keyword type code = ‘place.’ Geographic names
        may be duplicated in EX_Extent/description.-->
    <gmd:description>
      <gco:CharacterString>Some spatial/temporal description.</gco:CharacterString>
    </gmd:description>
    <gmd:geographicElement>
         <!-- (O-C) Resource content extent bounding box -USGIN profile requires a
        geographic bounding box with bounding latitude and longitude expressed using
        WGS 84 decimal de-grees. The corner coordinates for the geographic bounding
        box must not coincide in one point, because this may result in fatal errors with
        some CSW implementations. Point loca-tions must thus be represented as tiny
        rectangles. USGIN recommended practice is to place the actual point location in
        the lower left corner of the rectangle. -->
      <gmd:EX_GeographicBoundingBox>
        <gmd:extentTypeCode>
          <gco:Boolean>1</gco:Boolean>
        </gmd:extentTypeCode>
        <gmd:westBoundLongitude>
          <gco:Decimal>-109.911001</gco:Decimal>
        </gmd:westBoundLongitude>
        <gmd:eastBoundLongitude>
          <gco:Decimal>-109.910999</gco:Decimal>
        </gmd:eastBoundLongitude>
        <gmd:southBoundLatitude>
          <gco:Decimal>34.772899</gco:Decimal>
        </gmd:southBoundLatitude>
        <gmd:northBoundLatitude>
          <gco:Decimal>34.772901</gco:Decimal>
        </gmd:northBoundLatitude>
      </gmd:EX_GeographicBoundingBox>
    </gmd:geographicElement>
  </gmd:EX_Extent>

            <!-- (C-X) Resource content extent geographic description - Not used by
        USGIN profile, use keyword with type code = ‘place’ (with thesaurus if necessary).
        -->
            <!-- (C-X) Resource content extent bounding polygon - Not used by USGIN
        profile. To improve interoperability, USGIN mandates the use of Geographic
        Bounding Box; if bounding polygons are available they may be included, but
        clients may ignore content in this element. -->
  <!--
    <gmd:geographicElement>
      <gmd:EX_BoundingPolygon/>
    </gmd:geographicElement>      -->
</gmd:extent>
              <!-- (O-O) Resource temporal extent -   -->
```

```
      <gmd:extent>
        <gmd:EX_Extent>
          <gmd:temporalElement>
            <gmd:EX_TemporalExtent>
              <gmd:extent>
              <!-- Geologic time frame example -->
                <gml:TimePeriod gml:id="IdJurassic">
                  <gml:name>Jurassic</gml:name>
              <!-- USGIN requires the beginPosition and endPosition's frame property to be
              defined. urn:cgi:trs:CGI:StandardGeologicTimeMa identifies the standard Geologic
              time scale frame. -->
                  <gml:beginPosition
                    frame="urn:cgi:trs:CGI:StandardGeologicTimeMa">203</gml:beginPosition>
                  <gml:endPosition frame="urn:cgi:trs:CGI:StandardGeologicTimeMa
                    ">135</gml:endPosition>
                </gml:TimePeriod>
              </gmd:extent>
            </gmd:EX_TemporalExtent>
          </gmd:temporalElement>
        </gmd:EX_Extent>
      </gmd:extent>
                  <!-- (O-X) Resource spatio-temporal extent - Not used. USGIN mandates
              encoding space time location with EX_TemporalExtent and
              EX_GeographicBoundingBox -->
                  <!-- (O-O) Resource vertical extent   -->
      <gmd:extent>
        <gmd:EX_Extent>
          <gmd:verticalElement>
            <gmd:EX_VerticalExtent>
              <gmd:minimumValue>
                <gco:Real>-100</gco:Real>
              </gmd:minimumValue>
              <gmd:maximumValue>
                <gco:Real>200</gco:Real>
              </gmd:maximumValue>
              <!-- Use EPSG register of geodetic parameters such as at
              http://www.epsg-registry.org/. The default VerticalCRS is World mean sea level
              (MSL): http://www.spatialreference.org/ref/epsg/5714/ (replaces
              urn:ogc:def:crs:EPSG::5714) -->
              <gmd:verticalCRS xlink:href="http://www.spatialreference.org/ref/epsg/5714/">
              </gmd:verticalCRS>
            </gmd:EX_VerticalExtent>
          </gmd:verticalElement>
        </gmd:EX_Extent>
      </gmd:extent>
    </gmd:MD_DataIdentification>
  </gmd:identificationInfo>
                  <!-- ================================================================ -->
                  <!-- (O-O) Content information - Characteristics describing the feature
              cataloguecatalog, coverage, or image data. Not applicable to this resource. -->
  <gmd:contentInfo gco:nilReason="inapplicable"/>
```

```
                <!-- (O-O) Resource distribution information - This element provides
            information to inform users how to obtain or access the described resource.
            NOTE: there are several ways el-ements can be nested within MD_Distribution,
            please read the MD_Distribution sections in the USGIN profile document.   -->
<gmd:distributionInfo>
  <gmd:MD_Distribution>
                <!-- (O-O)   Resource distribution format - Information on the format or
            physical manifesta-tion of the resource. If the resource is a physical resource, like
            a book, rock sample, paper document, the distributionFormat/MD_Format/name is
            mandatory, and must be from the USGIN distribution format codelist. Value is
            optional for this document resource. -->
    <gmd:distributionFormat>
      <gmd:MD_Format>
        <gmd:name>
          <gco:CharacterString>application/pdf</gco:CharacterString>
        </gmd:name>
        <gmd:version>
          <gco:CharacterString>8.0</gco:CharacterString>
        </gmd:version>
      </gmd:MD_Format>
    </gmd:distributionFormat>
                <!-- (O-C) Resource distributor information - USGIN differs from NAP in this
            case (but not with ISO19115) by allowing multiple distributors, and binding
            between distributors, transfer options, and formats.   -->
    <gmd:distributor>
              <!-- For USGIN profile, each distributor/MD_Distributor is a binding between
            one or more transfer options and the distributor formats that are available through
            that/those transfer options (MD_DigitalTransferOptions/onLine/CI_OnlineResource
            in particular). If different formats are available from the same distributor, or have
            different transfer options, these should be represented as different
            distributor/MD_Distributor instances. See the USGIN Profile section 'Use of
            MD_Distribution and MD_Distributor' for instructions on use of these elements. -->
      <gmd:MD_Distributor>
        <gmd:distributorContact>
            <!-- contact point for questions, comments, problems with accessing the resource
            through this distribution -->
          <gmd:CI_ResponsibleParty>
            <gmd:organisationName>
              <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
            </gmd:organisationName>
            <gmd:contactInfo>
              <gmd:CI_Contact>
                <gmd:address>
                  <gmd:CI_Address>
                    <gmd:electronicMailAddress>
                      <gco:CharacterString>metadata@usgin.org</gco:CharacterString>
                    </gmd:electronicMailAddress>
                  </gmd:CI_Address>
                </gmd:address>
              </gmd:CI_Contact>
            </gmd:contactInfo>
```

```
      <gmd:role>
      <!-- CI_RoleCode codes applicable in this context include: {resourceProvider,
      dis-tributor, pointOfContact} -->
        <gmd:CI_RoleCode
          codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
          CI_RoleCode" codeListValue="distributor">distributor</gmd:CI_RoleCode>
      </gmd:role>
    </gmd:CI_ResponsibleParty>
</gmd:distributorContact>
      <!-- (O-C) Resource distributor format - USGIN profile specifies that the
      distributionIn-fo/MD_Distribution/distributionFormat may be included in the
      document (its schema val-id...), but distribution format information must be
      duplicated in a distributionInfo/distributor/MD_Distributor/distributorFormat element
      or the content can be lost -->
<gmd:distributorFormat>
  <gmd:MD_Format>
    <!-- Use USGIN distribution format code values. See "Online resource format
    names" in USGIN Profile -->
    <gmd:name>
    <-- registered MIME type IETF RFC-3778 -->
      <gco:CharacterString>application/pdf</gco:CharacterString>
    </gmd:name>
    <gmd:version>
      <gco:CharacterString>8.0</gco:CharacterString>
    </gmd:version>
  </gmd:MD_Format>
</gmd:distributorFormat>
<gmd:distributorTransferOptions>
  <gmd:MD_DigitalTransferOptions>
    <gmd:onLine>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
                <!-- The linkage element should contain the complete URL to access
          the resource direct-ly. CI_OnlineResource requires a Linkage element that is
          a gmd:URL. -->
          <gmd:URL>http://azgs.az.gov/resource/00C02E67-F1ED-473D-A240-068CC
          B041A73/borehole_report.pdf</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
    <!-- The protocol element defines a valid internet protocol used to access the
    resource. Recommended best practice (NAP) is that the protocol mnemonic
    should be taken from the controlled list of Official Internet Protocol Standards at
    http://www.rfc-editor.org/rfcxx00.html or the Internet Assigned Numbers Authority
    (IANA) at http://www.iana.org/protocols. ‘ftp’ or ‘http’ are common values. -->
          <gco:CharacterString>http</gco:CharacterString>
        </gmd:protocol>
        <gmd:name>
    <!-- The CI_OnlineResource/name element may duplicate the file name if the URL
    is a link to a file, but it is recommended to provide a user-friendly label for the file
    for presenta-tion in a user interface.   -->
          <gco:CharacterString>borehole_report.pdf</gco:CharacterString>
```

```
            </gmd:name>
          <!-- The CI_OnlineResource/description element may include a [[CDATA] section
          with key values pairs to provide additional necessary information. USGIN
          recommendataion is to encode these using JSON syntax. -->
            <gmd:description>
              <gco:CharacterString>Downloadable PDF document</gco:CharacterString>
            </gmd:description>
          <!-- (O-C) Resource distributor online distribution function -
          CI_OnlineResource/function is required by USGIN to indicate how linkage is to be
          used. If the resource is accessible as a web service, the metadata for the service
          should be separate metadata record with the dataset(s) exposed through the
          service identified in the service metadata record as cou-pledResources. -->
            <gmd:function>
          <!-- CI_OnlineFunctionCode names: {download, information, offlineAccess, order,
          search} -->
              <gmd:CI_OnLineFunctionCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.
                xml#CI_OnLineFunctionCode"
                codeListValue="download">download</gmd:CI_OnLineFunctionCode>
            </gmd:function>
          </gmd:CI_OnlineResource>
        </gmd:onLine>
      </gmd:MD_DigitalTransferOptions>
    </gmd:distributorTransferOptions>
  </gmd:MD_Distributor>
</gmd:distributor>
          <!-- (C-C) Resource distribution transfer options - MD_DigitalTransferOptions
          provides in-formation on digital distribution of resource. See USGIN Profile 'Use of
          MD_Distribution and MD_Distributor' for instructions on use of this element.
          Details on encoding for MD_DigitalTransferOptions are above in the
          distributorTransferOptions elements description. -->
<gmd:transferOptions>
  <gmd:MD_DigitalTransferOptions>
    <gmd:onLine>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
        <!-- The linkage element should contain the complete URL to access the resource
        directly. CI_Online-Resource requires a Linkage element that is a gmd:URL. -->

          <gmd:URL>http://azgs.az.gov/resource/00C02E67-F1ED-473D-A240-068CC
          B041A73/borehole_report.pdf</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
        <!-- The protocol element defines a valid internet protocol used to access the
        resource. Recommended best practice (NAP) is that the protocol mnemonic
        should be taken from the controlled list of Official Internet Protocol Standards at
        http://www.rfc-editor.org/rfcxx00.html or the Internet Assigned Numbers Authority
        (IANA) at http://www.iana.org/protocols. 'ftp' or 'http' are common values. -->
          <gco:CharacterString>http</gco:CharacterString>
        </gmd:protocol>
        <!-- (C-C) Resource distributor online distribution application profile -
```

```
            applicationProfile is required if the CI_OnlineResource/linkage does not connect to
            a web page, and another software application is needed to use the indicated file
            resource. The applicationProfile character string should specify the software using
            the following recommended syntax: “vendor:application name/application version”,
            e.g. “Microsoft:Word/2007”, or “ESRI:ArcGIS/9.3” -->
            <gmd:applicationProfile>
              <gco:CharacterString>Adobe:Acrobat/8.0</gco:CharacterString>
            </gmd:applicationProfile>
            <gmd:name>
            <!-- The CI_OnlineResource/name element may duplicate the file name if the URL
            is a link to a file, but it is recommended to provide a user-friendly label for the file
            that could be presented in a user interface.   -->
              <gco:CharacterString>borehole_report.pdf</gco:CharacterString>
            </gmd:name>
            <gmd:description>
              <gco:CharacterString>Downloadable PDF document</gco:CharacterString>
            </gmd:description>
            <!-- (O-C) Resource distributor online distribution function -
            CI_OnlineResource/function is required by USGIN to indicate how linkage is to be
            used. If the resource is accessible as a web service, the metadata for the service
            should be separate metadata record with the dataset(s) exposed through the
            service identified in the service metadata record as cou-pledResources. -->
            <gmd:function>
            <!-- CI_OnlineFunctionCode names: {download, information, offlineAccess, order,
            search}   -->
              <gmd:CI_OnLineFunctionCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
                CI_OnLineFunctionCode"
                codeListValue="download">download</gmd:CI_OnLineFunctionCode>
            </gmd:function>
          </gmd:CI_OnlineResource>
        </gmd:onLine>
      </gmd:MD_DigitalTransferOptions>
    </gmd:transferOptions>
  </gmd:MD_Distribution>
</gmd:distributionInfo>
<gmd:dataQualityInfo>
            <!-- In this profile, data quality information is optional, but if included, the
            mandatory content for data quality information is a free text explanation of data
            quality considerations. The
            DQ_DataQuality[1]//report[1]//explanation/CharacterString element should contain
            a free text discussion/description of data quality considerations for the indicated
            scope. The use of any specific data quality element to contain this explanation is
            arbitrary and should not be considered meaningful in this context.   -->
  <gmd:DQ_DataQuality>
    <gmd:scope>
      <gmd:DQ_Scope>
        <gmd:level>
          <!-- MD_ScopeCode values for discovery metadata: {collectionHardware,
          collec-tionSession, dataset, series, dimensionGroup, fieldSession, software,
          service, model, tile} -->
```

```
      <gmd:MD_ScopeCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
        _ScopeCode" codeListValue="dataset">dataset</gmd:MD_ScopeCode>
    </gmd:level>
          <!-- (C-C) Data quality scope level description - Not used by USGIN -->
  </gmd:DQ_Scope>
</gmd:scope>
<report>
          <!-- This is a container for a free text discussion/description of data quality
        considerations for the indicated scope. The use of any specific data quality
        element (e.g. DQ_DomainConsistency here) to contain this explanation is arbitrary
        and should not be considered meaningful in this context.   -->
  <DQ_DomainConsistency>
    <result>
      <DQ_ConformanceResult>
        <specification gco:nilReason="notApplicable"/>
        <explanation>
          <gco:CharacterString>Quality Statement: Original paper quality was poor quality,
            parts of scan are difficult to read. Scan resolution is 300 dpi.
          </gco:CharacterString>
        </explanation>
        <pass gco:nilReason="inapplicable"/>
      </DQ_ConformanceResult>
    </result>
  </DQ_DomainConsistency>
</report>
          <!-- (C-C) Resource lineage -   count(lineage/LI_Lineage/source +
        line-age/LI_Lineage/sourceStep + lineage/LI_Lineage/statement ) >0   for spatial
        dataset and spatial dataset series. Not applicable to services. -->
<gmd:lineage>
  <gmd:LI_Lineage>
          <!-- (C-C) Data quality lineage statement - General explanation of the data
        producer's knowledge of the dataset lineage. This is probably the most useful thing
        to supply. -->
    <gmd:statement>
      <gco:CharacterString>
        This dataset is maintained by the Arizona Geological Survey. Paper logs were
        obtained from the permitee and are on file at the Arizona Geological Survey. Logs
        were scanned in 2005 by S. J. Rauzi. Digital files are archived at the Arizona
        Geological Survey.
      </gco:CharacterString>
    </gmd:statement>
          <!-- (C-C) Data quality lineage process step - An event in the development of the
        dataset. Best practice recommended for USGIN is that source association from a
        process step is to inputs to a process, and processStep associations from a
        source element link an output resource to a process step that produced it.   -->
    <gmd:processStep>
      <gmd:LI_ProcessStep>
        <gmd:description>
          <gco:CharacterString>A detailed description of the scanning process could be
            included here. </gco:CharacterString>
```

```
            </gmd:description>
          </gmd:LI_ProcessStep>
        </gmd:processStep>
            <!-- (C-C) Data quality lineage source - not applicable in this example   -->
        <gmd:source/>
      </gmd:LI_Lineage>
    </gmd:lineage>
  </gmd:DQ_DataQuality>
</gmd:dataQualityInfo>
                  <!-- (O-O) Metadata constraint information - This element specifies use
            constraints for access to the metadata record.   -->
<gmd:metadataConstraints>
                  <!-- Constraints -->
  <gmd:MD_Constraints>
                <!-- metadataConstraints/MD_Constraints/useLimitation is mandatory when
            MD_Constraints is used to specify metadataConstraints.   -->
    <gmd:useLimitation>
      <gco:CharacterString>fair use</gco:CharacterString>
    </gmd:useLimitation>
  </gmd:MD_Constraints>
</gmd:metadataConstraints>
                  <!-- Legal constraint -->
<gmd:metadataConstraints>
  <gmd:MD_LegalConstraints>
                <!-- When one of the subtypes MD_LegalConstraints or
            MD_SecurityConstraints is used, useLimitation is optional. -->
    <gmd:useLimitation>
      <gco:CharacterString>one</gco:CharacterString>
    </gmd:useLimitation>
    <gmd:accessConstraints>
                <!-- MD_RestrictionCode names: {copyright, patent, patentPending, trademark,
            license, intellectualPropertyRights, restricted, otherRestrictions} - NAP expands
            with {licenseUnrestricted, licenseEndUser, licenseDistributor, privacy, statutory,
            confidential, sensitivity}. -->
      <gmd:MD_RestrictionCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_R
        estrictionCode" codeListValue="otherRestrictions">other
        restrictions</gmd:MD_RestrictionCode>
    </gmd:accessConstraints>
    <gmd:useConstraints>
                <!-- MD_RestrictionCode names: {copyright, patent, patentPending, trademark,
            license, intellectualPropertyRights, restricted, otherRestrictions} -->
      <gmd:MD_RestrictionCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_R
        estrictionCode" codeListValue="otherRestrictions">other
        restrictions</gmd:MD_RestrictionCode>
    </gmd:useConstraints>
                <!-- (C-C) otherConstraints is a free text element required by NAP if
            accessConstraints or useConstraints is set to "otherRestrictions." -->
    <gmd:otherConstraints>
      <gco:CharacterString>
```

```
            Data only to be used for the purposes for which they were collected.
          </gco:CharacterString>
        </gmd:otherConstraints>
      </gmd:MD_LegalConstraints>
    </gmd:metadataConstraints>
    <gmd:metadataConstraints>
                <!-- Security constraints; in general these will be in applicable to USGIN resources
                -->
      <gmd:MD_SecurityConstraints>
        <gmd:classification>
                    <!-- MD_SecurtyConstraints has various optional free text values, and a
                required MD_SecurityConstraints/classification from MD_ClassificationCode
                names: {unclassified, restricted, confidential, secret, topSecret}   -->
          <gmd:MD_ClassificationCode
            codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Cl
            assificationCode"
            codeListValue="unclassified">unclassified</gmd:MD_ClassificationCode>
        </gmd:classification>
      </gmd:MD_SecurityConstraints>
    </gmd:metadataConstraints>
                        <!-- (O-O) Metadata maintenance information - This element provides
                information about the maintenance schedule or history of the metadata record.
                -->
    <gmd:metadataMaintenance>
      <gmd:MD_MaintenanceInformation>
        <gmd:maintenanceAndUpdateFrequency>
                    <!-- Only one MD_MaintenanceInformation element may be included, with a
                required MD_MaintenanceFrequencyCode names: {continual, daily, weekly,
                fortnightly, monthly, quarterly, biannually, annually, asNeeded, irregular,
                not-Planned, unknown} -->
          <gmd:MD_MaintenanceFrequencyCode
            codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_M
            aintenanceFrequencyCode" codeListValue="asNeeded">as
            needed</gmd:MD_MaintenanceFrequencyCode>
        </gmd:maintenanceAndUpdateFrequency>
        <gmd:maintenanceNote>
          <gco:CharacterString>This metadata record has been processed by the
              iso-19115-to-usgin-19115-data XSLT to ensure that all mandatory content for
              USGIN profile has been added.2013-04-04T12:00:00</gco:CharacterString>
        </gmd:maintenanceNote>
      </gmd:MD_MaintenanceInformation>
    </gmd:metadataMaintenance>
  </gmd:MD_Metadata>
```

8.3 ISO19110 Feature Catalog example

Example encoding of entity attribute information using the xml implementation of ISO 19110. Comments in green provide some explanation.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<gfc:FC_FeatureCatalogue xmlns="http://www.isotc211.org/2005/gfc"
xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:gfc="http://www.isotc211.org/2005/gfc"
xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gml="http://www.opengis.net/gml/3.2"
xmlns:gmx="http://www.isotc211.org/2005/gmx" xmlns:gsr="http://www.isotc211.org/2005/gsr"
xmlns:gss="http://www.isotc211.org/2005/gss" xmlns:gts="http://www.isotc211.org/2005/gts"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" id="FC001"
xsi:schemaLocation="<http://www.isotc211.org/2005/gfc
http://www.isotc211.org/2005/gfc/gfc.xsd>" uuid="99ffdfd0-775a-11e0-a1f0-0800200c9a34">

    <!-- mandatory -->
    <gmx:name>
        <gco:CharacterString>Feature Catalogue for Provisional Processed CTD Data from the
Deepwater Horizon Incident Response in the Gulf of Mexico May 08, 2010 through November
15 ,2010
</gco:CharacterString>
    </gmx:name>
    <!-- mandatory 1..* -->
    <gmx:scope gco:nilReason="unknown"/>

    <!-- mandatory -->
    <gmx:versionNumber gco:nilReason="unknown"/>
    <!-- mandatory -->
    <gmx:versionDate gco:nilReason="unknown"/>
    <gmx:language>
        <gco:CharacterString>eng</gco:CharacterString>
    </gmx:language>
    <gmx:characterSet>
        <gmd:MD_CharacterSetCode
codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_CharacterSe
tCode" -"utf8">UTF-8</gmd:MD_CharacterSetCode>
    </gmx:characterSet>
    <!-- can put subCatalog here, but its unclear what the use case is -->
    <!-- mandatory -->
    <gfc:producer>
        <!-- could use xlink:href to resource contact -->
        <gmd:CI_ResponsibleParty>
            <gmd:individualName>
                <gco:CharacterString>No Name Was Given</gco:CharacterString>
            </gmd:individualName>
            <gmd:organisationName>
                <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
            </gmd:organisationName>
            <gmd:contactInfo>
                <gmd:CI_Contact>
```

```
                        <gmd:phone>
                            <gmd:CI_Telephone>
                                <gmd:voice>
                                    <gco:CharacterString>520-770-3500</gco:CharacterString>
                                </gmd:voice>
                            </gmd:CI_Telephone>
                        </gmd:phone>
                        <gmd:address>
                            <gmd:CI_Address>
                                <gmd:deliveryPoint>
                                    <gco:CharacterString>
                                    416 W. Congress St., Suite 100
                                </gco:CharacterString>
                                </gmd:deliveryPoint>
                                <gmd:city>
                                    <gco:CharacterString>Tucson</gco:CharacterString>
                                </gmd:city>
                                <gmd:administrativeArea>
                                    <gco:CharacterString>AZ</gco:CharacterString>
                                </gmd:administrativeArea>
                                <gmd:postalCode>
                                    <gco:CharacterString>85701</gco:CharacterString>
                                </gmd:postalCode>
                                <gmd:electronicMailAddress>
<gco:CharacterString>metadata@azgs.az.gov</gco:CharacterString>
                                </gmd:electronicMailAddress>
                            </gmd:CI_Address>
                        </gmd:address>
                    </gmd:CI_Contact>
                </gmd:contactInfo>
                <gmd:role>
                    <gmd:CI_RoleCode
codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_RoleCode"
codeListValue="curator">curator</gmd:CI_RoleCode>
                </gmd:role>
            </gmd:CI_ResponsibleParty>
        </gfc:producer>
        <!-- mandatory 1..* -->
        <gfc:featureType>
            <!-- the id attribute should be used as a fragment identifer in URLs for the feature
catalog citation in ISO 19115 metadata -->
            <gfc:FC_FeatureType id="fragmentIdentifer">
                <!-- mandatory -->
                <gfc:typeName>
                    <gco:LocalName>Name of the feature</gco:LocalName>
                </gfc:typeName>
                <gfc:definition>
                    <gco:CharacterString>CTD Header Information</gco:CharacterString>
                </gfc:definition>
                <gfc:code>
```

```
                    <gco:CharacterString>URI for feature type--handy if there is a
registry</gco:CharacterString>
                </gfc:code>
                <!-- mandatory. Indicates if the feature type is abstract or not. default is FALSE-->
                <gfc:isAbstract>
                    <gco:Boolean>false</gco:Boolean>
                </gfc:isAbstract>
 <!-- mandatory. pointer to containing catalog; UUID is defined in root element for this catalog-->
                <gfc:featureCatalogue uuidref="87ffdfd0-775a-11e0-a1f0-0800200c9a66"/>
                <gfc:carrierOfCharacteristics>
                    <gfc:FC_FeatureAttribute>
                        <!-- mandatory -->
                        <gfc:memberName>
                            <gco:LocalName>Name of attribute1</gco:LocalName>
                        </gfc:memberName>
                        <gfc:definition>
                            <gco:CharacterString>explanation of attribute
1</gco:CharacterString>
                        </gfc:definition>
                        <!-- mandatory -->
                        <gfc:cardinality>
                            <gco:Multiplicity>
                                <gco:range>
                                    <gco:MultiplicityRange>
                                        <gco:lower>
                                            <gco:Integer>0</gco:Integer>
                                        </gco:lower>
                                        <gco:upper>
   <!-- I guess this is how one would say as many as you want... -->
                                            <gco:UnlimitedInteger isInfinite="true">
                                                5</gco:UnlimitedInteger>
                                        </gco:upper>
                                    </gco:MultiplicityRange>
                                </gco:range>
                            </gco:Multiplicity>
                        </gfc:cardinality>
                        <!-- optional -->
                        <gfc:definitionReference>
                            <gfc:FC_DefinitionReference>
                                <gfc:definitionSource>
                                    <gfc:FC_DefinitionSource>
            <!-- most likely this will be the same as the citation for the catalog producer -->
                                        <gfc:source
                                            xlink:href="#UUID of citation for containing catalog"/>
                                    </gfc:FC_DefinitionSource>
                                </gfc:definitionSource>
                            </gfc:FC_DefinitionReference>
                        </gfc:definitionReference>
                    <!-- optional -->
                        <code>
                          <gco:CharacterString>
```

```
          URI for this attribute if it is registered somewhere
                            </gco:CharacterString>
                          </code>
                            </gfc:valueMeasurementUnit>
                    <!-- optional -->
                          <valueType>
                              <gco:TypeName>
                    <!-- its unclear how this scheme is supposed to be used for complex datatypes
-->
                                      <gco:Name>
                                          <gco:CharacterString>
          datatype name for quantifying this attribute suggest using xs:types if applicable,
but might be URI for a complex element
                                          </gco:CharacterString>
                                      </gco:Name>
                                  </gco:TypeName>
                          </valueType>
<!-- if valueType is an enumeration, there may be collection of FC_listedValues documenting
the choices    -->
                      </gfc:FC_FeatureAttribute>
                  </gfc:carrierOfCharacteristics>
                  <carrierOfCharacteristics>
                      <FC_FeatureAttribute>
                          <memberName>
                            <gco:LocalName
codeSpace="http://stategeothermaldata.org/uri-gin/aasg/xmlschema/activefault/1.2">FeatureU
RI</gco:LocalName>
                          </memberName>
                          <definition>
                            <gco:CharacterString>
                            unique identifier for this feature. Best practice is to define an http
                              URI that will dereference to a normative description of the feature
                            </gco:CharacterString>
                          </definition>
                          <cardinality>
                              <gco:Multiplicity>
                                  <gco:range>
                                      <gco:MultiplicityRange>
                                          <gco:lower>
                                              <gco:Integer>1</gco:Integer>
                                          </gco:lower>
                                          <gco:upper>
                                              <gco:UnlimitedInteger>1</gco:UnlimitedInteger>
                                          </gco:upper>
                                      </gco:MultiplicityRange>
                                  </gco:range>
                              </gco:Multiplicity>
                          </cardinality>
                          <valueType>
```

```
                                <gco:TypeName>
                                    <gco:Name>
                                      <gco:CharacterString>xs:string</gco:CharacterString>
                                    </gco:Name>
                                </gco:TypeName>
                            </valueType>
                        </FC_FeatureAttribute>
                    </carrierOfCharacteristics>
                    <!-- other attributes   -->
                </gfc:FC_FeatureType>
            </gfc:featureType>
            <!-- the catalog may document many feature types -->
</gfc:FC_FeatureCatalogue>
```

Extensible Markup Language (XML)

A Primer

1/3/2013

Table of Contents

Section 1: Introduction 2

Section 2: An XML Primer 2

Section 3: XML Attributes 4

Section 4: XML Namespaces and Prefixes 6

Section 5: XML Review 8

Section 1: Introduction

The United States Geoscience Information Network (USGIN) specifies the use of interchange formats to exchange and present data in an interoperable manner. One such interchange format is XML.

This tutorial is designed to familiarize the reader with the basics of XML, thereby permitting the reader to interact more easily with the ubiquitous XML documents used by USGIN.

Section 2: An XML Primer

XML stands for *Extensible Markup Language*, referring to the fact that XML is designed to be readily adapted for a variety of uses.

Because XML is a markup language, both the *content* of an XML document and the *structure* applied to that content are visible to the user. The visible structure of markup language documents is provided markup language elements.

A markup language *element* is demarcated by markup language *tags*. Opening tags *open* the element, and closing tags *close* the element. The *content* of a markup language element exists between the opening and closing tags of the element.

In a markup language document, elements appear like so:

```
<tag>content</tag>
```

For some, the above example may seem reminiscent of HTML. That is because HTML is a markup language that uses the same syntax as XML (HTML stands for "Hypertext Markup Language").

So XML and HTML are both markup languages. Both XML and HTML documents are structured by *elements*, and both XML and HTML elements are demarcated by *tags*.

Markup language elements are *hierarchical*: elements can be nested within elements. Once an element has been opened, everything between the opening tag and the closing tag, including elements, is part of the element demarcated by the surrounding tags.

For example, an XML document containing the contact information for a State Geothermal Data web service might appear as follows:

```
<ContactInfo>

    <Phone>
```

```
            <Facsimile>(987) 654-3210</Facsimile>
            <Voice>(123) 456-7890</Voice>

        </Phone>

    </ContactInfo>
```

The above example of nested markup language elements is analogous to folders within folders on a computer (Figure 1). But markup languages documents don't use folders; they use *elements*.

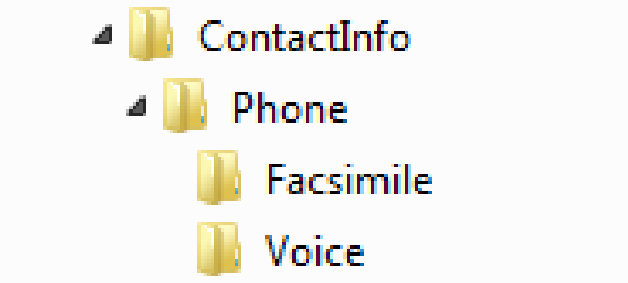


Figure 1: An example of nested folders within a computer filing system

All markup language documents begin with a *root element*: an element that contains all other elements in the document. In the case of HTML documents, this is the **HTML** element (`<html>`).

It is important to remember that, *unless otherwise specified*, elements *inherit* the characteristics of any elements within which they are nested; you will learn more about this in subsequent sections.

Markup language elements can close themselves. For example, the **horizontal rule** (`<hr>`) HTML element can be closed without a closing tag just by adding a **space** followed by a **forward slash (/)** to the opening tag of the element, like so:

```
<hr />
```

When interpreted by a browser, the **horizontal rule** tag produces a horizontal line, like so:

XML documents sometimes employ self-closing tags, usually when *attributes* modifying the element are more important than any content *within* the element.

Section 3: XML Attributes

Most XML documents tend to look a bit more complicated than previous examples suggest, owing to the presence of attributes.

Markup language attributes specify additional parameters that modify a given element; multiple attributes can modify a single element, and each attribute can specify multiple values.

In general, markup language attributes appear within the opening tag of a given element:

```
<tag attribute="value">content</tag>
```

Attributes can also be present within self-closing elements:

```
<tag attribute="value" />
```

As a practical example, the **paragraph** (`<p>`) element is used to designate a paragraph of text in an HTML document:

```
<p>The quick brown fox jumps over the lazy dog.</p>
```

When interpreted by a web browser, the sentence between the `<p>` tags of the **paragraph** element would appear as its own paragraph, like so:

The quick brown fox jumps over the lazy dog.

But the **paragraph** element can be modified by the **style** attribute, to produce all kinds of stylistic effects:

```
<p style="font-family:cursive; font-weight:bold;">The quick
brown fox jumps over the lazy dog.</p>
```

Here, the **style** attribute (highlighted in red), as applied to the **paragraph** element, specifies multiple values that will cause all content within this particular **paragraph** element to use the **cursive** font family; the text will also be **bolded**. When interpreted by a web browser, the above **paragraph** element would appear like so:

The quick brown fox jumps over the lazy dog.

Recall that elements can be nested; this means that any given element can include other elements.

For example: the **unordered list** (`<ul>`) and **list item** (`<li>`) elements are used to create lists in HTML documents. Once the **unordered list** element is opened, any **list item** elements within are interpreted by web browsers to create bullet points on a list, like so:

```
<ul>
     <li>Item 1</li>
     <li>Item 2</li>
     <li>Item 3</li>
</ul>
```

Interpreted by a web browser, this produces a bulleted list that would appear as follows:

- Item 1
- Item 2
- Item 3

As each of the three **list item** elements are nested within the **unordered list** element, anything that modifies the **unordered list** element *also* modifies the **list item** elements. For example:

```
<ul style="font-family:cursive; font-weight:bold;">
     <li>Item 1</li>
     <li>Item 2</li>
     <li>Item 3</li>
</ul>
```

Again, the **style** attribute has been used to modify the **unordered list** element (and everything nested within it), specifying **bolded** text using the **cursive** font family. This produces the following results

- **Item 1**
- **Item 2**
- **Item 3**

Note that attributes can also be used to further modify or even override the attributes of parent elements on an individual basis. For example, the **style** attribute could be used to make **Item 1** blue or Italicized, or it could be used to specify a different font family. These parameters would take precedence over those specified by the **style** attribute modifying the **unordered list** element.

XML elements can also be modified by attributes in the exact same way. For example:

```
<OperationsMetadata>
     <Operation name='GetCapabilities' />
```

```
</OperationsMetadata>
```

The **Operations Metadata** element (`<OperationsMetadata>`) contains metadata about the operations performed by a web service. Everything within this element, therefore, constitutes web service operation metadata.

In the above example, the **Operation** element (`<Operation>`) is modified by the **name** attribute (`name='GetCapabilities'`).

Here, the **name** attribute indicates that the name of one operation performed by the web service is **GetCapabilities**. Any content within the **Operation** element thereby constitutes metadata about the **GetCapabilities** operation. In this case, though, the **Operation** element is closed by a **forward slash** '/' and therefore contains no metadata.

Having made it this far, you are well on your way to understanding the basics of XML.

Section 4: XML Namespaces and Prefixes

Within XML documents, it sometimes becomes necessary to differentiate one element from another based on the context in which these elements are used.

For example, the `<table>` element is used differently between XML and HTML. Within an XML document in which both XML and HTML `<table>` elements are present, *namespaces* and *prefixes* can be used to differentiate XML table elements from HTML table elements.

Namespaces identify the context in which a given element is used, thereby differentiating that element from otherwise-identical elements.

Theoretically, *any* label can be used as a namespace, but XML namespaces should be URIs. Ideally, an XML namespace is an HTTP-capable URI that dereferences to a web-accessible description of the context identified by the namespace. Consequently, most XML namespaces in common use look like standard web addresses.

Going back to the above example, a namespace identifying HTML5 as the context for a given `<table>` element might appear as follows:

```
http://www.w3.org/TR/html5/
```

When entered into a web browser, this namespace dereferences to a description of HTML5.

Before a namespace can be used to differentiate a given element, the namespace must first be *declared* within the desired XML document. This is accomplished by means of the **XML namespace** attribute (`xmlns`); the namespace can be declared in two ways:

1. Declaring the namespace within the desired element, like so:

```
<table xmlns="http://www.w3.org/TR/html5/" />
```

2. Binding the namespace to an XML *prefix* that can be used as a proxy for the bound namespace within the element in which the binding took place (recall that nested elements inherit the characteristics of the parent elements in which they are nested)

Prefixes are short text strings that precede and differentiate markup language *tags*, *attributes*, and even the content of an element, like so:

```
<prefix:tag prefix:attribute="value">prefix:content</prefix:tag>
```

As indicated previously, prefixes serve as a proxy for a full namespace, but they can only do so within an element in which they have been bound to a namespace. In other words, a prefix cannot refer back to a bound namespace unless the prefix is used in an element in which the prefix has been bound to a namespace or which has inherited the binding from a parent element.

To give an example, let us assume that the `<table>` element we wish to differentiate via a bound namespace is nested within a `<body>` element. If we wished to declare a binding between a prefix and an appropriate namespace, we could do so either in the `<table>` element *or* within the `<body>` element within which the `<table>` element is nested, like so:

1. Binding takes place in the `<body>` element:

```
<body xmlns:html5="http://www.w3.org/TR/html5/">
     <html5:table>content</html5:table>

</body>
```

2. Binding takes place in the `<table>` element:

```
<html5:table xmlns:html5="http://www.w3.org/TR/html5/" />
```

In both examples, the **xmlns** attribute is used to bind the chosen prefix, html5, with the desired namespace, http://www.w3.org/TR/html5/; the prefix is then attached to the `<table>` element to invoke the namespace and thereby differentiate the `<table>` element from other `<table>` elements.

Typically, namespaces are bound to prefixes in parent elements and invoked via prefix where necessary in nested elements, except where it is absolutely necessary to bind a prefix to a namespace within the same element in which the namespace is invoked.

To give a more concrete example, color-coded in exactly the same manner as the previous example:

```
<wfs:WFS_Capabilities xmlns:wfs="http://www.opengis.net/wfs"
xmlns:ows="http://www.opengis.net/ows">
```

This is the root element (<WFS_Capabilities>) of the capabilities document that describes a National Geothermal Data System web feature service. Here, two namespaces are declared and bound to two different prefixes via the **xmlns** attribute:

- The **wfs** prefix is bound to the following namespace: http://www.opengis.net/wfs
- The **ows** prefix is bound to the following namespace: http://www.opengis.net/ows

Later on in the document, the following elements are nested within the root element in which the binding took place:

```
<ows:Abstract>Well header information for oil and gas wells
in Arizona.</ows:Abstract>

<wfs:Name>aasg:Wellheader</wfs:Name>
```

These prefixes indicate that the **Abstract** element is used in the context of an Open Geospatial Consortium OWS Common Implemetation Specification XML schema; likewise, the **Name** element is used in the context of an Open Geospatial Consortium web feature service. Information about the context of these elements can be found at the web location indicated by the namespaces to which the prefixes have been bound.

Section 5: XML Review

In this tutorial, we have established the following:

- XML is a markup language
- XML uses elements as basic structural components; elements are comprised by an opening tag, content, and a closing tag:

```
<tag>content</tag>
```

- Markup language elements can close themselves:

```
<tag />
```

- Elements can be *nested* within each other, much like folders on a computer:

```
<tag1>
     <tag2>

          <tag3 />
```

```
        </tag2>
    </tag1>
```

- Unless otherwise specified, nested elements inherit the characteristics of elements within which they are nested
- Markup language documents typically begin with a *root element* within which all other elements are nested
- Markup language elements can be modified by **attributes**; there can be multiple attributes per tag and multiple values per attribute:

```
<tag attribute="value" />
```

- Markup language elements can be differentiated from one another by *namespaces,* which identify a specific context in which a given element should be interpreted
- Namespaces must be declared using the **xmlns** attribute, either within a specific element or bound to a *prefix* that acts as a proxy for the full namespace
- Prefixes refer to and invoke a namespace that has been bound within a given element or the parents thereof; prefixes can modify tags, attributes, and content:

```
<prefix:tag prefix:attribute="value">prefix:content</prefix:tag>
```

With this background, it should now be possible to understand and interact with USGIN XML documents more easily.

This concludes the USGIN XML tutorial.

XML schema for Simple Feature WFS services in the National Geothermal Data System

January 23, 2015 SMR

In order to maximize the accessibility of WFS services deployed by the AASG geothermal data project, the first round of service deployments have been implemented as OpenGeospatial Consortium Simple Feature level 0 services. This document provides guidelines for construction of XML schema that implement geothermal domain feature and observation content models specified abstractly (independent of implementation approach) using lists of elements. AASG Geothermal data WFS Features consist of fields defined in a collection of Microsoft Excel content model workbooks accessible at http://www.stategeothermaldata.org/data_delivery/content_model_templates. These content models are in the process of being submitted for review and adoption for use by the wider DOE National Geothermal Data System and are presented as prototypes for US Geoscience Information Network (GIN) data services.

Implementation of the content models introduces some practical considerations that must be addressed to maximize their utility. The ESRI ArcGIS server that is being used by many data providers inserts an integer OBJECTID field as the first field in the feature content. This field should be populated with unique integers; although it can be omitted, ArcGIS WFS clients work better when an objectID field is included, so we are added this field in each feature class.

The feature geometry is included in an XML element named 'shape' that is always the last field in the XML. The geometry type for the shape is indicated in the FieldList entry for shape in the content model workbook.

The intention is that the XML schema is uniquely determined by the information listed in the FieldList worksheet in the Content model workbook. This includes the element name (which should be identical to the column heading in the template table or worksheet), a logical data type, and the xml data type that is used to implement the element in the schema. The field List includes rows for OBJECTID and shape elements, even though these are not (in general) populated in the template, but are supplied by the GIS into which the template data is loaded to deploy the WFS service.

This document provides a brief summary of provision in OGC specifications for conformance to the Simple Feature level 0 feature encoding using XML. There are two versions of the simple feature profile: 06-049r1 that uses GML 3.1.1, and 10-100r2 that uses GML 3.2.1.

Version 1 provisions

Abstracted from OGC 06-049r1 Geography Markup Language (GML) simple features profile, version 1.0

This profile prohibits use of gml:metaDataProperty elements for referencing metadata in instance documents

XML annotation required to specify the SF compliance level, placed at the top most nesting level of the schema such that the XPath expression:
`/xsd:schema/xsd:annotation/xsd:appinfo/gmlsf:ComplianceLevel`
can be used to determine the compliance level of the application schema

The following schema fragment shows how this annotation shall be declared in an application schema:

```
<xsd:annotation>
       <xsd:appinfo source=
"http://schemas.opengis.net/gml/3.1.1/profiles/gmlsfProfile/1.0.0/gmlsfLevels.xsd">
              <gmlsf:ComplianceLevel>0|1|2</gmlsf:ComplianceLevel>
              <gmlsf:GMLProfileSchema>
                     http://schemas.opengis.net/gml/3.1.1/profiles/gmlsfProfile/1.0.0/gmlsf[2].xsd
              </gmlsf:GMLProfileSchema>
       </xsd:appinfo>
</xsd:annotation>
```

The following XML fragment shows how to encode the GML application schema document's root element:

```
1 <xsd:schema
2 targetNamespace="target_name_space"
3 xmlns:prefix="target_name_space"
4 xmlns:xsd="http://www.w3.org/2001/XMLSchema"
5 xmlns:gml="http://www.opengis.net/gml"
6 xmlns:gmlsf="http://www.opengis.net/gmlsf"
7 elementFormDefault="qualified"
8 version="99.99.99">
```

In this profile, the gml:id attribute in the AbstractFeatureType is changed from optional in the GML specification [OGC 03-105r1] to mandatory in this profile.

In addition to the GML geometric property types defined in Table 6, this profile supports the following list of basic data types:
a) Integers (with optional specified value range and number of digits)
c) Measurements (that include a reference to units of measurement)
d) Character strings (with optional specified maximum length or fixed length)
e) Date
f) Boolean
g) Binary data
h) URI references to local or remote resources
i) References to other resources
j) Character strings whose values are from a code list
k) Reals (with optional specified number of digits, decimal precision and value range)

For compliance level SF-0, the attribute **minOccurs** can have the values 0 or 1 and **maxOccurs** can only have the value 1.

The following XML-Schema fragments show how to define property elements with content that is a measurement and includes a unit of measure in a GML application schema that complies with this profile:

```
1 <xsd:element name="propertyName" [minOccurs="0|1"] [maxOccurs="1"]>
2 <xsd:complexType>
3 <xsd:simpleContent>
4 <xsd:restriction base="gml:MeasureType">
5 <!-- … Optional facets described in sub-clause 8.4.4.2 … -->
6 </xsd:restriction>
7 </xsd:simpleContent>
8 </xsd:complexType>
9 </xsd:element>

SHORT FORM: <xsd:element name="propertyName" type="gml:MeasureType" [minOccurs="0|1"] [max-
Occurs="1"]/>
```

In all cases, geometry coordinates shall only be specified using the **gml:pos** for **gml:Point** or **gml:posList** elements for all other types

Simple feature v2 provisions:

Abstracted from OGC 10-100r2 Geography Markup Language (GML) simple features profile, version 2.0

Uses GML 3.2

Prohibits use of **gml:metaDataProperty** elements for referencing metadata in instance documents.

The maximum cardinality that can be declared in an application schema that complies with level SF-0 is 1. This means that the attribute **minOccurs** can have values of 0 or 1 and the attribute **maxOccurs** can only have the value of 1.

SF-0 does not allow the use of the **nillable** attribute when declaring elements in an application schema.

Application schemas that comply with level SF-0 can only declare, with the exception of geometric property types, feature properties that have simple or scalar values.

XML annotation required to specify the SF compliance level, placed at the top most nesting level of the schema such that the XPath expression:
`/xsd:schema/xsd:annotation/xsd:appinfo/gmlsf:ComplianceLevel`
can be used to determine the compliance level of the application schema

root element:

```
<xsd:schema
2   targetNamespace="target_name_space"
3   xmlns:prefix="target_name_space"
4   xmlns:xsd="http://www.w3.org/2001/XMLSchema"
5   xmlns:gml="http://www.opengis.net/gml/3.2"
6   xmlns:gmlsf="http://www.opengis.net/gmlsf/2.0"
7   elementFormDefault="qualified"
8   version="99.99.99">
```

As a GML application schema is revised, all first and second level revision changes shall require a new namespace identifier and thus a new value for the *target_name_space* placeholder

The following XML fragment imports the GML schema:

```
1   <xsd:import namespace="http://www.opengis.net/gml/3.2"
2    schemaLocation="http://schemas.opengis.net/gml/3.2.1/gml.xsd"/>
```

The following XML fragment imports the schema that declares the **ComplianceLevel** element used to declare the compliance level of an application schema:

```
1   <xsd:import namespace="http://www.opengis.net/gmlsf/2.0"
2        schemaLocation="http://…schema repository path…/gmlsfLevels.xsd"/>
```

An application schema compliant with this profile may import or include other application schemas based on this profile as long as the imported schemas have a compliance level declaration that is less than or equal to the importing schema's compliance level and are also compliant with this profile.

The following XML-Schema fragment shows how to define a feature type in a GML application schema that complies with level SF-0 of this profile:

```
1  <xsd:element name="FeatureTypeName"
2              type="[prefix:]FeatureTypeNameType"
3              substitutionGroup="gml:AbstractFeature"/>
4
5  <xsd:complexType name="FeatureTypeNameType">
6     <xsd:complexContent>
7        <xsd:extension base="gml:AbstractFeatureType">
8           <xsd:sequence>
9
10          <!-- … zero or more element declarations as described in 8.4.4 … -->
11
12          </xsd:sequence>
13       </xsd:extension>
14    </xsd:complexContent>
15 </xsd:complexType>
```

In addition to the GML geometric property types, this profile supports the following list of basic data types:

a) Integers (with optional specified value range and number of digits)
a) Measurements, that include a reference to units of measurement
b) Character strings (with optional specified maximum length or fixed length)
c) Date
d) Boolean
e) Binary data
f) URI references to local or remote resources
g) References to other resources
h) Character strings whose values are from a code list
i) Reals (with optional specified number of digits, decimal precision and value range)

The use of the attribute **nillable** is prohibited in compliance level SF-0 but is allowed for levels SF-1 (see 9.2.4) and SF-2.

Although XML-Schema [W3C XML-1] allows **minOccurs=maxOccurs=0**, this profile prohibits this combination.

The following XML Schema fragment defines the encoding of a property element whose content shall be from a list of values stored in one or more external dictionaries.

```
1 <xsd:element name="propertyName" [minOccurs="0|1"] [maxOccurs="1"]>
2    <xsd:complexType>
3       <xsd:simpleContent>
4          <xsd:restriction base="gml:CodeType">
5             <xsd:attribute name="codeSpace" type="xsd:anyURI" [use="optional"]
6                            default="…some uri…" | fixed="… some uri…"/>
7          </xsd:restriction>
8       <xsd:simpleContent>
9    </xsd:complexType>
10 </xsd:element>
```

default -- an instance document may point to an alternate value dictionary at runtime using the **codeSpace** attribute. **fixed** -- an instance document shall use values from the specified dictionary.

If the type is **xsd:decimal** then the facets *totalDigits* and *fractionDigits* listed in lines 5 and 6 shall be specified. If the type is **xsd:double**, then the facets *totalDigits* and *fractionDigits* listed in lines 5 and 6 shall not be specified.

Application schemas that comply with level SF-0 of this profile may freely use XML comments and annotations, using the **annotation** element, wherever they are legally allowed by XML-Schema and GML

NGDS Catalog Documentation

About

The National Geothermal Data System (NGDS) is a catalog of documents and datasets that provide information about geothermal resources located primarily within the United States (although information from other parts of the world is also included).

This complete and current catalog of available data, which is funded by the Department of Energy's Geothermal Technology Office, is designed to accelerate the development of U.S. geothermal resources, and can be used to:

- Determine geothermal potential
- Guide exploration and development
- Make data-driven policy decisions
- Minimize development risks
- Understand how geothermal activities affect your community and the environment
- Guide investments.

If you are new to this site, learn more about how you can get started using NGDS.

Below you will find more information about how the NGDS is typically used, who contributes data, and where it is hosted.

How is NGDS Typically Used?

The NGDS can be used in many ways, depending on your needs and interests. Generally, the NGDS is used by:

- Agencies, businesses, and researchers who wish to use the documents and datasets
- Individuals who want to contribute additional data
- Web developers who want to create custom applications that interact with NGDS data.

Learn more about typical uses and how you can get started using NGDS.

Who Contributes Data to NGDS?

NGDS data records are contributed by academic researchers, private industry, and state and federal agencies (including all fifty State Geological Surveys). Initially, five groups digitized their data and made their data records visible through the NGDS.

View a complete list of our NGDS data contributors on our Partners page or learn more about how to submit your data to the NGDS.

Where is NGDS Data Hosted?

Those who contribute data records to the NGDS host their data using their own computing resources according to the standards and protocols proposed by the U.S. Geoscience Information Network (USGIN).

The NGDS aggregates information about each of the documents and datasets, and makes these information records, which include direct access to most of the datasets, available to the public through the [[Map]] and [[Library]] interfaces in the NGDS.

This system allows NGDS contributors to retain complete control of their data, and allows the NGDS to be comprehensive yet scalable and decentralized.

If you're thinking of contributing documents or datasets to NGDS, learn more about the how to submit your data or read about our data models, exchange methods, and metadata.

What Data Formats are Available?

The documents and datasets you can find through the NGDS are delivered in the following ways:

Standardized, interoperable datasets that conform to interchange specifications are delivered through web services. These datasets are considered to be NGDS "Tier 3" structured data can be processed automatically by computer software with minimal user input. This type of data can be used by many different programs and in conjunction with any compatible datasets. For more information on data tiers and web services, please see the Data Models, Exchange Methods, and Metadata page.

- Documents and non-standardized datasets.

Although some of the data is available in non-standardized formats currently, the goal of the NGDS is to increasingly offer data that has been converted to standardized formats. To do this, we have adopted standards, protocols, and formats whenever possible.

What Standards Does the NGDS Meet?

NGDS metadata, which is used to describe each document or dataset, complies with International Organization for Standardization (ISO) standards. NGDS data and metadata can be transferred according to Open Geospatial Consortium (OGC) protocols such as Web Map Service (WMS), Web Feature Service (WFS), and Catalog Service for the Web (CSW).

Learn more about the NGDS data and metadata standards.

What Tools are Available to Access NGDS Data?

NGDS supports the popular ESRI Map Service protocol, so people may access the data through a variety of proprietary and free/open-source applications and software. See our Resources section for a listing of [[resources, models, and tools]] you can use to interact with NGDS data.

Getting Started

The National Geothermal Data System (NGDS) can be used in a variety of ways, but is typically used for:

- Finding U.S. geothermal documents and datasets
- Contributing new data
- Developing custom applications based on NGDS data

Below you'll find more information that will help you get started using the NGDS.

Finding and Using Data

The NGDS is commonly used by state and federal agencies, businesses interested in geothermal resource development, and individuals who are interested in knowing more about geothermal resources. Here's how you can use the NGDS to find and use our data:

- Use our Map to [[locate data records geographically]].
- Visit the Library to [[filter data records by keywords and other criteria]] such as category, author, or data type.
- The Resources section provides a list of [[other free and open-source applications]] that can also be used to interact with the NGDS data.
- The Help section is available to answer questions on the specifics of how to use the site Search, Map, or Library.

Contributing Your Data to NGDS

If you have information on geothermal resources, please consider making it available through the NGDS. Documents, datasets, or services that you catalog in the NGDS can be made openly accessible to the public, can remain privately accessible, or can be visible only to others within the contributing agency.

Here's how the process works:

- When you catalog your data with the NGDS, the data itself remains hosted on your own computing resources (or you can make arrangements for hosting with other NGDS nodes).
- A record of your data will be entered into the NGDS catalog through metadata records; these metadata records can then be accessed via the [[Map]] and [[Library]]. NGDS metadata describes a given data resource, any restrictions on its use, and how to obtain it.

Learn more about how to submit your data to the NGDS or read about our data models, exchange methods, and metadata.

Developing a Custom Application Using NGDS Data

NGDS is built on United States Geoscience Information Network (USGIN) standards, which allows Web developers to easily build innovative custom applications that use NGDS data. Some examples of third party applications might include data gap analysis, financial analysis, and other useful data manipulation.

Learn more about our [[data models, exchange methods, and metadata]] or [[find tools, forums, and developer communities]].

NGDS Exchange Methods

Metadata

All NGDS data records are organized and accessed using metadata. Metadata describes NGDS resources to support finding resources, evaluating them for a user's intended purpose, and obtaining the data. Each metadata record describes a resource that is considered an individual product with an author, editor or compiler, created with some particular purpose. The metadata record may describe a variety of 'distributions' or procedures for accessing the resource, such filling out an order form, clicking on a link, or linking to a web service with an appropriate client.

After hosting data using their own computing resources, NGDS data contributors submit one or more metadata records describing their data to an NGDS node. Metadata from NGDS nodes is harvested by the NGDS portal, which provides a single entry point to search the entire NGDS catalog (the catalog contains metadata describing information resources distributed across the various nodes in the network). Metadata records provide information necessary to obtain data through various distributions (file formats, web service protocols), allowing you to select the most appropriate way to access the software you are using. Metadata records in this catalog can be used to locate and access the data resources they describe.

Metadata records published by NGDS nodes are harvested by the central portal, which provides a web site for searching across all NGDS content, enabling users to discover NGDS data resources and learn how to access them. For more information about metadata, see the USGIN Metadata Tutorial.

Once the desired NGDS data resource is located using metadata that appears in the NGDS catalog, a user has several options for accessing that resource. Access options depend on the structure of the data resource. Typically, highly structured resources offer broader options for access and automated analysis.

Data Access Options

NGDS is uses the World Wide Web as a platform for accessing and exchanging data. Typical web-based data exchange methods are as follows:

- **File-based approaches**: Data is accessed in discrete digital packages called *files*. A file is located at specific URL (Web location). The content of a file can be structured to enable machine processing, but the file is only accessible as a unit.
- **Web application-based approaches**: Data is made available to users via Web-browser-based software. These applications typically provide a human user with functions to explore data with various kinds of visualizations, and in some cases to select particular data and download in files. These applications are typically tightly coupled to specific data sources.
- **Service-based approaches**: Data is available through *web services,* which provide interfaces for machine-to-machine interaction for finding, selecting and accessing data. Web services

require more sophisticated client and server software stacks and rigid quality control, but allow a greater level of automation in the data access and utilization process, fully taking advantage of interoperable interchange protocols and formats.

What Are Web Services?

A web service is a way to communicate data over the internet. Using web services allows NGDS data to be used in a variety of free-and-open-source applications, making it more freely available and easier to use.

NGDS purports to use web services with well-documented community specifications for service protocols and data interchange formats. Though web services require more sophisticated client and server software stacks and rigid quality control, web services have a number of advantages that meet the requirements of NGDS:

- Web services are web-accessible
- Data available as a web service can be consumed by web applications
- Web services are read-only and preserve data ownership
- A number of web service protocols are free-and-open-source, meaning that web services using these protocols can be hosted and accessed without proprietary software

Most NGDS web services provide functions to get data or metadata. The NGDS Node-in-a-Box (NIAB) software stack supports the creation and publication of Open Geospatial Consortium (OGC) compliant web services from data assembled in Excel spreadsheets.

Web Service Protocols in use with NGDS

- **WMS**: OGC Web Map Service (WMS). Provides raster images of georeferenced features for display in map views.
- **WFS**: OGC Web Feature Service (WFS). Provides structured data for georeferenced features including descriptive attributes.
- **ESRI Map Service**: Provides georeferenced rasters and features described by attributes. Requires ESRI ArcGIS server software.
- **CSW Applications**: OGC Catalog Service for the Web (CSW). Helps applications access metadata records used to locate NGDS information resources.
- **OPeNDAP**: Provides raster datasets as numeric array data.

As implemented in NGDS, these services provide the user read-only access to data published through the services. Transactional implementations allowing users to change content of the services are possible, but these are outside the scope of current NGDS specifications.

Interoperability

NGDS facilitates access to data conforming to the three aforementioned exchange methods through publication using interoperable protocols and formats. This enables greater automation of information

exchange, processing, and analysis, allowing human users to focus on understanding and learning from the data.

Data is categorized in one of three *interoperability tiers*; the higher the tier, the greater the level of interoperability. These tiers roughly conform to whether data can be accessed through a web service or must be downloaded or viewed as a file.

Tier 1: Unstructured data (text, images, or sound)

The simplest and most common access to resources is provided by simple Web links that result in a file download. Information contained in files can be accessed by users who have software that can recognize and open these files. This is the standard model for files accessible on the web, supported by HTTP servers and desktop web browser software.

Unstructured data requires user interpretation before it can be used for analysis. Users can utilize the information if they can understand the encoding and language, but the system provides no support for this understanding, and little or no automation is possible. Audio files must be transcribed; text files must be parsed and mined for data that is then broken down and structured in ways that can be processed by computers; images must be scanned, interpreted, and often georeferenced. Preparing Tier 1 data for analysis can be a painstaking and time-consuming process.

Tier 2: Structured data that does not conform to an NGDS schema

Tier 2 interoperability indicates that information content is structured (consistently organized) in a spreadsheet or database file such that it is amenable to computer processing; that said, Tier 2 data does not use a shared, documented interchange format. Data in this tier must be transformed by the data consumer on a case-by-case basis for integration with other datasets, requiring them to study each new data source to figure out how to extract the information they need. Obtaining data in a structured format is a step towards interoperability because once the format is understood, computer programs can be instructed to extract the desired information.

Tier 3: Structured data that conforms to an NGDS schema

Tier 3 data is structured data that conforms to an NGDS information exchange. Data that is published according to the exchange specification (content model, interchange format, service protocol) is interoperable with any other data published using that exchange. This is referred to as Tier 3 interoperability.

Creating Interoperable Data

Tier 3 data structured using a content model is among the most highly interoperable data found on the web. Tier 3 data can be accessed and analyzed automatically through a number of web services. NGDS uses content models based on USGIN-adopted standards and protocols to structure data to the Tier 3 format.

USGIN: Powering NGDS

The goal of USGIN is to facilitate the development of interoperable geoscience data-sharing networks that improve public access to geoscience information and aid scientific queries. For more information visit the USGIN website.

USGIN proposes:

- Protocols and standards.
- Profiles for the utilization of USGIN protocols and standards.
- Development tools to build distributed data-sharing networks using open-source software and existing World Wide Web infrastructure.

USGIN Architecture

To promote interoperability, USGIN adopts existing standards, protocols, and formats whenever possible. These include:

- Standards for metadata: ISO 19115, ISO 19119, and ISO 19139
- Protocols for data exchange:
  - Open Geospatial Consortium (OGC), Web Map Service (WMS), Web Feature Service (WFS), and Catalog Service for the Web (CSW)
  - ESRI Map Service
  - OPeNDAP
- GeoSciML-Portrayal schemas for geologic map data.

NGDS is based on protocols and standards adopted by USGIN. All NGDS content models are structured according to USGIN standards.

Publish Data

How to Submit Data to the NGDS

NGDS Node in a Box

The NGDS Node-in-a-Box (NIAB) is a software package designed to enable quick deployment of an NGDS node. The software installation includes a CKAN instance along with NGDS-specific extensions for conformance with NGDS metadata requirements, support for NGDS information exchanges, and publication of data through OGC services using a linked GeoServer instance. The NGDS NIAB CKAN extension provides users with a web-accessible user interface designed to help NGDS contributors submit and manage metadata records.

Installing the NIAB

Instructions on installing and configuring a NIAB instance are available with the following document:

http://ngds.github.io/install-and-run/NGDS_Installation_Documentation.htm

This document and other helpful information, including community conversation, are available at NGDS CKAN GitHub.

Finding an Existing NIAB

In addition to installing and configuring a specific Node-in-a-box instance, a registered user may contribute data using a public account on an existing Node-in-a-box if permission is granted from the Node-in-a-box owner. The data contributor is responsible for contacting the owner of the NIAB for permission. A listing of available NIABs, owners, and contact information is in development.

Overview: Becoming an NGDS Node

To become an NGDS node, one must place file-based information resources in a web-accessible location and then submit metadata describing these resources to the NGDS metadata catalog. Various workflows are available to create and submit metadata:

- Log in to an existing registered instance of the NGDS Node-in-a-Box (NIAB) and add metadata records through the provided interface.
- Set up and register an NGDS NIAB instance. This requires:
    1. Installing and configuring the NGDS NIAB software (see Installing the NIAB section above).
    2. Registering an NGDS NIAB instance with the NGDS portal (main aggregator) via email: metadata@usgin.org.
    3. Creating metadata records using the interface provided by your NGDS NIAB instance.
- Deploy some other metadata management tool that provides a CSW interface to NGDS-conformant ISO metadata for registered resources, and registering the catalog service with the NGDS central portal for harvesting. This will require contacting system adminstrators via email at metadata@usgin.org to harvest your CSW into the main aggregator at http://www.geothermaldata.org.
- Create NGDS-conformant ISO metadata records using any other tool and place them in a web-accessible folder that is registered with the NGDS portal for harvesting. Contact system adminstrators via email at metadata@usgin.org for more information.

How to Submit Your Metadata Records via a pre-installed NGDS Node-in-a-Box (NIAB)

Data providers can publish their data to NGDS via any NGDS Node-in-a-Box instance for which they have an account with data publication permission.

Connect to an NGDS NIAB using a web browser and click **Contribute** and populate the appropriate forms to submit a metadata record to the catalog. During this process, either upload a resource or specify the web location of an existing web-accessible data resource. Those providing data to a NIAB that is managed by another entity should make appropriate arrangements with the managers of the node

regarding future edits to that data; those who manage their own NIAB have complete control over the data submitted then harvested to the main aggregator for larger use in the system. The nodes (NIAB installations or other node types) will be harvested at intervals, so any edits made at a node will be reflected in the system with the next scheduled harvest.

Step 1: Create a Collection

The **Contribute** page leads to the **Create a Collection** form; as the name implies, populating this form will create a collection for a single or multiple metadata records. This is the first step in the file or resource upload process, which will create metadata in the system for that resource. The fields on this form are as follows:

- **Title**: Collection title
- **URL**: The URL at which this collection will be found. This URL is automatically generated when the **Title** field is populated, but it can be modified manually
- **Description**: A description of the collection and the metadata records stored in it
- **Tags**: Keywords that can be used to search for this collection and any metadata records stored in it
- **License**: The license under which any dataset(s) described by any metadata record(s) within this collection are published; among other things, the license determines the manner in which the dataset described by the metadata in this collection can be reused. Information on supported licenses can be found at the following locations:
    - Creative Commons
    - GNU Free Documentation License
    - Open Data Commons
    - UK Open Government License

When finished, click **Next: Add Data**.

Step 2: Add Data

Clicking **Next: Add Data** on the **Create Collection** form leads to the **Add Data** form. Populating this form creates a metadata record describing a specified data resource; this metadata record is added to the collection created in Step 1.

Four types of data may be upload. See Data Models, Exchange Methods, and Metadata page for helpful text on determining categories (tiers) of data which may be uploaded.

- **Structured Resource**: Tier 3 structured and standardized data (CSV files). Files submitted here must be “schema-valid”, which refers to the file as is conforming to the standards of a given information exhange schema. The blank Excel files for a given information exchange are found at http://schemas.usgin.org/models/. When data conforms to these specifications (field headings, data types, etc.) the file can be validated at http://schemas.usgin.org/validate/cm before up-loading to the node as a tier 3, structured dataset. After file upload, the option exists to publish the data as OGC web services (WFS and WMS), which have the greatest utility in the system. Once a “schema-valid” file is

uploaded, one must choose the appropriate information exchanges used from the **Content Model** drop-down lists.

- **Unstructured Resource**: Tier 1 and 2 data (text, images, structured (but not standardized) files like CSV or Excel, PDF files, etc.)
- **Offline Resource**: Metadata record would describe resources which do not exist online such as a core repository.
- **Data Service**: Register already published web services such as WMS, WFS, CSW, and OPeNDAP by providing the GetCapabilites or other URL in the **Resource** field.

When the appropriate fields have been populated:

- Click **Save & add another** to add another metadata record to the collection; this will repeat the steps described in Step 2.
- Click **Next: Additional Info** to finalize process by adding collection metadata.

Step 3: Collection Metadata

Clicking **Next: Additional Info** leads to the **Collection Metadata** form. Here, provide metadata that applies to the collection created in Step 1 and the metadata records within it created in Step 2.

The **Collection Metadata** form includes the following fields:

- **Visibility**: Determines whether or not this collection and the metadata record(s) within it are visible to the public
- **Authors**: The authors of the data resource(s) described by metadata record(s) in the collection
- **Maintainer**: The individual, agency, company, or organization responsible for the data resource(s) described by metadata record(s) in the collection
- **Spatial Location**: A rough approximation of the features in the data resource(s) described by metadata record(s) in the collection
- **Category**: A description of the type of data resource(s) described by metadata record(s) in the collection
- **URI**: A *uniform resource identifier* for the collection; see the USGIN URI Tutorial for more information
- **Quality**: A rough assessment of the quality of the data resource(s) described by metadata record(s) in the collection
- **Lineage**: The provenance of data resource(s) described by metadata record(s) in the collection
- **Status**: The workflow status of data resource(s) described by metadata record(s) in the collection
- **Publication Date**: The date on which data resource(s) described by metadata record(s) in the collection were published
- **Language**: The language in which the data resource(s) described by metadata record(s) in the collection are written

- **Dataset Extent**: The physical extent of features in the data resource(s) described by metadata record(s) in the collection; click the **Draw a rectangle button** ( ) to draw a box indicating the dataset extent

When done, click **Finish**. Doing so will finalize this collection and all metadata records within it; these records will be added to the catalog.

For **Structured Resources**, continue with the following one-click step to publish the resource as OGC WFS and WMS services:

Click **Publish as OGC.**

Curate your data contributions

Anyone can contribute data to NGDS by establishing web-accessible data resource(s) and then submitting metadata record(s) to a registered NGDS node. Metadata submitted to a registered node will be harvested by the NGDS portal, thereby becoming part of the NGDS catalog. Participants may also contribute data by installing and configuring a Node-in-a-box built on the CKAN platform. Users who choose to install the Node-in-a-box will independently self-host their own data, making it available to the NGDS network through common standards and protocols. NGDS offers numerous opportunities to developers who want to use this data to create specific applications. Please see the Develop with NGDS page for more information.

Those who wish to submit data to the NGDS should view the How to Submit Data to the NGDS page.

Those who wish to learn more about NGDS exchange methods, standards and protocols should view the Data Models, Exchange Methods, and Metadata page.

Retain Data Ownership

NGDS is designed from the ground up to allow contributors/nodes to maintain ownership and control of their data. It is the responsibility of the data provider to maintain the information they publish and look after its long-term stewardship.

Metadata records listed in the NGDS metadata catalog are publicly accessible, allowing anyone to learn of the existence of an NGDS information resource and the license under which the described resource can be viewed, distributed, and reused. Information on supported licenses can be found at the following locations:

- Creative Commons
- GNU Free Documentation License
- Open Data Commons
- UK Open Government License

Tools & Applications

Once you've used the National Geothermal Data System (NGDS) Map or& Library to locate the data records you are looking for, you can use the resources, models, and tools described here to interact with the data.

Here are the categories of tools and resources available:

- Geothermal exploration
- Financial
- Geographic Information Systems (GIS) and data applications

Have you developed a custom application to analyze NGDS data? If you have used NGDS data to develop a third party application that could also benefit others, pleaseshare it with us so we can include it here.

Geothermal Exploration

The following tools and resources can be used to facilitate your geothermal exploration.

- NREL Geothermal Prospector: A Web-accessible mapping tool developed for the geothermal power industry. This tool is designed to provide easy access to geothermal resource datasets and other data relevant to utility-scale geothermal power projects. Geothermal Prospector can load NGDS Web Map Services (WMS) and Web Feature Services (WFS) to facilitate exploration - e.g., Reservoir models.
- NGDS Data Explorer: A Web application designed to load and display NGDS WFS from a text or location search. When a WFS has been loaded, a user can then export attributes (standardized data) from this service as an HTML table, CSV file, or Microsoft Excel workbook. Please note that the user can load more than one WFS at a time.

Financial

The following tools, models, and resources can be used to predict and provide information about the costs of geothermal development.

- GeoFRAT: Geothermal Financial Risk Classification and Assessment Tool (GeoFRAT) is an NGDS application that helps users assess the resource potential and financial risks that may be associated with producing electricity from a given geothermal resource. When used in conjunction with other widgets and tools such as GETEM, GeoFRAT can help you conduct detailed risk analysis for a geothermal resource development project.
- Geothermal Electricity Technology Evaluation Model (GETEM): This tool models the estimated performance and costs of currently available U.S. geothermal power systems. GETEM can analyze and evaluate currently available technologies and estimate the cost of certain technologies five to 20 years in the future. The model is intended to help the

Geothermal Technologies Office (GTO) determine which proposed RD&D programs and projects offer the greatest prospective return on investment (ROI).

Geographic Information Systems (GIS) and Data Applications

These Web, desktop, and GIS software applications can be used to access, explore, visualize, and analyze NGDS data. Some GIS software is free or open source.

Free and Open-Source Desktop GIS Software

- gvSIG: Desktop GIS application designed for storing and analyzing data in compliance with Open Geospatial Consortium (OGC) standards.
- GrassGIS: Desktop GIS software suite used for geospatial data management and analysis, image processing, graphics and maps production, spatial modeling, and visualization. GrassGIS runs on Mac OSX, Linux, and Windows.
- QGIS: Desktop GIS software available under the GNU General Public License. QGIS is an official project of the Open Source Geospatial Foundation (OSGeo). QGIS runs on Linux, Unix, Mac OSX, Windows, and Android, and supports numerous vector, raster, and database formats and functionalities.
- uDig: Desktop GIS software available under the Eclipse Public License 1.0 (EPL-1.0) and BSD 3-Clause License. uDIG is a standalone application that provides a Java solution for desktop GIS data access, editing, and viewing. uDIG runs on Windows, Mac OSX, and Linux.Proprietary Desktop GIS Software.
- ArcGIS: Desktop GIS software designed to facilitate geoprocessing, geospatial data editing, and cartographic production.
- ArcGIS Explorer: ArcGIS Explorer Desktop is a free GIS viewer that gives you an easy way to explore, visualize, and share GIS information. This free software supports adding layers from ESRI's ArcGIS online catalog, KML files, and GIS Services.
- ArcGIS Online - ArcGIS Online is a cloud-based, collaborative content management system for maps, applications, data, and other geographic information. ArcGIS Online supports ArcGIS Server Web Services, OGC WMS, KML and CSV.

Data Applications

- GeoScene 3D: GeoScene3D is an application for 2D and 3D visualizing, interpreting, editing and publishing geological data, including drill or bore holes, terrain, and geologic units. This application is intended for 3D viewing, so there is not an extensive base map selection. Although additional steps are required to view Web map services in this application, it is WMS-aware.
- Geozilla: A free application that provides 3D terrain views with Web map service overlays.
- Google Earth: A free application that provides 3D terrain views with Web map service overlays.
- Worldwide Telescope: A free application that provides 3D terrain views. This application does not yet support WMS, but it can consume .CSV files to provide visualizations events as data points over time. These data points can be assigned different colors and sizes for

further visualization of data, though doing so can require extensive preparation. See more information on importing data.

State Geological Survey Web Map Applications

- NREL Geothermal Prospector - http://maps.nrel.gov/gt_prospector
- GETEM, DOE Geothermal Technologies Office - http://www1.eere.energy.gov/geothermal/getem.html
- GeoFrat, Energy and Geoscience Institute at the University of Utah - http://geofrat.egi.utah.edu/
- University of Nevada, Reno Interactive Geothermal Maps - http://www.nbmg.unr.edu/Geothermal/InteractiveMaps.html
- Washington State DNR Resources Map Application - https://fortress.wa.gov/dnr/geology/?Theme=geothermal
- Oregon DGMI Geothermal Data Map Application - http://www.oregongeology.org/sub/gtilo/index.htm
- West Virginia Geothermal Interactive Map - http://ims.wvgs.wvnet.edu/WVGeothermal/viewer.htm
- AASG - http://search.stategeothermaldata.org/
- SMU Geothermal Databases - http://geothermal.smu.edu/gtda/
- USGS Geothermal - http://energy.usgs.gov/OtherEnergy/Geothermal.aspx, http://viewer.nationalmap.gov/viewer/
- Geothermal Data Repository - http://gdr.openei.org/

Help

Here you will find additional information that can help you successfully use the following site features to access data through the National Geothermal Data System (NGDS):

- Search feature
- [[Map interface]]
- [[Library]]
- Managing Datasets
- [[Software for NGDS]]
- Frequently asked questions

Search Feature

This site's search feature provides access to a catalog of available geothermal resource data by indexing a database of metadata about those resources.

Metadata is basically "data about the data" – it is keywords and terms that the data contributors have used to describe each data resource. When you type a search word or phrase into the search box, the

NGDS compares it to the metadata records and returns the most appropriate records. Information on how to access a given data resource is contained in the metadata records returned in the search results.

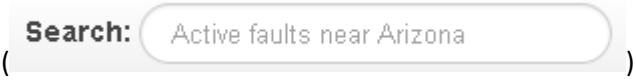
Map Interface

The [[map interface]] is typically used when you want to see a geographic distribution of data sites that may be of interest to you.

You can use the map to search the NGDS either by typing a keyword into the search box or by using your mouse to draw a "bounding box." A bounding box delineates geographic parameters on the map and the system retrieves the records that fall within those parameters.

Search results are displayed in list form and as colored circles on the map.

Map controls are as follows:

- **Search by Keyword** (

): Enter search terms and press the **Enter** key; the results will appear in list form beneath the search field as well as on the map.
- **Zoom In** ( ): Zooms in on the map.
- **Zoom Out** ( ): Zooms out.
- **Draw a Rectangle (or "Bounding Box")** ( ): Use this tool to search for results within a given area by drawing a rectangle on the map. Any record describing features within the bounding box you drew will appear as search results.
- **Zoom to Extent** ( ): Zoom to the original extent of the map.
- **Active Layer Menu** ( ): Mouseover for a toggle-able list of active layers in the map. This menu is populated as users choose to add available WMS resources to the map from their search queries.

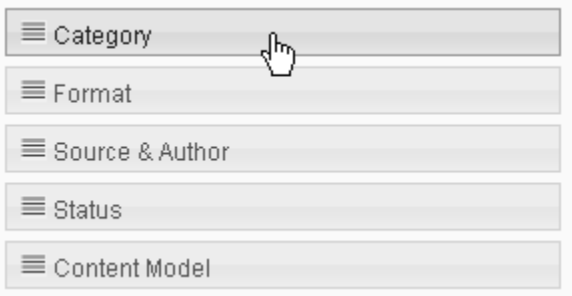
- **Search Results Menu** (): Toggle a menu of search results. This widget has some automatic capabilities; if a user makes a query by bounding box or keyword, the search results menu will automatically toggle itself.

Library

The [[Library interface]] allows you to use keywords and phrases to search the data records. Once you have results, you can narrow them to find the records you are looking for by choosing a set of criteria by which to filter your results.

Narrowing Your Search Results

1. Type a keyword or phrase into the search box to receive an initial set of results.
2. Choose one or more of the filter topics from the left side of the screen:



3. Within each topic, check the box by the criteria you wish to use as your filter(s):

   Selecting a check box will automatically update your search results to show only data that match the criteria you have selected. Note that you can use multiple criteria simultaneously to filter your results.

4. To remove a filter, select the name of the filter at the top of the screen:

You can filter your search results using the following topics:

- **Category**: Filter results by category and subcategory. For example, the category “drilling technology” includes a variety of subcategories related to drilling, such as the presence or absence of casing diagrams.
- **Format**: Filter results by format of the documents or datasets. Formats include data types and online resources, such as Web services.
- **Source & Author**: Filter results by data source and/or author.
- **Status**: Filter results by the status of the data within the context of the NGDS project. Data status tends to reflect the state of completion of that data. For example, typical NGDS documents or datasets will be either completed, ongoing, or deprecated.
- **Content Model**: Filter results by the content model to which the data conforms. All Tier 3 NGDS data conforms to a specific content model, which reflect properties of a given data set. A search for a specific content model would return results of data structured to conform to that content model.
- **Public/Private**: Filter results by the visibility of the data. When published, data contributors can designate whether or not a given document or dataset is publicly accessible or private. A setting of “private” renders the dataset invisible in public searches.

Managing Datasets

Delete a Dataset

In order to delete a dataset, click the **Edit** button on the page for the dataset. At the bottom of the page click **Delete.**

This does not permanently delete the dataset but does remove it from the Library. Since the dataset is not permanently deleted, the dataset's URL is not available for use by a new dataset.

To permanently delete a dataset go to the [[Node Admin]] page and purge the datasets. Now the dataset's URL is available for use again.

USGIN URIs

A Primer

11/21/2012

Table of Contents

Section 1: Introduction 2

Section 2: What is a URI? 2

Section 3: URNs and URLs 2

Section 4: USGIN URIs 3

Section 5: USGIN URI Syntax 4

Section 1: Introduction

Millions of database records and numerous datasets are shared according to USGIN specifications in a distributed, interoperable network. Organizing and maintaining uniqueness among these records and datasets requires a robust system of identifiers that is well-suited for the digital medium.

This document provides a general introduction to URIs and demonstrates URI applications within USGIN.

Section 2: What is a URI?

A Uniform Resource Identifier (URI) is a compact string of characters for identifying an abstract or physical resource (Berners-Lee ta al, 1998, rfc2396). Breaking this down into manageable chunks:

- URIs are *uniform:* they always use or conform to specific Internet Engineering Taskforce (IETF) URI syntax
- URIs identify a given resource.

Consequently: at their most basic level, URIs use a *consistent syntax* and *identify a resource*. Though multiple URIs can identify a given resource simultaneously, it is *imperative* that a single URI should not identify multiple different resources. Hence, each URI should be *globally unique.*

Generic IETF URI syntax may be found here; World Wide Web-specific IETF syntax may be found here.

Section 3: URNs and URLs

URIs come in two flavors: URNs and URLs. URNs and URLs are both URIs, and a URI can be a URN and a URL simultaneously.

- If a URI identifies a resource regardless of the location of that resource, then the URI is a **URN (Uniform Resource Name)**
- If a URI identifies a (web) location at which a bound resource may be found, then the URI is a **URL (Uniform Resource Locator)**

Though URIs (including URNs and URLs) must follow the appropriate IETF syntax to be considered URIs, most people are familiar with a number of URI analogs.

For example: Social Security numbers and car license plate numbers function as URNs, in that they use consistent syntax and identify resources without specifying that resource's location. Each individual Social Security number uses consistent syntax (NNN-NN-NNNN) and provides an identifier for a human resource (an individual person). Similarly, a car license plate number uses a combination of 6-7 letters and numbers and identifies a specific vehicular resources (a car). In both cases, the identifier does not indicate the location of the resource: a Social Security number does not indicate the location at which a

specific person may be found, nor does a license plate number indicate the location at which an individual car may be found.

Telephone numbers and street addresses are analogous to URLs, in that they follow a consistent syntax and identify a location at which a resource may be found. For example, street addresses use consistent syntax (NNNN [Street Name] [City] [State] [Zip Code] in the United States) and identify the location at which a resource (a shopping center, business, or office) may be found. Telephone numbers also function as URLs: they follow consistent syntax (NNN-NNN-NNNN) and identify a location (on the switchboard) at which a human resource (a person) may be found. In both cases, resources are associated with an identified location by means of a binding: a business resource is bound to a street address in the same way that a human resource is bound to a telephone number. If a business moves to another address or a person is no longer associated with a given phone number, then the binding is invalid.

All web addresses entered into your web browser are URLs. The web site http://www.google.com is a URL that identifies the location at which the Google search engine (a computing resource) may be found.

As with URIs, IETF-compliant syntax is available for both URNs and URLs.

Section 4: USGIN URIs

Like URLs, USGIN URIs can be entered into a web browser. Entering a USGIN URI into a web browser produces a *representation* of the resource identified by the URI. A representation is a symbolic proxy for something.

For example, a fault in the middle of the Mojave Desert can be represented by a symbol on a map, a digital photograph, or a text description of the fault. A USGIN URI identifying this fault would typically dereference to an XML representation of the fault (Figure 1). This XML document would provide structured information about the fault's attributes, including the fault's age, dip, and slip. These attributes can be used for analysis.

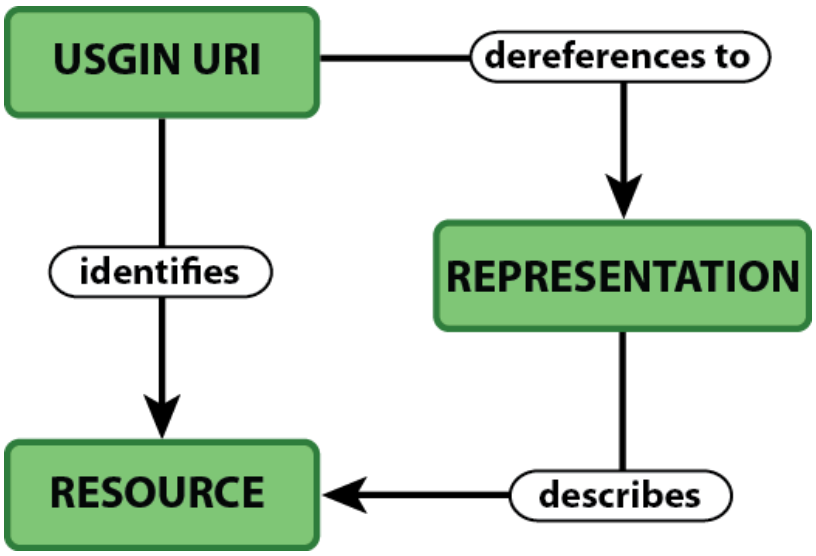


Figure 1: The relationship between USGIN URIs, XML documents, and the resource identified by the USGIN URI

However: even though they dereference to a web page when entered into a web browser, USGIN URIs are **not** URLs.

URLs identify a location at which a bound resource may be found. When entered into a web browser, a URL instructs the web browser to go to the indicated location and retrieve whatever resource is there.

By contrast: USGIN URIs are not URLs because the resource they identify is neither the representation found at the web location to which the USGIN URI dereferences, nor the web location at which said representation may be found

Returning to the example of a fault in the middle of the Mojave Desert: the *fault* is the resource identified by the USGIN URI. The USGIN URI does not identify the *representation* of said fault, nor the location at which said representation may be found. The USGIN URI identifies only the fault; the representation of the fault can be considered an intentional incidental.

Section 5: USGIN URI Syntax

As described above, USGIN URIs identify resources and use HTTP-compatible syntax to produce representations of said resources in web browsers. A USGIN URI can be broken down into a number of component tokens, which are represented in simplified form below:

http://host/uri-gin/authority/resource-type/resource-specific/

The tokens that comprise USGIN URIs are as follows:

- **http://** This token specifies the HTTP protocol for the URI.

- **host/** This token is a standard World-Wide Web domain name that is used to locate a particular server that has the necessary software to dereference the URI and return a representation of the resource identified by the URI. Currently, all USGIN URIs are dereferenced through the host located by:

    resources.usgin.org

- **uri-gin/** By USGIN convention, this token indicates that the URI is a USGIN URI. All URI components following this token are terminal components of the USGIN URI. The terminal components of the USGIN URI often reflect or incorporate existing resource identifiers that are not URIs, such as American Petroleum Institute (API) numbers. This is only practical, as existing resource identifiers are often sufficient for the purposes of USGIN - why reinvent the wheel?

- **authority/** This token identifies the local naming agent responsible for issuing and maintaining any resource identifiers that are present within the the terminal components of the USGIN URI. For example, if a USGIN URI identifies an oil well with an API number, and if the URI incorporates that API number into the URI, then the **authority/** token would reflect the authority responsible for issuing and maintaining API numbers: the American Petroleum Institute (or their agent in a particular state), abbreviated API. Examples of authority abbreviations in the context of USGIN URIs are as follows:

    - **American Petroleum Institute (API)**: http://host/uri-gin/api/
    - **United States Geological Survey (USGS)**: http://host/uri-gin/usgs/
    - **Arizona Geological Survey (AZGS)**: http://host/uri-gin/azgs/
    - **Department of the Interior (DOI)**: http://host/uri-gin/doi/
    - **Department of Energy (DOE)**: http://host/uri_gin/doe/

- **resource-type/** This token indicates the type of resource identified by the URI. Note that this token may itself be hierarchical, with resource subtypes delimited by the path separator character '/' Sample USGIN URI conventions can be found below:

    - Active or quaternary fault resource: http://host/uri-gin/authority/fault/
    - Borehole temperature resource: http://host/uri-gin/authority/bhtemp/
    - Thermal spring resource: http://host/uri-gin/authority/thermalspring/
    - Well header resource: http://host/uri-gin/authority/well/

- **resource-specific/** This token uniquely identifies a resource. Literally anything is possible within this component of the USGIN URI, although USGIN recommends constructing the string systematically, using existing resource identifiers and URL-safe characters. A list of URL-safe characters for resource-specific strings is as follows:

- 0-9
- a-z
- A-Z
- /
- - (dash)
- _ (underscore)

Consequently, a finished URI that follows USGIN specifications and identifies a specific car (vehicular resource) might appear as follows:

http://resources.cars.org/uri-gin/azdmv/vehicle/ABC123

Likewise, a finished URI that follows USGIN specifications and identifies a military aircraft (military resource) might appear as follows:

http://resources.airforce.gov/uri-gin/usaf/vehicle/aircraft/f22/1234567

This concludes the USGIN URI tutorial.

Archived USGIN Lab Discussions

Best Practices

Best Practices for USGIN Web Service Hosting

This blog is designed to encompass technical questions and troubleshooting for deploying and hosting USGIN services, particularly related to the AASG State Geothermal Data Project. Topics covered will included basic deployment procedures with WFS and WMS specifics, custom capablilites, handling large volume datasets, service and layers naming conventions, URIs, hosting online related files, metadata in the USGIN repository, and validating WFS and WMS services. The technical team at AZGS has encountered a need for such discussion as individual institutions and states begin hosting thier own USGIN AASG Geothermal Data Project services.

Convert File GDB to PostGIS Database

The goal of this post is to get our data from a File GDB to a format that GeoServer can use, specifically a PostGIS GDB. If we try and use a shapefile, the field names will be truncated and this will invalidate your schema that you have worked so hard to conform to the USGIN schema standards.

*Only works for 10.x and up geodatabases. You need to upgrade your gdb to 10.x if you are using an older geodatabase.

Geoserver App-Schema Set-Up: In Progress

Submitted by ccaudill on Fri, 02/01/2013 - 8:53am

One hurdle to using Geoserver to deploy AASG Geothermal Data services is the resolution of schemas. Creating additonal configuration files for Geoserver allows the use of "app-schemas" which will automatically download the indicated web-accessible schema (.xsd) and be referenced in the GetFeature request. This would solve a problem with service deployment using ArcGIS, where custom schema locations is necessary.

This blog will document the problems and break-throughs I've had so far in implementing Geoserver app-schemas. I have not yet been able to get schema resolution, but would like to make a post so that others can take into account my work thus far and chime in about other ideas or advances. This work is based on Tomcat running Geoserver, set up on an Oracle VM VirtualBox. The data for the services are pulled from a PostGIS database.

Service Deployment Using ArcGIS Server for WMS/WFS services

Once you have your data in the appropriate format, and the required fields populated, the next step is to create a service for the data. When it comes to AASG or OneGeology services, there are naming conventions and metadata rules.

If you are deploying a OneGeology WMS, please visit: http://onegeology.org/wmsCookbook/home.html for more detailed requirements. For AASG naming conventions, please visit this link

Web Service and Layer Naming Conventions

Submitted by ccaudill on Thu, 09/13/2012 - 11:06am

The AASG State Geothermal Data project provides OGC (Open Geospatial Consortium, opengeospatial.org/standards) compliant WMS and WFS services per the specifications outlined by the USGIN (see the USGIN site for more detials). AASG adopted the requirements and recommendations of USGIN for networking, data sharing, and enhanced interoperability. Such recommondations include naming conventions; the following outlines AASG-adopted naming conventions which require a specific format for naming Web Services and the Layers that contain the data within that service.

Creating WFS / WMS capabilites using ArcGIS service publishing wizard

Submitted by ccaudill on Mon, 08/27/2012 - 8:30am

When deploying a web service using ArcGIS publishing service, there are a few specifics that need to be considered when constructing the WFS and WMS Capabilites in the ArcGIS deployment wizard. This post is geared toward publishing a service for the AASG Geothermal Data Project, but the following practices are similar when deploying any USGIN-specification services using ArcGIS publishing products.

In the ArcGIS service publishing wizard or "Add New Service" wizard, you will be promted to choose Capabilities. Capabilites for both WMS and WFS are provided in AASG Geothermal Data Project services.

Validating your WFS and WMS services

Submitted by ccaudill on Fri, 08/17/2012 - 3:56pm

After a web service has been deployed, testing must be done to ensure that the service will work in the system, that is, be interoperable and available to end users. This group post will discuss using WFS GetFeature requests and WMS GetCapabilities requests for service validation. All OGC (Open Geospatial Consortium) services offer these requests which return an XML representation of the data. Some of the following information is taken from GeoSciML Portrayal Cookbook (available in full here http://usgin.org/sites/usgin.org/files/GeoSciML-PortrayalCookbook.pdf).

GetCapabilities Requests

If your service is deployed using ArcMap products, the request starts from your service rest page. Scroll to the "WFS" and "WMS" links on the service rest page. The URL given from these links is the Get Capabilities request, which consists of a service endpoint part, and the OGC request part:

ArcGIS-generated fields and AASG content model fields order and type

Submitted by ccaudill on Fri, 08/17/2012 - 10:02am

When deploying your service using ArcMap products, two fields are automatically generated. These are the OBJECTID and SHAPE fields. The XML schemas (http://schemas.usgin.org/schemas/) that are

associated with each AASG content model have a very specific ordering of the fields from that content model, and included in this specific order are the ArcMap-generated fields.

To acheive validation of your WFS service, and for your service's data to be interoperable and available within the system, the fields must be in the specific spelling, order, and data type that are assigned in the XML schema. These are usual errors we find when checking the WFS validation of hosted services.

Even though every content model has different fields, the ArcMap-generated fields are reliably in the same place in the field order. OBJECTID must come just **before** all the content model fields, and SHAPE must come directly **after** all the content model fields. (A "lax" function does allow Shape_length or Shape_Area or other fields to come after SHAPE field.)

Implementation of a Content Process "lax" function Element for AASG Schemas

Submitted by ccoleman on Fri, 03/23/2012 - 1:38pm

Discussion of using the "lax" function in Schema Documents

We've recently changed how we are structuring our XSD schemas by adding a "lax" process function element at the end of the sequence. The 'lax' function, or 'any Element' function, can be specified by the author of the schema document, and enables additional elements not specified by the schema to be included.

Data Interchange Document XML Schema Validation

Submitted by ccoleman on Thu, 03/22/2012 - 11:20am

Validation of XML data interchange documents with normative schemas using free, open-source software (XML Explorer)

One of the objectives of the US Geoscience Information Network (USGIN) is to establish a community of practice for developing, documenting, adopting, and using standard document encoding for data interchange. The technology for encoding schemes is evolving continuously, and current practice is to use XML encoding with data schema defined by XML schema.

Troubleshooting

Data Interchange Document XML Schema Validation

Submitted by ccoleman on Thu, 03/22/2012 - 11:20am

Validation of XML data interchange documents with normative schemas using free, open-source software (XML Explorer)

One of the objectives of the US Geoscience Information Network (USGIN) is to establish a community of practice for developing, documenting, adopting, and using standard document encoding for data

interchange. The technology for encoding schemes is evolving continuously, and current practice is to use XML encoding with data schema defined by XML schema.

How to Create a Styled Layer Descriptor (SLD) using Arc2Earth

Submitted by ccoleman on Thu, 03/22/2012 - 10:39am

If you decide to use GeoServer to deploy your Web Services, you will find that part of the process includes creating an SLD or Styled Layer Descriptor for portrayal of the data you intend to publish. A natural symbology for such layers is an ESRI layer file format. Using this format initially presented a problem, though; creating an SLD for data with a complicated symbology scheme, such as a geologic map, is not trivial using XML editing software-- even if you have a template to work from. Many times, conversion from RGB to Hexidecimal is required for each entry in your symbology layer. Even if one uses a built-in formula converter in Excel (or from various websites), a large amount of hand entry is still required.

Making Web Maps

Using WFS with MANY features in ArcGIS

Submitted by srichardAzgs on Fri, 05/10/2013 - 12:35pm

If you're using a WFS in ArcMap, it downloads and indexes ALL the data from the service. If you're connecting to a service that has ALOT of data (>?150,000 features), my experience is that you don't actually get all the data. One solution is to put a bounding box filter on the WFS connection to restrict the data request to the area you're actually interested in.

To do this, when you set up the WFS connection properties, and are in the 'Parameters' dialog (which you'll need to do to select the feature type from the service), put an XML bounding box filter expression in the 'XML Filter Expression'

here's a bounding box filter that works (WFS 1.0.0):

```
<ogc:Filter xmlns="http://www.opengis.net/ogc">
 <ogc:BBOX>
  <ogc:PropertyName>shape</ogc:PropertyName>
 <gml:Box srsName="EPSG:4326">
   <gml:coord>
    <gml:X>-109</gml:X>
    <gml:Y>37.00</gml:Y>
```

Using and Customizing Leaflet

Submitted by Ryan Clark on Wed, 03/14/2012 - 1:51pm

Leaflet is a simple javascript framework for building web-mapping applications. In thier words:

Leaflet is a modern, lightweight open-source JavaScript library for mobile-friendly interactive maps. It is developed by CloudMade to form the core of its next generation JavaScript API. Weighing just about 21kb of gzipped JS code, it still has all the features you will ever need for your web mapping needs, while providing a fast, smooth, pleasant user experience.

It is built from the ground up to work efficiently and smoothly on both desktop and mobile platforms like iOS and Android, utilizing cutting-edge technologies included in HTML5 and CSS3. The focus is on usability, performance, small size, A-grade browser support, flexibility and an easy-to-use API. The OOP-based code of the library is designed to be modular, extensible and very easy to understand.

When you're choosing a javascript framework to build your web-mapping application, you have a myriad of options. Why choose Leaflet? Well this is why I liked it:

ArcGIS Server and OGC Services

SDE-Trickery and Awesomeness: Data Transformations Behind-The-Scenes

Submitted by Ryan Clark on Thu, 06/02/2011 - 3:26pm

Imagine that you collected geographic data in one schema and you want to provide it in some other schema. For many of us, this only happens pretty much *all of the time*.

For example, maybe your organization has an enterprise database which stores all the information that you have about recent, active faults. Now you want to provide that data as a WFS service with fields that conform to the ActiveFaults template being used by the AASG Geothermal Data System. Perhaps you store all your geologic map data in such a database and you want to convert it into FeatureClasses that fit into NCGMP09-style geodatabases.

Here's a nice way to convert your data formats in such a way that you only have to do it once. You can continue to make changes to your data just like you always have, and there will always be a featureclass hanging out there in the new schema which conveys the most up-to-date information that you have. The pre-requisites are:

OGC Services don't work outside local network!

Submitted by Ryan Clark on Thu, 05/26/2011 - 5:35pm

A commonly overlooked step in the configuration of an ArcGIS Server is that you need to tell the server what its name is to the outside world. You see, on an internal network your computers might have names. As geologists we prefer names like "Eclogite", "Kyanite", "Harzburgite", etc... However, on the

internet computers generally have more complicated names. These names are things like "services.azgs.az.gov" or "www.google.com" or "lab.usgin.org".

By default, your ArcGIS Server knows what the computer's internal name is. However, once you want to share your services with people outside that local network, you need to tell the ArcGIS Server how those other people will find it. You do this by editing a file called "rest.config". If you installed IIS and ArcGIS Server in all the default locations, you'll find this file at C:\inetpub\wwwroot\ArcGIS\rest\rest.config.

Look through the file and find all the URLs in it. Make sure that instead of using the computer's internal name, those URLs use the computers external name.

For more information, here are a couple of links:

Dynamic Content as WFS

Submitted by Ryan Clark on Thu, 04/28/2011 - 3:57pm

There could be a number of reasons why you'd like to use a database view as a backend for a WFS. Perhaps your data is split across multiple tables and updated regularly. Perhaps the data is in a table without an explicit ESRI-style SHAPE field. That's the case in this example, where the geometry information is stored in a table as UTM coordinates.

Using Django for USGIN

Thought and ideas about using Django to accomplish USGIN-related... things.

Updated URI Redirection Engine

Submitted by Ryan Clark on Fri, 05/25/2012 - 11:44am

I've made a new version of the URI redirection engine that runs in Django (see the original post here). The tool allows you to provide a list of URI patterns that identify specific resources, and redirect traffic that matches that pattern to a URL where a representation of the resource exists. It allows for content-negotiation, and the new version also begins to build infrastructure by which various instances of the application can communicate with each other.

Each instance of the application is informed about the existence of various registers, which are essentially sets of URIs. The application may or may not know how to resolve URIs that are a part of any given register, but can send incoming requests to others servers capable of resolving that register's URIs.

PyGreSQL: This will be useful!

Submitted by Ryan Clark on Thu, 02/03/2011 - 6:49pm

PyGreSQL is a python library that allows you to write python code to interact with a PostgreSQL database. I can see some utility here if I start wanting to write Django methods that do things like create views of some of my tables.

May be useful in not too long! Don't forget about it!

Documentation: http://pygresql.org/pg.html

Installation:

sudo apt-get install python-pygresql

Postgres COPY for CSV importing

Submitted by Ryan Clark on Wed, 02/02/2011 - 2:55pm

Working on some integration between Geoserver, Django and NCGMP geodatabases, I often need to use PostgreSQL's COPY command to import data from CSV files into PostgreSQL. Here are some tips:

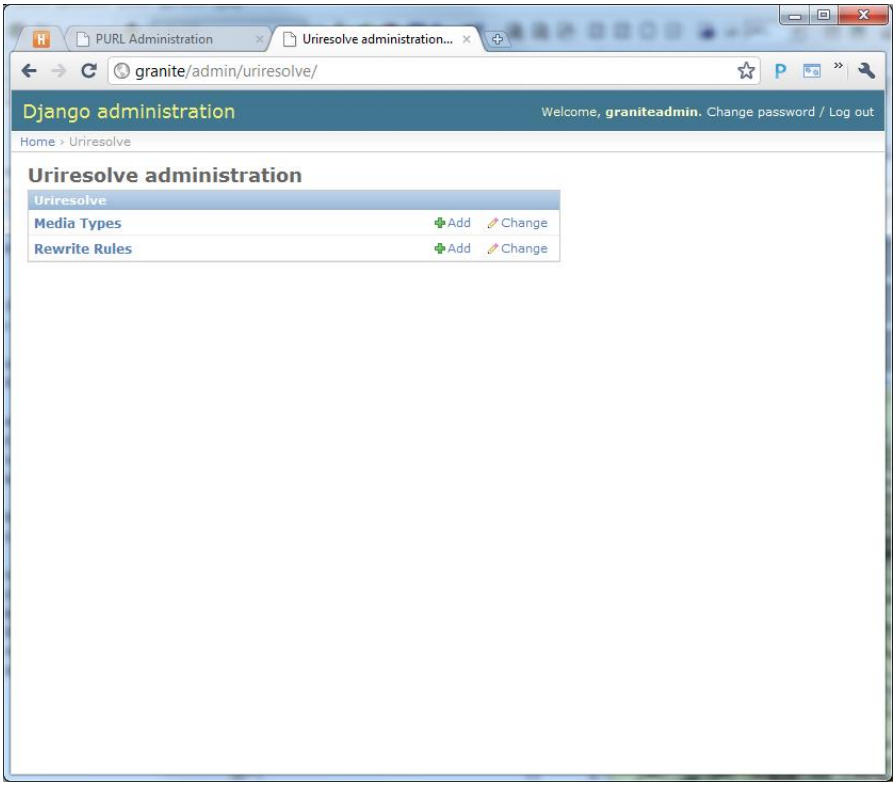
URI Redirection Engine (with pictures!)

Submitted by Ryan Clark on Mon, 01/24/2011 - 2:10pm

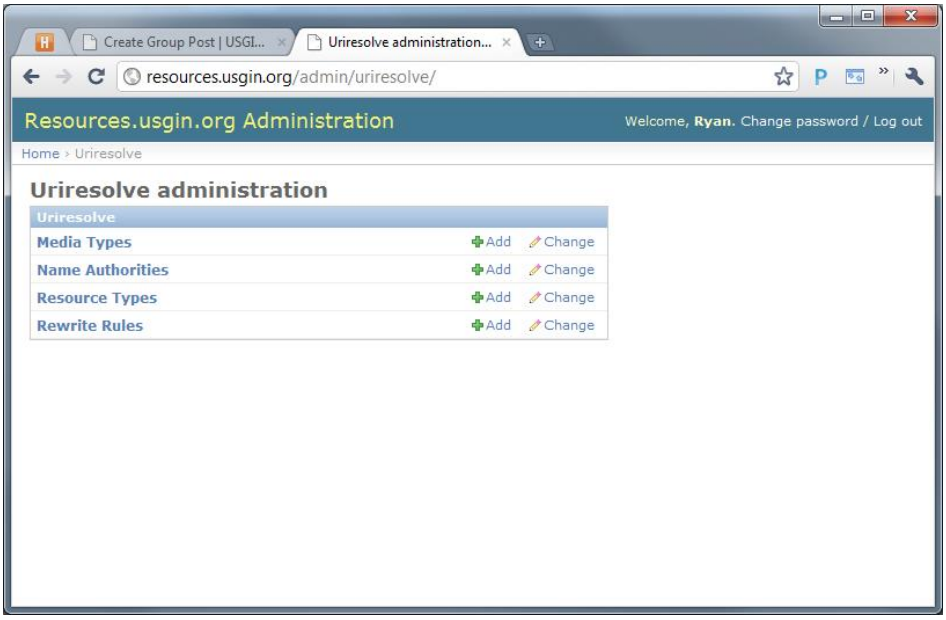
Here are a few screenshots to give a taste of what it is like to use our new Django app to manage a set of URI redirections:

Administration Overview:

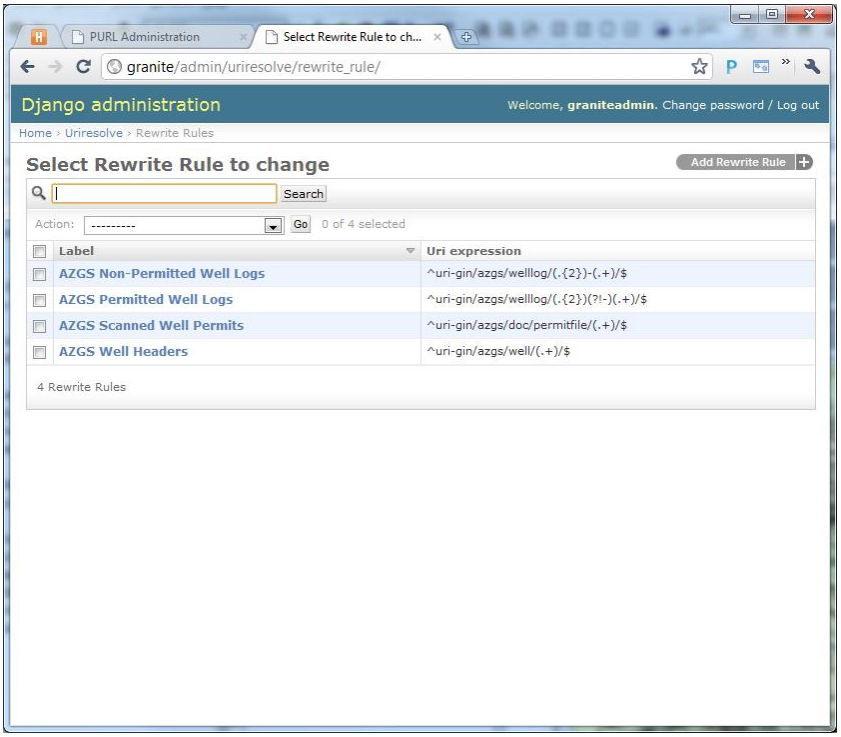
For any URI-Scheme: Basically, the Django app allows you to control a listing of MIME Media Types, and Rewrite rules.



USGIN URI-Scheme: The same basic structure, but the USGIN-specific app allows you to control a listing of Name Authorities and Resource Types, as described in [this post.]



Listing of Rewrite Rules: A nice overview of what rewrite rules you have in place, and a mechanism for searching for the one you might need to adjust.



URI dereferencing with Django / HTTP Content Negotiation

Submitted by Ryan Clark on Fri, 01/21/2011 - 9:46am

I've been building a Django app that essentially allows for the dereferencing of URIs that follow the USGIN Scheme. The engine also does some content-negotiation by reading HTTP GET accept-headers and responding with an appropriate file located somewhere on the internet. I had never even heard about content negotiation before embarking on this, and here are a few things that I learned:

Redirection Engine for USGIN-Compliant URI Resolution

Submitted by Ryan Clark on Fri, 01/07/2011 - 2:40pm

A common problem we run into is the unique identification of resources. This is a very broad statement because it represents a very broad problem. We have identification schemes for all kinds of things: people have names, web pages have URLs, books have ISBN numbers, etc. For the USGIN system to work smoothly, we need an identification scheme for resources that are part of the network. In USGIN we use Universal Resource Identifiers (URIs) which follow a specific format. For more information about the format, and for more background, see this post on lab.usgin.org.

Once we have this identification scheme built, it is important that the URIs can be dereferenced, or in other words that they are resolvable. We would like a user to be able to put the URI of a particular feature into the address bar of their web browser, push enter, and be presented with some sort of representation of their resource of interest. This is the niche that our Django-based redirection engine it trying to fill.

Django Setup

Submitted by Ryan Clark on Mon, 12/27/2010 - 11:03am

Introduction

I'm relatively new to Django, and maybe to servers in general, but setting up Django is a chore. Or at least figuring out how to set it up is. There is no installer. You have to have a pretty good understanding of how your HTTP server (IIS, Apache, whatever) works before you're going to be able to get anywhere. Then you'll have to know how to set up quite a few ancillary applications (e.g. Python, WSGI, FastCGI, Flup, MySQL, PostgreSQL). The details of which other applications you want to set up depend entirely on the evironment you're building - so that means it is a little different every time - and that means it is pretty hard to write one single "walkthrough" for how you should do it.

What I want to do here is detail my setup, and along the way make some suggestions about what seems to make things easier. So here goes:

Suggestion: Use Ubuntu (Most recent stable version, always.)

I say this because that apt system really makes installations a breeze. I fumbled for a while to try and setup WSGI (a module for Apache) to work on my Windows machine, to no avail. In Ubuntu it is so easy:

Geoportal on an Amazon Virtual Machine

PostgreSQL Trigger - make harvested/uploaded metadata records editable in Geoportal

Submitted by Ryan Clark on Mon, 12/27/2010 - 12:00pm

By default, metadata records that are harvested into the Geoportal are not editable. This includes any records in the Geoportal that were not created explicitly with interface provided by the web application ("Use dedicated editor to create metadata manually"). The SQL below creates a database trigger that allows harvested records to be edited.

Be aware that there is the potential for a lossy data translation when you do this. The harvested record is completely rewritten using the schema definitions that your Geoportal implements. If the original harvested record has any data that is not included in your schema implementation, that data will be lost after the original is edited. Another way to put that is, if you edit a harvested record, the edited version will only contain information that you see in the editor interface. If the original started with anything else in it, that will be lost.

Using jQuery for basic ISO Metadata record form validation

Submitted by averill.cate on Mon, 11/29/2010 - 11:50am

ESRI's Geoportal metadata record editing has some validity checking, but error checking does not occur until after a metadata record form is submitted. This can be a wasteful process because it involves having the client (the web browser) send a full request to the application server. If the submitted form has data validity errors then the whole form is transmitted back to the client for editing. One way to save the use of server resources (e.g., bandwidth) is by using jQuery to monitor information typed in html input boxes and check value validity before submitting the form. Since, ESRI's Geoportal comes with jQuery already enabled it is relatively to add validity checks. The first thing to do is create a custom javascript file (e.g. [geoportal_install_dir]/web/geoportal/catalog/js/custom.js). The contents of my script are:

```
$(document).ready(function()

{
  $('#mdEditor\\:identification_date').mask('9999-99-99T99:99:99');
  $('#mdEditor\\:general_datestamp').mask('9999-99-99');
```

Install Apache Directory Server

Submitted by Ryan Clark on Thu, 10/28/2010 - 3:38pm

ESRI's Geoportal wants an LDAP server, and suggests using Apache Directory Server.

I started a new VM using the image usgin2-atpms. Installation was as easy as:

wget http://apache.imghat.com//directory/apacheds/unstable/1.5/1.5.7/apacheds-1.5.7-i386.deb

sudo dpkg -i /home/ubuntu/apacheds-1.5.7-i386.deb

It was easy to connect to via SSH Tunnel to port 10389. The next step would be to fix some issues with the instance and its data-store, remove the extra MySQL installation, and then move on to configuration of the PostgreSQL database.

Drupal Development

Discuss ways to bend the Drupal content management system to your needs.

Creating a Format Agnostic Metadata Wizard

Submitted by wgrunberg on Fri, 05/28/2010 - 12:22pm

The USGIN Metadata Wizard is a metadata format agnostic web tool for creating, managing, and exporting new metadata records. This Drupal based service was conceived as a means to lowering the brier to metadata creation for those who are not able or willing to become proficient with metadata standards and informatics. The tool is based on a format agnostic metadata model which then can be exported into multiple metadata standards.

Creating ISO 19139 metadata through Drupal Views and Views Bonus Pack

Submitted by wgrunberg on Fri, 04/16/2010 - 2:00pm

Following is how we modified the Views Bonus Pack module for Drupal 6 to generate ISO 19139 XML metadata files for the USGIN Document Repository. This hack can be be used to relatively easily generate other XML metadata (FGDC, CSW records) or most complex XML. Ideally, this hack will evolve into a separate module. If you are interested or have questions, please participate in the Generate ISO 19139 metadata record XML files Views Bonus Pack issue thread.

Overview

The goal is to generate minimum ISO 19139 dataset metadata XML files (that conform to the USGIN profile) from the core and CCK fields used in the repository's Collection (ct_collection) and Document-like Information Object (DLIO, ct_dlio) content types.

Drupal Content Types used in the USGIN Document Repository

Submitted by wgrunberg on Fri, 04/16/2010 - 11:20am

Following are the two Drupal 6 content types that define the USGIN Document Repository (you may want to familiarize yourself with general Drupal concepts):

Collection (ct_collection)

The Collection content type acts as an arbitrary logical unit (or container) that groups the Document-like Information Object (DLIO) nodes which hold the actual data resource. At this time, Collections fulfill 3 objectives at once. Perhaps, they need to be separated at a later point.

Drupal Modules used in the USGIN Document Repository

Submitted by wgrunberg on Thu, 04/15/2010 - 4:28pm

Following is a break-down of the Drupal 6 modules used in the USGIN Document Repository site (you may want to familiarize yourself with general Drupal concepts):

Running Drupal on a Windows Server 2008 machine

Submitted by Ryan Clark on Wed, 04/14/2010 - 4:43pm

We recently struggled through figuring out how to make Drupal run on a Windows Server 2008 machine that was running IIS. Every Drupal site that we had dealt with prior to this was hosted on a Linux machine running Apache, MySQL and PHP, and so we needed to figure out what could be done differently to make it work on our in-house server, which was already running a production instance of ArcGIS Server under IIS7.

From the outset, we had two options:

Creating a Document Repository in Drupal

Submitted by wgrunberg on Wed, 04/14/2010 - 2:56pm

We just finished putting together a Drupal based Document Repository site (http://repository.usgin.org) as a testbed and as an alternative for other Repository systems such as DSpace. The goal was to fulfilled the following requirements in a ca. 2 week development period (200+ man hours):

GeoNetwork Configuration and Development

This group will focus upon GeoNetwork setup, configuration, and development. Any questions or comments are welcome!

Following are some resources that can help you learn about GeoNetwork:

- dev list
  - registration https://lists.sourceforge.net/lists/listinfo/geonetwork-devel
  - archive http://n2.nabble.com/GeoNetwork-developer-f2014756.html
- user list
  - registration https://lists.sourceforge.net/lists/listinfo/geonetwork-users
  - archive http://n2.nabble.com/GeoNetwork-users-f2013074.html
- SourceForge http://sourceforge.net/projects/geonetwork/
- Wiki http://trac.osgeo.org/geonetwork/
- Trac Tickets http://trac.osgeo.org/geonetwork/report
- USGIN Lab pages that contain "GeoNetwork" http://lab.usgin.org/search/node/GeoNetwork

Python Pylint/Pydev Code Checking in Eclipse

Submitted by lwolfe on Wed, 09/15/2010 - 6:16pm

Pylint is a good python code checking tool. Pydev is a good plugin for Eclipse for editing and running python programs. Much of this info is taken from Ryan Fraser's page . Wolfgang has an alternative description for installing pylint in the Komodo python IDE here

Windows 7 64-bit guide for GeoNetwork development within Eclipse

Submitted by wgrunberg on Wed, 07/28/2010 - 11:38am

This setup guide for downloading GeoNetwork using Subversion, building using Maven 2, developing, deploying and debugging on Tomcat and Eclipse on Windows 7 64-bit still uses the MckoiDB and should be updated to use MySQL.

Introduction

This document was developed for GeoNetwork's new Maven 2 based setup. See http://trac.osgeo.org/geonetwork/wiki/Maven for more information.

Note that several installation and configuration options may not be necessary for GeoNetwork development but can be necessary for other software development projects.

This document was compiled from experiments, notes, and web sites by the following contributors and sources (in alphabetical order):

Converting FGDC XML metadata records for WMS and WFS services to ISO 19139 via Python and GeoNetwork

Submitted by wgrunberg on Wed, 03/10/2010 - 1:04pm

This post describes a series of Python scripts to convert FGDC XML metadata for WMS and WFS services to ISO 19139 service metadata via GeoNetwork's OGC harvest service. Here are the steps:

1. Extract WMS and WFS GetCapabilities URLs from FGDC XML metadata and write them to a file.
2. Create a GeoNetwork Harvest Node from extracted GetCapabilities URLs and let GeoNetwork's OGC WMS/WFS harvester create ISO 19139 metadata from GetCapabilities response. Note: the resulting ISO 19139 metadata is only as good as the one in the GetCapabilities response. In addition, GetCapabilities responses do not include all fields necessary for minimum ISO 19139 metadata.
3. Copy some FGDC metadata entries (Title, Abstract, etc.) to newly created ISO 19139 metadata.

Our goal is to have working WMS and WFS metadata records in our CSW catalog that can be used to add the described services to analytical software such as ESRI's C-SW Client for ArcGIS Desktop. Following are the reasons why I am currently choosing this convoluted process instead of a direct FGDC to ISO 19139 metadata conversion:

Create a GeoNetwork OGC Harvest Node (WMS or WFS GetCapabilities to ISO 19139 metadata) through xml.harvesting.add request

Submitted by wgrunberg on Fri, 03/05/2010 - 11:05am

GeoNetwork offers a harvest service that, among other, follows a OGC WMS or WFS GetCapabilities URL and transforms the response into an ISO 19139 metadata record. Bear in mind that the resulting ISO metadata record is only as good as the GetCapabilities response and that the metadata can never be entirely ISO 19139 conformant without dummy values due to limitations of the OGC GetCapabilities schema.

GeoNetwork offers a handy user interface to add those "GeoNetwork Harvesting Nodes" and also exposes their functionality through a even handier XML service. See chapter 19.3 on "Harvesting Services" of the GeoNetwork opensource V 2.4 The Complete Manual. Following are my notes on creating an OGC harvesting node through GeoNetwork's harvesting service.

Tomcat configuration in Eclipse

Submitted by wgrunberg on Thu, 02/04/2010 - 1:35pm

Tomcat kept running out of memory when running it as a service from within Eclipse. Following is a solution:

Double-click on your Tomcat service under the Servers tab, and click on the "Open launch configuration" to open the Tomcat launch configuration window. On the Arguments tab, append the two following lines in the "VM arguments" form window:

-XX:MaxPermSize=128m
-Xmx1024m

How-to on running Tomcat within Eclipse:

- Eclipse Help: Creating a server
- IBM: Developing Web applications with Tomcat and Eclipse

Import MEF metadata archive into GeoNetwork through Python

Submitted by wgrunberg on Fri, 11/06/2009 - 1:53pm

This is a Python example script for importing GeoNetwork Metadata Exchange Format 1.1 (MEF) archives to GeoNetwork 2.4.2's mef.import service. The mef.import service requires a multipart/form-data POST through a modified MultipartPostHandler.py library which now supports Unicode (urllib2). It has been tested in Windows XP and Python 2.6.

Create GeoNetwork MEF files from ISO 19139 XML through Python

Submitted by wgrunberg on Fri, 11/06/2009 - 10:11am

This is an example Python script that showcases the creation of GeoNetwork Metadata Exchange Format 1.1 (MEF) archives from ISO 19139 metadata XML files.It has been tested in Windows XP and Python 2.6.

Metadata Interest Group

general forum for questions, comments, discussion about metadata

Discussion of Use of CI_Online Resource

Submitted by srichardAzgs on Fri, 02/21/2014 - 12:49pm

The CI_OnlineResource element is used in a variety of contexts in ISO19115 content. This complex-content-element provides a linkage to some online resource related to the metadata or the described context resource. It is thus analogous to atom:link (IETF RFC-4287) or the link element described in IETF RFC-5988. For links that locate online documents accessible using standard browser and file type resolution technology, the link can be as simple as a single URL element that retrieves a representation of the resource. There are many other kinds of related resources that the onlineResource element may point to, including web services that provide access to a dataset resource, metadata for related or source resources, specifications for standards or extensions to standards.

Lucene search syntax

Submitted by srichardAzgs on Mon, 02/03/2014 - 12:58pm

Useful guide to search term syntax for any applicaiton that's hitting a lucene index

http://lucene.apache.org/core/2_9_4/queryparsersyntax.html

JSON template for a 'complete' metadata record

Submitted by srichardAzgs on Mon, 12/12/2011 - 10:42pm

Incorporates concepts from ISO 19115, 19115-1, 19119, FGDC, 19110, and FRBR along with some additional stuff like recording the lineage of the metadata record, which is useful in a metadata harvesting federated system. Builds on the conceptual model presented in a previous discussion here.

This is an experiment to see what it would look like, trying to bring in just about any metadata property from those various specs... Lot's of comments in line.

The JSON is not perfect -- I'm a newbie.

Scenario for a community metadata editor

Submitted by srichardAzgs on Sat, 12/10/2011 - 12:19pm

Introduction

This is a scenario for a generic metadata edit tool, intended for brainstorming development of a community component that is transportable and reusable. Use of a standardized tool for editing metadata will reduce the cost of entry into the metadata game by reducing training and increasing

productivity, and should help improve the quality and consistency of metadata. The attached word document has the same text. Please add any other features or requirements that would be useful.

Here is a run down of some functionality.

The Scenario

Start: User opens the editor application.

Machine actionable links-- metadata, linked data, contexts...

Submitted by srichardAzgs on Mon, 10/03/2011 - 5:01pm

I spent some time researching this issue in the context of some other projects I've been following or involved in (OWS context Standards Working Group, energy industry ISO19115 metadata profile, and USGS CGI Web Application Integration Framework group). I would like to see this considered in a larger context to come up with a solution that can be used in these various contexts, because it's essentially the same problem.
Here' s the gist of the proposal for link attributes outlined in the attached doc:

href -- the link URL

type -- MIME type of response. Specifies encoding and optionally the native software application environment. We're sort of stuck with this because of existing usage. One big question is how much of serviceType, puprpose, and outputScheme it makes sense to conflate in the MIME type. I favor keeping the MIME type limited to low-level file formatting--to answer the question 'what software can read this file in a useful way' Understanding the content and intention of a document or the interface to a service should be specified by separate properties.

Googley Metadata

Submitted by Ryan Clark on Tue, 09/27/2011 - 10:19pm

We can all agree that search engines like Google have really defined the way that we search for information on the internet. So the question is, why aren't we using Google to search for our geologic datasets? Do we have to do all that formal, structured metadata? What the hell is CSW anyways?

As an experiment, I built a new site: metadata.usgin.org. This site is buit with Node.js, a really cool way to build a web server by running server-side javascript code. Its so easy. Here's my code. So what does it do?

It acts as a proxy server. You ask it for http://metadata.usgin.org/record/a386a4ba-e892-11e0-9e4a-0024e880c1d2, and it reads that file identifier out and issues a CSW GetRecordByID request to a CSW server. The response to that CSW request is an XML document, and nobody (including Google) really cares about that, so the server formats the record as a very simple HTML page.

Really Generic Functional Requirements for Metadata

Submitted by Ryan Clark on Thu, 09/08/2011 - 12:49pm

We write metadata so that we can search through it to find datasets we're interested in using for some purpose. I've recently started building a map-centric user-interface for doing this that has highlighted some simple functional requirements that metadata should meet in order to be helpful in accomplishing the task.

What I've done here is essentially listed the conceptual elements that seem neccessary in order to make a functional user-interface for searching. This has nothing to do with metadata syntax, just with metadata concepts.

Energy Industry Profile of ISO/DIS 19115-1 v1.0 Release Candidate is Available for Public Review

Submitted by srichardAzgs on Fri, 09/02/2011 - 7:51am

FYI-- this is a draft profile against the upcoming revision of ISO19115 for metadata, scoped to energy industry, but the scope overlaps requirements for USGIN (I've been very active in developing profile and looked after that). Please have a look at the docs if you're interested in metadata. Given the standards process time lines, it probably won't directly impact us until next year.

steve

Conceptual model for general purpose resource metadata

Submitted by srichardAzgs on Thu, 06/17/2010 - 11:29am

I have built a UML model (using Enterprise Architect 7.5 tool) of metadata content that is based on ISO19115/19119, but attempts to clean up some of the logical incoherence and inscrutable element labeling in those specifications, as well as adding hooks for more detailed information on specific resource types and resource lineage.

The attached files include a zip archive of the Enterprise Architect-generated html report on the UML model; unzip to a directory and start on the index.htm page.

The EA document (zipped into GenericCompleteMetadataConceptualModel.zip) is the actual UML model. You'll need EA 7.5 or later to open it. The project contains a Metadata package, within which the 'GenericMetadataModel' package contains the actual model. The other packages include UML models for related ISO specifications, elements of which have been incorporated (copied, renamed, modified -- not imported) into the generic model. They are included for reference.

The html documentation generated by Enterprise Architect can also be viewed at

http://usgin.org/GenericMetadataModelDoc/index.htm

Files:

Recommendations for metadata content to describe geoscience resources

Submitted by wgrunberg on Wed, 05/05/2010 - 3:31pm

Now that we have a better grasp of the IT and informatics challenges of a distributed and interoperable information network, we can think more again about its adoption - how to make it work and useful on the ground. Following was our rough chain of thought on metadata:

ISO 19139 service metadata record srv:DCPList codelist confusion

Submitted by wgrunberg on Tue, 11/24/2009 - 12:16pm

It appears that the GeoNetwork WMS harvest selects an incorrect codeListValue for a service metadata record's srv:DCPList (Distributed Computing Platform). Currently, it enters HTTP-GET and HTTP-POST. However, there is a bit of confusion within GeoNetwork and among standards and profiles. Specifically, one could interpret a conflict between DCP in a service's GetCapabilities (OGC) and a service's metadata record (ISO):

Useful site for CSW testing and resources

Submitted by srichardAzgs on Thu, 10/29/2009 - 10:29am

CSW requirements

appears to mostly be compiled by Armin Retterath, Kompetenz- und Geschäftsstelle GDI-RP

Has listing of active CSW servers, some requirements listings

http://www.mapbender.org/CSW_Requirements

ISO metadata editor tool comparisons

Submitted by srichardAzgs on Thu, 10/22/2009 - 4:41pm

The geospatial metadata community has benefitted greatly from the CSDGM Metadata Editor reviews created by Hugh Phillips and posted at: http://www.sco.wisc.edu/wisclinc/metatool/. As we make the transition toward implementation of an international metadata standard, the community requested that FGDC compile information about applications capable of creating ISO 19115 metadata. The information on this page was compiled to address that need.

Presentations and Posters

Post your posters and presentations related to USGIN topics.

Presentation for European Geosciences Union meeting, 2012-04-25

Submitted by srichardAzgs on Mon, 05/07/2012 - 9:55pm

Slides for talk at EGU, Vienna, April 2012, presented by SM Richard.

Presentation for NGDS Kickoff meeting 2009-11-03

Submitted by srichardAzgs on Mon, 05/07/2012 - 9:51pm

A rather large power point stack, for a one hour presenation on USGIN web services and their use for the National Geothermal Data System. Presented at the kick off meeting in Boise, ID, November 3, 2009 by SM Richard.

AGU Fall 2011 -- 2 posters: GeoSciML v3 release, and USGIN metadata

Submitted by srichardAzgs on Mon, 12/12/2011 - 10:54am

GeoSciML v3 'release' announcement poster

USGIN metadata poster

State Geothermal Data and the US Geoscience Information Network; SDSC Summer institute in Geoinformatics

Submitted by srichardAzgs on Mon, 12/12/2011 - 8:04am

Outline: Background; Approach: Metadata, Services, Interchange formats; Current status; Architecture and artifacts; Way forward

GSA national meeting 2011: USGIN Data integration framework

Submitted by srichardAzgs on Mon, 12/12/2011 - 7:54am

Progress 2010-2011, service inventory, promoting transition to open data and reusuable applications.

USGIN/NGDS Presentations at ESRI User Conference

Submitted by Ryan Clark on Thu, 07/14/2011 - 1:53pm

I gave two presentations at the ESRI User conference this year. One was geared mostly towards what we do with metadata in our systems, while the other was more of a broad overview of USGIN and a little bit of NGDS progress report.

Brief USGIN Overview: WRP GIS Committee Meeting, Mar. 8th, 2011

Submitted by Ryan Clark on Fri, 03/18/2011 - 11:52am

A brief overview of USGIN.

From WRP Meeting Notes:

III. Geoscience Information Network

Usgin Draft Specifications For Geoscience Resource Metadata And Identifiers

Submitted by srichardAzgs on Thu, 11/18/2010 - 5:11pm

Poster presented at Geological Society of America Meeting, Monday, Nov. 1, 2010 in Denver, Colorado. Overview of metadata recommendations, catalog system architecture, and proposed URI scheme.

NGDS & USIN Poster at DMT '10

Submitted by Ryan Clark on Wed, 05/26/2010 - 9:11am

This poster outlines the general goals of the National Geothermal Data System, describes its relationship to the US Geoscience Information Network, and shows some of the ways that the Arizona Geological Survey plans to make it easier for individual state surveys to contribute information to NGDS. It was presented in Sacramento by Ryan Clark at the Digital Mapping Techniques 2010 conference.

Presentation, USGS Menlo Park, 04/15/2010

Submitted by srichardAzgs on Fri, 04/16/2010 - 4:23pm

Powerpoint stack for presentation at USGS, Menlo Park, CA, 04/16/2010, by Steve Richard

USGIN posters at 2009 Geological Society of America (GSA) annual meeting

Submitted by wgrunberg on Wed, 10/21/2009 - 2:55pm

We presented the following posters at the 2009 GSA Annual Meeting in Portland, Oregon. Note, the PDF files are large.

XML ETL presentation at Data Preservation Workshop

Submitted by wgrunberg on Mon, 09/14/2009 - 6:30pm

Following is a presentation on XML Extract-Transform-Load (ETL) we did at the Geoscience Data Preservation Techniques Workshop, Indiana Geological Survey, Bloomington, IN on July 2009.

Although this presentation was geared towards National Geological and Geophysical Data Preservation Program (NGGDPP) metadata, the same tools can be used for other ETL needs.

ETL Debug Blog

Post to this group blog about your Extract-Transform-Load (ETL) efforts.

Helpful Regular Expression (regex) links

Submitted by wgrunberg on Wed, 05/12/2010 - 1:43pm

Following are helpful Regular Expression links

- Regular Expression HOWTO for Python
- Regular Expression Basic Syntax Reference
- Wikipedia: Regular Expression

Debugging Python scripts with Pylint

Submitted by wgrunberg on Wed, 03/17/2010 - 11:10am

I am currently using Pylint to check for bugs and formatting errors in my Python scripts. I used Brandon Corfman's nice Komodo Hacks: Integrating Pylint tutorial to set up Pylint and how to integrate it with the Komodo Edit IDE (see Easy setup of a Python 2.6 development environment on Windows about setting up Python).

Here is my **updated** **Pylint execution macro for Komodo Edit 5** which includes a verbose and an errors only reporting command.

Converting FGDC XML metadata records for WMS and WFS services to ISO 19139 via Python and GeoNetwork

Submitted by wgrunberg on Wed, 03/10/2010 - 1:04pm

This post describes a series of Python scripts to convert FGDC XML metadata for WMS and WFS services to ISO 19139 service metadata via GeoNetwork's OGC harvest service. Here are the steps:

1. Extract WMS and WFS GetCapabilities URLs from FGDC XML metadata and write them to a file.
2. Create a GeoNetwork Harvest Node from extracted GetCapabilities URLs and let GeoNetwork's OGC WMS/WFS harvester create ISO 19139 metadata from GetCapabilities response. Note: the resulting ISO 19139 metadata is only as good as the one in the GetCapabilities response. In addition, GetCapabilities responses do not include all fields necessary for minimum ISO 19139 metadata.
3. Copy some FGDC metadata entries (Title, Abstract, etc.) to newly created ISO 19139 metadata.

Our goal is to have working WMS and WFS metadata records in our CSW catalog that can be used to add the described services to analytical software such as ESRI's C-SW Client for ArcGIS Desktop. Following are the reasons why I am currently choosing this convoluted process instead of a direct FGDC to ISO 19139 metadata conversion:

Create a GeoNetwork OGC Harvest Node (WMS or WFS GetCapabilities to ISO 19139 metadata) through xml.harvesting.add request

Submitted by wgrunberg on Fri, 03/05/2010 - 11:05am

GeoNetwork offers a harvest service that, among other, follows a OGC WMS or WFS GetCapabilities URL and transforms the response into an ISO 19139 metadata record. Bear in mind that the resulting ISO metadata record is only as good as the GetCapabilities response and that the metadata can never be entirely ISO 19139 conformant without dummy values due to limitations of the OGC GetCapabilities schema.

GeoNetwork offers a handy user interface to add those "GeoNetwork Harvesting Nodes" and also exposes their functionality through a even handier XML service. See chapter 19.3 on "Harvesting Services" of the GeoNetwork opensource V 2.4 The Complete Manual. Following are my notes on creating an OGC harvesting node through GeoNetwork's harvesting service.

Authoritative fgdc-std-001-1998 (FGDC CSDGM) schema (XSD) and data definition document (DTD) URLs

Submitted by wgrunberg on Thu, 02/18/2010 - 3:51pm

Following are authoritative FGDC Content Standard for Digital Geospatial Metadata (CSDGM) XML schema document (XSD) and document type definition (DTD) URLs for FGDC-STD-001-1998. These documents are used to validate a FGDC XML metadata file.

- **http://www.fgdc.gov/metadata/fgdc-std-001-1998.dtd** - Usage:
  - <?xml version="1.0" encoding="ISO-8859-1"?>
    <!DOCTYPE metadata SYSTEM "http://www.fgdc.gov/metadata/fgdc-std-001-1998.dtd">
- **http://www.fgdc.gov/metadata/fgdc-std-001-1998.xsd** - Usage:
  - <metadata
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:noNamespaceSchemaLocation="http://www.fgdc.gov/metadata/fgdc-std-001-1998.xsd">

There are also annotated and modified FGDC-STD-001-1998 schema available for download:

FGDC XML schema woes

Submitted by wgrunberg on Wed, 02/10/2010 - 2:59pm

I am working on transforming FGDC XML metadata records to the USGIN 1.1 version of ISO 19139 metadata and having a hard time finding formal FGDC schema, name spaces, and schema locations. This is what I found so far:

XSLT transformations in Python through the Gnome libxml C parser

Submitted by wgrunberg on Mon, 11/09/2009 - 10:29am

- C
- ETL
- Python
- XSLT

This is an example script on how to do XSLT transformations in Python 2.6 through the Gnome libxml XML C parser and toolkit. The relative fast C library is available on multiple platforms but you will also need Python bindings for libxml2 and libxslt. There is a handy libxml2 and libxslt Python bindings installer for Windows which also includes the C libraries in DLL form.

Import MEF metadata archive into GeoNetwork through Python

Submitted by wgrunberg on Fri, 11/06/2009 - 1:53pm

This is a Python example script for importing GeoNetwork Metadata Exchange Format 1.1 (MEF) archives to GeoNetwork 2.4.2's mef.import service. The mef.import service requires a multipart/form-data POST through a modified MultipartPostHandler.py library which now supports Unicode (urllib2). It has been tested in Windows XP and Python 2.6.

Create GeoNetwork MEF files from ISO 19139 XML through Python

Submitted by wgrunberg on Fri, 11/06/2009 - 10:11am

This is an example Python script that showcases the creation of GeoNetwork Metadata Exchange Format 1.1 (MEF) archives from ISO 19139 metadata XML files.It has been tested in Windows XP and Python 2.6.

GeoNetwork authentication and CSW transactions through Python

Submitted by wgrunberg on Tue, 10/20/2009 - 10:51am

This is an example Python script that showcases GeoNetwork authentication, session handling, and CSW transactions.

XSLT to Transform WMS GetCapabilities response to CSW Insert transaction XML

Submitted by wgrunberg on Thu, 10/01/2009 - 10:13am

Tested with deegree-csw 2.3pre
Read the XSLT file for more information.

Attached is an example XSLT1 script to transform a WMS GetCapabilities 1.1.1 response to a CSW Insert transaction. The script is based on deegree's wms2iso19119.xsl (http://www.deegree.org/).
Note that currently it only supports WMS version 1.1.1 (<WMT_MS_Capabilities>) responses because it chokes on 1.3.0 (<WMS_Capabilities>) responses.

XML ETL presentation at Data Preservation Workshop

Submitted by wgrunberg on Mon, 09/14/2009 - 6:30pm

Following is a presentation on XML Extract-Transform-Load (ETL) we did at the Geoscience Data Preservation Techniques Workshop, Indiana Geological Survey, Bloomington, IN on July 2009.

Although this presentation was geared towards National Geological and Geophysical Data Preservation Program (NGGDPP) metadata, the same tools can be used for other ETL needs.

Building a GeoSciML WFS Server

I want to help make the process of building a GeoSciML WFS server easier. I'll be collecting links to relevant information, cookbooks, walkthroughs, etc. Hopefully this will chronicle my process of going from knowing practically nothing about how to implement a service to making one work.

Updated App-Schema Documentation

Submitted by Ryan Clark on Thu, 10/15/2009 - 1:51pm

The official documentation for the app-schema extension has been revamped! There is a lot more to look through here - absolutely a must-read for anyone interested in using GeoServer to serve GeoSciML features.

http://docs.geoserver.org/trunk/en/user/data/app-schema/index.html

A Very Useful Conversation

Submitted by Ryan Clark on Fri, 09/18/2009 - 3:22pm

I've been in the midst of a daily back-and-forth email discussion with the app-schema developers on the GeoServer-Users mailing list. Reading through the conversation may be helpful, as they've already resolved a number of issues I've had. Here's a link to the thread:

http://www.nabble.com/Geoserver-2.0-availability---included-app-schema-capabilities-td24624783.html

NCGMP to GeoSciML: MappedFeature

Submitted by Ryan Clark on Fri, 09/18/2009 - 3:20pm

I posted an example of the <typeMappings> section of a GeoServer/app-schema mapping file here: A Sample GeoServer Mapping file for NCGMP09 to GeoSciML. Checkout this other post to see where your mapping file should be, and how your datastore will refer to it: GeoServer Complex-Feature WFS Configuration: Directory Structure.

The <typeMappings> section is where you indicate what fields in your source data (NCGMP09) go into which fields in your output data (GeoSciML). Here I want to describe the different levels of complexity you can undergo when making a mapping file.

The Simplest Case: From one Field to Another

From NCGMP09 to GeoSciML (and back again)

Submitted by Ryan Clark on Fri, 09/18/2009 - 3:07pm

I'm almost done building a complete mapping file that will output GeoSciML MappedFeatures from polygon features in a database that follows the NCGMP09 schema. This schema, developed as part of the USGS National Cooperative Geologic Mapping Program, is a proposed standard database schema for

the publication of single digital geologic map. An ESRI Personal Geodatabase implementation of the NCGMP09 schema is available on the website linked above.

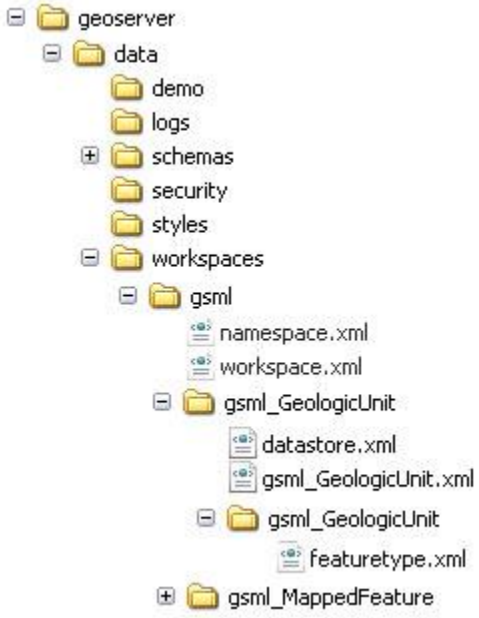
GeoServer Complex-Feature WFS Configuration: Data Connections in the Mapping File

Submitted by Ryan Clark on Fri, 09/18/2009 - 3:06pm

When you configure a complex feature service, you have to manually map your data connection parameters into an XML document called a mapping file. I found that the easiest way to make this connection is as follows:

GeoServer Complex-Feature WFS Configuration: Directory Structure

Submitted by Ryan Clark on Fri, 09/18/2009 - 2:57pm



At present, there is no user-interface component that helps you configure a complex-feature WFS in GeoServer. I'm sure that it is being worked on, but in the meantime, the configuration is done manually, by building the appropriate directory structure and putting XML configuration files in the right places.

- This is an example of the directory structure that has to be in place.
- Login to post comments
- Read more

It Works!

Submitted by Ryan Clark on Tue, 09/15/2009 - 10:58am

It works! Once you get the app-schema extension installed and the tutorial data in place, it works just fine. Make sure that in your GetFeature requests you specify the outputFormat as GML3, as it won't work at all in GML2 (which GeoServer defaults to if no outputFormat is specified).

Also, be aware that the app-schema extension is only supported at WFS version 1.1.0.

Now I can move on to beginning to map my data into GeoSciML features. I'll write some more walkthrough as I put it together.

GeoServer: Installing the App-Schema Extension

Submitted by Ryan Clark on Tue, 09/15/2009 - 10:56am

I've been spending the last couple of days working with GeoServer version 2.0-beta2, which introduces the ability to serve complex features. Serving complex features is a critical component of providing geologic data in GeoSciML format.

GeoServer accomplishes this through the use of what's called the "app-schema" extension. Unfortunately, there is not yet a user-interface for configuring the extension, and so you set up your complex features manually, by creating folders and configuration files in the right places that say the right things.

Complex Features in GeoServer

Submitted by Ryan Clark on Tue, 09/15/2009 - 10:55am

GeoServer has released a 2nd Beta of version 2.0. You can read about it on their blog.

This is exciting because they've introduced what they're calling the "app-schema extension". Essentially, it represents real support for complex features. This will open the door to GeoSciML delivery through the friendly GeoServer interface.

Groundwork: Basic software installations

Submitted by Ryan Clark on Mon, 09/14/2009 - 12:03pm

There have already been successful implementations of GeoSciML Web Feature services, using a few different approaches.

Getting Started

Submitted by Ryan Clark on Mon, 09/14/2009 - 11:57am

I'm starting the process of implementing a WFS that delivers GeoSciML data. I intend to make Group Posts here detailing what I'm trying, what works, what doesn't, in the hopes that it wil both help me keep track of what I've done, and also be of use to someone else attempting to do something similar. If I find things that are of particular use, I'll end up making more formal walkthroughs. I'm coming at this without having much background in open-source software, OGC services or XML, so I'll have a lot to learn.

CSW Debug Blog

Post to this group blog about your metadata Catalog Service for the Web (CSW) client and server implementation efforts and debugging sagas. I encourage occasional rants of the cathartic nature.

ISO 19139 service metadata record srv:DCPList codelist confusion

Submitted by wgrunberg on Tue, 11/24/2009 - 12:16pm

It appears that the GeoNetwork WMS harvest selects an incorrect codeListValue for a service metadata record's srv:DCPList (Distributed Computing Platform). Currently, it enters HTTP-GET and HTTP-POST. However, there is a bit of confusion within GeoNetwork and among standards and profiles. Specifically, one could interpret a conflict between DCP in a service's GetCapabilities (OGC) and a service's metadata record (ISO):

OGC Schema Issues in Relation to CSW

Submitted by Ryan Clark on Wed, 11/18/2009 - 11:38am

In order to create metadata for both static datasets and dynamic, online services and for use with CSW, the OGC created an xml schema that merges the schema for ISO19115 (dataset metadata) and ISO19119 (service metadata) (see secion D.1.5, page 105 in OGC 07-045). The way that was accomplished was by creating a schema located at http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd. The contents of that schema are quite simple -- it looks like this:

Connecting to GeoNetwork's MckoiDB through the Druid Database Manager

Submitted by wgrunberg on Fri, 11/13/2009 - 12:45pm

Here is a tutorial on how to connect to GeoNetwork's MckoiDB through the Druid Database Manager: http://geonetwork-opensource.org/documentation/how-to/druid_dbms

Note that the GeoNetwork 2.4.2 installer will not prompt you to provide a MckoiDB user name and password. Those are generated automatically and stored in

\geonetwork\web\geonetwork\WEB-INF\db\account.prop

Useful site for CSW testing and resources

Submitted by srichardAzgs on Thu, 10/29/2009 - 10:29am

CSW requirements

appears to mostly be compiled by Armin Retterath, Kompetenz- und Geschäftsstelle GDI-RP

Has listing of active CSW servers, some requirements listings

http://www.mapbender.org/CSW_Requirements

GeoNetwork 4.2.1 exploration notes v3

Submitted by wgrunberg on Thu, 10/08/2009 - 2:59pm

BUG: what I consider a bug
FEATURE: what I consider a missing feature
TODO: possibilities that need to be explored or decide on

GeoNetwork 4.2.1 exploration notes v2

Submitted by wgrunberg on Wed, 10/07/2009 - 9:22am

BUG: what I consider a bug
FEATURE: what I consider a missing feature
TODO: possibilities that need to be explored or decide on

GeoNetwork 4.2.1 exploration notes v1

Submitted by wgrunberg on Wed, 10/07/2009 - 9:08am

BUG: what I consider a bug
FEATURE: what I consider a missing feature
TODO: possibilities that need to be explored or decide on

CSW metadata editors we are exploring

Submitted by wgrunberg on Thu, 10/01/2009 - 3:20pm

Now that we know what kind of CSW services we can pull off, we need to pay more attention to creating and maintaining metadata. Right now, I just do a bunch of ETLs from our in-house data repositories (ArcSDE, Access, Excel, etc.) to CSW Insert Transaction XML files. That method is not user friendly and editing individual metadata records are a pain.

So, here are some CSW metadata editors I am looking at:

Ruminations on XML and large datasets

Submitted by wgrunberg on Wed, 09/16/2009 - 11:57am

While working on thoroughly understanding CSW & metadata related OGC and ISO standards, I can not but start questioning XML and its use to carry and format data. After all, USGIN's goal is not just to implement interoperable USGIN metadata and data profiles but also to encourage their widespread adoption. The following blog entry "How XML Threatens Big Data" by Michael Driscoll discusses some of the limitations XML poses to dealing with large datasets and ease of use.

Metadata about Metadata for a Service That Provides Metadata

Submitted by Ryan Clark on Wed, 09/16/2009 - 11:56am

... kind of sounds like overkill, but there's reason behind the confusion!

Imagine that you're trying to find some data. In the USGIN scheme, in order to find data you search through CSW services, which are catalogues full of ISO1939 metadata documents. Each of these documents points you to a dataset or service that is somewhere, hopefully online, for you to take a look at. These CSW services are like the card catalogues of the digital era.

List of interesting metadata implementation recommendations and profiles

Submitted by wgrunberg on Wed, 09/16/2009 - 11:55am

Following is a list of some metadata implementation recommendations and profiles I am looking at.

Configuring deegree-csw

Submitted by wgrunberg on Wed, 09/16/2009 - 11:54am

This is a list of our configuration changes too deegree-csw - specifically version 2.3-pre (missing link) with an PostgreSQL backend.

Follow instructions in deegree's documentation

See deegree_csw_2.0.2_documentation_en.pdf. Make sure the document is in a version corresponding to your version of deegree-csw (in this case "deegree Web Catalogue Service v.2.2").

Geospatial data and metadata protocols, services, standards, and profiles

Submitted by wgrunberg on Wed, 09/16/2009 - 11:50am

The beginnings of a list of geospatially enabled data and metadata protocols, services, standards, and profiles we are getting exposed to. Please, comment with corrections and additions.

What is the difference between CSW, CS-W, and CS/W?

Submitted by wgrunberg on Wed, 09/16/2009 - 11:49am

Well .. I couldn't find any difference between CSW, CS-W, and CS/W.

CSW Clients and Servers we are testing

Submitted by wgrunberg on Wed, 09/16/2009 - 11:47am

Following is a list of Web Service client, server, and hybrid applications we are aware of that are supposed to handle the CSW 2.0.2 protocol. Check out OGC's registry of products for more CSW implementations.

Public CSW services to test against

Submitted by wgrunberg on Wed, 09/16/2009 - 11:46am

Here are some CSW services I use to test CSW clients against

CatalogConnector CSW Client

Submitted by wgrunberg on Mon, 06/15/2009 - 12:17pm

USGIN CatalogConnector test client **http://catalog.usgin.org/catalogConnector/** (version 0.5)

This promising, stand-alone, Java-based, CSW web client is currently being developed by SDI Catalonia (IDEC) and USGIN at http://sourceforge.net/projects/catalogconnecto/

Appendix C: USGIN Schema Development

1. Content Models List……2
2. Development Tutorials
    a. Content Model Guidelines……12
    b. Define New Information Exchange……16
    c. Content Model Workbook……19
3. Tools-submitted to OSTI as software
    a. Log ASCII standard Format Tools: https://github.com/usgin/WellLogDataTools - Available to public at http://schemas.usgin.org/tools/ under Well Log Tools and here https://github.com/usgin-models/BoreholeLASLogData
        i. LAS converter
        ii. Automated Well log entry tool
4. Websites – submitted to OSTI as software
    a. Content model development site: https://github.com/usgin-models
    b. Metadata and documents capture and management: http://repository.stategeothermaldata.org/repository/
    c. Content model management site: https://github.com/usgin/modelmanager; public URL http://schemas.usgin.org/admin/

Content Model	Description	Version	Link
Abandoned Mines	This information exchange is for delivery of abandoned underground mine features in the AASG geothermal data project. The AbandonedUndergroundMine (AUM) worksheet specifies content elements for an interchange format for location, area, volume, heat capacity, temperature and other measurement data obtained for AUMs. Fields in that spreadsheet will become XML elements in interchange documents for WFS simple	0.3	http://schemas.usgin.org/files/abandoned-mines/0.3/AbandonedMines.xls
Active Fault/Quaternary Fault	Faults are typically represented on maps by the linear outcrop trace where the fault intersects the Earth Surface. Thus the feature service for fault uses a line feature class for the geometry. This linear geometry cannot simply be represented in a spreadsheet, so a template for service deployment provides as an ESRI personal geodatabase (which is a Microsoft Access .mdb file). In addition, for consistency with point-based content models, an Excel Workbook is provided with work sheets for the data provider metadata, description of fields in the template, and vocabularies recommended for use in interchange documents. For those that do not have software that will work with the geodatabase file, a zip archive containing a shape file feature class is included. The shape file truncates field names and text values that are strings longer than 255 characters. If some field contain text longer than 255 characters, fault descriptions should be loaded into the	1.2	http://schemas.usgin.org/files/active-faultquaternary-fault/1.2/ActiveFaults.xls

Content Model	Description	Version	Link
Aqueous Chemistry	This is a data interchange content model for an observation of the chemical composition of an aqueous fluid. A header content model for sample characterization, location, and analysis metadata is included to assist users finding analyses for specific samples, locations, time intervals, etc. Typically water temperature at the time of sampling is recorded along with sample data. Several 'suites' of analytes representing common analysis results are defined, with the intention that an implementation of the content model would offer several observation feature types, each of which consists of the header fields combined with the analytes for a suite. A content type for reporting results for a single analyte is also proposed, with an abbreviated	1.1	http://schemas.usgin.org/files/aqueous-chemistry/1.10/AqueousChemistry.xls
Borehole Lithology Intercepts	This workbook contains the BoreholeLithIntercept content model that specifies content elements for an interchange format for lithology log intercepts associated with a borehole. Fields in the BoreholeLithIntercept Content model will become XML elements in interchange documents for WFS simple features served by AASG-geothermal-data web services. Each intercept feature has a depth coordinate measured from the ground surface along	1.1	http://schemas.usgin.org/files/borehole-lithology-intercepts/1.1/BoreholeLithIntercepts.xls
Borehole Lithology Interval Feature	This is a content model for features representing of lithology log intervals associated with a borehole. Each interval has a top and a bottom, measured from the ground surface (convention for interoperability). Each	0.9	http://schemas.usgin.org/files/borehole-lithology-interval-feature/0.9/BoreholeLithIntervals.xls
Borehole Temperature Observation	This is a data interchange content model for a subsurface temperature measurement made in a borehole. Header content that characterizes the borehole sampling feature is included with each observation to assist users finding measurements based on location, depth, particular borehole, etc. This information will be provided through a	1.5	http://schemas.usgin.org/files/borehole-temperature-observation/1.5/BoreholeTemperature.xls

Content Model	Description	Version	Link
Contour Lines	This information exchange specifies content elements for an interchange format for linear spatial features representing contour (isoline) data. These lines connect locations registering the same value for some variable. Examples include structure contour maps that represent lines of constant elevation on some surface, or gravity	1	http://schemas.usgin.org/files/contour-lines/1.0/ContourLines_1.xls
Direct Use Feature	Direct use features describe facilities that utilize geothermal energy directly without transformation to electricity. See the NREL Geothermal Direct Use web page	1.5	http://schemas.usgin.org/files/direct-use-feature/1.5/DirectUseSites.xls
Fluid Flux Injection and Disposal	This workbook documents a content model for exchange of information about fluid fluxes into or out of a well. A generic model is intended to allow for oil, gas, or water injection (negative flow). The focus of this observation type is on injection (any fluid flowed into well) or disposal. Because of the wide variety of possible flow measurements	1.1	http://schemas.usgin.org/files/fluid-flux-injection-and-disposal/1.1/FluidFluxInjectionDisposal.xls
Geologic Contact Feature	The ContactFeature tab of this workbook indicates the content for Geologic Contact features. The entirety of the workbook is an information exchange workbook for GeoSciML-Portrayal, in which three features for WFS implementation are described (see alsohttp://schemas.usgin.org/models/#geosciml-portrayal-units and http://schemas.usgin.org/models/#geosciml-portrayal-faults). The model is based on the IUGS CGI interoperability working group. See https://www.seegrid.csiro.au/wiki/CGIModel/GeoSciMLPortrayalViewModel. This schema is a view of GeoSciML data that de-normalizes the data and concatenates complex property values into single, human-readable, labels and returns single, representative values from controlled	2	http://schemas.usgin.org/files/geologic-contact-feature/2.0/GeoSciML.xlsx

Content Model	Description	Version	Link
Geologic Fault Feature / Shear Displacement Structure	The ShearDisplacementStructure tab of this workbook indicates the content for Geologic Fault features. The entirety of the workbook is an information exchange workbook for GeoSciML-Portrayal, in which three features for WFS implementation are described (see alsohttp://schemas.usgin.org/models/#geosciml-portrayal-contacts and http://schemas.usgin.org/models/#geosciml-portrayal-units). The model is from the IUGS CGI interoperability working group. See https://www.seegrid.csiro.au/wiki/bin/view/CGIModel/GeoSciMLThematicView... . This schema is a view of GeoSciML data that denormalizes the data and concatenates complex property values into single, human-readable, labels and	2	http://schemas.usgin.org/files/fault-feature-shear-displacement-structure/2.0/GeoSciML.xlsx
Geologic Reservoir	This information exchange documents information useful to characterize a subsurface reservoir for geothermal potential. The content model was created to provide a template for reservoir data gathered by the Texas BEG to be hosted by the NGDS. Provide a format that is clear and	0.2	http://schemas.usgin.org/files/geologic-reservoir/0.2/GeologicReservoirs.xls

Content Model	Description	Version	Link
Geologic Units	The GeologicUnitView Feature tab of this workbook indicates the content for Geologic Unit features. The entirety of the workbook is an information exchange workbook for GeoSciML-Portrayal, in which three features for WFS implementation are described (see alsohttp://schemas.usgin.org/models/#geosciml-portrayal-contacts and http://schemas.usgin.org/models/#geosciml-portrayal-faults). The model is based on the IUGS CGI interoperability working group. See https://www.seegrid.csiro.au/wiki/bin/view/CGIModel/GeoSciMLThematicView.... This schema is a view of GeoSciML data that denormalizes the data and concatenates complex property values into single, human-readable, labels and returns single, representative values from controlled vocabularies for multi-valued properties that can be used when generating thematic maps, or portrayals, of the data. It is separate to, but harmonized with, GeoSciML and conforms to the level 0 of the Simple Features Profile for GML (link). Labels will be 'free-text' fields that will be, in robust services, well-structured summaries of complex	2	http://schemas.usgin.org/files/geologic-units/2.0/GeoSciML.xlsx
Geothermal Area	This package contains both an Excel spreadsheet and an accompanying personal geodatabase (PGD) for data delivery of a Geothermal Area dataset (polygons). Design is focused on providing characteristics of sites identified as having geothermal potential. Geothermal area content for the state geothermal data system is intended as a tool to identify areas that may be associated with geothermal resource potential. Thus, the content model does not	0.7	http://schemas.usgin.org/files/geothermal-area/0.7/GeothermalAreas_1.xls

Content Model	Description	Version	Link
Geothermal Metadata Compilation	This template is for use in compiling metadata content in a structured tabular format that can be automatically transformed into any of several metadata encoding and content schemes. The content is based on the USGIN Recommended Metadata Model The goal of this template is to make metadata creation easier while still complying to metadata requirements for the: National Geothermal Data	1.3.5	http://schemas.usgin.org/files/geothermal-metadata-compilation/1.3.5/MetadataCompilation.xlsx
Geothermal Power Plant Facility	This information exchange indicates content requested for features representing Geothermal Power Plant locations. Design is focused on power plants portrayed as geospatial points with an associated plant type, capacity in megawatts, and use application. Power plant data for the geothermal data system is intended as a tool to identify sites with existing power plants. Thus, the content model	0.2	http://schemas.usgin.org/files/geothermal-power-plant-facility/0.2/PowerPlantFacility.xls
Gravity Stations	This information exchange defines data model for observations of the acceleration of gravity. Fields in the content model will become XML elements in interchange	0.1	http://schemas.usgin.org/files/gravity-stations/0.1/GravityStations.xls
Heat Flow	This spreadsheet indicates the content requested for observation features representing Heat Flow measurements. The content model is based on specifications gathered from the SMU Heat Flow database, and has been extended to include the specifics of Heat	1.23	http://schemas.usgin.org/files/heat-flow/1.23/Heatflow.xls
Heat Pump Facility	This spreadsheet indicates the content requested for features representing heat pump facility locations. Design is focused on sites/facilities portrayed as geospatial points with associated facility type, use application, and operational date. Heat Pump facility data for the geothermal data system is intended as a tool to identify	0.6	http://schemas.usgin.org/files/heat-pump-facility/0.6/HeatPumpFacility.xls

Content Model	Description	Version	Link
Hydraulic Properties	This information exchange is an aid for compiling data for delivery as hydraulic property observations for the National Geothermal Data System. The HydraulicPropertyObservation worksheet specifies content elements for an interchange format for data obtained from boreholes. Fields in that spreadsheet will become XML elements in interchange documents for WFS simple features/geothermal data web services. Typically hydraulic	1	http://schemas.usgin.org/files/hydraulic-properties/1.0/HydraulicProperties_1.xls
Physical Sample	This workbook documents a content model for basic data characterizing a physical sample. The content model is based on consideration of content requested for SESAR, EarthTime, GeoSciML, Storet, ISO19115. Focus is on terrestrial (land based) samples. Content model includes	0.8	http://schemas.usgin.org/files/physical-sample/0.8/PhysicalSample.xls
Powell and Cumming Geothermometry	Specific liquid and gas analyte suites in this workbook have been tailored for use in the analysis formulated by Powell and Cumming, 2010. Analytes reported with this content model match those in the Powell and Cumming, 2010 spreadsheets and can be directly inputted into their model, which creates numerous ternary diagrams and other Excel graphics to create a geothermomerty report. Citation: Powell, Tom and Cumming, William 2010. Spreadsheets for		http://schemas.usgin.org/files/powell-and-cumming-geothermometry/2.0/PowellCummingGeothermometry.xls
Power Plant Production	This workbook documents a content model for reporting electrical power production. In the context of NGDS, these will normally be geothermal power plants, but the content	1	http://schemas.usgin.org/files/power-plant-production/1.0/PowerPlantProduction.xls
Radiogenic Heat Production	This workbook documents a content model for interchange documents with data from measurements of Radiogenic Heat Production based on analysis of individual rock samples. Calculation of heat production is based on measured U, Th, and K content, which may derive from chemical analysis, gamma ray spectral analysis or other	0.5	http://schemas.usgin.org/files/radiogenic-heat-production/0.5/RadiogenicHeatProduction.xls

Content Model	Description	Version	Link
Rock Chemistry	This is a data interchange content model for an observation of the chemical composition of whole rock samples, including trace elements, gases, and isotopes. SamplingFeatureURI gives a unique identifier describing the geologic unit/formation described by the record, while ParentSampleURI indicates the unique identifier for the parent of the sample used in the analysis. Several 'suites' of analytes representing common analysis results are defined, with the intention that an implementation of the content model would offer several observation feature types, each	0.4	http://schemas.usgin.org/files/rock-chemistry/0.4/RockChemistry.xls
Seismic Event Hypocenter	This spreadsheet indicates the content requested for features representing Seismic event hypocenter observations. Design is focused on hypocenters portrayed as points with an associated depth. Hypocenter data for geothermal data system is intended as a tool to identify seismically active areas that are often associated with	1.7	http://schemas.usgin.org/files/seismic-event-hypocenter/1.7/SeismicHypocenters.xls
Thermal Conductivity Observation	This spreadsheet indicates the content requested for features representing thermal conductivity observations. Design is focused on thermal conductivity measurements/observations portrayed as geospatial points with an associated depth, TC value, headerURI, construction information, and related resources. Thermal	2.1	http://schemas.usgin.org/files/thermal-conductivity-observation/2.1/ThermalConductivity.xls
Thermal/Hot Spring Feature	This spreadsheet indicates the content requested for basic data characterizing a hot spring feature for the AASG geothermal data project. Typically, water temperatures are recorded with other information such as water quality or chemical analysis from a particular spring. The temperature and flow rate reported here are meant to be generalized characterization. Other observation services should be	1.8	http://schemas.usgin.org/files/thermalhot-spring-feature/1.8/ThermalSprings.xls
Volcanic Vents	This spreadsheet indicates the content requested for features representing active volcanic vent data. This simple content model is intended to identify recently active volcanic vents that may indicate areas of active hydro-	1.4	http://schemas.usgin.org/files/volcanic-vents/1.4/VolcanicVents_1.xls

Content Model	Description	Version	Link
Well Fluid Production	This workbook documents a content model for exchange of information about fluid production for a given well. A generic model is intended to allow for a variety of flows from a well including oil, gas, and water as well as steam production. Because of the wide variety of possible flow measurements (commodities, aggregation type), the model adopts a thin approach, which may result in many records being returned for a single well. Thus most data about the well is linked through the HeaderURI instead of included inline to keep the redundant data volume down. This template defines the content model for a service that	1.1	http://schemas.usgin.org/files/well-fluid-production/1.1/WellFluidProduction.xls
Well Header Observation	A well is a facility defined by its function to extract fluids from within the earth. A well may be a simple hole in the ground (generally not the kind of wells we're interested in here...), but in general will consist of a dug shaft or drilled borehole. Drilled wells may include one or more boreholes (well bores) that are accessed from the surface at an origin collar, which defines the origin of the well. Individual boreholes in the well my have origin points (collars) that	1.5	http://schemas.usgin.org/files/well-header-observation/1.5/WellHeader.xls
Well Log Observation	This workbook describes a content model for well log observation features. These will be point features that locate the well origin for a well bore from which a well log is available. The well log observation record describes the log (type, top and bottom of logged interval) and provides links to representations of the actual log data, which typically will be either a paper copy, scanned image, or LAS file. In some cases the log may be represented in an Excel Spreadsheet. The content model includes basic information describing the well and well bore from which the log was obtained for use in search and filtering to discover logs of	0.8	http://schemas.usgin.org/files/well-log-observation/0.8/WellLog.xls

Content Model	Description	Version	Link
Well Tests	This template defines the content model for a service that delivers well stem test data collected during development of Oil and Gas wells. Normal drilling procedures control formation pressures and fluids through the use of a hydrostatic head. Well testing brings these formation	1	http://schemas.usgin.org/files/well-tests/1.0/WellTests.xls

Simple Feature Content Model Guidelines

See https://github.com/usgin/usginspecs/wiki/Content-Model-Guidelines which is the 'living document' that will be revised and updated as practice evolves. For more information on how to propose and submit a new information exchange (content model), see https://github.com/usgin/usginspecs/wiki/Define-New-Information-Exchange.

Introduction

This document is intended to provide guidance for creating an Excel workbook that defines a content model for a flat file data interchange format. The content model defines the information that will be associated with a feature or observation type; the content model may be implemented in a variety of ways, but USGIN is currently implementing these interchange formats as GML Simple Features to be served by an OGC WFS.

For point data, the Excel worksheet can be converted to a GIS feature class that can be used to deploy the WFS. The worksheet can thus serve as a means to package and check datasets for distribution through a service.

If the data require line or poly feature classes, then the actual data interchange template is NOT in the spreadsheet, but in a geodatabase feature class because the geometry or shape field can not be implemented (in a useful way…) in the spreadsheet. Recommended practice is that the Excel workbook should be named xxxContentElementsN.N.xls, where xxx is the feature name. This worksheet should contain a field list that defines the thematic (not spatial geometry) fields in the content model. The data compilation/delivery package for line or polygon features (Fault, contact, geologicUnitOutcrop in GeologicMapData, activeFault) is a folder that includes feature name and version number in the folder name. The folder should contain the content elements workbook, a personal geodatabase containing a single feature class that is the template for deploying the service, and a zip archive with shape file and excel workbook with one sheet for the template and a readme saying this is a last resort if you can't work with geodatabase. – see activeFault (http://repository.usgin.org/uri_gin/usgin/dlio/168) as an example.

Worksheet/Tab names in the workbook should be consistent: About, Notes, DatasetMetadata, xxxTemplate (where xxx is the featureName in the WFS), FieldList, DataValidTerms, ReviewerComments. These are discussed in the sections below. To inspect templates currently in use by the AASG geothermal data NGDS project, visit http://repository.usgin.org/search/node/template. These will be posted for review by the Geothermal Data System Development and Population Working Group.

Worksheets

About

This is the title page for the workbook, and includes title of the content model, version number, description of the intention of the model, list of editors (contributors), and revision history information for both the template (from NGDS developers) and for data loaded in the template (from the data provider)

Notes

Instructions for use of template, explanation of sheets included in the template, especially if there are any non-standard sheets.

DatasetMetadata

This worksheet includes information about the data provider that will be used to create the metadata record describing the dataset loaded into the spreadsheet. Required values are a title, description (abstract), and originator for the dataset, along with a telephone number or e-mail address for each one.

Template

All templates should have xxxURI, xxxName, Notes (not remarks, or description—only one free text additional information field unless there's a compelling reason for more…), and Source. 'xxx' is a prefix that ideally is the same as the feature name in the service to be deployed, and the same prefix is used for this worksheet name and the field containing the URI for the feature in the content model. xxxURI, xxxName and Source are always mandatory. Note second place version number for the content model will need to increment if field names are changed or fields added.

FieldList

The FieldList worksheet provides a listing of all elements (fields) in the content model, specifies data types, xml implementation, and provides information to define the content and explain usage of the content element. The FieldList worksheet should have these columns:

Interchange content element:

This is the field name that appears in column headings in the template worksheet, and will become the XML element name in the interchange format. Note that many of these element names are longer than 10 characters and will be truncated if the content model is implemented in a dBase table (e.g. as an ESRI Shapefile)

DataTypeName:

Logical data type. Valid Values:

- free text: any XML valid alpha numeric characters, no limitation on length of content.
- term: element value is a word from a controlled vocabulary. Implication is that there will be a list of terms in the DataValidTerms worksheet that may be used to populate this field. If users add vocabulary terms, they should be added to the DataValidTerms sheet and defined.
- date: a date and time value
- URI: a unique identifier as defined by IETF 3986 (http://tools.ietf.org/html/rfc3986)
- Decimal: a number that quantifies a value along a continuum.
- Integer: a number representing a discrete, countable quantity.

Implementation:

XML data type used to implement the element.

- string: sequence of alphanumeric characters, no length restriction
- string ISO 8601: string conforming to ISO8601 syntax for date/time information. "yyyy-mm-ddThh:mm:ss"
- string nnn: sequence of alphanumeric characters with a maximum length of nnn characters
- double: a floating point number, used to represent real numbers
- integer: a number in the set {…-2, -1, 0, 1, 2,…}

Cardinality:

Restriction on the number of instances of an element that may be present in a valid interchange document. For GML simple features, the valid values are '1' (required) or '0..1' (optional). To encode an attribute that has multiple values using a GML simple feature, the adopted approach is to concatenate the multiple values into a single string value with '|' (pipe) character delimeters separating individual values. The individual values may be a tuple; syntax for these tuples is defined in the Element Description column for that attribute.

Element Description:

Text description of content element. This text is required for each element. The text should define the intention of the element content for feature description, any limitations on the range of acceptable values, conventions to be followed in string syntax for content in the field.

Element Instructions:

Content element usage instructions

Element Notes:

Other information about usage

If there is confusion on the use of Description, Instructions, and Notes, put it all the text in the description column.

DataValidTerms

Worksheet contains ranges of cells containing lists of terms for use in the template worksheet. Each attribute that has data type = term should have a corresponding cell range on this sheet, clearly labeled to associate it with the appropriate field or fields in the field list. In some cases a complete, closed vocabulary may be supplied, in other cases the template may provide some sample values and data providers are expected to add any other terms they use, along with definitions of the terms.

ReviewComments

Worksheet for reviewers to add comments on data included in the template sheet.

Conventions

Use of conditional formatting in the workbook

On the DatasetaMetadata and Template sheets, users will enter content used to create interchange documents. Required content is indicated by convention using Microsoft Excel conditional cell format-

ting, such that the cell has pink fill if it is required but there is no content, and the fill color disappears when a value is entered in the cell. To apply the conditional formatting, select the cell or column of cells to format, click on 'Conditional Formatting' menu (in Excel 2010, Main Tabs ribbon, Styles tab), select 'new Rule'. In the dialog that opens, select 'Format only cells that contain' in the 'Select a Rule Type' section (top of dialog box, second row). Then in the bottom part of the dialog, select 'Blanks' in the combo box underneath 'Format only cells with:'. Click the 'Format..' button next to the Preview (bottom of the dialog box), select the 'Fill' tab (top right in 'Format Cells' dialog), and pick the light orange-pink color (upper right corner). Click 'OK' in the 'Format Cells' dialog, 'OK' in the 'New Formatting Rule' dialog, and cells in the selected region that are blank should turn pink.

Field names

Field names may not contain spaces. Use CamelCase if name has more than one word (first letter of each word capitalized, spaces removed. Underscores may be used if multiple capital letters appear in sequence. Special characters !@#$%^&*(){}[]\|~`"'<>,/? may not be used in field names. Valid non-alphanumeric characters are '-', '_' and '.' XML field names may not begin with a numeral.

If a field has dataType 'URI', the field name should have the suffix 'URI'.

Other conventions

Excel comes with a number of pre-defined cell formats. It is recommended that the 'Heading 3' format be applied to the field names in the template worksheet.

The Excel outline feature may be used to hide groups of fields that have related content that is optional. Users are free to modify the outlining to hide cells that they are not using, but a column in the template spreadsheet should never be renamed or deleted.

Define new information exchange

The decision to define a new information exchange should be based on the likelihood that others will want to publish similar datasets in the future. Members of the USGIN community propose specifications for data sharing exchanges. Exchange documents are developed and reviewed using a publicly accessible repository on GitHub (https://github.com/usgin-models). Each exchange has a separate repository associated with the usgin-models pseudo organization. A proposed model must have an identified steward, and a working group of at least three participants with relevant domain knowledge and understanding of the interchange technology. There is no formal process for defining workgroup membership; normally the challenge is finding a sufficient number of qualified individuals to provide meaningful reviews and comment. The exchange steward is responsible to assemble the workgroup and assure sufficient expertise in the group to generate a sound content model and implementation. The exchange steward requests creation of a new model repository at the usgin-models gitHub from the organization members, and identifies workgroup members who will have commit privleges on the repository. Any community member can create a repository branch to propose changes using standard GitHub procedures, and request consideration for merging back into the developing model.
After review and approval by the workgroup, a call goes out to a USGIN technical review e-mail list or by RSS feed (http://notifications.usgin.org/) for comments from the community. An open review period of 4 weeks is normal, after which any comments from the community must be resolved to the satisfaction of the commenter. When issues are resolved to the satisfaction of the stakeholders (workgroup and engaged community), the exchange specification is adopted.

When a specification is adopted, all associated documents are copied to a 'tag' branch in the gitHub repository and are not changed after they are 'tagged'. The Specification documents are also copied to the exchange repository at http://schemas.usgin.org, which is a web site set up to provide public access to exchange specifications and any related xml schema documents or other artifacts required for the deployment and operation of the information exchange.

Here is a more detailed outline for setting up a USGIN information exchange:

1. Steward assembles workgroup, defines scope of model, and gets repository set up on usgin-models gitHub at https://github.com/usgin-models.
2. Workgroup defines content model. The recommended procedure is to scope the model based on stated target use scenarios, and on example datasets that the interested parties want to share. When creating a content model draft, please see 'ContentModelGuidelines' document at https://github.com/usgin/usginspecs. To gather interested individuals, the USGIN Notifications system via RSS feed may be used at https://github.com/ngds/service-notifications/issues where new a new issue is posted to introduce the proposal.
3. Get review of workgroup draft from community of expected users. Revise content model as necessary and post drafts in the GitHub repository.

4. Select interchange service protocol. This should be based on the availability of server and client software in the community of users. Some example possibilities include OGC WFS, WMS, WCS, OpenDAP/NetCDF/THREDDS, Microsoft OData, ESRI Geoservices API.
5. Implement the content model using an encoding scheme compatible with the chosen service protocol, e.g. XML, JSON, turtle, csv.
6. Define validation rules for instance documents
7. Document the content model, interchange format, service protocol, and any special conventions or profile. Specify how data access links to data exposed using this exchange will be described using the metadata fields in USGIN profile ISO19139 CI_OnlineResource elements. Use existing identifiers where possible to identify service and MIME types.
8. Deploy an example service and test with client software. Iterate 1-5. Register a dataset distributed using the exchange in a USGIN catalog.
9. Have documentation reviewed by target users and technical experts; respond to comments, updating 1-7 as necessary. If using USGIN Notifications, post comments on the open issue to alert the working group and/or community to ensure all parties have commented.
10. Adopt content model for implementation. Tag the version of the content model and specification documents in the GitHub repository, and deposit specification documents in USGIN exchange repository at http://schemas.usgin.org/models at which time it is made available to the public.

Outline for creating a new version (editing an existing content model):

1. Any member of working group makes edits to content model in the given repository at https://github.com/usgin-models. Any interested party may also make a new branch and make edits, but will need appropriate privileges from the steward of that repository to push changes to the master branch for the community to view.
2. Notifications regarding changes in the content model must be posted at https://github.com/ngds/service-notifications/issues to be read by the RSS feed at http://notifications.usgin.org/. This issue must indicate reasons for the change and a reasonable comment period.
3. Comments and edits continue in the GitHub repository.
4. Following approval by the working group and the steward, a new version of the content model is indicated on the "About" tab of the content model Excel file and tagged in the GitHub repository, rendering the previous version of the content model deprecated.
5. The new version of the content model is then uploaded by an administrator at http://schemas.usgin.org/models/ to be made available to the public.

Checklist for exchange steward:

Before exchange specification documents are tagged in GitHub and put online at schemas.usgin.org/models, the following checks should be made:

- Field headings are consistent in spelling, capitalization, and order in all tabs of the Excel workbook (Data and FieldList tabs).

- Field headings spelling, capitalization, order, and cardinality are exactly the same in the Excel workbook (as described in the FieldList tab) and in the schema (XSD).
- The schema (XSD) has the first field as “OBEJCTID”.
- The schema (XSD) has the last field as “Shape”. Where both “ShapeLength” and “Shape” fields are used, “ShapeLength” is second to last, and “Shape” is the last field. [xs:any, ShapeArea]

Repository management:

A good first step is to download the GitHub GUI at https://github.com/ where you will click “Set up Git” and follow the steps. Be sure to always create a “readme.md” file to explain the purpose of the information exchange, indicate the content in the given repository, and provide any updates regarding the version in use or other pertinent information.

Create a new repository by going to the USGIN GitHub site at https://github.com/usgin-models and logging into GitHub. Next to the user name in the upper right-hand corner, click “+” and choose “New repository”. Once created, click the “Clone in Desktop” button in the repository to get a local copy and begin entering documents into your local folder. Once ready to push the changes to the “master” to be viewed by others, commit the changes and “sync” in the GUI. Edit existing content models in much the same way; clone the repository in your desktop then edit and sync your changes back to the master branch.

File naming conventions:

The name of the schema (XSD) and Excel (XLS) represenations of the information exchange should be identical and reflect the exact name of the eventual service.

Content Model Workbook: description

Content model information available at: http://schemas.usgin.org/home/

Worksheets

About

This is the title page for the workbook, and includes title of the content model, version number, description of the intention of the model, list of editors (contributors), and revision history information for both the template (from NGDS developers) and for data loaded in the template (from the data provider)

Notes

Instructions for use of template, explanation of sheets included in the template, especially if there are any non-standard sheets.

DatasetMetadata

This worksheet includes information about the data provider that will be used to create the metadata record describing the dataset loaded into the spreadsheet. Required values are a title, description (abstract), and originator for the dataset, along with a telephone number or e-mail address for each one.

Template

All templates should have xxxURI, xxxName, Notes (not remarks, or description—only one free text additional information field unless there's a compelling reason for more…), and Source. 'xxx' is a prefix that ideally is the same as the feature name in the service to be deployed, and the same prefix is used for this worksheet name and the field containing the URI for the feature in the content model. xxxURI, xxxName and Source are always mandatory. Note second place version number for the content model will need to increment if field names are changed or fields added.

FieldList

The FieldList worksheet provides a listing of all elements (fields) in the content model, specifies data types, xml implementation, and provides information to define the content and explain usage of the content element. The FieldList worksheet should have these columns:

Interchange content element:

This is the field name that appears in column headings in the template worksheet, and will become the XML element name in the interchange format. Note that many of these element names are longer than 10 characters and will be truncated if the content model is implemented in a dBase table (e.g. as an ESRI Shapefile)

DataTypeName:

This column defines the Logical data type. Valid values are summarized in the following table.

Table 5-1. Logical data types for use in NGDS content model field definitions.

free text	any XML valid alpha numeric characters, no limitation on length of content.
term	element value is a word from a controlled vocabulary. Implication is that there will be a list of terms in the DataValidTerms worksheet that may be used to populate this field. If users add vocabulary terms, they should be added to the DataValidTerms sheet and defined.
date	a date and time value
URI	a unique identifier as defined by IETF 3986 (http://tools.ietf.org/html/rfc3986)
Decimal	a number that quantifies a value along a continuum.
Integer	a number representing a discrete, countable quantity.

Implementation:

This field specifies the XML data type used to implement the field in XML schema. Types are summarized in the following table.

Table 5-2. XML implementation data types

string	sequence of alphanumeric characters, no length restriction
string ISO 8601	string conforming to ISO8601 syntax for date/time information. "yyyy-mm-ddThh:mm:ss"
string nnn	sequence of alphanumeric characters with a maximum length of nnn characters. if the maximum length is missing, the text may be any length.
double	a floating point number, used to represent real numbers
integer	a number in the set {…-2, -1, 0, 1, 2,…}

Cardinality:

Restriction on the number of instances of an element that may be present in a valid interchange document. For GML simple features, the valid values are '1' (required) or '0..1' (optional). To encode an attribute that has multiple values using a GML simple feature, the adopted approach is to concatenate the multiple values into a single string value with '|' (pipe) character delimeters separating individual values. The individual values may be a tuple; syntax for these tuples is defined in the Element Description column for that attribute.

Element Description:

Text description of content element. This text is required for each element. The text should define the intention of the element content for feature description, any limitations on the range of acceptable values, conventions to be followed in string syntax for content in the field.

DataValidTerms

Worksheet contains ranges of cells containing lists of terms for use in the template worksheet. Each attribute that has data type = term should have a corresponding cell range on this sheet, clearly labeled to associate it with the appropriate field or fields in the field list. In some cases a complete, closed vocabulary may be supplied, in other cases the template may provide some sample values and data providers are expected to add any other terms they use, along with definitions of the terms.

ReviewComments

Worksheet for reviewers to add comments on data included in the template sheet. This is for use in the data review process if the content model template worksheet is actually used to compile data for release.

Conventions

Use of conditional formatting in the workbook

On the DatasetaMetadata and Template sheets, users will enter content used to create interchange documents. Required content is indicated by convention using Microsoft Excel conditional cell formatting, such that the cell has pink fill if it is required but there is no content, and the fill color disappears when a value is entered in the cell. To apply the conditional formatting, select the cell or column of cells to format, click on 'Conditional Formatting' menu (in Excel 2010, Main Tabs ribbon, Styles tab), select 'new Rule'. In the dialog that opens, select 'Format only cells that contain' in the 'Select a Rule Type' section (top of dialog box, second row). Then in the bottom part of the dialog, select 'Blanks' in the combo box underneath 'Format only cells with:'. Click the 'Format..' button next to the Preview (bottom of the dialog box), select the 'Fill' tab (top right in 'Format Cells' dialog), and pick the light orange-pink color (upper right corner). Click 'OK' in the 'Format Cells' dialog, 'OK' in the 'New Formatting Rule' dialog, and cells in the selected region that are blank should turn pink.

Field names

Field names may not contain spaces. Use CamelCase if name has more than one word (first letter of each word capitalized, spaces removed. Underscores may be used if multiple capital letters appear in sequence. Special characters !@#$%^&*(){}[]\|~`"'<>,/? may not be used in field names. Valid non-alphanumeric characters are '-', '_' and '.' XML field names may not begin with a numeral.

If a field has dataType 'URI', the field name should have the suffix 'URI'.

Other conventions

Excel comes with a number of pre-defined cell formats. It is recommended that the 'Heading 3' format be applied to the field names in the template worksheet.

The Excel outline feature may be used to hide groups of fields that have related content that is optional. Users are free to modify the outlining to hide cells that they are not using, but a column in the template spreadsheet should never be renamed or deleted.

Appendix D: NGDS Catalog System

1. Data Validation Tutorials and Tools
    a. QA/QC documentation
        i. GeoSciMLPortrayalCookbook
        ii. Data Development Cycle (graphic)
    b. QA/QC tools- Uploaded separately to OSTI
        i. USGIN Models API: https://github.com/usgin/usginmodels
        ii. Content model Spreadsheet Validator: http://schemas.usgin.org/validate/cm
        iii. Web Feature Service validator at http://schemas.usgin.org/validate/wfs
        iv. Excel to NGDS Services Tool: https://github.com/usgin/ExcelToNGDSServiceTool
        v. NGDS data archiver: https://github.com/ngds/data-archiver
2. Conformance Tutorials, Documentation, and Tools
    a. Tutorials & Documentation
        i. USGIN metadata profile documentation (USGIN_ISO_Metadata.pdf)
        ii. XML Tutorial
        iii. XML Schema Profile for Simple Feature WFS Services
        iv. Import of Metadata to NGDS catalog (CatalogDocumentation.pdf)
        v. USGIN Identifier (URI) Conventions Tutorial
        vi. USGIN Lab Discussions

        Uploaded separately to OSTI:

        vii. Metadata repository: https://github.com/usgin/metadata-repository
        viii. Metadata compilation: (see http://schemas.usgin.org/models/#Metadata)
    b. Tools
        i. Django, Node.js, and CouchDb metadata standardization and staging system: https://github.com/usgin/metadata-server
        ii. CSV to XML python tool: https://github.com/usgin/usgincsv

GeoSciML-Portrayal

An introduction to a simple feature schema for supporting map services and guidelines for service deployment

A product of the Arizona Geological Survey

Table of Contents

Table of Contents 1
1 Introduction 3
1.1 What is GeoSciML? 3
1.1.1 What is GeoSciML-Portrayal? 3
1.1.2 GeoSciML Specifications 3
1.2 What is USGIN? 4
1.3 Cookbook Prerequisites 4
1.3.1 Prerequisite Notes 5
1.4 Cookbook Workflow 5
2 GeoSciML-Portrayal Feature Types 6
2.1 ContactView Features 6
2.1.1 ContactView Feature Specifications 7
2.2 ShearDisplacementView Features 7
2.2.1 ShearDisplacementStructureView Feature Specifications 9
2.3 GeologicUnitView Features 9
2.3.1 GeologicUnitView Feature Specifications 11
3 Integration of Web Map Services 12
4 Deployment Cookbook 12
4.1 Mapping from source data 12
4.1.1 Interchange formats 13
4.1.2 Schema mapping in ESRI ArcGIS 13
4.1.3 Schema mapping in SQL 15
4.1.4 Notes on schema mapping in SQL 16
4.2 Vocabulary mapping 17
4.3 Excel Workbook template 18
4.4 Configuring OGC services 18
5 Styling 19
5.1 How to Create a Styled Layer Descriptor (SLD) using Arc2Earth 19
6 Deploying Your Web Service 20
7 Testing Your Web Service 21
7.1 Web Service Requests 21
7.2 Adding a Web Service in ArcCatalog 22
7.3 Adding a Web Service in uDig 22
1

8 Registering with OneGeology and Submitting your Service 23
9 Appendix A: Deploying GeoSciML-Portrayal Web Services in GeoServer 23
9.1 Logging in 23
9.2 Service-Level Metadata 24
9.3 Creating a Workspace 24
9.4 Connecting to your PostGIS Database 25
9.5 Adding Layers to a Workspace 26
9.5.1 The Data tab of the Edit Layer page 27
9.5.2 The Publishing tab of the Edit Layer page 27
9.5.3 Importing Layer Styles from an SLD File 27
9.5.4 Finishing Up 27
9.6 GeoServer Troubleshooting 27
10 Appendix B: Deploying GeoSciML-Portrayal Web Services in ArcGIS Server 29
10.1 Connecting to your ArcGIS Server instance 29
10.2 Create an ArcMap Project for your web service 29
10.3 Publishing on ArcGIS Server 29
10.4 Troubleshooting ArcGIS Server 30
11 Appendix C: OneGeology Web Service Requirements 30
11.1 Service-Level Metadata Requirements 30
11.2 Service Title and Service Name Requirements 31
11.2.1 OneGeology Service Title Examples: 31
11.3 Layer Naming and Title Conventions 31
11.3.1 Layer Title Examples 32
11.3.2 Layer Name Examples 32
11.3.3 A sample WMS GetCapabilities document demonstrating Layer Names and Layer Titles 32
11.4 Additional Layer-level Metadata Requirements 32
12 Glossary 34

1 Introduction

This document is a cookbook designed to facilitate the deployment and maintenance of geoscience information as interoperable, publicly available OGC web services in GeoSciML-portrayal in accordance with USGIN and OneGeology specifications.

To facilitate interoperability, it is possible to perform all of the steps indicated in this cookbook with free-and-open-source software.

1.1 What is GeoSciML?

GeoSciML stands for Geoscience Markup Language.

GeoSciML was developed to support interoperable map services, these services delivering geologic map unit, contact, and shear displacement structure (fault and ductile shear zone) portrayals. Ideally, use of standard vocabulary enables map portrayal using shared legends to achieve basic visual harmonization of maps, even when these maps are provided by different services.

In other words: GeoSciML is a specialized markup language by which geoscience data can be shared and portrayed interoperably, as web services, using shared legends.

1.1.1 What is GeoSciML-Portrayal?

GeoSciML-portrayal is a simplified **view** of GeoSciML (Figure 1).

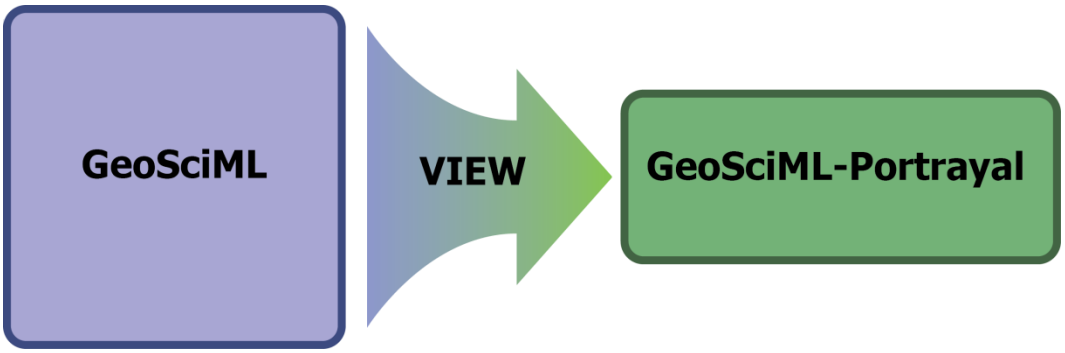


Figure 1: A visual representation of the relationship between GeoSciML and GeoSciML-Portrayal

In the context of a database, a **view** is a selection of fields. For example: given a database with twelve fields, an arbitrary grouping of four such fields would constitute a **view** of the database. A more concrete example is a card catalog: if one chose to look *only* at the **Title** field of each card in the catalog, this choice would constitute a discrete **view** of the catalog.

Markup languages can and do function similarly to databases, wherein markup language tags are analogous to database fields. Consequently, it is possible to create the equivalent of a **view** for a markup language.

The **GeoSciML-portrayal** is the equivalent of a **view** of the GeoSciML markup language: a specialized, distilled form of the full GeoSciML markup language.

Consequently, GeoSciML-portrayal can be considered a derivative markup language. As such, documents written using the GeoSciML-portrayal markup language can be structured by schemas. In part, the GeoSciML-portrayal cookbook will describe the GeoSciML-portrayal schema that structures documents written using the GeoSciML-portrayal markup language.

1.1.2 GeoSciML Specifications

GeoSciML is based on XML and developed as a GML v3.2 application for encoding a wide variety of geoscientific information (Richard et al., 2007); GeoSciML defines a collection of complex GML features for describing geoscience entities. Use of standard vocabulary enables map portrayal using shared legends to achieve some basic visual harmonization of maps provided by different services. In addition the scheme includes text information for human users when browsing a geologic map, a link to a full GeoSciML feature element if available, and a symbol identifier field to enable a user-defined symbolization scheme in each map service. Most WMS implementations simplify deployment of WFS service, as the same schema is used. Linking the simple feature WMS and WFS allows clients to acquire basic text geologic feature descriptions that can be used in web-mapping applications to construct custom legends. Linking to full GeoSciML features allows the portrayal scheme to be used in a map browsing and query interface to identify and select features for further processing that can be acquired as highly structured, information-rich, complex GML features.

The GeoSciML-portrayal schema is separate from, but harmonized with, GeoSciML v3.0. GeoSciML-portrayal conforms to the level 0 of the Simple Features Profile for GML (OGC 10-100r3 - van den Brink et al., 2011; OGC 06-049). The simple features profile supports only a limited subset of possible GML geometry types that may be used to describe feature geographic location and shape. For the purposes of GeoSciML simple features, these include gml:Point, gml:LineString, gml:Curve, gml:Polygon, gml:Surface, gml:MultiPoint, gml:MultiCurve, gml:MultiSurface and multi-geometry types consisting of collections of these base types. For a useful discussion of GML simple vs. complex features, see this Geoserver documentation page.

GeoSciML-Portrayal features are analogous to GeoSciML mapped features, with additional text attributes for human consumption, a flatted-relation view of the age, and assignment to a single lithology. The portrayal scheme consists of 'free-text' fields and identifier fields. In robust services the free-text fields will contain well-structured summaries of complex GeoSciML data in a format suitable for reading by the intended users. Identifier fields should contain identifiers for concepts in a controlled vocabulary (for example CGI Simple Lithology) that specify representative thematic properties. Inclusion of these standardized identifiers enables interoperability across services. Ideally these should be URIs that can be dereferenced to obtain machine-processable or human-readable representations of the identified concepts.

In addition, each GeoSciML-portrayal feature includes an (optional) identifier for a specification, which is a resource containing a description of that particular feature. In many cases, the descriptions will be the same for all polygons assigned to the same map unit or classified as the same kind of contact or structure. If more complete information is available, different descriptions may be associated with subsets of features of the same type that are portrayed with the same symbol. In the most extreme case, each feature might have a unique description that captures the full spatial variability of a geologic unit or structure. Following the standard patterns of web architecture, the specification_uri should be dereferenceable to obtain one or more representations of that description. For maximum interoperability, one of these representations should be a GeoSciML encoded description of the feature, but other encodings might also be available, for example html web pages, other XML schema, or JSON. For those familiar with full GeoSciML v3.0, the specification_uri property is equivalent to the specification association from MappedFeature to GeologicFeature.

Deployment of an interoperable dataset requires first that the interchange format used for interoperability is well understood. The following sections are intended to provide the necessary background understanding of GeoSciML-portrayal, and should be studied carefully. The deployment process consists of determining how the information in the dataset to be published can best be represented using the elements in the GeoSciML-portrayal model, populating the feature collections that will be served, and configuring a WMS server to display the data.

1.2 What is USGIN?

The United States Geoscience Information Network (USGIN) initiative is the product of a partnership between the Association of American State Geologists (AASG) and the United States Geological Survey (USGS) created to facilitate discovery of, and access to, geoscience information provided by state and federal geological surveys of the United States.

USGIN is:

- conceptual framework for sharing geoscience data
- a distributed data-sharing network
- a collection of open-source applications, standards, procedures, and protocols for sharing geoscience data
- web-based, distributed, open-source, and interoperable

In other words: USGIN is a distributed data-sharing network that uses open-source software and existing World Wide Web infrastructure and browsers. Anyone can access the data shared over USGIN, and any data that is shared according to USGIN specifications is automatically a part of the network.

For more information about USGIN, see the USGIN website (http://usgin.org/).

1.3 Cookbook Prerequisites

To comply with this cookbook, you will need the following:

1. Geologic data in tabular form; valid data formats include:
    a. File geodatabases
    b. Microsoft Access databases
    c. Personal geodatabases
    d. PostGIS databases

    e. SDE databases
    f. Shapefiles
2. Server software capable of fielding web service requests; examples include:
    a. GeoServer and accompanying software
    b. ESRI ArcGIS Server (proprietary)
3. GIS software; examples include:
    a. ArcGIS
    b. Grass GIS
    c. gvSIG
    d. OpenLayers
    e. PostGIS plugin for PostgreSQL DBMS
    f. Quantum GIS
    g. uDIG
4. GeoSciML-portrayal cookbook (this document)
5. An Internet connection to deploy your data as a publicly accessible web service

1.3.1 Prerequisite Notes

The following section provides supplementary background information about topics introduced in Section 1.3:

- As indicated in Section 1.3, Cookbook Prerequisites, **GeoServer** and **ESRI ArcGIS Server** both constitute server software capable of providing web services
    - **ESRI ArcGIS Server** is proprietary server software
    - **GeoServer** is actively maintained, free-and-open-source server software
    - **GeoServer** can be found at the following web locations:
        - http://geoserver.org/
        - http://opengeo.org/technology/GeoServer/
    - **GeoServer** can ingest a **PostgreSQL** database that is **PostGIS**-enabled
- **PostGIS** is a plugin for **PostgreSQL** database management system (DBMS)
- **PostgreSQL** DBMS is a free-and-open-source database management system
- A database management system is a program that controls the creation, maintenance, and use of a database; numerous examples exist, including:
    - Relational DBMS, such as MRDS, Rel, and Micro DBMS
    - SQL DBMS, such as INGRES, DB2, and Microsoft SQL Server
        - SQL DBMS are so named because they use Standardized Query Language (SQL)
    - NoSQL DBMS, such as memcached, Redis, and CouchDB
        - NoSQL DBMS do not use SQL as their primary **query** language (NoSQL means "not only SQL")

For back-end server management, the Arizona Geological Survey uses pgAdmin III (included in the software bundle from opengeo.org). pgAdmin III provides the user with an interface to create and edit databases as well as data tables. An introduction to pgAdmin as well as documentation for using the software can be found on their website here.

For information on PostGreSQL and deploying services using multiple geodatabases, click here.

1.4 Cookbook Workflow

This cookbook will encompass the following overall steps:

1. Map your data from its source schema to an appropriate GeoSciML-portrayal schema; this step can be further broken down into the following sub-steps
    a. Map your data from its source database to an interchange format structured according to the appropriate GeoSciML-portrayal schema
    b. Export your data from the interchange document to a database format appropriate for the server software you will use to deploy your web service
2. Define appropriate symbology for your web service
3. Deploy your web service and, if desired, register it with OneGeology

2 GeoSciML-Portrayal Feature Types

Within the scope of the GeoSciML-portrayal markup language, three feature classes are defined:

- **ContactView**: contains features that represent the mapped traces of boundaries between two geologic units
- **ShearDisplacementStructureView**: used to represent the mapped trace of any type of fault or shear zone that is treated as a single surface for map portrayal
    - The concept of 'Shear displacement structure' includes all fault types or shear zones along which displacement has occurred, from a simple, single 'planar' brittle or ductile surface to a fault system com-prising multiple strands of both brittle and ductile nature. This structure may have some significant thickness (a deformation zone) and have an associated body of deformed rock that may be considered a DeformationUnit. However ,because this feature class is constrained to have a linear geometry, representation is limited to shear displacement structures that are considered single surfaces at the scale of portrayal.
- **GeologicUnitView**: contains polygon features that represent the outcrop of a geologic unit mapped on some outcrop surface called the 'map horizon'. Example map horizons include:
    - Earth surface: the most common map horizon for a geologic map
    - Top of basement: bedrock on which a stratographic section of interest is deposited. For example, in the Grand Canyon, the Precambrian crystalline rocks in the inner gorge (e.g. Vishnu schist) would be considered the basement.
    - Mission-Pima pit, 6/20/1990

  Since the map horizon is not specified in each feature's properties, it must be described in the metadata for the feature collection. Overlapping polygons representing outcrops on different horizons (e.g. 'Earth surface' and 'bedrock surface') will be represented as distinct features in different GeoSciML-portrayal services. Each GeoSciML-portrayal service provides geologic unit outcrop polygons, associated contacts between the units, and fault traces on a single map horizon at any particular location in the extent of the feature collection.

Each of the above feature classes contains features that provide a simplified **view** of existing GeoSciML features. In other words: the above GeoSciML-portrayal feature classes contain features that are also contained by GeoSciML features; but since GeoSciML-portrayal is a simplified **view** of GeoSciML, GeoSciML-portrayal features are simpler than their GoeSciML counterparts.

For more information about the relationship between GeoSciML and GeoSciML-portrayal, refer back to the GeoSciML-Portrayal section.

2.1 ContactView Features

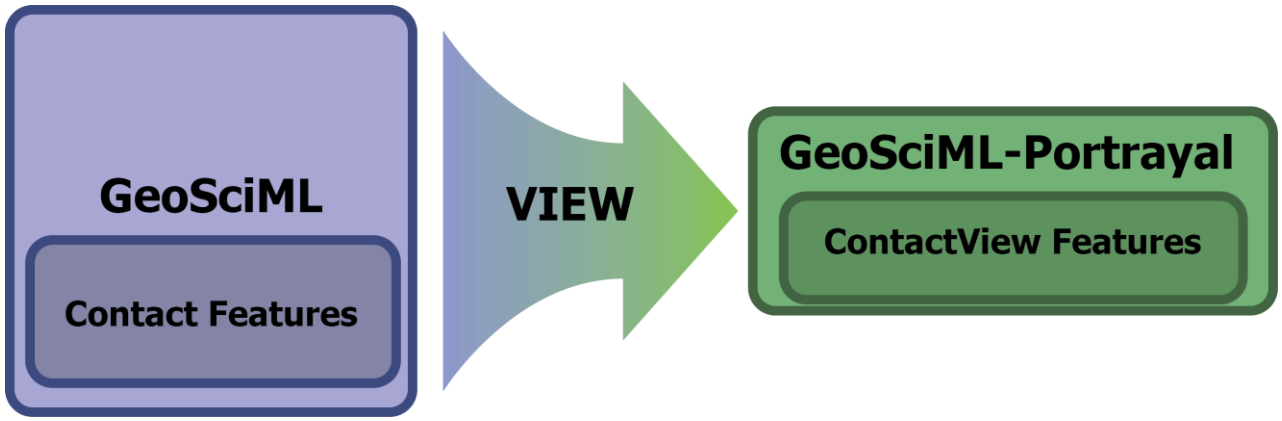


Figure 2: The relationship between GeoSciML Contact Features and GeoSciML ContactView Features

These features provide a simplified **view** of GeoSciML Contact Features (Figure 2).

Table 1 describes each of the elements that describe **ContactView Features**; in this table:

- **Name** is the name of an element that describes a given feature
- **Implementation data type** indicates the specifications for a given element:
    - AnyURI: This element will contain an appropriate URI
        - URIs can be used to refer to other resources, such as controlled concepts in published vocabularies

- String: This element will contain a free-text series of alphanumeric characters
- Decimal: This element will contain a numeric value

Note that these values are modified by the **xs:** prefix; the xs: prefix indicates conformity to the W3C XML Schema

- **Notes** contains a general description of a given element

Table 1: ContactView feature scheme

Name	Implementation data type	Notes
identifier	xs:AnyURI	Globally unique identifier for the individual feature. Recommended practice is that this identifier be derived from the primary key for the spatial objects in the source data in case information needs to be transferred from the interchange format back to the source database. This identifier is analogous to the identifier for a GeoSciML MappedFeature.
name	xs:string	Display name for the Contact. Examples: 'depositional contact', 'unconformity', 'Martin-Escabrosa contact.'
description	xs:string	Text description of the contact; may be a generic description of a contact type taken from an entry on a geological map legend or a more specific description of the particular contact.
contactType	xs:string	Text label specifying the kind of surface separating two geologic units including primary boundaries such as depositional contacts, all types of unconformities, intrusive contacts, and gradational contacts, as well as faults that separate geologic units. Ideally this would be the preferred label for the concept identified by contactType_uri.
observationMethod	xs:string	Metadata snippet indicating how the spatial extent of the feature was determined. ObservationMethod is a convenience property that provides a quick and dirty approach to observation metadata.
positionalAccuracy	xs:decimal	Quantitative values define the radius of an uncertainty buffer around a mappedFeature (e.g. a positionAccuracy of 100 m for a line feature defines a buffer polygon of total width 200 m centered on the line).
source	xs:string	Text describing feature-specific details and citations to source materials, and if available providing URLs to reference material and publications describing the geologic feature. This could be a short text synopsis of key information that would also be in the metadata record referenced by metadata_uri.
contactType_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the Contact types. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
specification_uri	xs:AnyURI	URI referring the GeoSciML Contact feature that describes the instance in detail. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
metadata_uri	xs:AnyURI	URI referring to a formal metadata record describing the provenance of data.
genericSymbolizer	xs:string	Identifier for a symbol from standard (locally or community defined) symbolization scheme for portrayal. There should be an accompanying SLD file. available defining the symbol associated with each genericSymbolizer value.
shape	GM_Object	Geometry defining the extent of the feature of interest. This is the only element with complex content, and must contain a GML geometry that is valid for the Geography Markup Language (GML) simple features profile (OGC 06-049r1.)

2.1.1 ContactView Feature Specifications

In GeoSciML terms, GeoSciML-portrayal ContactView Features will be an instance of a GeoSciML MappedFeature with key property values from the associated ContactFeature summarized in text (data type xs:string) fields as well as properties suffixed with '_uri' containing URIs referring to other resources (for example, controlled concepts in published vocabularies).

2.2 ShearDisplacementView Features

7

These features provide a simplified **view** of GeoSciML ShearDisplacementStructure Features (Figure 3).

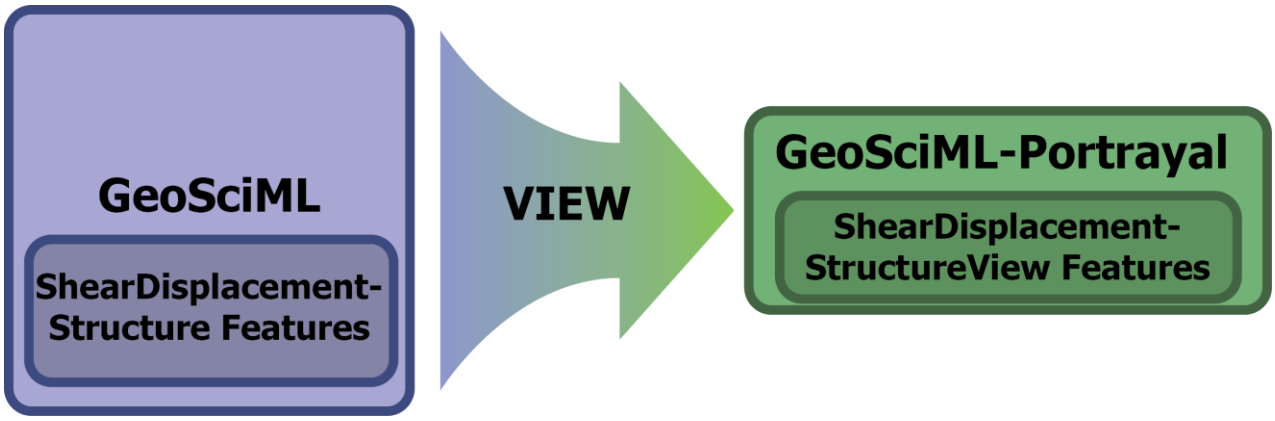


Figure 3: The relationship between GeoSciML ShearDisplacementStructure Features and GeoSciML ShearDisplacementStructureView Features

Table 2 describes each of the elements that describe **ShearDisplacementStructure Features**; in this table:

- **Name** is the name of an element that describes a given feature
- **Implementation data type** indicates the specifications for a given element:
    - AnyURI: This element will contain an appropriate URI
        - URIs can be used to refer to other resources, such as controlled concepts in published vocabularies
    - String: This element will contain a free-text series of alphanumeric characters
    - Decimal: This element will contain a numeric value

  Note that these values are modified by the **xs:** prefix; the xs: prefix indicates conformity to the W3C XML Schema
- **Notes** contains a general description of a given element

Table 2: Elements in ShearDisplacementStructureView feature scheme

Name	Type	Notes
identifier	xs:AnyURI	Globally unique identifier for the individual feature. Recommended practice is that this identifier be derived from the primary key for the spatial objects in the source data in case information needs to be transferred from the interchange format back to the source database. This identifier is analogous to the identifier for a GeoSciML MappedFeature.
name	xs:string	Display name for the the ShearDisplacementStructure. This may be a generic fault type (e.g. 'thrust fault', 'strike-slip fault') or a particular fault name (e.g. 'Moine thrust', 'San Andreas Fault').
description	xs:string	Text description of the ShearDisplacementStructure, typically taken from an entry on a geological map legend.
faultType	xs:string	Type of ShearDisplacementStructure (as defined in GeoSciML).
movementType	xs:string	Summary of the type of movement (e.g. dip-slip, strike-slip) on the ShearDisplacementStructure.
deformationStyle	xs:string	Description of the style of deformation (e.g. brittle,ductile etc) for the ShearDisplacementStructure.
displacement	xs:string	Text summary of displacement across the ShearDisplacementStructure.
geologicHistory	xs:string	Text (possibly formatted with formal syntax) description of the age of the ShearDisplacementStructure (where age is a sequence of events and may include process and environment information).
observationMethod	xs:string	Metadata snippet indicating how the spatial extent of the feature was determined. ObservationMethod is a convenience property that provides a quick and dirty approach to observation metadata when data are reported using a feature view (as opposed to observation view).

Name	Type	Notes
positionalAccuracy	xs:string	Quantitative values which define the radius of an uncertainty buffer around a mappedFeature (eg: a positionAccuracy of 100 m for a line feature defines a buffer polygon of total width 200 m centered on the line).
source	xs:string	Text describing feature-specific details and citations to source materials, and if available providing URLs to reference material and publications describing the geologic feature. This could be a short text synopsis of key information that would also be in the metadata record referenced by metadata_uri.
faultType_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the fault type (ShearDisplacementStructure). Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
movementType_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the ShearDisplacementStructure movement type. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
deformationStyle_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the ShearDisplacementStructure deformation style. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
representativeAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative stratigraphic age interval for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing the all or part of the feature's history.
representativeLowerAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative lower value in a range of stratigraphic age intervals for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing the all or part of the feature's history.
representativeUpperAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative upper value in a range of stratigraphic age intervals for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing the all or part of the feature's history.
specification_uri	xs:AnyURI	URI referring the GeoSciML ShearDisplacementStructure feature that describes the instance in detail. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
metadata_uri	xs:AnyURI	URI referring to a metadata record describing the provenance of data.
genericSymbolizer	xs:string	Identifier for a symbol from standard (locally or community defined) symbolization scheme for portrayal.
shape	GM_Object (GM_curve)	Geometry defining the extent of the feature of interest.

2.2.1 ShearDisplacementStructureView Feature Specifications

In GeoSciML terms this will be an instance of a MappedFeature with key property values from the associated ShearDisplacementStructure feature summarized in text fields (data type xs:string) as well as fields containing identifiers (URI) for fault type, deformation style, movement type, geologic age, and a formally-encoded (ideally in GeoSciML) specification for interoperability. The latter are the properties suffixed with '_uri' and will contain URIs referring to other resources, for example controlled concepts in published vocabularies.

2.3 GeologicUnitView Features

9

These features provide a simplified **view** of GeoSciML GeologicUnit Features (Figure 4), which are used to provide properties associated with a body of material in the earth.

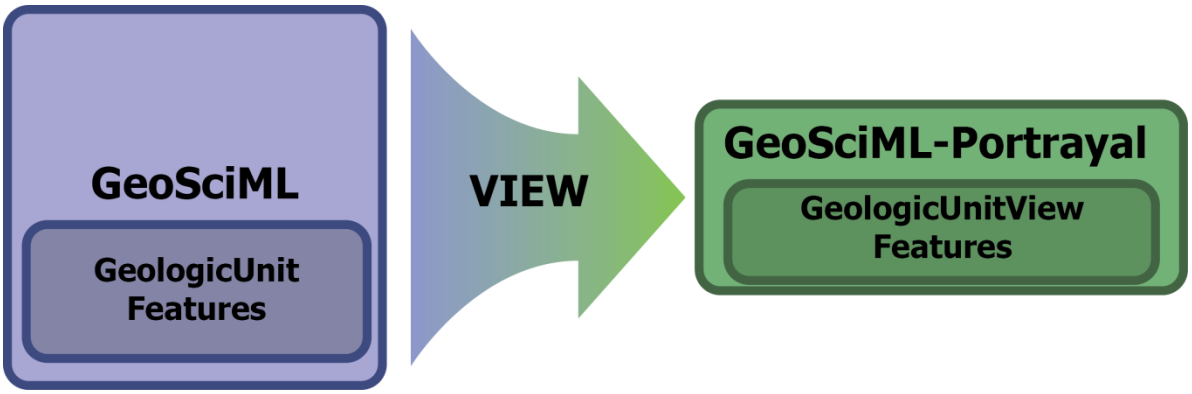


Figure 4: The relationship between GeoSciML GeologicUnit Features and GeoSciML GeologicUnitView Features

Table 3 describes each of the elements that describe **ShearDisplacementStructure Features**; in this table:

- **Name** is the name of an element that describes a given feature
- **Implementation data type** indicates the specifications for a given element:
  - AnyURI: This element will contain an appropriate URI
    - URIs can be used to refer to other resources, such as controlled concepts in published vocabularies
  - String: This element will contain a free-text series of alphanumeric characters
  - Decimal: This element will contain a numeric value

  Note that these values are modified by the **xs:** prefix; the xs: prefix indicates conformity to the W3C XML Schema
- **Notes** contains a general description of a given element

Table 3: Elements in GeologicUnitView feature class

Name	Type	Notes
identifier	xs:AnyURI	Globally unique identifier for the individual feature. Recommended practice is that this identifier be derived from the primary key for the spatial objects in the source data in case information needs to be transferred from the interchange format back to the source database. This identifier is analogous to the identifier for a GeoSciML MappedFeature.
name	xs:string	Display name for the GeologicUnit; this may be populated by a geologic unit name, or more likely, an abbreviation used to label outcrops of the unit in a map display.
description	xs:string	Text description of the GeologicUnit, typically taken from an entry on a geological map legend.
geologicUnitType	xs:string	Type of GeologicUnit (as defined in GeoSciML).
rank	xs:string	Stratigraphic rank of GeologicUnit (as defined in GeoSciML). Examples include formation, member, group, or supergroup.
lithology	xs:string	Text (possibly formatted with formal syntax) description of the GeologicUnit's lithology.
geologicHistory	xs:string	Text (possibly formatted with formal syntax) description of the age of the GeologicUnit (where age is a sequence of events and may include process and environment information).
observationMethod	xs:string	Metadata snippet indicating how the properties of the feature were determined. ObservationMethod is a convenience property that provides a quick and dirty approach to observation metadata. Example values might include 'field observation by author', 'compilation from published maps', 'air photo interpretation'.

Name	Type	Notes
positionalAccuracy	xs:string	Quantitative values define the radius of an uncertainty buffer around a mappedFeature (eg: a positionAccuracy of 100 m for a line feature defines a buffer polygon of total width 200 m centered on the line).
source	xs:string	Text describing feature-specific details and citations to source materials, and if available providing URLs to reference material and publications describing the geologic feature. This could be a short text synopsis of key information that would also be in the metadata record referenced by metadata_uri.
geologicUnitType_uri	xs:AnyURI	URI referring to a controlled concept from a vocabulary defining the GeologicUnit types. Mandatory property - if no value is provided then a URI referring to a controlled concept explaining why the value is nil must be provided.
representativeLithology_uri	xs:AnyURI	URI referring to a controlled concept specifying the characteristic or representative lithology of the unit. This may be a concept that defines the super-type of all lithology values present within a GeologicUnit or a concept defining the lithology of the dominant CompositionPart (as defined in GeoSciML) of the unit. Typically this identifier might be used as the symbol key for a lithologic map portrayal of the geologic unit features.
representativeAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative stratigraphic age interval for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing all or part of the feature's history. Typically, this identifier might be used as a symbol key for a geologic-age-based portrayal of the geologic unit features.
representativeLowerAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative younger value in a range of stratigraphic age intervals for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing all or part of the feature's history.
representativeUpperAge_uri	xs:AnyURI	URI referring to a controlled concept specifying the most representative older value in a range of stratigraphic age intervals for the GeologicUnit. This will be defined entirely at the discretion of the data provider and may be a single event selected from the geologic feature's geological history or a value summarizing all or part of the feature's history.
specification_uri	xs:AnyURI	URI for a complete description of the geologic unit cropping out within the extent of the feature's geometry. Preferred representation is a GeoSciML GeologicUnit feature instance. Mandatory property - if no value is provided then a nil reason URI explaining why the value is nil must be provided.
metadata_uri	xs:AnyURI	URI referring to a metadata record describing the provenance of data.
genericSymbolizer	xs:string	Identifier for a symbol from standard (locally or community defined) symbolization scheme for portrayal.
shape	GM_Object, (GM_polygon)	Geometry defining the extent of the feature of interest.

2.3.1 GeologicUnitView Feature Specifications

In GeoSciML terms GeologicUnitView features are instances of MappedFeature with key property values from the associated GeologicUnit feature summarized in text fields (data type xs:string) for human data consumers, with standard identifiers for geologic unit type, representative lithology, and geologic age. The specification_uri identifies a description resource specific to the geologic unit cropping out in the extent of the polygon (or other) geometry of the feature. The specification_uri should dereference to yield a formally-encoded representation of the geologic unit, ideally in GeoSciML for interoperability. Properties populated by identifiers are suffixed with '_uri' and contain URIs referring to other resources, for example controlled concepts in published vocabularies.

3 Integration of Web Map Services

The key to interoperable map services is the ability to portray adjacent maps using the same symbolization scheme such that when the maps are displayed, the visual discontinuity at the boundary is minimized. Because of the discrepancies in mapping interpretation and intention, there are commonly differences in the definition of map units between geologic maps produced by different authors or at different times. Resolving such differences is a compilation process that often requires additional field work; this is outside the scope of service deployment.

What can be avoided is visual discontinuity due to different portrayal schemes that assign different colors to similar units on adjacent maps. GeoSciML-portrayal addresses this issue by requesting that geologic unit mapped features be categorized using the CGI Simple Lithology vocabulary (representativeLithology_uri) and the ICS stratigraphic time scale (representativeAge_uri) using registered identifiers.

The use of community portrayal schemes associated with these categorization schemes provides a basis for first order lithologic and age harmonization between maps provided by different services. The genericSymbolizer property is intended to enable preservation of the original legend units and portrayal from the source map.

4 Deployment Cookbook

Once geologic map data is automated into a structured digital form, publication of the data for access as an Open Geospatial Consortium (OGC) web map service (WMS) or web feature service (WFS) is a three-step process:

1. Extract the data from its source data set
2. Transform the data to an interchange format
    - In the context of this cookbook, the GeoSciML-portrayal schema constitutes the appropriate interchange format
3. Load the transformed data into whatever sort of data store is most convenient for the server implementation of the OGC WMS and WFS

This process is commonly called 'ETL'.

There exists a broad spectrum of approaches to the ETL process. We focus in this tutorial on what we deem to be a common situation in which the source data is in a database feature class (possibly in an ESRI geodatabase, shape file, or PostGIS table) with one record for each geometric feature (line or polygon), each feature has either a link to a map unit description or includes fields specifying the age and lithology of the unit. Source information for the map data and unit descriptions is also required.

The most difficult part of the ETL process is typically determining what representative lithology, age, younger age and older age categories from a standard vocabulary to assign to each unit using URIs. OneGeology conformance proscribes that these vocabularies should be the CGI SimpleLithology vocabulary for lithology and the International Stratigraphic Commission Geologic Time Scale 2009 for age.

4.1 Mapping from source data

The following outline describes the basic ETL workflow:

1. Determine which fields in the source data contain the information that is to be delivered in the interchange format fields. Multiple source fields may be combined into single interchange fields, and a single source field may impact values in multiple interchange fields.
2. Determine what steps are necessary to get the content into the interchange format. This may involve some calculation, such as concatenating text from multiple fields to populate requisite fields in the interchange format. Use of the standard vocabularies for interoperability will likely require mapping vocabulary terms in the source data to identifiers for concepts in the controlled vocabularies for the fields that require URIs (see next section).
3. Set up a query to generate a table with **field names exactly matching the field names in the interchange schema**. In some cases it may be convenient to generate the interchange schema table in several steps, populating subsets of the fields each time. It may also be useful to generate a table with unique combinations of contact, fault, or geologic unit properties from the source data, and map each combination to corresponding properties in the interchange format. Depending on individual situations, the unique descriptions can be identified using the specification_uri, name, or genericSymbolizer field; identifiers in this field can then be used to join the interchange format properties with individual features.

4. The table of unique descriptions can then be **joined** with the geometry elements to generate the final feature classes for the portrayal view service. ESRI Shapefiles and PostGIS tables are useful representations for the final feature class that the WMS server will use. Shapefiles will truncate field names longer than 10 characters, so in order to use them, the map server must support defining XML element names distinct from data table field names that are longer than 10 characters.
5. Set up the configuration for the WMS server. This procedure will be specific to the particular server implementation that is being used. Procedures for deployment using GeoServer and ArcGIS server are discussed in Appendices A and B respectively.
6. Create metadata for the dataset and its service distribution and load it into a catalog server to make the serviced discoverable.

4.1.1 Interchange formats

For a variety of reasons, this cookbook specifies GeoSciML-portrayal XML documents as interchange formats. These documents should be structured according to an appropriate GeoSciML-portrayal schema.

The GeoSciML-portrayal feature classes detailed in Section 2 constitute **representations** of three appropriate schemas.

- Table 1 represents the schema for the GeoSciML-portrayal ContactView schema by describing each element and the data that should be entered therein
- Table 2 represents the schema for the GeoSciML-portrayal ShearDisplacementStructureView by describing each element and the data that should be entered therein
- Table 3 represents the schema for the GeoSciML-portrayal GeologicUnitView by describing each element and the data that should be entered therein

By entering geoscience data into an XML document structured according to any of the above schemas, one is creating an interchange document.

4.1.2 Schema mapping in ESRI ArcGIS

This section provides an example of schema mapping in ESRI ArcGIS, using ArcMap, ArcCatalog, and ArcToolbox.

1. Open ArcCatalog
2. Locate or create a **spatial database** that will be used to host your web services in accordance with the GeoSciML-portrayal schema
    a. Sometimes it's easier to create a new database in accordance with the GeoSciML-portrayal schema; other times, it's better to create a new table within an existing database. The conditions needed to make this determination are multifarious, highly dependent on individual context, and therefore beyond the scope of this document.
3. Connect to the desired spatial database under **Database Connections**. To connect to a database, you will need to click **Add Spatial Database Connection** (Figure 5) in ArcCatalog (below **Database Connections**). This will open the **Spatial Database Connection** window.

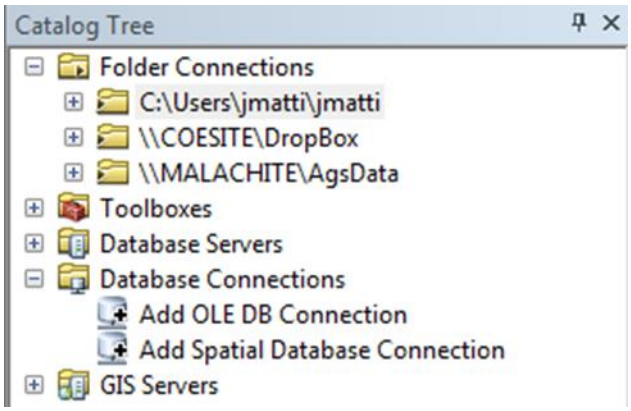


Figure 5: Adding a Spatial Database Connection in ArcCatalog

4. In the **Spatial Database Connection** window, specify the location and (if necessary) the credentials needed to access the desired spatial database. Click **OK** when finished.
5. After you have connected to the spatial database you intend to use for your services, right-click the desired database; in the context menu that appears, click **New** > **Feature Class**…

   This will bring up the **New Feature Class** dialogue box.

a. In the **New Feature Class** dialogue box, type the name for your feature class (this is *not* the name of the web service). Use the pulldown menu to select the appropriate **Type** of feature geometry for the data you are going to publish. When you are finished filling out the form, click **Next**.
b. Choose the **coordinate system** for your feature class.
    i. For the purposes of interoperability, WGS 1984 is the recommended coordinate system. To select the WGS 1984 coordinate system, navigate to the **Geographic Coordinate System** > **World** directory and click the WGS 1984 coordinate system.

When you have located the desired coordinate system, click **Next**.

c. Specify the resolution for your feature class by entering an appropriate amount of degrees into the **XY Tolerance** field.
    i. The XY tolerance defaults to 0.000000008983153, which is fine for most applications, but should be changed if necessary.

When finished, click **Next**.

d. Specify the **Configuration keyword** for your feature class.
    i. The default Configuration keyword is fine for most database storage architectures.

When finished, click **Next**.

e. Enter appropriate field headings under the **Field Name** column; specify the parameters for each field by entering appropriate values in the **Data Type** column and in the **Field Properties** section (Figure 6).
    i. **To comply with GeoSciML-portrayal, all field names and associated parameters must conform exactly to the GeoSciML-portrayal schema that is appropriate for your dataset**. This includes capitalization, spacing, underscores, field lengths, and cardinality.

    Appropriate values can be found in Section 4, GeoSciML-Portrayal Feature Types.

When done, click **Finish**.

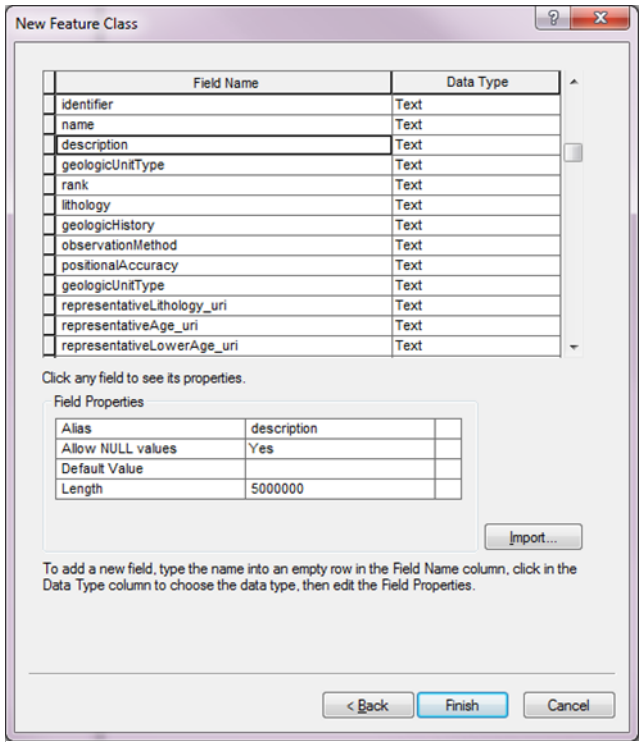


Figure 6: Step 5E - Entering GeoSciML-portrayal schema field names and parameters

Your new **Feature Class**, structured by the appropriate **GeoSciML-portrayal schema**, will appear in the spatial database you selected or created above.

Next, you will use **ArcToolbox** to import and map your existing data into the new feature class you just selected or created above.

1. Open **ArcCatalog**

2. On the menu bar of ArcCatalog, click the **ArcToolbox** button
3. In ArcToolbox, click **Data Management Tools > General > Append**
4. In the **Append** window (Figure 7), you need to perform a number of tasks:
    a. Specify the **Input Dataset**. This will be your source data that does not conform to the GeoSciML-portrayal schema
    b. Specify the **Target Dataset**. This will be the Feature Class you created above
    c. Change the default **Schema Type** to **NO_TEST**. This last instruction is very important; it allows you to perform the schema mapping
    d. For each field in the **Field Map (optional)** table, map your **Input Dataset** to the **Target Dataset**:
        i. Right-click a target field in the **Field Map (optional)** table
        ii. In the context menu that appears, click **Add Input Field**
        iii. In the dialog box that appears, select the corresponding field from your **Input Dataset**
    e. Repeat steps **i-iii** for every field in the **Target Dataset**
    f. When finished, click **OK**
5. Repeat Step 4 in this list for each dataset you wish to deploy as a [web service](#).

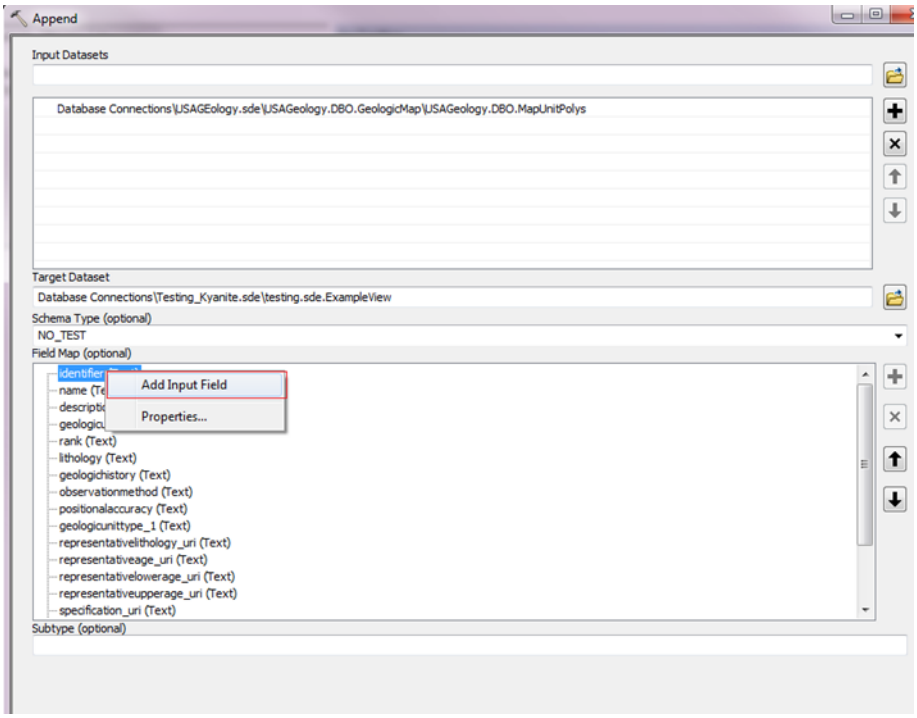


Figure 7: The Append window

You are now ready to deploy your services.

Note: If you have multiple tables that you need to map into a single schema, it may be easier to write a SQL query to do the mapping for you, as described in Section 4.1.3. For our GeologicUnitView features, we originally mapped the NCGMP-09 data into GeoSciML-portrayal in PostGIS using an SQL query by creating a view. For more information, see page XX of the accompanying document.

For an in-depth, technical description of how our tables were originally mapped, visit our lab.usgin blog by following this link.

4.1.3 Schema mapping in SQL

The following example constitutes a [schema](#) mapping query, written in Postgres-flavor SQL (minor modifications may be necessary for other SQL environments). The source schema is the USGS/AASG National Cooperative Geologic Mapping Program NCGMP09 database schema (Haugarud et al., 2010), and the target is a GeoSciML-portrayal GeologicUnitView. This is meant as an illustrative example of the kind of processing that may be required. The level of difficulty in the schema mapping is largely determined by how similar the internal data model schema is to the GeoSciML conceptual model, and how much content the data provider wishes to make available in the portrayal view.

Code Example 1: PostGIS SQL query to produce GeoSciML-portrayal view for map unit polygons.

```
1  CREATE TABLE sde.geologicunitview As
2  SELECT
3  mup.objectid as objectid,
4  mup.mup_id AS identifier,
5  dmu.description AS name,
6  mup.notes AS description,
7  'Geologic Unit'::text AS "geologicUnitType",
8  'Not Specified'::text AS rank,
9  polyextattr.lith6name AS lithology,
10 dmu.age AS "geologicHistory",
11 (datasources.source::text || ' '::text) || datasources.notes AS source,
12 'http://…/cgi/geologicunittype/0008'::text AS "geologicUnitType_uri",
13 polyextattr.lithuri AS "representativeLithology_uri",
14 polyextattr.ageuri AS "representativeAge_uri",
15 mapunitages.ageyoungerterm AS "representativeLowerAge_uri",
16 mapunitages.ageolderterm AS "representativeUpperAge_uri",
17 'http://www.opengis.net/def/nil/OGC/0/missing'::text AS specification_uri,
18 'http://catalog.usgin.org/geoportal/…'::text AS metadata_uri,
19 mup.mapunit AS "genericSymbolizer",
20 shape::geometry as shape
21 FROM mapunitpolys AS mup
22 LEFT JOIN polyextattr ON mup.mapunitpolys_id = polyextattr.ownerid
23 LEFT JOIN datasources ON mup.datasourceid = datasources.datasources_id
24 LEFT JOIN descriptionofmapunits as dmu ON mup.mapunit = dmu.mapunit
25 LEFT JOIN mapunitages ON mup.mapunit = mapunitages.mapunit;
```

This example SQL query creates a table that has fields conforming to the GeologicUnitView for the geosciml portrayal scheme. The SQL is a PostGIS dialect, so field names that are not all lower case are enclosed in quotes, and constant values are explicitly types (e.g. '::text'). 'mapunitpolys' (mup),' datasources', 'descriptionof-mapunits' (dmu) are tables from the source NCGMP09 database, abbreviated as indicated in the query. 'poly-extattr' and 'mapunitages' are Postgres views that aggregate lithology and age properties correlated with polygons or map units respectively. The source data is a regional dataset that correlates lithology and age categories at the polygon level to document lithologic and age variation within the map units. This information is contained in the 'extendedattributes ' and 'geologicevent' tables in the source dataset (See NGCMP, 2010 or http://ngmdb.usgs.gov/Info/standards/NCGMP09/ for details on the input data structure).

4.1.4 Notes on schema mapping in SQL

Line numbers refer to numbered lines of text in Code example 1.

- Line 1. Result table is placed in ESRI sde schema in PostGIS database.
- Line 3. For ESRI Spatial Database Engine (SDE) to recognize the table as a feature class that may be used to source a service, an integer unique-value field must be present. In this case the field from the source dataset is used because it serves the same purpose there. This field does not need to be set up as an autoincrement field because the resulting table is a read-only view of the data. The ObjectID field will likely be placed at the last column in the table for better interoperability with systems that do not require such a field.
- Line 4. Unique polygon identifiers from the source dataset are carried into the interchange format view.
- Line 7. Source data map units are not categorized into specific unit types (e.g. lithostratigraphic, chronostratigraphic, geophysical, etc.), so this field gets a constant value 'Geologic Unit', which is the preferred label for the corresponding concept in the CGI Geologic Unit Type vocabulary.
- Line 8. Likewise, rank is not assigned for the regional geologic units in the source dataset, so this is a constant that is the preferred label ('Not Specified') for the corresponding concept in the CGI Strati-graphic Rank vocabulary.

- Line 9, 10. Text descriptions of lithology and the age of the unit are derived from corresponding fields in the source data. The lithology description is mapped at the individual polygon level (via the polyextattr Postgres view), note the JOIN in line 22 that uses the .mapunitpolys_id as the foreign key. The age text description is mapped at the map unit level, note the JOIN in line 24 that uses the .mapunit field as the foreign key.
- Line 11. The text for the source field in the portrayal scheme is calculated by concatenating two fields (.source and .notes) from the datasources table in the input dataset.
- Line 12. Since all the map units are classified simply as 'Geologic Unit', this URI is also a constant value from the CGI Geologic Unit Type vocabulary.
- Line 13, 14. Representative lithology and age categories are assigned on a polygon by polygon basis, thus the polyextattr table is the source of these URIs. Note the JOIN in line 22 that uses the .mapunitpolys_id as the foreign key.
- Line 15, 16. Younger and older age bounds are assigned by map unit (not at the individual polygon level), so these fields are sourced from the mapunitages aggregation query in the database; this query uses the extendedattributes table to join map-units with geologicevent records that contain the lower and upper age bounds for the unit.
- Line 17. specification_uri is a link to a more complete description of the geologic unit that crops out in the polygon's extent. This description can be viewed as a resource in the context of web architecture. The intention of the GeoSciML design team was that this URI should dereference to return a full Geo-SciML GeologicUnit element instance. With the use of content negotiation on the web, this description might also have representations as a web page or other structured description (rdf, owl…). In the example instance, a more complete description is not available and an OGC nil URI is used to indicate that the resource is missing (does not exist).
- Line 18. metadata_uri identifies a metadata resource that contains more complete information on the provenance of the information in the feature element. This may be a metadata record that is scoped to the individual feature, or may identify a metadata record describing some collection of polygons that have similar enough provenance to document together. The text in the source field (Line 11) should be a succinct summary of the information in this metadata record pertinent to the containing feature.
- Line 19. The genericSymbolizer field should contain an identifier for the symbol used to display this poloygon in the default portrayal (legend) chosen by the data provider for the gsmlp:GeologicUnitView instance in the containing feature collection. This field can be used to capture the map unit assignment and portrayal color scheme from the original data from which the GeologicUnitView polygon was digitized. The legend may be encoded in an accompanying Styled Layer Descriptor (SLD) file, and if such an SLD exists it should be recorded as a related resource in the metadata record specified by metadata_uri.
- Line 20. The shape field contains a representation of the geometry of the outcrop area described by the GeologicUnitView instance. The content of this field will be managed by the GIS and WMS server, and thus will generally not need to be manipulated by users outside the GIS environment. The shape field may need to be cast into a recognized geometry field in order to get PostGIS and ESRI SDE to recognize the output as a feature class. In our configuration at AZGS, we ended up using this:

  'st_geometryfromtext(ST_AsText(shape)::text, 4326)::geometry as shape'

  More expert PostGIS users may have a better solution for this problem.

4.2 Vocabulary mapping

There are many possible approaches to mapping terms from one vocabulary to another.

One situation in which a standard process can be defined involves a controlled vocabulary used to populate a field in the source data, and that field maps directly to a field in the interchange format. For instance: if the source data contains a 'dominant lithology' field, the information in this field can be used to populate the 'representativeLithology_uri' for a GeologicUnitView feature.

Recommended procedure:

1. Produce a table of the unique values in the source data
2. Add a column to this table for the corresponding element in the target interchange scheme
3. Determine the best matching value from the interchange vocabulary for each term in the source vocabulary

In general: the most specific term from the interchange vocabulary that completely subsumes (encompasses) the meaning of the term in the source vocabulary should be used.

If the source vocabulary has terms that are more specific than the controlled vocabulary, there will be some information lost in this process, but the original source terminology can be preserved in free-text description fields in the interchange document. Sample free-text fields that may be used to preserve source vocabulary information are as follows:

- Lithology
- geologicHistory

Remember, the primary purpose of the controlled vocabulary fields is for data integration and use as search criteria, likely by non-expert users.

In some cases, unique values from a combination of multiple source data fields may be necessary to define mapping to interchange concepts, particularly for the following interchange fields:

- representativeAge_uri
- representativeLithology_uri
- genericSymbolizer

The procedure is the same, but any unique values query will involve more than one source field and multiple-field **joins** will be necessary to construct queries generating the output schema content.

Standard vocabularies are listed in the accompanying Excel Workbook.

4.3 Excel Workbook template

The accompanying Microsoft Excel Workbook provides spreadsheets containing the content models for the description properties associated with geologic contacts, faults or shear zones, and geologic units in appropriate GeoSciML-Portrayal schemas.

These spreadsheets do not include geometry fields that are necessary for an actual GIS feature class for the corresponding ContactView, ShearDisplacementStructureView, and GeologicUnitView features. Rather, they provide a template for compiling the necessary descriptions (combinations of attributes) that can then be **joined** with a GIS feature class to produce the dataset for feature web service deployment.

The specification_uri in each feature class provides a link to structured representation of the geologic feature that is intended to be a GeoSciML feature.

The metadata_uri provides a link to a structured metadata record maintaining complete provenance information for features.

See the notes tab in the workbook for additional information describing the GeoSciML Portrayal feature classes and the work-sheets included in this document.

4.4 Configuring OGC services

GeoSciML-Portrayal is intended to support map services for online viewing and exploration of geologic data, and the creation of mash-ups integrating data from different servers and possibly different thematic domains.

The OGC WMS specification includes a 'getFeatureInfo' operation that returns a description of a map feature at a user-identified point on the map. The operation returns a document determined by the server configuration. A variety of different formats may be offered, and specified by a request parameter. For GeoSciML-Portrayal services, an XML document conforming to the GeoSciML-Portrayal XML schema should be offered. Other getFeatureInfo response formats may also be offered.

Server software that implements the OGC Web Map Service will typically also allow parallel deployment of an OGC Web Feature Service with no additional effort. The XML provided by the feature service is the same as that provided in the WMS getFeatureInfo response formats.

The GML simple feature profile requires that each feature include a gml:id attribute. According to the GML specification, this identifier must be an alphanumeric string that has an alphabet letter as its first character. These ID's can be generated from the primary key in the data table used to compose the features for the map service. If data are coming from an ArcGIS environment, the ObjectID required for any ArcGIS feature class can be used to construct the gml:id. If no primary key is present in the feature class for the service, one will have to be added to source the gml:id property. Technically, the gml:id is meant to be unique within the scope of a feature collection

from a particular service to enable internal cross referencing within a single WFS response document or to allow queries to retrieve a feature using the gml:id.

Server configuration using GeoServer and ESRI ArcGIS Server is described in Appendices A and B as examples of the workflow necessary to set up a GeoSciML-Portrayal service. Other implementations exist, and the details of the procedure will change as new versions of the server software are released.

5 Styling

Part of the process of deploying a web service includes defining a map legend.

ESRI ArcGIS Server provides a number of tools to define a map legend, but if you are using GeoServer to deploy your data as a web service, you will need an OGC SLD file to define your map legend.

OGC Styled Layer Descriptors (SLDs) are XML documents that define a mapping from some property of an OGC feature to a symbol to use for portrayal. By mapping a property of a feature to a symbol that will be used to represent a feature with a given property, you are essentially defining a map legend.

SLD files for portrayals using the CGI simple lithology and ICS 2009 time scale URIs are available at the following web location:

> http://schemas.usgin.org/schemas/slds/

Aside from those SLDs made available at the above web location, it is possible to use a basic text editor to manually create an SLD. Unfortunately, manually creating an SLD for data with a complicated symbology scheme, such as a geologic map, is not a trivial proposition, for a variety of reasons. For example, colors must be converted from RGB to hexadecimal on a per-symbol basis, which is a labor-intensive process. Instructions for creating an SLD file manually are out of the scope of this document.

Alternatively, software exists that permits the user to create an SLD file based on symbology extant in ArcGIS. See the following section for an example.

5.1 How to Create a Styled Layer Descriptor (SLD) using Arc2Earth

Arc2Earth is a plugin for ArcGIS that is available in several different editions. The **Community** edition is free and provides access to the SLD generator. Arc2Earth is available at http://www.arc2earth.com/.

1. Install the Arc2Earth plugin
2. Open ArcMap
3. Open the ArcMap project containing the layer or layers to export as SLDs
4. Select the desured layer in the catalog tree
5. In the **Arc2Earth** toolbar, click **Export** > **Export Layer Style to SLD**
6. Navigate to the appropriate location
7. Click **Export**

After your SLD has been exported, open the .sld in any XML or text editor. Note that the Arc2Earth plug-in adds an outline (stroke) to polygon layers. If you are working with polygon layers, the stroke may need to be tailored or removed manually. Also, it may be necessary to customize the heading for your SLD, or change the schema location.

> **Note**: though .sld files, like any XML document, may be edited by a generic text editor, such as the **Notepad** or **Wordpad** text editors included with Windows operating systems, some text editors are better suited to working with XML documents than others. One example of a free-and-open-source text editor designed with XML support is Notepad++.

Below is a snippet of an SLD created using Arc2Earth (Code Example 2).

Code Example 2: Styling in XML

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<StyledLayerDescriptor xmlns:xlink='http://www.w3.org/1999/xlink' xmlns:gml='http://www.opengis.net/gml'
xmlns:ogc='http://www.opengis.net/ogc' xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance' version='1.0.0'
xsi:schemaLocation='http://www.opengis.net/sld StyledLayerDescriptor.xsd' xmlns='http://www.opengis.net/sld' >
        <NamedLayer>
                <Name><![CDATA[GeologicAge]]></Name>
                <UserStyle>
                        <FeatureTypeStyle>
                                <Rule>
                                        <Name><![CDATA[Holocene]]></Name>
                                        <Title><![CDATA[Holocene]]></Title>
                                        <ogc:Filter>
                                                <ogc:PropertyIsEqualTo>
                                                        <ogc:PropertyName>representativeage_uri</ogc:PropertyName>

        <ogc:Literal><![CDATA[http://resource.geosciml.org/classifier/ics/ischart/Holocene]]></ogc:Literal>
                                                </ogc:PropertyIsEqualTo>
                                        </ogc:Filter>
                                        <PolygonSymbolizer>
                                                <Fill>
                                                        <CssParameter name="fill" >#f9f97f</CssParameter>
                                                        <CssParameter name="fill-opacity" >1</CssParameter>
                                                </Fill>
                                                <Stroke>
                                                        <CssParameter name="stroke" >#f0f0f0</CssParameter>
                                                        <CssParameter name="stroke-width" >0.4</CssParameter>
                                                        <CssParameter name="stroke-opacity" >1</CssParameter>
                                                </Stroke>
                                        </PolygonSymbolizer>
                                </Rule>
                                <Rule>
                                        <Name><![CDATA[Quaternary]]></Name>
                                        <Title><![CDATA[Quaternary]]></Title>
                                        <ogc:Filter>
                                                <ogc:PropertyIsEqualTo>
                                                        <ogc:PropertyName>representativeage_uri</ogc:PropertyName>

        <ogc:Literal><![CDATA[http://resource.geosciml.org/classifier/ics/ischart/Quaternary]]></ogc:Literal>
                                                </ogc:PropertyIsEqualTo>
                                        </ogc:Filter>
                                        <PolygonSymbolizer>
                                                <Fill>
                                                        <CssParameter name="fill" >#f9f97f</CssParameter>
                                                        <CssParameter name="fill-opacity" >1</CssParameter>
                                                </Fill>
                                                <Stroke>
                                                        <CssParameter name="stroke" >#6e6e6e</CssParameter>
                                                        <CssParameter name="stroke-width" >0</CssParameter>

                                                        <CssParameter name="stroke-opacity" >1</CssParameter>
```

6 Deploying Your Web Service

If you have completed the following steps, you are ready to deploy your web service in a GeoServer environment.

1. You meet the prerequisites detailed in Section 1.3, Cookbook Prerequisites
2. You have successfully mapped your data to an appropriate GeoSciML-portrayal schema, as described in Section 4, Deployment Cookbook
3. You have symbolized your data, as described in Section 5, Styling.

If you have followed all previous steps in this cookbook, you should now be ready to deploy your GeoSciML-portrayal web service.

The overall process of web service deployment is as follows:

1. Place your data in a web-accessible folder
2. Set up and/or access server software that is capable of providing a web service, such as:
    a. GeoServer
    b. ArcGIS Server
3. Use the desired server software to deploy your data as a web service
4. Test your web service (see Section 7, Testing Your Web Service, for more details)

Unfortunately, there are as many ways to deploy a web service as there are individual contexts; to attempt to account for all possible scenarios is beyond the scope of this cookbook.

Instead, this cookbook will provide two generic examples of web service deployment:

- Web service deployment in a GeoServer environment (Appendix A)
- Web service deployment in an ArcGIS Server environment (Appendix B)

7 Testing Your Web Service

Testing your web service is very simple procedure: you attempt to access it using your preferred GIS application. If you can access your web service successfully, and if your web service displays properly, then nothing more need be done.

If, on the other hand, you cannot access the web service, or if it displays improperly, then you need to troubleshoot your web service.

7.1 Web Service Requests

In order to access your web service, you must perform a **GetCapabilities** request in your web browser. A **GetCapabilities** request is a request for data from a web service that returns an XML document describing the capabilities of the web service (Figure 8).

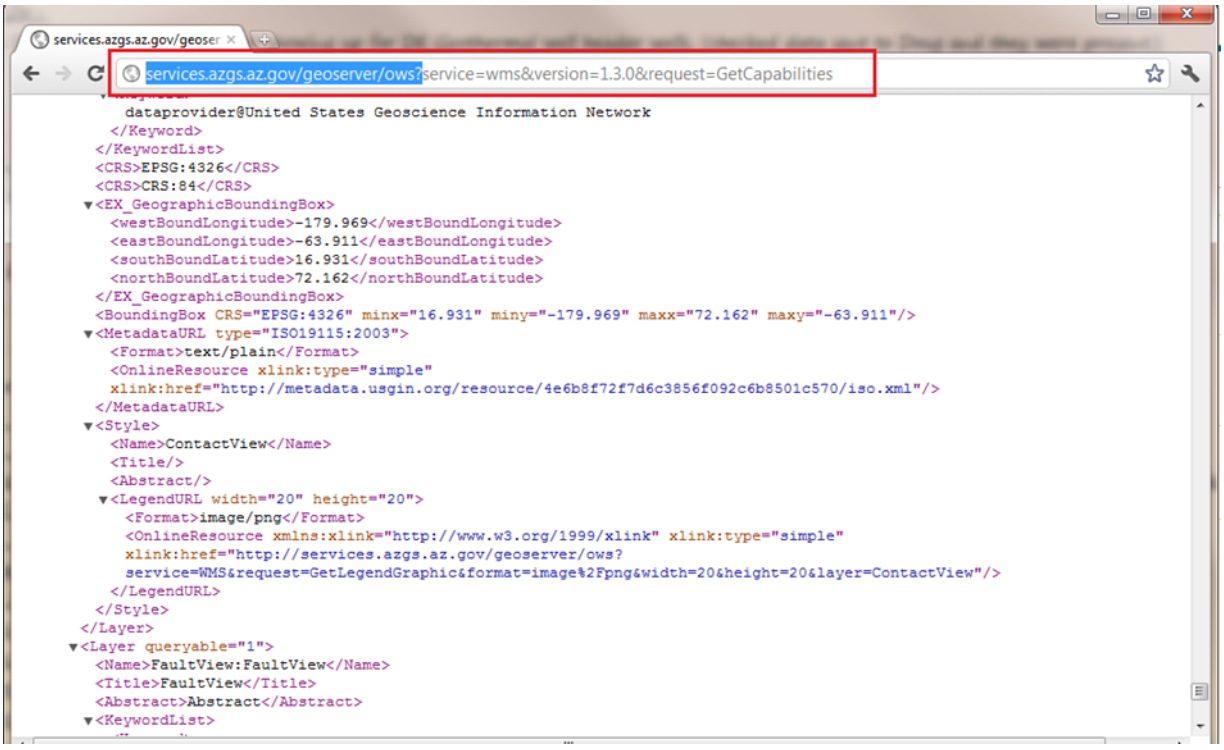


Figure 8: A typical capabilities document returned by a GetCapabilities request in a web browser

To perform a **GetCapabilities** request, simply use your web browser to navigate to the appropriate **service directory** and click your web service. This will return a **capabilities document** in your web browser; the web address of this capabilities document is your **service request**.

Sample GetCapabilities requests:

21

http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/ALBoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities&service=WFS

http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/ALBoreholeTemperatures/MapServer/WMSServer?request=GetCapabilities&service=WMS

In both examples, **bolded** text constitutes the **service endpoint**, whereas all text afterwards constitutes the **service request**. **In order to test your web service, you will need to select and copy the service endpoint of your web service and append a service request to it**.

Once you have accessed the capabilities document of your web service and copied the service endpoint, you can proceed to the Section 7.2.1 or 7.2.2.

Sections 7.2.1 and 7.2.2 contain instructions for accessing a web service via GIS applications. As 90% of the USGIN user base uses ArcGIS, Section 7.2.1 will describe the process of accessing a web service via ArcCatalog. Section 7.2.2, however, will describe the process of accessing a web service via the free-and-open-source desktop GIS application uDig.

If you are not sure how to access your web service, here is a short list of GIS software that supports web services:

- ArcGIS Desktop
- ArcGIS Explorer
- ArcGIS Online
- Grass GIS
- gvSIG
- OpenLayers
- Quantum GIS
- uDig

7.2 Adding a Web Service in ArcCatalog

1. Open **ArcCatalog**
2. In the Catalog Tree, expand **GIS Servers**
3. Click **Add WMS Server** or **Add WFS Server**, depending on the kind of web service you wish to access.
    - See the glossary for more information about WMS and WFS.
4. Paste the **service endpoint** (see Section 7, Testing) for your web service into the URL field
5. It may be necessary to navigate back to GeoServer for your URL. On the GeoServer Welcome page, Service Capabilities are listed on the right. Click on the latest version of WMS or WFS, whichever service you wish to use.
6. When redirected, copy the URL (from beginning up to the question mark) and paste into the URL field of Add WMS Server in ArcCatalog:
7. http://services.azgs.az.gov/geoserver/ows?service=wms&version=1.3.0&request=GetCapabilities
8. Click Get Layers
9. Click OK
10. Add the desired layer(s) to ArcMap

7.3 Adding a Web Service in uDig

1. Open uDig
2. Software available at: http://udig.refractions.net/
3. Create a new map, or open an existing map to which you would like to add the web service
4. On the upper menu bar, click Layer > Add…
5. In the window that appears, click Web Map Server or Web Feature Server, as appropriate; click Next
6. Paste the service endpoint in the URL field; click Next
7. In the Resource Selection window that appears, select all layers you wish to add. When you are done, click Finish

8 Registering with OneGeology and Submitting your Service

To submit your service to OneGeology, you will first need to register your organization on their website by filling out their registration form. If the data provider is able to host their own service, visit this link:

http://www.onegeology.org/technical_progress/data_coordination.cfm

If the data provider would prefer a 'buddy' serve their data, fill out this form:

http://www.onegeology.org/technical_progress/buddy_coordination.cfm

Next, you will need to send an email to onegeology@bgs.ac.uk with the draft URL of the proposed WMS service. Include this information: Name of Geographic Area, Name of Data Provider, Name of Service Provider.

You will be contacted by the OneGeology secretariat with feedback, and confirmation of receipt.

9 Appendix A: Deploying GeoSciML-Portrayal Web Services in GeoServer

This appendix provides a tutorial of GeoSciML-portrayal web service deployment in a GeoServer environment.

9.1 Logging in

To deploy a web service in the GeoServer environment, your first step is to log in to the **Web Administration Interface** for the GeoServer instance you will be using to deploy your data as a web service (Figure 8). The Web Administration Interface is, as the name suggests, a browser-based user interface capable of administering server software remotely via your Internet connection.

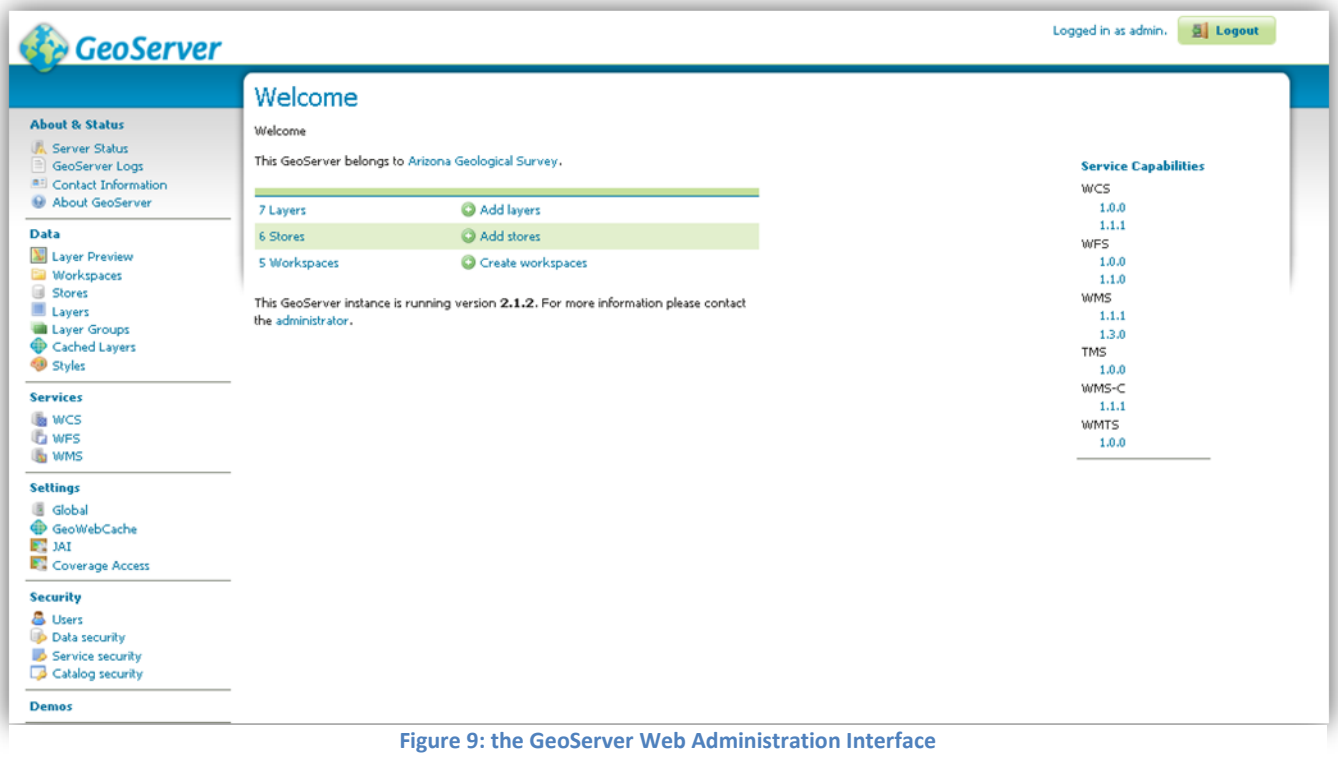


Figure 9: the GeoServer Web Administration Interface

To access the Web Administration Interface of a GeoServer instance, open your web browser and enter the appropriate web address into the navigation bar.

GeoServer instances located on your physical machine can be accessed by entering the following address into your web browser:

http://<**host**>:<**port**>/**geoserver/web/**, where <**host**> is usually **localhost** and <**port**> is usually **8080**

Unfortunately, remote users will be unable to access your GeoServer instance at this location via the World Wide Web. You will need to contact your network administrator to configure the host and port of your GeoServer instance in such a way that your GeoServer instance can be accessed remotely.

The default account settings for GeoServer are as follows:

Username: admin
Password: geoserver

For security reasons, it is recommended that you change your password to something more secure as soon as possible.

9.2 Service-Level Metadata

Within the GeoServer **Web Administration Interface** (Figure 8), click **Contact Information**, under **About & Status**.

This brings you to a **Contact Information** form (Figure 9) in which you can provide contact information for your GeoServer instance. The information entered here becomes service-level metadata for each web service you host from this GeoServer instance. Consequently, Contact Information entered here should be as precise and comprehensive as possible.

If you wish to deploy your service in compliance with OneGeology standards, your server-level metadata will need to meet very specific

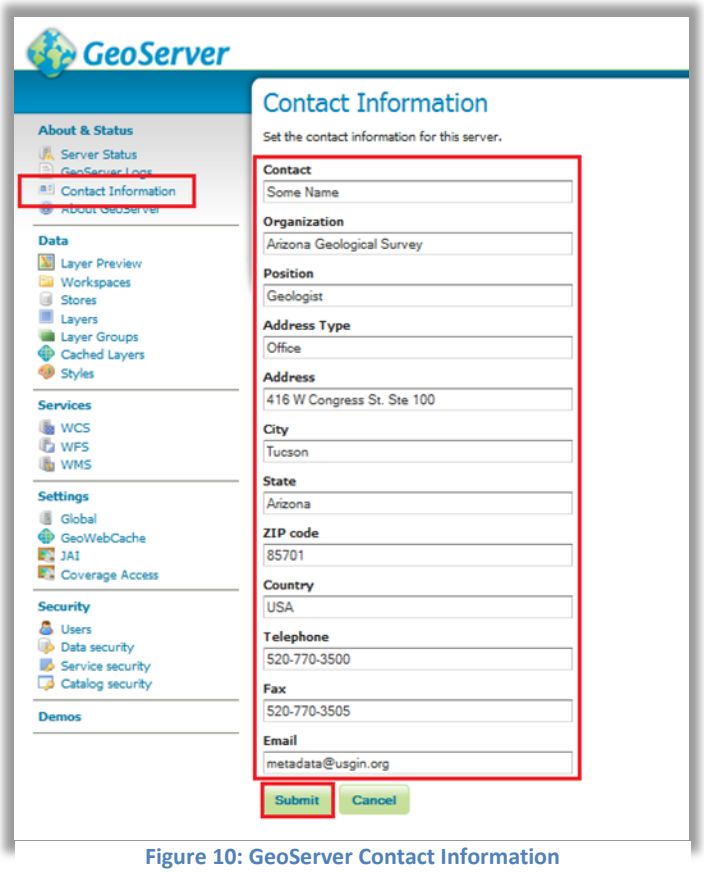


Figure 10: GeoServer Contact Information

requirements. See Appendix C.1 for details.

9.3 Creating a Workspace

After entering contact information for your GeoServer instance, you will need to create a workspace for your web service. The workspace is the equivalent of the **service name** when deploying a web service in the ArcGIS Server environment.

If you wish to deploy your service in compliance with OneGeology standards, your workspace will need to meet very specific requirements. See Appendix C.2 for details.

On the left side of the GeoServer **Web Administration Interface**, under **Data**, click **Workspaces**. This will bring you to the **Workspaces** page, wherein you can manage existing workspaces and create new workspaces.

Click **Add New Workspace**. This will bring you to the **Edit Workspace** page for your new workspace.

Two fields are present on the Edit Workspace page:

Name: The **service title**; may contain spaces or special characters; usually named according to the following convention:

GeographicArea_Theme, where GeographicArea is the overall region containing the map extent of the web service Theme is an indicator of the kinds of features

Namespace URI: A URI associated with your project

When you are finished, click **Save**.

9.4 Connecting to your PostGIS Database

Having created a new workspace, you will now create a **store**, or **database connection**, for your service.

On the left side of the GeoServer **Web Administration Interface**, under **Data**, click **Stores**. This will bring you to the **Stores** page. On the **Stores** page, click **Add New Store**.

This will bring you to the **New Data Source** page (Figure 10).

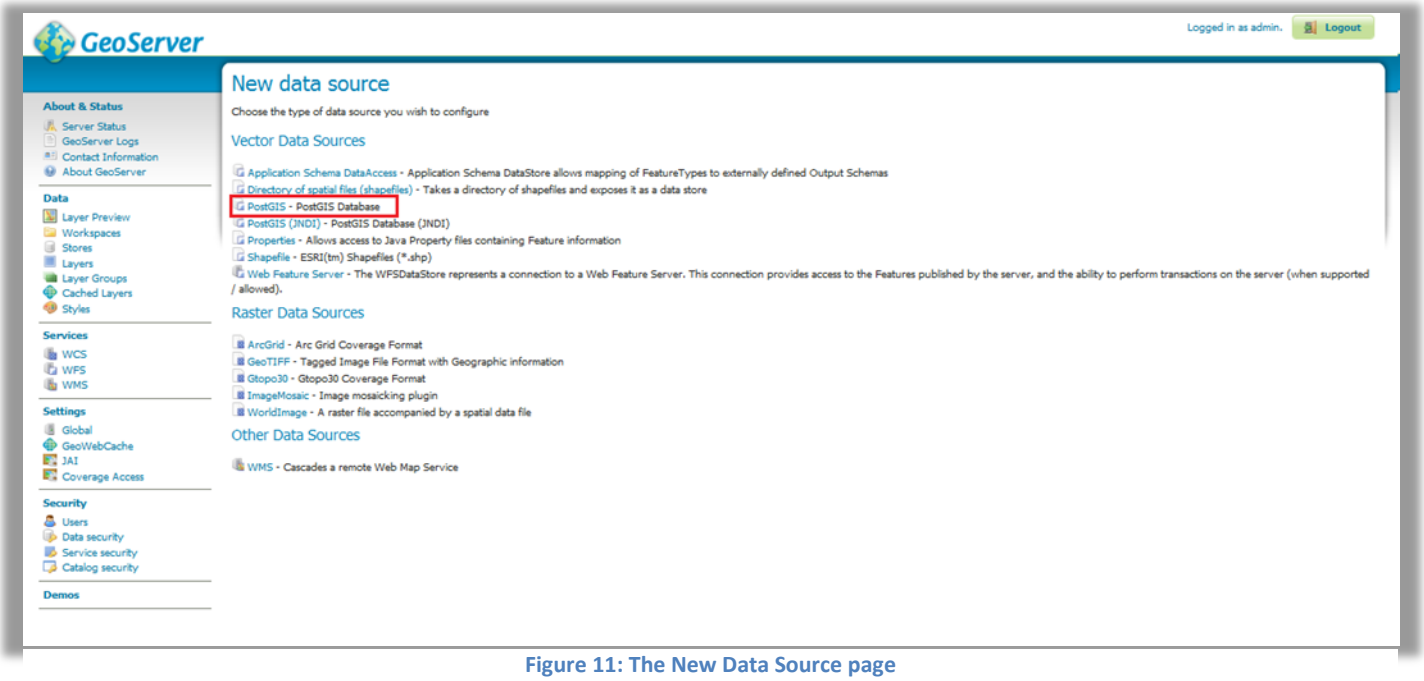


Figure 11: The New Data Source page

On the **New Data Source** page, choose **PostGIS** as the data source by clicking **PostGIS**. This will bring you to the **New Vector Data Source** page. Complete the following steps:

- Select a **Service Title** from the **Workspace** drop down menu. Select the workspace you created in Appendix A.3
- Type a name for your data store in the **Data Source Name field** (spaces are acceptable here); add a description if desired
- Make sure that the **Enabled** checkbox is checked
- Set the **Connection Parameters** for your PostGIS data source; if the PostGIS data source is located on a remote server, you will need to provide the appropriate host, port, database name, user name, and password to access it:
    - **Host**: use "localhost" if the PostGIS data source is on the same machine as your GeoServer instance; more specific host information will be necessary if your PostGIS data source is on a remote server
    - **Port**: default is 5432

- **Database name**: this information will depend on the PostGIS data source
- **Schema**: this information will depend on the PostGIS data source
- **User name**: this information will depend on the PostGIS data source
- **Password**: this information will depend on the PostGIS data source

When finished, click **Save**.

9.5 Adding Layers to a Workspace

Having created a workspace and specified a PostGIS data source for your web service, you will now populate your web service with data layers from your workspace.

On the left side of the GeoServer **Web Administration Interface**, under **Data**, click **Layers**. This will bring up the **Layers** page.

On the **Layers** page, click **Add a new resource**. This will take you to the **New Layer** page.

On the **New Layer** page, use the pulldown menu at the top of the page to select the data source you specified in Appendix A.4. Doing so will populate the **New Layer** page with a list of layers that may be published; click **Publish** to make the associated layer publicly accessible to anyone who connects to your web service.

Note: You may publish the same layer multiple times. To do so, click **Publish again.

After clicking **Publish**, the **Edit Layer** page for the corresponding layer automatically appears (Figure 11). The **Edit Layer** page contains two tabs, **Data** and **Publishing**.

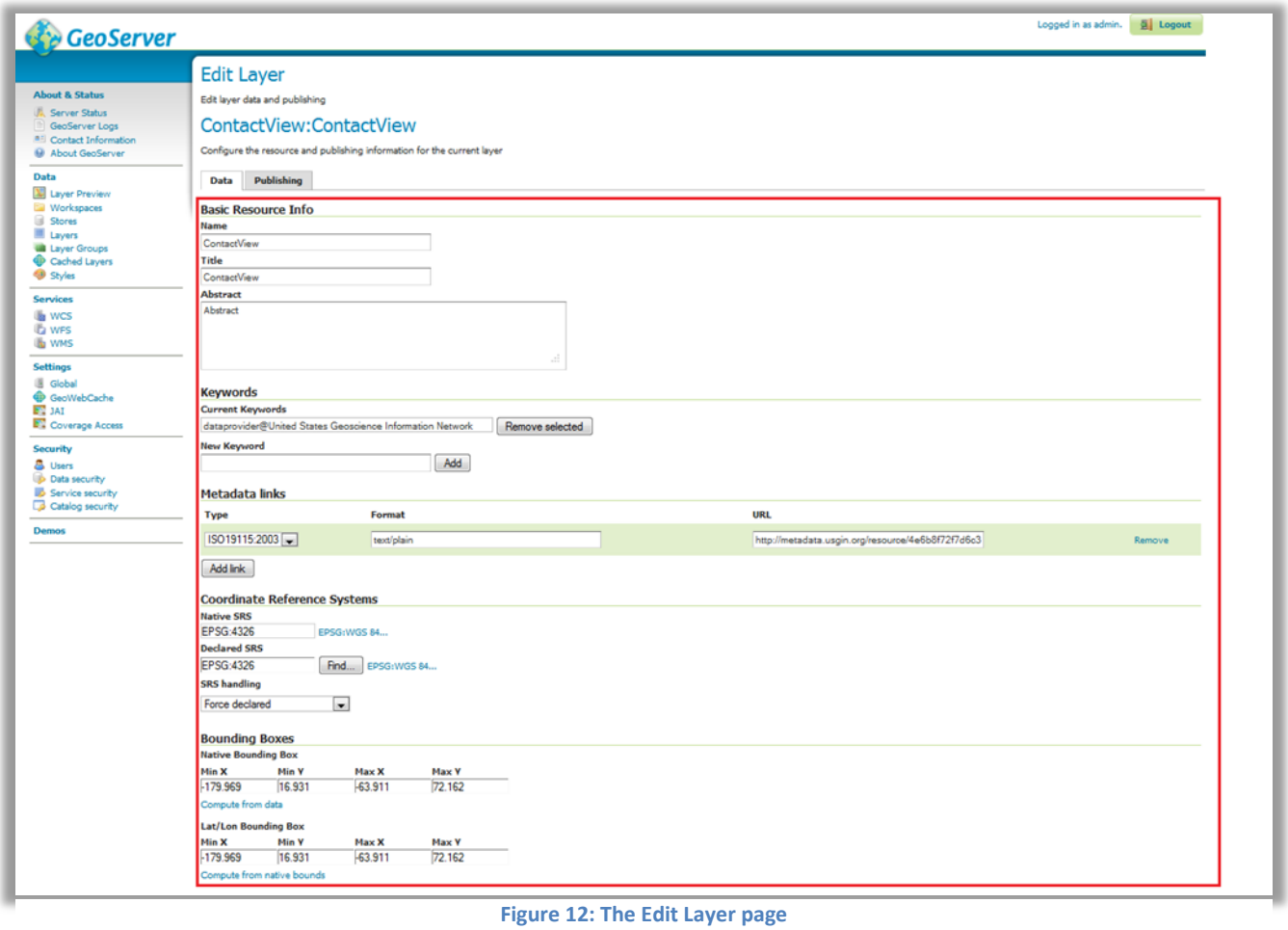


Figure 12: The Edit Layer page

9.5.1 The Data tab of the Edit Layer page

The **Data** tab contains fields within which you may specify the title, abstract, bounding box, spatial reference system, keywords, and metadata links for each layer in your web service. This information will be present within the **Capabilities document** produced by your web service in response to a **GetCapabilities** request, so it is very important to enter this information carefully for each layer in your web service.

If you wish to deploy your service in compliance with OneGeology standards, the information you enter for your layers will need to meet very specific requirements. See Appendix C.3 for details.

It is recommended to enter the bounding boxes for your service manually, as doing so permits you to provide a more useful bounding box for your web service.

**Note: GeoServer occasionally enters a loop of web errors. To fix this, we recommend restarting the service on Apache-Tomcat. It also helps to use Mozilla Firefox or Google Chrome for debugging (rather than Internet Explorer).

9.5.2 The Publishing tab of the Edit Layer page

After populating the fields in the **Data** tab, click the **Publishing** tab.

The **Publishing** tab contains **Layer Style** settings for the corresponding layer of your web service. **Layer Style** settings are dependent on the geometry of the layer (point, line, polygon). See Section 5, Styling, of this document for more details.

Generally, it is faster and more precise to import **Layer Style** settings from an existing style than it is to manually specify values for each field in the **Publishing** tab. To import **Layer Style** settings from an existing style, select the desired style from the **Available Styles** list.

To populate the **Available Styles** list, you might need to import layer styles from an SLD file. For further instructions, see Appendix A.5.3, Importing Layer Styles from an SLD File.

When you have populated the fields in the **Publishing** tab, click **Save**.

9.5.3 Importing Layer Styles from an SLD File

On the left side of the GeoServer **Web Administration Interface**, under **Data**, click **Styles**. This will open the **Styles** page.

If the Style you want is not listed on the **Styles** page, you will need to add it to the list by clicking **Add a New Style**. This brings you to the **Style Editor** page.

On the Style Editor page, you have the choice to copy/paste an SLD or upload a .sld XML document. The **Validate** button may be used to validate the SLD file against a schema before using it. Click **Submit** to add the SLD to the list on the **Styles** page.

9.6 Finishing Up

Repeat Appendix A.5 until all data you wish to publish has been added to the desired workspace. When finished, return to Section 7, Testing Your Web Service.

9.7 GeoServer Troubleshooting

Q: I made a change in the database on my server and now the service is not working.

A: Try clearing the cache and reloading GeoServer on the Server Status page (Figure 22). If that doesn't work, try hard restarting the service through Apache Tomcat.

Q: Can I use a replicated Feature Class to create a service on Geoserver?

A: No.

Q: I set up my services under three separate workspaces. When I connected to the WMS in ArcCatalog, all the layers appeared as one bundle. Is there a way to separate them out so I can add them individually?

A: Yes.

Though setting up your services under different workspaces seems to imply that they can be accessed as discrete services, Geoserver defaults to providing one capabilities document containing the information for all of the services set up on your instance of GeoServer. To access workspaces individually, you will need customize your Get request to specify the desired workspace.

For example: the Arizona Geological Survey runs three services on Geoserver (http://services.azgs.az.gov/geoserver/web/):

- GeologicUnitView
- FaultView
- ContactView

To perform a GetCapabilities request for GeologicUnitView, your GetCapabilities request will appear as follows:

http://services.azgs.az.gov/geoserver/GeologicUnitView/ows?service=wms&version=1.3.0&request=GetCapabilities

This URL opens the WMS **capabilities document** for the GeologicUnitView workspace only. A generic form of the service endpoint for the request is as follows:

http://[host server]/geoserver/[Workspace Name]/ows

Individual workspaces can be added in ArcGIS, as well.

Q: Is it possible to configure GeoServer so that I do not need to use PostGIS?

A: Try installing the ArcSDE plug-in for GeoServer. To do this, you will need to download the extension from GeoServer's website. Make sure to match the versions of the extension and GeoServer.

Q: All of my data are in Shapefiles. Can I deploy a shapefile as a GeoSciML-portrayal service?

A: The problem you will run into is the truncation of field names that occurs in shapefiles. Ideally you will have a full version of the data in PostGIS. As mentioned in the above document, to be compliant with GeoSciML-portrayal, you will need to make sure there is no truncation in field names; they must be an exact match for the GeoSciML-portrayal schema.

Q: I would like to deploy a service using full GeoSciML v.3.0 schema; where can I go?

A: Visit the GeoSciML website at http://www.geosciml.org/.

Q: Why would I want to create a WMS or WFS at all?

A: To display your data in a map format. Map format shows: data relative to its geographical surroundings, data relative to other data on the same map, data points relative to other data points in the same dataset (size, quantity, order, etc).

Q: How do I use the ArcCatalog Interoperability Extension?

A: Here is a brief tutorial on connecting to the WFS in ArcCatalog.

1. Open ArcCatalog
2. Enable the Interoperability Extension
    - For more information visit http://www.esri.com/software/arcgis/extensions/datainteroperability/common-questions.html
3. Click **Add Interoperability Connection**; this will bring up the **Interoperability Connection** dialog box:
    1. Click the '**…**'button next to **Format** and select the appropriate format
    2. Paste **the service endpoint** for the service you want to add into the appropriate field
    3. Click the **Parameters** button. In the window that appears, click the '**…**' button next to **Table List** and select the tables you wish to load. Confirm that the **Max Features** box is scaled appropriately
        - If you are loading a dataset with 10 million features, the Max Features box must be scaled to ten million. If you are loading a dataset with three (3) data points, the Max Features box can be scaled down quite a bit.
    4. Click **OK** in the Interoperability Connection dialogue box.
4. Depending on the size of the dataset, it may take some time to load. When it has loaded, expand the Connection (you can also rename the connection, which can be helpful).

5. Add your service to ArcMap

10 Appendix B: Deploying GeoSciML-Portrayal Web Services in ArcGIS Server

This appendix provides a tutorial of GeoSciML-portrayal web service deployment in a GeoServer environment.

10.1 Connecting to your ArcGIS Server instance

1. Open **ArcCatalog**
2. In the catalog tree, expand **GIS Servers** by clicking the nearby +
3. Double-click **Add ArcGIS Server**
4. Click the Manage GIS Services radio button in the **Add ArcGIS Server** window and then click **Next**.
5. Type in the URL for your services and the name of the server, then click **Finish**.

10.2 Create an ArcMap Project for your web service

1. Open an instance of ArcMap
2. Create a blank map
    a. Note: there is a one-to-one relationship between .mxd files and services: each .mxd file, along with all layers it contains, is equal to one web service
3. Add your source data to the map as a layer
    a. If you wish to deploy your service in compliance with OneGeology standards, your layers will need to meet very specific requirements. See Appendix C.3 for details
4. Save your map as a .mxd file

10.3 Publishing on ArcGIS Server

1. Open ArcCatalog.
2. Navigate to the database connection you set up in Appendix B.1, Connecting to your ArcGIS Server instance
3. In the catalog tree on the left side of the screen, expand the connection to ArcGIS Server
4. Right-click your server and select **Add New Service**; this opens a window containing fields for **Service Title** and **Service Description; if** you wish to deploy your service in compliance with OneGeology standards, your **Service Title** will need to meet very specific requirements. See Appendix C.2 for details
    a. Specify the appropriate **Service Title** based on content and map extent
    b. Using the pulldown menu, specify whether you are deploying a Web Map Service or a Web Feature Service
    c. Enter an appropriate description for your service in the **Description** field
    d. Click **Next**
5. Specify the capabilities of your web service
    a. Use the checkboxes to indicate which capabilities your web service will support
        i. OneGeology requires either WMS or WFS capabilities
            1. Specifying WMS or WFS will reveal a form in the **Properties** box; if you wish to deploy your service in compliance with OneGeology standards, this information must be provided in accordance with very specific requirements. See Appendix C for more details.
            2. Alternatively, if you know of a web service capabilities document that is already compliant with OneGeology standards, you may save this capabilities document as an XML file and import it
                a. To import an external XML document, click **Use external capabilities files** and then specify the location of the XML document you will use
                b. To save a capabilities document as an XML document, access the appropriate capabilities document (described in detail in Section 7) in your web browser and click **File > Save Page As...**
    b. Click the **Use layer names from the map document** checkbox
    c. Click **Next**
    d. Click **Next** again
    e. Click **Finish** to start your service
        i. Note: you will need to log in to your service and **Stop** your service in order to edit the capabilities document. Service title cannot be edited after publishing; you will need to create a new service

Proceed to Section 7, Testing Your Web Service.

10.4 Troubleshooting ArcGIS Server

Q: How do I connect to the WFS in ArcCatalog?

A: Here is a brief tutorial on connecting to the WFS in ArcCatalog.

1. Open **ArcCatalog**
2. Enable the **Interoperability Extension** in ArcCatalog.
    a. For more information visit http://www.esri.com/software/arcgis/extensions/datainteroperability/common-questions.html
3. Click **Add Interoperability Connection**
4. A Dialog box appears
    a. Click the '…' button next to the **Format** pulldown menu
        i. Select the appropriate type of web service from the dialog
        ii. Click **OK**
    b. In the **Dataset** field, paste the URL of the capabilities document for the service you wish to add
        i. For more information, see Section 7, Testing Your Web Service
    c. Click the **Parameters** button.
        i. Click the '…' button next to **Table List**
        ii. Select the tables you wish to load
        iii. Confirm that the **Max Features** box is scaled appropriately
            1. If you are loading a dataset with 10 million features, the "Max Features" box must be scaled appropriately to ten million
            2. If you are loading a dataset with three (3) data points, the "Max Features" can be scaled down quite a bit
        iv. Click **OK**
    d. Once the layers have been added under the interoperability connections, you will see something like this in ArcCatalog
    e. Click **OK**
5. Depending on the size of the dataset, it may take some time to load. When it is finished, expand the Connection (you can also rename the connection, which can be helpful).
6. Drag and drop the layer into your ArcMap Project.

Q: All of my data are in Shapefiles. Can I deploy a shapefile as a GeoSciML-portrayal service?

A: The problem you will run into is the truncation of field names that occurs in shapefiles. Ideally you will have a full version of the data in an sde databse (or file/personal geodatabase). As mentioned in the above document, to be compliant with GeoSciML-portrayal, you will need to make sure there is no truncation in field names; they must be an exact match for the GeoSciML-portrayal schema.

Q: I would like to deploy a service using full GeoSciML v.3.0 schema; where can I go?

A: Visit the GeoSciML website

Q: Why would I want to create a WMS or WFS service at all?

A: To display data in a map format. Map format shows: data relative to its geographical surroundings, data relative to other data on the same map, data points relative to other data points in the same dataset (size, quantity, order, etc).

11 Appendix C: OneGeology Web Service Requirements

If you wish to submit your web service to the OneGeology portal, your web service must comply with OneGeology standards.

This appendix summarizes the service-level and layer-level requirements that must be respected for compliance with OneGeology standards.

11.1 Service-Level Metadata Requirements

The following section describes OneGeology standards as they relate to service-level metadata.

When you submit your web service to OneGeology, service-level metadata is harvested into the portal in the GetCapabilities response document. The minimum requirements for inclusion with OneGeology (source: http://www.onegeology.org/wmscookbook/2_4_1.html)

Table 4: One-Geology Service-Level Metadata Requirements

Field	Description	Conditions
Version (1.3.0, 1.1.1)	This will be automatically populated by GeoServer	N/A
Title	Required for 2-star service (2-stars); see Section 7.2 for more information	Required for 2 Stars
Abstract	May include data owner organization, goals, scale of data and other information	Required for 2 Stars
Access Constraints	Definition of access constraints, otherwise "none" (2 to 3 Stars)	Required for 2-3 Stars
Keywords	"OneGeology", "MD_DATE@2012-02-02", "MD_LANG@ENG" (3 Stars)	Required for 3 Stars
Data (owner) provider	Full name of data owner organization (not service provider) (2 Stars)	Required for 2 Stars
Image Format	This will be automatically populated by GeoServer	N/A
On-line Resource	Required by the WMS specification (2 Stars)	Required for 2 Stars
Contact Person	Named Individual or contact within data organization (2 to 3 Stars)	Required for 2-3 Stars
Country	Country of the data owner organization (2 to 3 Stars)	Required for 2-3 Stars
Email Address	Email address of the named contact (2 to 3 Stars)	Required for 2-3 Stars
Fees	Any fees required to use the WMS Service; otherwise "none" (3 Stars)	Required for 3 Stars

11.2 Service Title and Service Name Requirements

This section describes the OneGeology standards that govern web service **Title** and **Name**.

Service Title: Conforming to the naming conventions for service title is a 2 star requirement with OneGeology. **Service Titles** should contain the following tokens (in order):

[Geographical extent] [Data owner organization] [Language Code] [Scale] [Theme], where:

- **Geographic extent:** a rough approximation of the extent of the portrayal provided by the web service
  - Generally an ISO three letter country code (such as "USA") and, if necessary, a province, such as "USA-KY" for the state of Kentucky
  - See http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3 for ISO three-letter country codes
- **Data owner organization code**: an acronym or abbreviation for the entity that owns the data portrayed by the web service
  - The **data owner organization** *can be*, but is *not necessarily*, the name of the organization responsible for deploying the data as a web service
- **Language**: a two-letter ISO 639-1 code for the language in which the web service is portrayed
  - See http://en.wikipedia.org/wiki/List_of_ISO_639-1_codes for ISO language codes
- **Scale:** expresses the scale of the portrayal as a ratio (X:Y); large-scale portrayals should be abbreviated using SI symbols for million "M" and thousand "k"
  - An example of a valid scale in the service title is "1:1M"
- **Theme**: a short or abbreviated description of the data contained in the layer

Service Name: For version 1.3.0 services, the service name should be 'WMS'; for version 1.1.1, the service name should be 'OGC:WMS'

11.2.1 OneGeology Service Title Examples:

USA USGIN 1:3M Geology
USA-AZ AZGS ES 1:2k Faults

***Note: Language is omitted from the examples above, as they are in the default language English.

11.3 Layer Naming and Title Conventions

Each layer in a OneGeology-compliant web service should have a **Layer Title** and a **Layer Name**. The **Layer Name** should be identical to the **Layer Title**, using underscores instead of spaces and an abbreviated **Scale** token (see below)

Both **Layer Titles** and **Layer Names** should contain the following tokens (in order):

[Geographical Extent] [Data Owner Organization][Language Code][Scale][Theme], where

- **Geographic extent:** a rough approximation of the extent of the portrayal provided by the web service
  - Generally an ISO three letter country code (such as "USA") and, if necessary, a province, such as "USA-KY" for the state of Kentucky
  - See http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3 for ISO three-letter country codes
- **Data owner organization code**: an acronym or abbreviation for the entity that owns the data portrayed by the web service
  - The **data owner organization** *can be*, but is *not necessarily*, the name of the organization responsible for deploying the data as a web service
- **Language**: a two-letter ISO 639-1 code for the language in which the web service is portrayed
  - See http://en.wikipedia.org/wiki/List_of_ISO_639-1_codes for ISO language codes
- **Scale:** expresses the scale of the portrayal as a ratio (X:Y); large-scale portrayals should be abbreviated using SI symbols for million "M" and thousand "k"
  - An example of a valid scale in the service title is "1:1M"
  - **Layer Names** should abbreviate the **Scale** token by removing the **antecedent** and the **colon**. For example, a 1:23k scale would be abbreviated as 23k
- **Theme**: a short or abbreviated description of the data contained in the layer

11.3.1 Layer Title Examples

GBR BGS 1:625k Bedrock Age
USA USGIN 1:3M Geologic Age

11.3.2 Layer Name Examples

World_25M_Geol_Units
US-KY_KGS_24k_Faults
USA_USGIN_3M_Lithostratigraphy

11.3.3 A sample WMS GetCapabilities document demonstrating Layer Names and Layer Titles

Code Example 3: OneGeology-compliant Layer Name and Layer Title in a WMS GetCapabilities document

```
<Layer queryable="1" >
        <Name>USA_USGIN_3M-Geologic_Age</Name>
        <Title><![CDATA[USA USGIN 1:3M Geologic Age]]></Title>
        <Abstract><![CDATA[This layer provides a portrayal in which units are categorized according to the representative geologic age of the
        unit.]]></Abstract>
        <KeywordList>
                <Keyword>continent@North America</Keyword>
                <Keyword>geographicarea@USA</Keyword>
                <Keyword>dataprovider@USGIN</Keyword>
```

11.4 Additional Layer-level Metadata Requirements

Aside from **Layer Title** and **Layer Name** (discussed in Section 7.3), OneGeology standards require additional layer-level metadata, including the following:

Abstract: A description of your layer data; should be as comprehensive and interoperable as possible. You may wish to include other metadata in the abstract, such as information about your organization and other data you make available. You may also wish to include a statement on access constraints (source: http://www.onegeology.org/wmscookbook/2_6.html).

Keywords: Tags used to locate and categorize your layer. Required and conditional keywords are described in Table 5 (source: http://www.onegeology.org/wmscookbook/2_6.html):

Table 5: OneGeology Keywords

Keyword Description	Appearance in GetCapabilities Document	Example	Cardinality
Continent	continent@value	continent@North America	Required
Subcontinent	subcontinent@value	subcontinent@India	Conditional
Geographic area (country)	geographicarea@value	geographicarea@USA	Required
State (region or province)	subarea@value	subarea@Arizona	Conditional
Data Provider	dataprovider@value	dataprovider@USGIN	Required
Service Provider	serviceprovider@value	serviceprovider@AZGS	Required

Bounding Box: A rough indicator of the extent of your data; can be computed from data in the GeoServer 'Edit Layer' interface (Figure 17), or customized.

SRS: The spatial reference system used by your web service, indicated by an EPSG code.

Spatial reference systems specify a datum, a reference ellipsoid, and a coordinate system for cartographic representations. Most known spatial reference systems are associated with an EPSG code number; EPSG stands for European Petroleum Survey Group.

The default SRS is WGS 1984 (EPSG:4326). For the purposes of interoperability, your web service must be deployed using EPSG:4326.

MetadataURL: A link to an external ISO-valid metadata document. A **MetadataURL** is required for services deployed on OneGeology. To create an ISO Metadata record using ISO19115:2003, use the USGIN Metadata Wizard.

Readable Legend Graphic: URL(s) for the legend graphic. Readable Legend Graphic URLs should be present in the getCapabilities document by default when an SLD is assigned to a layer. For more information on SLDs, see Section 5, Styling.

12 Glossary

Attribute (GIS)

Within the context of geographic information systems, an attribute describes a feature. For example: attributes of a fault feature might include the latitude and longitude coordinates for each fault, as well as the fault's age, dip, and slip.

Database

A method of storing data. In a database, data is divided up into database records; in turn, database records are divided up into database fields. The advantage of a database is that it can be sorted and searched by field contents.

Though modern databases are usually digital, a physical example of a database is a card catalog in a public library. In a card catalog, data is divided up into individual cards, which are directly analogous to database records. Each card (record) in the catalog corresponds with and describes a book. The information about each book is divided up into fields: title; author; subject; publication date; etc. Digital databases can be in tabular format (that is, a table) in which rows represent individual records and columns constitute fields; or they can be viewed record-by-record.

Database Field

A subdivision of a database record in which a specific type of data is entered. Using the analogy of a card catalog in a public library: if the card catalog is directly analogous to a database, and if cards in the catalog are directly analogous to database records, then the different subdivisions of information found in each card in the catalog (title, author, publishing date, etc.) are all database fields.

Database fields are functionally similar to markup language elements.

Database Record

A subdivision of a database. Using the analogy of a card catalog in a public library, each record in a database is analogous to an index card in the card catalog; each card (record) corresponds with and describes an individual book. Database records are further subdivided into fields, which contain specific kinds of data (title, author, etc).

Element

Elements are logical document components found in markup languages such as XML and HTML. Elements simultaneously define the structure and content of a document; they are constituted by two markup language tags and the content between the two. For example:
<tag>content</tag>

The two tags and their content form an element. In a more concrete example, HTML uses the tag to demarcate text that should be Italicized. So, an Italicized element of an HTML document would appear as follows:

Italicized content

In a web browser, this element would produce the following result:

Italicized content

For a more detailed overview of elements, see the USGIN XML Tutorial.

Markup language elements are functionally similar to database fields.

Feature

The word *feature* may indicate a *geologic* feature, such as a fault, formation, or dike. Alternatively, a feature can be a GIS representation of a real-world resource. GIS features do not always correspond with geologic features because GIS software can be used to represent anthropogenic objects such as buildings, roads, or canals. The definition of the term feature therefore depends on its context: it can be used to mean either a geologic feature or a GIS feature.

GIS features are often described by attributes.

Interchange Format

An interchange format is an application-neutral format used to exchange data between systems. For example, if one needed to transfer data between an Oracle database and an Access database, an interchange format could be used to transfer this data from one database to the other.

Interoperability

The capability to communicate, execute programs, or transfer data in a manner that requires the user to have little or no specialized knowledge. Interoperability is based on a shared data schema and the use of standard vocabulary terms.

Markup Language

Markup languages, such as HTML and XML, use elements to annotate text in such a way that the structure of the document is distinguishable from the contents of the document. Consequently, markup language documents are very good for storing data, since the distinction between structure and content is transparent and immediately discernible.

Metadata

Literally, data about data. Metadata is used to organize, categorize, browse, and discover data resources. For more information, see the USGIN Metadata Tutorial.

Schema

A schema is the structural framework for a data model or collection of attributes that describe something. For practical purposes, schemas *structure* documents by dictating where and how data from a dataset should be entered into a document.

Uniform Resource Identifier (URI)

Uniquely identifies a resource. See the USGIN URI Tutorial for more information about URIs.

Web Service

Under the client-server relationship, a web service is a server-side protocol that specifies syntax used by client software applications to interact with web-accessible resources on the server. As long as a request for web-accessible resources from the server uses the appropriate syntax, web services can provide a response. Web service syntax defines required input parameters, operation output, and the results of any operations performed.

Web services decouple the client and the server, allowing the two systems to develop independently: as long as the web service is in place and specified syntax is used, changes may be made to either server or client without jeopardizing the server's ability to make web-accessible data available upon request or the client's ability to request web-accessible resources from the server. Consequently, web services can field requests from multiple different client platforms, including mobile devices.

Etymological Note: "hosting data as a service" is considered synonymous with "hosting data in accordance with a web service protocol." Likewise, data hosted in accordance with a web service protocol is often referred to as a web service.

Web services are well suited to the task of sharing digital geoscience information in accordance with **USGIN** specifications (for a number of reasons: owing to standardized data requests, they can be accessed from a variety of platforms; they are read-only and preserve data ownership; and they come in a number of different flavors for different purposes.

The Open Geospatial Consortium has provided specifications for several different flavors of web service that are relevant to geographic information systems, **USGIN**, and the State Geothermal Data project. These include, but are not limited to, **web map services (WMS)** and **web feature services (WFS).**

Web Map Services (WMS) provide a georeferenced map image within a geographic bounding box. Existing georeferenced map images are acquired by **getMap** requests sent by client software to the web map service; **getMap** requests specify the geographic area of the bounding box and the map that is desired. A variety of other parameters allow control over image size, map projection, and other details. The server responds by returning an image file (typically .tif, .jpg or .png) for the requested area.

For example, a **getMap** request could be used to request a georeferenced satellite photo of Arizona; if the satellite photo exists on the server as a web-accessible image file, the web service will respond to the **getMap** request by providing the client with the desired image.

Web map service image files are often based on a source shapefile or feature class. Map services support **FeatureInfo** requests at point locations; **FeatureInfo** request returns are not standardized and depend on server configuration. Map services are most useful for client applications designed to visually explore geographic relationships.

Web Feature Services (WFS) provide collections of georeferenced features described by attributes. Attributes are essentially a collection of database records that describe the location and characteristics of bound features.

For example a feature service containing data about river systems might provide features with linear geometry representing river segments; each feature might be described by attributes such as average flow rate, width, and depth for the segment.

Feature services are useful for geospatial analysis performed using data describing individual geographically located entities. For example, a client application used to calculate the most efficient route between points in a city utilizing real-time traffic information would require a representation of the streets as features described by attributes.

Owing to the large amount of data associated with WFS, these web services tend to require fast Internet connections.

XML

XML stands for Extensible Markup Language. XML documents are often used as interchange formats for database records, owing to the ability for XML elements to emulate database fields. To use XML documents as an interchange format, USGIN defines XML schemas in which each XML document corresponds with a specific database record and each element corresponds with data entered in a database field. In these documents, tags define the database fields.

For example: well_name is a field within the WellHeader content model. Any known oil and gas well names go in the well_name field. An XML schema for expressing database records within the WellHeader content model might appear as follows:

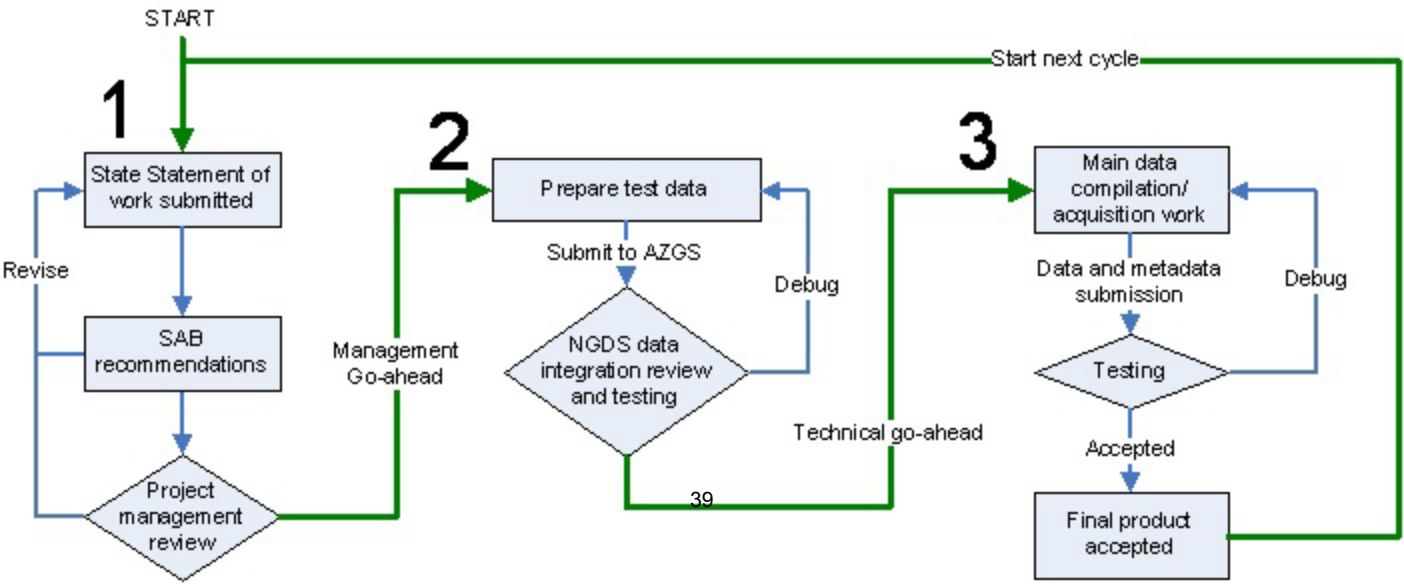
<content_model_name>WellHeader</content_model_name>

<database_record>

<well_name>Union Oil Company of California-South 7 Veysey</well_name>

In this example, the Union Oil Company of California-South 7 Veysey well name constitutes the content delineated by the <well_name> tag.

For a more detailed overview of XML, see the USGIN XML Tutorial.


START
Start next cycle
1
State Statement of work submitted
Revise
SAB recommendations
Project management review
Management Go-ahead
2
Prepare test data
Submit to AZGS
Debug
NGDS data integration review and testing
Technical go-ahead
3
Main data compilation/ acquisition work
Data and metadata submission
Debug
Testing
Accepted
Final product accepted

USGIN
U.S. Geoscience Information Network

USGIN Metadata Profile: Use of ISO metadata specifications to describe geoscience information resources

Version 1.2draft

Title:
USGIN Metadata Profile: Use of ISO metadata specifications to describe geoscience information resources

Latest released version:
http://lab.usgin.org/profiles/usgin-iso-19139-profile

Creator:
USGIN Standards and Protocols Drafting Team

Editors:
Stephen M. Richard and Wolfgang Grunberg

Creation date:
8/18/2009

Last revision date:
3/2/2014 12:11 PM

Document Status:
Live trunk version, post 1.2 edits

Publisher:
Arizona Geological Survey

Description:

This document describes recommended practices for using ISO19139 xml encoding of ISO 19115 and ISO 19119 metadata to describe a broad spectrum of geoscience resources. The document provides guidance for the population of ISO19139 encoded metadata documents to enable interoperability of catalog service clients with multiple servers conforming to this profile.

Contributor:

See acknowledgements

Document Identifier:

gin2010-009.1.2.x http://repository.usgin.org/uri_gin/usgin/dlio/337

Notices

Neither the USGIN project, nor any of the participating agencies, take any position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither do they represent that there has been any effort to identify any such rights.

This document and the information contained herein is provided on an "AS IS" basis and USGIN DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Revision History

Version	Date	Comments	By
0.1	2009-08-18	Initial draft document	Wolfgang Grunberg
0.7	2009-10-28	Revisions, addition of material, re-title, focus on use of ISO 19139. Revising adding material done, circulate for review, comment. No discussion of serviceIdentification section as yet; this will have to be added, at which point the file title will be revised.	Stephen Richard
0.7.1	2009-11-19	Revisions, crosschecked with ISO 19139 and NAP Profile, added dataset metadata example	Wolfgang Grunberg
0.8	2009-11-24	SMR review, editing based on W. Grunberg comments	Stephen Richard
0.8.1	2009-11-30	Codelist review, approved previous tracked changes	Wolfgang Grunberg
0.9	2009-12-03	Add section on distribution and distributor, review and accept changes	Stephen Richard
0.9.1	2009-12-04	First public release of draft dataset and dataset series profile, Word and PDF file preparation	Wolfgang Grunberg
0.9.3	2009-12-11	Review edits and comments. Split out data identification element into separate tables for MD_DataIdentification and SV_ServiceIdentification; add service metadata elements.	Stephen Richard
1.0.0	2010-01-11	Review dataset and service metadata sections; update metadata examples. Second public release of draft dataset and dataset series profile, Word and PDF file preparation	Wolfgang Grunberg
1.0.1	2010-01-20	Additional examples, codeList overhaul	Wolfgang Grunberg
1.1.0	2010-02-02	Change recommendation for codelist usage to use ISO identifiers; update codelist discussion. Update distribution format discussion. Review and formatting. Public release, Word and PDF file preparation	Stephen Richard, Wolfgang Grunberg
1.1.1	2010-03-04	Review, title page changes	Stephen Richard
1.1.2	2010-03-04	AZGS Contributed Reports release, Word and PDF file preparation	Wolfgang Grunberg
1.1.3	2010-11-16	Minor edits and bug fixes, modify title and title page text.	Wolfgang Grunberg, Stephen Richard
1.1.4	2010-11-18	Correct example document for distribution element conventions, add simple data quality statement, funding acknowledgement, update use case section	Stephen Richard
	2011-12-28	add ‘Recommended practices’ to title. Update language in distribution section.	Stephen Richard
	2013-06-13	Finish major review and edit of text to bring in alignment with practice that has evolved since 2009. Remove 'Recommended practices' in title, add 'USGIN Metadata Profile'. Remove 'service metadata example, it was too deeply flawed to fix. May add a service metadata example later.	Stephen Richard
	2014-01-23	Add hierarchyLevelName string values for use in metadata document in Table 1	Stephen Richard

Acknowledgement

Many individuals and organizations have contributed to or inspired the development of these Guidelines.

The USGIN Standards and Protocol Drafting Team include (alphabetically):

Ryan Clark - AZGS
Wolfgang Grunberg – AZGS
Dominic Owen - AZGS
Stephen M. Richard – AZGS

Funding Provided by (chronologically):

National Science Foundation under EAR-0753154 to the Arizona Geological Survey acting on behalf of the Association of American State Geologists; 2009.

U.S. Department of Energy under award DE-EE0001120 to Boise State University; 2009.

U.S. Department of Energy under award DE-EE1002850 to the Arizona Geological Survey acting on behalf of the Association of American State Geologists, 2010.

Special Thanks to (alphabetically):

ANZLIC
INSPIRE
MEDIN
NAP
NOAA

Contact Information

Arizona Geological Survey
416 W. Congress St., Suite 100.
Tucson, Arizona 85701-1381
Phone: 520.770.3500
Fax: 520.770.3505

Email

metadata@usgin.org

Online

http://usgin.org
http://lab.usgin.org

Table of Contents

1 Introduction 8
1.1 Normative References 8
1.2 Purpose 9
1.3 Terminology 9
1.4 ISO Schemas Location 10
2 Overview of the Profile 12
2.1 Quick Reference 12
2.1.1 Non-service resources 14
2.1.2 Service Resources 14
2.2 USGIN specification constraints and recommendations 15
2.3 USGIN specification extensions 15
2.4 General Objectives 15
2.5 Requirements 16
2.6 Use cases to be supported 16
2.7 Resources of interest 17
3 USGIN Usage of Metadata Elements 23
3.1 Core spatial dataset, dataset series, and service elements 23
3.2 Dataset Identification properties (MD_DataIdentification) 40
3.3 Service identification elements (SV_ServiceIdentification) 50
4 Usage notes for Metadata Elements 63
4.1 Metadata file identifier 63
4.2 Metadata hierarchy 63
4.3 Metadata Contact vs. Resource Citation vs. Resource Contact 63
4.4 Resource Title 64
4.5 Resource Abstract 64
4.6 Resource Type 64
4.7 Resource Locator 65
4.8 Unique Resource Identifier 65
4.9 Browse Graphics 65
4.10 Resolution and equivalentScale 66
4.11 Resource Language 66
4.12 Encoding of Vertical Extents 66
4.13 Content information: Entities and Attributes 66
4.14 Use of MD_Distribution and MD_Distributor 67
4.15 Distribution Format 69
4.15.1 Digital resources 69
4.15.2 Non digital resources 70
4.16 CI_OnlineResource 72
4.16.1 CI_OnlineResource protocol 73
4.16.2 CI_OnlineResource applicationProfile 73
4.16.3 CI_OnlineResource name and description 73
4.16.4 CI_OnlineResource function property 74
4.16.5 Scoped name in service distributions 75
4.17 Responsible parties and logos 75
4.18 Extensions to CharacterString 76
4.18.1 Web extensions 76
4.18.2 Language localization 76
4.18.3 Codelists 77

4.19 Geographic bounding box 84
4.20 Data Quality 84
4.20.1 Simple quality statement 84
4.21 Lineage 84
4.22 Temporal extents 86
4.23 Operation metadata 87
5 Abbreviations 88
6 References 89
6.1 Cited literature 89
7 Codelists 90
7.1 ServiceType 90
7.2 Linkage name conventions 92
8 Examples 94
8.1 USGIN ISO 19139 Minimum Dataset Metadata 94
8.2 USGIN ISO 19139 Dataset Metadata 102
8.3 ISO19110 Feature Catalog example 125

Tables and Figures

Table 1. Summary of resource types described by metadata for US Geoscience Information Network catalogs.. 14
Table 2. Description best practices for ISO19139 metadata elements in USGIN profile.. 18
Table 3. Dataset Identification properties (MD_DataIdentification) 34
Table 4. Service Identification properties (SV_ServiceIdentification) 43
Table 5. Example format strings for digital files.. 61
Table 6. USGIN Distribution formats for non digital resources. 62
Table 7. OnlineFunctionCode values from NAP (INCITS 453) and ESRI Geoportal toolkit v. 3.1.. 63
Table 8. Codelist crosswalk between ISO, NAP and USGIN. 69
Table 9. Usage of data quality scope description elements 75
Table 10. Inspire spatial data service type 82
Table 11. USGIN service type vocabulary.. 82
Table 12. USGIN Names to identify special linkage URL's for CI_Online Resource. 83
Figure 1. gmd:MD_Distribution_Type diagram 60

1 Introduction

A key component of a distributed information network is a catalog system, a collection of resources that allow data and service providers to register resources, and data consumers to locate and use those resources. Currently, many online catalogs are web pages with collections of URLs for services, or services are discovered accidently or by word of mouth. The vision is to enable a web client (portal) to search across one or more metadata registries without having to configure the client individually for each of the registries that will be searched. Thus, metadata providers can focus on data development, without having to also develop web clients to enable search of that metadata.

The Open Geospatial Consortium (OGC) Catalog Service for the Web (CSW) specification defines a collection of basic operations for searching catalogs of metadata via the web. Engineering the desired interoperability requires adding additional constraints on CSW operation; one of the major constraints is selection of the xml schema that will be used to encode metadata for the service. The core CSW specification requires use of a basic xml schema that includes content defined by the Dublin Core Metadata specification [Dublin Core, 2008-01-14]. This document concerns use of the ISO19115 and ISO19119 content models implemented using the ISO19139 xml schema for encoding of metadata content. Some more specific constraints on use of this implementation may be included in a separate document (USGIN, 2012) describing metadata constraints for different kinds of resources.

1.1 Normative References

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 19115 designates these two normative references:

- ISO 19115:2005, *Geographic information - Metadata*
- ISO 19115/Cor.1:2006, *Geographic information – Metadata, Technical Corrigendum*

ISO 19119 designates these normative references:

- ISO 19119:2005, *Geographic information - Services*
- ISO 19119:2005/Amd 1:2008, *Extensions of the service metadata model ISO 19108 designates*:
- ISO 19108:2005, *Geographic information – Temporal Schema*

ISO 639-2, Codes for the representation of names of languages - Part 2: Alpha-3 code control ISO 8601, Data elements and interchange formats - Information interchange - Representation of dates and times

ISO/TS 19139:2007, Geographic information - Metadata – XML Schema Implementation

OGC 07-006r1¸ OpenGIS Catalog Services Specification version 2.0.2, Corrigendum 2 release, 2007

OGC 07-045, OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile, Version 1.0.0, 2007

INCITS 453-2009, North American Profile of ISO 19115:2003 – Geographic Information – Metadata (NAP-Metadata), 2009, American National Standards Institute, Inc.

ISO 10646-1, Information technology — Universal Multiple-Octet Coded Character Set (UCS) — Part 1: Architecture and Basic Multilingual Plane

RFC 2119, Key words for use in RFCs to Indicate Requirement Levels, Network Working Group, 1997.

1.2 Purpose

The USGIN uses the ISO 19115/19119 specifications for metadata content, and the ISO 19139 XML schema for encoding this content search results provided by USGIN CSW services. This profile conforms to most of the provisions of the North American Profile of ISO metadata (INCITS 453-2009, referred to as NAP), except it allows multiple distributor-format-transferOptions bindings for resource distribution, and recommends use of ISO19115 codelist values. This USGIN document is meant to provide guidance on the use of the ISO19139 XML schema to encode metadata for geoscience resources, with sufficient detail that application developers can produce software clients that utilize the metadata for automated discovery, evaluation, and utilization workflows. The focus of the profile is to enable interoperable catalog services for discovery, evaluation, and access to information resource of interest to geoscientists.

1.3 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in Internet RFC 2119.

Application profile: a schema that consists of data elements drawn from one or more namespaces, combined together by implementers, and optimized for a particular local application. (Rachel Heery and Manjula Patel, 2000, http://www.ariadne.ac.uk/issue25/app-profiles/)

Catalog application: Software that implements a searchable metadata registry. The application must support the ability to register information resources, to search the registered metadata, to support the discovery and binding to registered information resources within an information community.

Codelist (also as Code list): a controlled vocabulary that is used to populate values for an XML element.

Data product specification: a definition of the data schema and value domains for a dataset. The data schema specifies entities (features), properties associated with each entity, the data type used to specify property values, cardinality for property values, and if applicable, other logical constraints that determine data validity. Value domains are specified for simple data types—strings or numbers, and may include controlled vocabularies for terminology required to specify some properties.

Dataset series: collection of datasets sharing the same product specification (ISO 19115). ISO 19115 does not define product specification. For the purposes of USGIN, a product specification defines a data schema, any required controlled vocabularies, and recommended practices for use of schema (see Data product specification).

Dataset: an identifiable collection of data (ISO19115). USGIN extends the concept of data items to include physical artifacts like books, printed maps and diagrams, photographs, and material samples--any identifiable resource of interest. DCMI definition is "Data encoded in a defined structure" with additional comment "Examples include lists, tables, and databases. A dataset may be useful for direct machine processing." A digital dataset is a collection of data items in which individual data items are identified and accessible. Metadata for the collection is a different type than metadata for individual items in the collection (dataset vs. features). Criteria for what unifies the collection are variable (topic, area, author...). Data items may represent intellectual content -- information content and organization (data schema) -- or may represent particular manifestations (formats) of an intellectual artifact.

Interoperability: "The capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units." ISO/IEC 2382-01 (SC36 Secretariat, 2003)

Metadata element: a discrete unit of metadata (ISO 19115), an attribute of a metadata entity. A metadata element contains some content specifying the value of the element; this content may be simple—a number or string, or may be another metadata entity.

Metadata entity: a named set of metadata elements describing some aspect of a resource.

Metadata register: an information store that contains a collection of registered metadata records, maintained by a metadata registry. (ISO 11179)

Metadata registry: an information system for assignment of unambiguous identifiers to administered metadata records. (ISO 11179)

Metadata section: Part of a metadata document consisting of a collection of related metadata entities and metadata elements (ISO 191115).

Metadata: data about a resource in some context. Generalize from ISO 11179 definition of metadata, which constrains the scope to data about data. For USGIN purposes, metadata may describe any resource—including electronic, intellectual, and physical artifacts. Metadata specify characteristics that can be queried and presented for evaluation and further processing by both humans and software.

Profile: set of one or more base standards and - where applicable - the identification of chosen clauses, classes, subsets, options and parameters of those base standards that are necessary for accomplishing a particular function [ISO 19101, ISO 19106]

Resource: An identifiable thing that fulfills a requirement. Usage here is closer to definition used in RDF (www.w3.org/TR/REC-rdf-syntax*),* generalized from ISO19115, which defines resource as an ‘asset or means that fulfills a requirement’ without defining asset or means. "An object or artifact that is described by a record in the information model of a catalogue" (OGC 07-006r1)
Service metadata: metadata describing the operations and information available from a server.

Source Specification: The specification or standard that is the subject of a profile.

User Community: A group of agents who decide to make a similar usage of a source specification in order to be able to interoperate.

Note that throughout this document, the names of XML elements are shown in `this typecase`. Long X-paths have been broken with non-breaking hyphen characters. Note that hyphens are not used in any XML attribute or element name, so if they appear in the text, they are strictly for better text wrapping. In Xpath expressions `/../` indicates that some elements have been omitted from the path.

1.4 ISO Schemas Location

ISO I9139 xml schemas are hosted by the International Organization for Standardization (ISO) at http://standards.iso.org/ittf/PubliclyAvailable-Standards/ISO_19139_Schemas/gmd/. The ISO Technical Committee 211 (TC211) also hosts the schemas at http://www.isotc211.org/2005/gmd/. These schema implement ISO Technical Specification 19139:2007 (dated 2007 Apr 17), which appears to include the changes from ISO 19115:2003 Cor 1;2006(E), but do not include any of the service metadata content.

The Open Geospatial Consortium (OGC) hosts a copy of the schemas in an online repository at http://schemas.opengis.net/iso/19139/. In the OGC repository, two versions are posted: 20060504 and 20070417. Unfortunately, these two directories both contain schema with the same target namespace, so there is no clear way to distinguish applications that are based on one or the other. The medatadaEntity.xsd in the two directories is identical; other schema have not been compared (but see discussion paper gin2009-005 at http://lab.usgin.org/node/269). The 20070417 directory contains schema implementing ISO Technical Specification 19139:2007 (dated 2007 Apr 17), apparently identical to that in the ISO repositories, but this is not declared in any included documentation (need metadata on the metadata schema!).

The 20070417 version of the ISO 19139 schemas references GML 3.2.1. However, there is no mention of the SRV namespace (http://www.isotc211.org/2005/srv) in this schema. The SRV namespace is required to specify information about dynamic, online services such as WFS and WMS, so the 20070417 version is not useful for metadata catalogs that register services.

In order to create metadata for both static datasets and dynamic, online services and for use with CSW, the OGC created an xml schema that merges the schema for ISO19115 (dataset metadata) and ISO19119 (service metadata) (see section D.1.5, page 105 in OGC 07-045). The way that was accomplished was by creating a schema located at http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd. This schema simply imports .. iso/19139/20060504/gmd/gmd.xsd and .. iso/19139/20060504/srv/srv.xsd. Thus for CSW 2.0.2 implementations, the 20060504 versions of the ISO19139 schema must be used.

2 Overview of the Profile

2.1 Quick Reference

This section provides a summary of the rules for an ISO 19139 XML document to be considered a valid USGIN metadata record

1. MD_Metadata/fileIdentifier/gco:CharacterString must be present, and may be interpreted by client applications as a string that is unique to each particular metadata record in the scope of the catalog from which it was obtained. This is typically a UUID autogenerated by the metadata editing tool or import process

2. MD_Metadata/language/gco:CharacterString will be assumed to be 'eng' (English) unless another value is specified. Any characters after the first three letters in this character string may be ignored by client applications.

3. Unless otherwise specified, the character set element will be assumed to be:

```
<gmd:characterSet>
      <gmd:MD_CharacterSetCode
codeListValue="utf8">UTF-8</gmd:MD_CharacterSetCode>    </gmd:characterSet>
```

Unless otherwise specified MD_Metadata/hierarchyLevel/MD_ScopeCode will be assumed to be:

```
<gmd:hierarchyLevel>
  <gmd:MD_ScopeCode  codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/
  ISO_19139_Schemas/resources/Codelist/gmxCodelists.xml#MD_ScopeCode"
    codeListValue="dataset">dataset</gmd:MD_ScopeCode>
</gmd:hierarchyLevel>
```

Metadata with scope code values that are not in the set {`collectionHardware`, `collectionSession`, `dataset`, `series`, `nonGeographicDataset`, `dimensionGroup`, `fieldSession`, `software`, `service`, `model`, `tile`} will be considered out of scope and may be ignored by a harvesting catalog.

4. MD_Metadata/hierarchyLevelName/gco:CharacterString is mandatory to identify the USGIN resource type from Table 1 (above). Default value is "Dataset". Encode hierarchy by including `hierarchyLevelName` elements for all broader resource categories. E.g. default should also include a `hierarchyLevelName`="Collection" element.

USGIN Hierarchy level name	**Broader Resource Type**
Collection	
Dataset	Collection
Catalog	Dataset
Physical artifact collection	Collection
Document	
Image	Document
StillImage	Image
Human-generated image	StillImage
Photograph	StillImage

Remote sensing Earth image	StillImage
Map	Human-generated image | StillImage
MovingImage	Image
Sound	Document
Text	Document
Hypertext document	Text
Model	
Physical artifact	
Service	
Software	
Stand-AloneApplication	Software
WebApplication	Software

5. USGIN requires at least one MD_Metadata/contact/CI_ResponsibleParty with role.CI_RoleCode@codeListValue = "originator" (CI_RoleCode element value = "originator") that identifies the original source of the metadata record. The CI_ResponsibleParty element must include a contact e-mail address (electronicMailAddress) or telephone number (voiceTelephone), and one of individualName, organisationName, or positionName must be present.

6. USGIN profile requires a MD_Metadata/dateStamp/gco:DateTime (Note this contrasts with INSPIRE mandate to use dateStamp/gco:Date) that specifies the date and time when the metadata record was created or last updated.

7. MD_Metadata/metadataStandardName/gco:CharacterString must be the string "ISO-USGIN" to identify a metadata record as conforming to this profile.

8. MD_Metadata/metadataStandardVersion/gco:CharacterString must be the string "1.2" to identify a metadata record as conforming to this profile.

9. MD_Metadata/identificationInfo/*MD_Identification*/citation/CI_Citation/title is mandatory. A meaningful title that is unique within the scope of the catalog should be provided for each resource that is described by a metadata record. This may be in a MD_DataIdentification or SV_ServiceIdentification element.

10. MD_Metadata/identificationInfo/*MD_Identification*/citation/CI_Citation/date is mandatory, and must be specified as an xml DataTime data type: `YYYY-MM-DDThh:mm:ssZ`. Some validators require time zone as well; recommend GMT indicated by 'Z', other time zones are indicated by "+" or "-" integer offset from GMT, e.g. "+7" for US MST.

11. MD_Metadata/identificationInfo/*MD_Identification*/citation/CI_Citation/responsibleParty is mandatory with role CI_RoleCode@codeListValue one of {originator, principalInvestigator, processor, author}. This element identifies the agent responsible for the intellectual content of the described resource. The CI_ResponsibleParty element must include a contact e-mail address (electronicMailAddress) or telephone number (voiceTelephone), and one of individualName, organisationName, or positionName must be present.

12. identificationInfo/*MD_Identification*/abstract is mandatory, but nilable. Should include a description of resource content and any other information that will help users evaluate and use the resource.

13. identificationInfo/*MD_Identification*/-extent/EX_Extent/-geographicElement/-EX_Ge

ographicBoundingBox is mandatory if the resource has a geographic footprint. Geographic extent must be represented by bounding box in WGS 84 decimal degrees. Non-geographic resources are given a location keyword 'non-geographic'.

2.1.1 Non-service resources

Any resource that is not a service is described by a metadata record with a MD_DataIdentification as the concrete implementation of the abstract *MD_Identification* element.

1. If the metadata describes a physical resource, identificationInfo/MD_DataIdentification/-pointOfContact/CI_ResponsibleParty must provide contact information for the agent who is the resource steward. Count of (individualName + organisationName + positionName) must be > 0. The CI_ResponsibleParty/role/CI_RoleCode is from CI_RoleCode codelist. ISO role codes for physical resource point of contact are {custodian, owner, pointOfContact}.

2. MD_Metadata/distributionInfo/MD_Distribution/distributor/MD_Distributor/distributorContact/CI_ResponsibleParty is mandatory with role.CI_RoleCode@codeListValue = "pointOfContact" (CI_RoleCode element value = "pointOfContact") that provides a contact point to report problems with accessing the resource through distributions associated with the distributor. The CI_ResponsibleParty element must include a contact e-mail address (electronicMailAddress) or telephone number (voiceTelephone), and one of individualName, organisationName, or positionName must be present.

3. MD_Metadata/distributionInfo/MD_Distribution/distributor/MD_Distributor/distributionOrderProcess/MD_StandardOrderProcess is mandatory if the described resource is not available online. This should be a free text explanation of how to access the described resource.

4. MD_Metadata/distributionInfo/MD_Distribution/distributor/MD_Distributor/distributorTransferOptions/-MD_DigitalTransferOptions/online/CI_OnlineResource/linkage/URL is mandatory if the described resource is available online. Note that if online access is through a service, additional CI_OnlineResource properties may be required to provide sufficient information to allow software agents to use the metadata to connect to the service.

2.1.2 Service Resources

Service metadata records in the USGIN profile are intended to describe a particular service instance from the point of view of a service implementation and enpoint. Each dataset exposed by the service should be described by a separate metadata record with MD_HierarachyLevel = 'Dataset'.

1. The srv:ServiceType element value must be populated from a string in the USGIN ServiceType codelist (Table 11).

2. Service status is required. Value is from MD_ProgressCode codelist. ISO Code names applicable to services include {`completed, obsolete, onGoing, planned, required, underDevelopment`}.

3. Service metadata records must provide a coupling type attribute specifying 'tight', 'mixed', or 'loose' coupling with particular data.
4. If coupling type is 'tight' or 'mixed' a at least one `coupledResource` element is mandatory.
5. At least one service operation MUST be described with the `operationDescription/-CharacterString` = 'serviceDescription'. The `identificationInfo/SV_ServiceIdentification/-containsOperations/SV_OperationMetadata/connectpoint` link for this operation MUST retrieve the service-specific self description document (e.g. WSDL, GetCapabilities, WADL). The `CI_OnlineResource` in this connectPoint elment must have `CI_OnlineResource/name` = "serviceDescription" (from the table in section 7.2 Linkage name conventions).

2.2 USGIN specification constraints and recommendations

Recommended practices

- DatasetURI if there is one. Nilable if not available.
- All resources should have a point of contact responsible party thaqt specifies the agent currently responsible for curation and maintenance of the resource.
- Allow multiple distributor-format-transfer option combinations for a single resource. This is critical information for enabling automated access to resources based on the metadata content. The ISO metadata model allows different approaches to binding between a distributor and resource distributions from that distributor when there are multiple distributors and distributions. This profile defines conventions for documenting those bindings.
- Representation of aggregated resources done using `identificationInfo/MD_Data-Identification/aggregationInfo/MD_AggregateInformation`, not `MD_Metadata/parentIdentifier`
- Vertical resource extent may use an elevation convention or a depth convention. CRS referenced to mean sea level, meters, measured positive up.
- Resolution expressed using `equivalentScale/MD_RepresentativeFraction/denominator` is mandatory.
- Recommended distribution format codes (Table 6) for `distributionFormat/name` introduced for categorization of physical resources, like a book, rock sample, paper document. USGIN recommends use of MIME types if they are registered for the format, and provides a recommended syntax for file formats that do not have corresponding MIME types.
- Table 12 is a list of recommended `CI_OnlineResource/name` strings used to identify special online resources link icons for branding.
- Use of a controlled vocabulary for a CharacterString data type value is indicated by including an xsi:type=" gml:CodeType" on the gco:characterString element, which then requires a codeSpace attribute (see 4.14.2-Non digital resources and 7.2-Linkage name conventions). This codeSpace should be the URI for the vocabulary used. The CharacterString element value MUST then be an identifier from that vocabulary. This convention turns the gco:CharacterString value into a GML scoped name or gco:LocalName element.

- If the metadata resource is not specific to a geographic extent, the *`MD_identification/`*`extent/-EX_Extent/geographicElement/` should have a

nilReason="inapplicable" and nil="true" attributes and the place keyword 'non-geographic' should be included in the descriptiveKeywords element.

2.3 USGIN specification extensions

Extensions to ISO19115, ISO119, ISO19139 introduced by USGIN profile:

Allow use of `identificationInfo/SV_ServiceIdentification/coupledResource/-SV_CoupledResource/ScopedName` defined by OGC 07-045 ISO profile for CSW 2.0.2, use to provide WMS layer names or WFS feature names for service requests.

2.4 General Objectives

The Profile defines:

- mandatory and conditional metadata sections, metadata entities, and metadata elements
- the minimum set of metadata elements for any resource in order to conform to the Profile
- the core metadata for geographic datasets
- optional metadata elements that allow for a more extensive standard description of resources
- some recommended practices to increase the utility and interoperability of metadata.

2.5 Requirements

M (mandatory). Metadata element must have a valid value.

C (conditional). Metadata element is mandatory based on values of other metadata elements in the metadata record.

O (optional). Metadata element may be null in a valid document.

X (not used). Metadata element is not used by a Profile. The element may be included where it is schema valid, but may be ignored by applications implementing the profile.

2.6 Use cases to be supported

This section includes a number of user scenarios for how we intend USGIN metadata to be used, and discussion of several basic approach requirements that guide metadata content recommendations. At its heart, the problem is to find resources of interest via the internet, based on criteria of topic, place, or time, evaluate resources for an intended purpose, and learn how to access those resources. Detailed metadata describing a resource data schema, describing service or application operation, or providing detailed descriptions of analytical techniques and parameter are outside the scope intended for USGIN metadata. Our contention is that this more domain/resource specific type information is better accounted for with linked documents utilizing schema appropriate to those specific resource. Some examples include OGC getCapabilities, WSDL, and ISO 19110 feature catalogs. For more in depth discussion of use cases, scenarios, and requirements, see Metadata Recommendations for Geoscience Resources (Richard and Grunberg, 2010).

- A user specifies a geographic bounding box or one or more text keywords to constrain the resources of interest, and searches a metadata catalog using these criteria. The user is presented with a web page containing a list of resources that meet the criteria, with links for each resource that provide additional detailed metadata, and direct access to the resource if an online version is accessible, e.g. as a web page, Adobe Acrobat document, or online application (see Accessing Resources, below).
- A client application provides user with a map window that contains some simple base map information (political boundaries, major roads and rivers). User wishes to assemble a variety of other data layers for a particular area for some analysis or data exploration, e.g. slope steepness, geologic units, bedding orientation, and vegetation type for a hazard assessment. User centers map view on area of interest, then using an 'add data' tab, accesses a catalog application that allows them to search for web services that provide the desired datasets. After obtaining the results and reviewing the metadata for the located services, user selects one or more to add to the table of contents for the client application. Response from catalog has sufficient information to enable the client application to load and use the resource (e.g. serviceType, OnlineResourceLinkage). More concrete instances of this case would be finding Web Map services to add as layers in an ESRI ArcMap project, borehole Web Feature Services to post borehole logs in a 3-D mapping application, or water chemistry data Web Feature Service to bring data into a spreadsheet or database.
- A catalog operator wishes to import and cache catalog records from a collaborating catalog that have been inserted or updated during the last month (harvest). This operation requires knowledge of the metadata standard and version used for the returned records.
- A user discovers an error in a metadata record for a resource that they have authored, and wishes to contact the metadata producer to request correction.
- A search returns several results that appear to contain the desired content, and user must select the most likely to meet their needs. Metadata should provide sufficient information to guide this decision.
- A project geologist at Company X is searching for data relevant to a new exploration target, and wishes to restrict the search to resources that are publicly available.

2.7 Resources of interest

Table 1 summarizes the geoscience information resources of interest to the community that can be registered and discovered using this metadata profile. Note that this collection of resource types includes several kinds of resources that are not typically associated with ISO19115/ISO19119, which were created specifically for geospatial resources. *Table 1. Summary of resource types described by metadata for US Geoscience Information Network catalogs. Resource type **names in bold** have been prioritized for implementation in version one catalogs. The Resource type names include the type hierarchy encoded with the broader (parent) resource type indicated in the Broader Resource Type column.*

esource ype erarchy	**Broader Resource Type**	**Source**	**Definition**	**string value for hierarchyLe elName**
ollection		DCMI resource Types http://dublinco	An aggregation of resources. A collection is described as a group; its parts may also be separately described. (from http://www.ukoln.ac.uk/metadata/-dcmi/collection-application-profile/): The term "collection" can be applied to any aggregation of physical or digital	collection

		re.org/documents/dcmi-type-vocabulary/	items. Those items may be of any type, so examples might include aggregations of natural objects, created objects, "born-digital" items, digital surrogates of physical items, and the catalogs of such collections (as aggregations of metadata records). The criteria for aggregation is arbitrary. Collections may contain any number of items and may have varying levels of permanence. A "collection-level description" provides a description of the collection as a unit: the resource described by a collection-level description is the collection, rather than the individual items within that collection. Collection-level descriptions are referred to in Heaney [2000] as "unitary finding-aids".	
ataset	Collection	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A collection of data items in which individual data items are identified and accessible. DCMI definition is "Data encoded in a defined structure." with additional comment "Examples include lists, tables, and databases. A dataset may be useful for direct machine processing." The container may be a stand-alone digital file (mdb, spreadsheet, table in a Word document), a web service, or an enterprise database; these are called different 'distributions'. Synonym: structured data collection. This resource type represents the intellectual artifact (FRBR 'work') -- the information content and organization; the dataset may have more than one manifestation (format) -- as a list, a table, databases, using different software implementations.	collection:dataset
atalog	Dataset	USGIN	A collection of data items that index resources, as in metadata records; a metadata registry. The resource represents the information content and organization. Catalogs are accessed using other resources, like an interactiveResource or Service, and may have different formats. This Resource type should be used to categorize web applications and services that are interfaces to metadata catalogs.	collection:dataset:catalog
hysical tifact ollection	Collection	USGIN	A collection of identifiable physical objects, unified based on some criteria. Criteria for defining a collection may be who collected, where curated, why collected, kind of material…	collection:physical artifact collection
ocument		USGIN	A packaged body of intellectual work; has an author, title, some status with respect to Review/authority/quality. USGS peer reviewed would be a 'status property'. A document may have a variety of physical manifestations (pdf file, hardbound book, tiff scan, Word processor document...), and versions may exist as the document is traced through some publication process. May be map, vector graphics, text. Sound, moving images are included as document types.	document
nage	Document	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A visual representation other than text. Examples include images and photographs of physical objects, paintings, prints, drawings, other images and graphics, animations and moving pictures, film, diagrams, maps, musical notation. Note that Image may include both electronic and physical representations.	document:image
tillImage	Image	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A static visual representation. Examples include paintings, drawings, graphic designs, plans and maps. Recommended best practice is to assign the type Text to images of textual materials if the intent of the image is to capture the textual content as opposed to the appearance of the medium containing the text. Instances of the type Still Image must also be describable as instances of the broader type Image. Subtype of Image.	document:image:stillimage
uman-generated image	StillImage	USGIN	Image produced by human drawing or painting, using any media. May be entirely product of human imagination, human perception of the world, or a human-modified photographic image.	document:image:stillimage:human-generate

				image
hotograph	StillImage	USGIN	Image produced by optical device with chemical or electronic image capture; represents things in the field of view directly as captured by the device. Photographs may be modified by human processing; there is a continuum between photographs and human-generated image. Distinction between the two is largely based on intention	document:ima e:stillimage:hu man-generate image:map
emote ensing arth image	StillImage	USGIN	Image of Earth surface acquired by an air born or earth-orbiting sensor. May be georeferenced such that location in the image directly corresponds to location on the earth.	document:ima e:stillimage:ph tograph
ap	Human-gen erated image	USGIN	Human-generated depiction of some part of the earth using a mathematical system of correspondence between geometry in the image and location on the earth.	document:ima e:stillimage:re mote sensing earth image
ovingImage	Image	DCMI resource Types http://dublinco re.org/docume nts/dcmi-type-vocabulary/	A series of visual representations imparting an impression of motion when shown in succession. Examples include animations, movies, television programs, videos, zoetropes, or visual output from a simulation. Instances of the type Moving Image must also be describable as instances of the broader type Image. Subtype of Image. Commonly include sound	document:ima e:moving imag
ound	Document	DCMI resource Types http://dublinco re.org/docume nts/dcmi-type-vocabulary/	A resource primarily intended to be heard. Various media and encoding schemes may be used, e.g. vinyl disks, digital files, magnetic tape recordings.	document:sou d
ext	Document	DCMI resource Types http://dublinco re.org/docume nts/dcmi-type-vocabulary/	A resource consisting primarily of words for reading. Examples include books, letters, dissertations, poems, newspapers, articles, archives of mailing lists. Note that facsimiles or images of texts are still of the genre Text.	document:tex
ypertext ocument	Text	USGIN	A body of text that contains navigable links allowing non-linear reading paths through the work. Links to documents or other resources outside of the document are possible; and the document itself may be packaged in discrete parts (e.g. pages, files). The criteria for determining document membership are somewhat arbitrary, but in general the document (analogous to a single web site) should contain related parts authored and managed by the same agent.	document:tex hypertext document
vent		DCMI resource Types http://dublinco re.org/docume nts/dcmi-type-vocabulary/	A non-persistent, time-based occurrence (perdurant). Metadata for an event provides descriptive information that is the basis for discovery of the purpose, location, duration, and responsible agents associated with an event. Examples include an exhibition, webcast, conference, workshop, open day, performance, battle, trial, wedding, tea party, and conflagration.	event
roject	Event	USGIN	A termporally extended event characterized by activity that has some stated purpose; projects typically have associated spatial extents, which represent the area of interest for the project. This extent serves as a mechanism to filter descriptions and concepts in the information system for those that may be related to the project based on spatial relationships. Projects in a large	event:project

			organization will likely have hierarchical (part-whole) relationships.	
odel		USGIN	Algorithm, workflow; an abstract representation of a collection of related processes, objects and relationships. A model resource may be related to various kinds of document that portray the model, or to software that implements the model, or with datasets as input or output.	model
hysical rtifact		DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	General category for physical resources that are indexed by metadata records; also root of an artifact type hierarchy. An identifiable physical object. Identification is always a function of some human intention, thus differentiating an artifact from other 'natural' things. Note that digital representations of these objects that are accessible on the WWW will be Images, Text or one of the other types.	physical artifa
ervice		DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A resource that provides one or more functions via a network interface designed for machine interaction. An implementation of an interface to some sort of digital resource, using either a 'pull' model in which client requests some content from the service, and receives that content in a single 'response' package, or a 'push' model in which client establishes connection and monitors for change events (update, new data…) from service. Difficult to draw line on when a service provides 'files' and when it provides 'data', because responses are always in a form that could be considered a file. Also includes interfaces to digital resources that provide a continuous (with some sampling interval and unit of packaging for distribution) feed of some sort of data.	service
oftware		USGIN	A computer program in source or compiled form. Examples include a C source file, MS-Windows .exe executable, or Perl script. Identity is associated with function, authorship, and programming environment.	software
tand-Alone-pplication	Software	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	Identifiable stand alone software application. Identity of resource is based on function performed, input and output requirements, and authorship. The same application may be packaged in different file formats to run in different software environments; thus an application might have one or more associated digital files. For this catalog scheme, stand alone applications are software that can be packaged in a single file and transferred between machines, unpackaged and compiled or installed on a computer meeting specified hardware and software environment conditions, to execute the described function on that computer, independent of any network connection.	software:stand alone-applicat on
teractive-esource	Software	DCMI resource Types http://dublincore.org/documents/dcmi-type-vocabulary/	A resource requiring interaction from the user to be understood, executed, or experienced. Comment: Examples include forms on Web pages, applets, multimedia learning objects, chat services, or virtual reality environments. Interactive resources are software driven. From the point of view of the catalog, they are accessed by a URL to a web site that is the interface for operating the application. The application operates by interaction with one or more human participants. The application requires network connection to operate, is accessible via the internet, and requires human interaction.	software:inter ctive resource
tructured gital data em		USGIN	An individually identifiable item in a structured digital data collection. Characterized by a schema, and some particular values. In ISO11179 terms, this is an instance of a data element. Tagging, commenting, reviewing, rating community interaction with catalog will probably require metadata records about particular data items in cataloged datasets (including metadata items in catalogs.)	structured digital data item
ampling oint, site, ation	Structured digital data item	From ScienceBase item types, SMR redux	A resource that is a location-based container/base for observation data. These represent OGC O&M sampling features, and can be generalized to include other sampling geometry (borehole, image footprint)... Analogous in function to a keyword, but carries metadata on who located, when, why, how...	sampling poin

3 USGIN Usage of Metadata Elements

3.1 Core spatial dataset, dataset series, and service elements

Table 2 is a listing of ISO19115 metadata elements used to describe any resource. Tables 3 and 4 provide specifics for describing datasets and services. Note that in the USGIN context, dataset is construed quite broadly to include any kind of georeferenced information resource, including physical samples and hard copy documents. The service metadata elements are defined by ISO19119. The root element of ISO XML-encoded metadata is MD_Metadata. Elements are discussed in this table in the order that they appear in the metadata document. Not all elements are discussed in detail. In a number of places where USGIN makes no specific provisions, we defer to recommendations in the North American Profile for ISO metadata (INCITS 453, referred to as NAP). Note that throughout this and the subsequent tables, the names of XML elements are shown in `this typecase`. Long X-paths have been broken with non-breaking hyphen characters. Hyphens are not used in any XML attribute or element name, so if they appear in the text, they are strictly for text wrapping.

Table 2. Description best practices for ISO19139 metadata elements in USGIN profile. This table includes base elements.Elements are in the order that they appear in a metadata instance.

ISO 19115 (M/C/O) xPath from MD_Metadata	**NAP-USGIN M/C/O**	**Comments**
Metadata file identifier (O) `fileIdentifier`	M-M	Identifier for the metadata record, as opposed to `DatasetURI`, which resource. A unique metadata record identifier must be included to such as GetRecordById or harvest transactions. This identifier shou harvest operations. Ideally there is one metadata record describing that there should be a one-to-one mapping between metadata fileI DatasetURIs. However, not all described resources will have a Dat metadata record is a different resource from the resource it describ have the same identifier. The protocol used to generate the identifi long as it generates globally unique identifier strings. Services that (e.g. serviceURL and layerID) are expected to put the key values in there is technically no limit on the length of the identifier string, sug to keep the string length less than 255 so the string will fit in legacy fields. USGIN, ANZLIC, and the OGC CSW profiles for ISO metadata (OG the use of the UUID (Universally Unique Identifier) for the `fileIdent` `fileIdentifier` is used to identify duplicate copies of metadata reco metadata record from another (via `MD_DataIdentification/aggregatio` metadata from a described resource (e.g. `DS_Dataset/has/MD_Metadat` difference between the two metadata records then one can determ version by the content of other elements in the metadata record. Th metadata record should be the only one made publicly available in systems such as a catalog service.
Metadata language (M) `language`	M-M	The language string is composed of a three lower-case letter langu Most CSW client and server applications only support the three lett (Note: The recommended USGIN practice is different from the NAP including an alpha3 country code (ISO3166-1), given in uppercase,

		"<ISO639-2/T three letter language code><;><blank space><ISO3 country code>", e.g. fra; CAN. Language searches for the three le wildcard character will work with either convention.)
Metadata character set (C) `characterSet`	M-M	USGIN recommends use of ISO codelists: codeListValue="utf8" co **Hyperlink reference not valid.**. See 4.17.3 Codelists for discussio USGIN requires that a character set code is specified, but this will r value for all records in a catalog.
Parent metadata record (O) `parentIdentifier`	O-X	Not used in USGIN profile. USGIN CSW service implementations c be able to navigate parent links to obtain inherited metadata proper using parent links, so this element is not used. To represent relatio described resources use `MD_Identification/aggregationInfo`.
Resource type (C) `hierarchyLevel`	M-M	Cardinality is 1...*. NAP and ISO codelists are equivalent. See 4.1 discussion of encoding of codelist values. Due to interoperability pr mandates use of ISO codelists. At least one MD_ScopeCode codelist value is required; default is 'd catalog metadata, the codelist is {`collectionHardware, collect series, nonGeographicDataset, dimensionGroup, fieldSes service, model, tile`}; other values from the ISO 19115 Scope considered out of scope for USGIN catalog metadata and the meta ignored by harvesting catalogs. The European INSPIRE Implement (MD_IR_and_ISO_20090218) proscribes the code list for the first h element in an MD_Metadata document to be one of {`dataset, se` metadata set will be considered out of scope for the directive (see s Type); in order to maintain compatibility with INSPIRE catalogs, it is scope code from this subset. ISO Example – dataset metadata: <gmd:hierarchyLevel> <gmd:MD_ScopeCode codeList="http://standards.iso.org/ittf/PubliclyAvailableStanda ISO_19139_Schemas/resources/Codelist/gmxCodelists.xml#M codeListValue="dataset">dataset</gmd:MD_ScopeCode> </gmd:hierarchyLevel>
Resource hierarchy level name (C) `hierarchyLevelName`	O-M	USGIN makes this property mandatory to identify the USGIN resou (above). Default USGIN `hierarchyLevelName.CharacterString` is "Data by including `hierarchyLevelName` elements for all broader resource ca should also include a `hierarchyLevelName`="Collection" element. For `hierarchyLevelName.CharacterString` is "Service". As use cases deve rationale for definition of sub-categories of service, the resource ca expanded. (ISO 19115 assumes that the metadata hierarchy level "dataset" if it is not documented. NAP does not use it.) Example – dataset metadata: <gmd:hierarchyLevelName> <gco:CharacterString>Dataset</gco:CharacterString> </gmd:hierarchyLevelName> <gmd:hierarchyLevelName> <gco:CharacterString>Collection</gco:CharacterString> </gmd:hierarchyLevelName>
Metadata point of contact	M-M	Cardinality on contact is 1..*. USGIN requires at least one `CI_Resp`

(M) contact/CI_ResponsibleParty		role.CI_RoleCode@codeListValue = "originator" (CI_RoleCode elemen that identifies the original source of the metadata record. If the poin report errors, updates to metadata, etc. is different than the originat contact/CI_ResponsibleParty element may be included with role.CI_ "pointOfContact" (CI_RoleCode element value="pointOfContact" the metadata records wishes to identify itself in result records, this included in an additional MD_Metadata/contact/CI_ResponsibleParty el role.CI_RoleCode@codeListValue = "distributor". See 4.17.3 *Code* encoding of codelist values. ISO Role codes applicable in this cont {distributor, originator, pointOfContact}. The CI_ResponsibleParty element must include a contact e-mail add (electronicMailAddress) or telephone number phone/CI_Telephone/voi (individualName + organisationName + positionName) > 0. If the contactInfo/CI_Contact/onlineResource/CI_OnlineResource element fo ponsibleParty with role.CI_RoleCode@codeListValue = "originator" h name = "icon", the CI_OnlineResource/linkage/URL will be assumed to file (e.g. tif, png, jpg) for the metadata originator. This Icon will be d results to credit the metadata originator. Metadata harvesters must metadata originator information so that the origin of metadata recor should harvest the point of contact information if it is different. Othe applying to the metadata record (not the described resource) may b catalogs.
Metadata date stamp (M) dateStamp	M-M	USGIN profile requires use of dateStamp/gco:DateTime (Note this con mandate to use dateStamp/gco:Date). This is the date and time when was created or updated (following NAP). The dateStamp is assume reflect any change in the metadata record that the metadata publish through the USGIN catalog system. This is the time stamp that will to determine if a metadata needs to be updated in a harvesting cat
Metadata standard name (O) metadataStandardName	M-M	USGIN profile conformant metadata is indicated by using "ISO-USG "ISO-NAP-USGIN" has been used in the past and will also be recog should be discontinued. Use is mandatory to indicate that the metadata record conforms to
Metadata standard version (O) metadataStandardVersion	O-M	For this version of the USGIN profile, Must use the string "1.2". Use is mandatory to specify the version of the profile used
DataSet Identifier (O) dataSetURI	O-O	This is a string that uniquely identifies the described resource. If the resou should be included here; if the resource will be referenced from other met identifier here. Any kind of resource (not only datasets) may have an iden identifier is not specified, but some sort of documented scheme to assure u (UUID, URN…). USGIN is attempting to make the semantics of iden provision (see Unique resource identifier in Table 3, below) that the MD_DataIdentification/citation/CI_Citation/identifier/MD_Identifie resource. This may be identical with the resource described by the its value is identical to MD_Metadata/dataSetURI, or it may be a publica intellectual source of the described resource, in which case it is a d The MD_Distribution/transferOptions/MD_DigitalTransferOptions/onli used to specify URLs for access to the resource, even if the datasetURI is dereference to obtain a representation of the described resource. The dataS considered an opaque identifier. This will avoid ambiguity about where to

		access to a described resource. If the dataset is coupled to a service, the value of the MD_Metadata/d unique resource identifier used by srv:coupledResource to link the The *OpenGIS® Catalogue Services Specification 2.0.2 - ISO Meta* (OGC 07-045) Annex F recommends that MD_DataIdentification/ci identifier/MD_Identifier/code match the identifiers specified by SV_ServiceIdentification/operatesOn and SV_ServiceIdentification/c linking a described service to datasets that the service operates on fileIdentifier (above), this requires that a MD_DataIdentification/c CI_Citation/identifier that identifies the described resource is inclu record; its value must be the same as MD_Metadata/dataSetURI.
Other languages (C) locale	C-X	Other languages used in metadata free text description. If description in m provided, this property should indicate what those languages are. The prin metadata description is identified with MD_Metadata/language and charact languages are identified by MD_Metadata/locale/PT_locale elements, in provided according to ISO 639-2/T three-letter terminology codes in lowe country is provided according to ISO 3166-1 three-letter codes in upperca characterEncoding. See 4.17.3 *Codelists* for discussion of encoding USGIN Catalogs must harvest and process metadata content in the specified by MD_Metadata/language and characterSet, and may ig localized content.
[role] Resource spatial representation (O) spatialRepresentationInfo	O-O	Best practice is to include metadata for spatial representation if the georeferenced dataset. Metadata for Spatial data representation are der Metadata is instantiated as one or more of MD_GridSpatialRepresentation MD_VectorSpatialRepresentation, MD_Georectified, or MD_Georeferencea profile follows NAP for spatial representation metadata. Vector Spatial Re provided if point or vector objects exist in the dataset. If MD_VectorSpatia either spatialRepresentationInfo/MD_VectorSpatialRepresentation/to spatialRepresentationInfo/MD_VectorSpatialRepresentation/geometric or both." (NAP) MD_GridSpatialRepresentation or one of its subtypes (MD MD_Georeferenceable) should be provided if dataset objects are gridded. M used if the grid (image) is georeferenced, and MD_Georeferenceable is use georeferenced. Follow NAP optionality if these elements are used.
Resource spatial representation vector topology (O) spatialRepresentationInfo/-MD_VectorSpatialRepresentation/topologyLevel	C-C	Code that specifies the degree of complexity of spatial relationships betwe Value is from ISO codelist MD_TopologyLevelCode. (Code names in this lis {geometryOnly, topology1D, planarGraph, fullPlanarGrapl fullSurfaceGraph, topology3D, fullTopology3D, abstract for discussion of encoding of codelist values. It is unclear precisely w terms of the topology encoding. To be useful, assertion that topology is pr topological relationships that may be implicit in the encoded vector geom represented (e.g. by correlation tables—left poly, right poly for a polyline clarification should be included in the resource abstract. Mandatory if MD_VectorSpatialRepresentation is present.
Resource spatial representation vector geometric objects (O) spatialRepresentationInfo/-MD_VectorSpatialRepresentation/geometricObjects	C-C	"Identification of the objects used to represent features in the dataset." Pro count for the number of objects of each type. Use the ISO MD_GeometricOb Code names in this list are: {complex, composite, curve, point, 4.17.3 *Codelists* for discussion of encoding of codelist values. Mand MD_VectorSpatialRepresentation is present.

[role] Resource's spatial reference system (O) `referenceSystemInfo`	O-O?	Description of the spatial and/or temporal reference systems used in the da (`identificationInfo/spatialRepresentationType/MD_SpatialRepresenta` or (`../MD_SpatialRepresentationTypeCode` = "`grid`") or (`../MD_SpatialR "`tin`") implies count `referenceSystemInfo` >= 1) }. See 4.17.3 Codelis encoding of codelist values. Use ISO codelist.
Reference System identifier code (O) `referenceSystemInfo/-MD_Refe renceSystem/-referenceSystem Identifier/-RS_Identifier/co de`	C-C	If `referenceSystemInfo` is included, then the `RS_Identifier` element m code value. For USGIN the code should be a value from the EPSG Dataset register (http://www.epsg-registry.org/) in the form "EPSG:nnnn" code number for the CRS. If the CRS is not defined in the EPSG registry, specified in the NAP profile should be followed, e.g. the CRS shall be des 19111 and ISO/TS 19127, assigned an identifier, and registered with an au referenced here. The `RS_Identifier/codespace` in this case should identify where the CRS definition is registered, such that the definition can be loca USGIN purposes is to provide georeferenced data using one of the EPSG reference systems if this is possible.
Metadata extension information (O) `metadataExtensionInfo`	X-X	Not used in this profile.
Resource identification information (M) `identificationInfo`	M-M	Cardinality 1..*. The content of this element identifies the described resources that are not services, use `MD_DataIdentification` (**see Tal `SV_ServiceIdentification` is required (**see Table 4)**.
[role] Content information (O) `contentInfo`	O-O	Characteristics describing the feature catalog, coverage, or image `Information` is an abstract class. One or more of `MD_FeatureCatalogu` `MD_CoverageDescription` or `MD_ImageDescription` elements may be use content. `MD_FeatureCatalogueDescription` describes content in a featu like an ESRI geodatabase that may have more than one feature, e. polygons, fault line features, and point observation locations for stri Entity-attribute metadata for dataset should be documented using I are located at http://www.isotc211.org/2005/gfc/gfc.xsd. The ISO1 implementation of ISO19115 does not allow inclusion of the feature metadata record, but provides a citation element in the `gmd:content` `CatalogueDescription` element. The USGIN profile provision is to link using an xlink:href attribute on the `gmd:featureCatalogueCitation`, ide accompanying xlink:title attribute with the value "Link to ISO19110 Example 2 shows recommended encoding. An example feature cat ISO19110 XML is included in section 8.4, ISO19110 Feature Catal `MD_CoverageDescription` is for datasets that are one of the types liste MD_CoverageContentTypeCode: image, thematicClassification, ph coverage is a data structure that acts as a function to return values direct position within its spatiotemporal domain (OGC 07-067r5). In values for light intensity in a given wavelength range, thematicClas return codes corresponding to some domain concept, and physical coverages return values representing some physical quantity like m density, resistivity.
[role] Resource distribution information (O) `distributionInfo`	O-M	This element provides information to inform users how to obtain or resource, which is considered part of the minimum useful content th by a metadata record. USGIN profile specifies that at least one `MD_` and one `MD_Distribution/transferOptions` element are required, and

		provides the specified transfer options. See section 4.13 '*Use of MD_Distribution and MD_Distributor*' for instructions for more complicated combinations transfer options, and ordering instructions.
Resource distribution format (O) `distributionInfo/-MD_Distribution/distributionFormat`	O-O	Information on the format or physical manifestation of the resource • If the resource is a physical (not electronic) resource, like a boo document, the `distributionFormat/MD_Format/name` is mandatory, Table 6. USGIN Distribution formats for non digital resources. U http://resources.usgin.org/registry/distributionFormatNames201 • For digital resources, the format specifies the file type, either us formatted string. Pattern for digital resources that do not have a `[vendor:applicationName]/fileExtension`. The vendor an may not be applicable, and could be omitted, but the '/' and file always be present. • If the format consists of a single file, the file extension is a three abbreviation assigned by the vendor.
Resource distributor information (O) `distributionInfo/-MD_Distribution/-distributor/-MD_Distributor/`	O-C	*USGIN differs from NAP* in this case (but not with ISO19115) by all distributors, and binding between distributors, transfer options, and profile, each `distributor/MD_Distributor` is a binding between one o and the distributor formats that are available through that/those tra (`MD_DigitalTransferOptions/onLine/CI_OnlineResource` in particular). I available from the same distributor, or have different transfer option represented as different `distributor/MD_Distributor` instances. See *MD_Distribution and MD_Distributor*' for instructions on use of thes
Resource distributor responsible party (O) `distributionInfo/-MD_Distribution/-distributor/MD_Distributor/-distributorContact/-CI_ResponsibleParty`	C-M	`MD_Distributor` is required, which requires one `CI_ResponsibleParty`. role.CI_RoleCode@codeListValue = "`pointOfContact` " (CI_RoleC "`pointOfContact` ") is mandatory, to provide a contact point to rep accessing the resource through distributions associated with the di CI_ResponsibleParty element must include a contact e-mail addres (electronicMailAddress) or telephone number (electronicMailAddres individualName, organisationName, or positionName must be pres distributorContacts with different roles may be ignored by USGIN c applicable in this context include: {`resourceProvider, distribu` See section 4.17.3 '*Codelists*' for details on codelist encoding.
Resource distributor order process (O) `distributionInfo/-MD_Distribution/distributor/-MD_Distributor/-distributionOrderProcess/-MD_StandardOrderProcess`	O-C	Information on the availability of the resource including at least one and time, ordering instructions, or turnaround. For physical resourc are mandatory as these will typically indicate the method of access
Resource distributor format (O) `distributionInfo/-MD_Distribution/distributor/-MD_Distributor/-distributorFormat/MD_Format`	O-O	See section 4.14 'Distribution Format ' for instructions on use of the
Resource distribution transfer options (O)	C-C	`MD_DigitalTransferOptions` provides information on digital distributio section 4.13 '*Use of MD_Distribution and MD_Distributor*' for instru

distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransferOptions		element. If distribution information is included (MD_Distribution is n one one MD_Distribution/transferOptions element is required.
Resource distributor online distribution linkage (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransferOptions/-online/-CI_OnlineResource/linkage	M-M	Digital transfer options are "Technical means and media by which a from the distributor." The CI_OnlineResource/linkage/URL element mu URL to access the resource directly (see sections 4.13 to 4.15 for r
Resource distributor online distribution linkage (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransfer-Options/-online/-CI_OnlineResource/protocol	M-O	The CI_OnlineResource/protocol element is mandatory if access to tl additional protocol information that is stacked on the protocol speci CI_OnlineResource/linkage/URL prefix (e.g. http:, ftp:). Typically this w WMS-1.3.0, OpenSearch-1.1, or DAP-2.0. Ideally these strings sho dereferenced to learn something about the service specification.
Resource distributor online distribution linkage (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransfer-Options/online/-CI_OnlineResource/name	O-O	The CI_OnlineResource/name element may duplicate the file name if t but it is recommended to provide a user-friendly label for the file tha a user interface. For links to resource distribution end points (distributionInfo//CI_OnlineResource) that are not simple file access online order forms, mailto links to e-mail a request for the resource, special values defined in Table 13.
Resource distributor online distribution application profile (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransferOptions/-online/CI_OnlineResource/-applicationProfile	C-O	applicationProfile is mandatory if the CI_OnlineResource/linkage do HTML web page (or other standard format that will be recognized b another software application is needed to use a linked file resource is a service instance conforming to a profile. The applicationProfile specify the software using the following recommended syntax: "ver name/application version", e.g. "Microsoft:Word/2007", or "ESRI:Ar documents for which the service type and base protocol prefix on tl provide sufficient information to guide client software, the applicatio used to indicate a profile on the serviceType or some variation in de content conventions. See section 4.15 CI_OnlineResource for more
Resource distributor online distribution function (O) distributionInfo/-MD_Distribution/-transferOptions/-MD_DigitalTransferOptions/-online/CI_OnlineResource/-function	O-O	CI_OnlineResource/function is mandatory if the linkage is not a simp downloadable file. USGIN recommended values for CI_OnlineFunc summarized in Table 7. If the resource is accessible as a web serv service should be separate metadata record with the dataset(s) exp service identified in the service metadata record as coupledResour
[role] Data quality information (O) dataQualityInfo	C-O	Either dataQualityInfo/DQ_DataQuality/report or dataQualityInfo/DQ_ mandatory if a dataQualityInfo element is present. dataQualityInfo/ required, with value from MD_ScopeCode: {collectionHardware dataset, series, nonGeographicDataset, dimensionGroup, software, service, model, tile}. See 4.17.3 *Codelists* for dis codelist values. See section 4.19 Data Quality for discussion of dat parts.

Data quality scope (O) dataQualityInfo/DQ_DataQuality/scope	C-C	Mandatory if DQ_DataQuality is not null. Specifies the extent of cha data quality information is reported. Value is from MD_ScopeCode: {collectionHardware, collectionSession, dataset, serie nonGeographicDataset, dimensionGroup, fieldSession, so model, tile}. See 4.17.3 *Codelists* for discussion of encoding of
Data quality scope level description (O) dataQualityInfo/-DQ_DataQuality/-scope/levelDescription	C-X	Not used by USGIN. See section 4.19 *Data Quality* for more dis levelDescription.
Data quality report (O) dataQualityInfo/-DQ_DataQuality/report	C-C	In this profile, data quality information is optional, but if included, th data quality information is a free text explanation of data quality co DQ_DataQuality[1]//report[1]//explanation/CharacterString element text discussion/description of data quality considerations for the ind of any specific data quality element to contain this explanation is ar be considered meaningful in this context. Other data quality information should be preserved by USGIN catal metadata that includes these elements, and presented when the us metadata record. Each DQ_DataQuality/report element requires that possible data quality elements must be present, and multiple repor within each DQ_DataQuality element. Each of these *AbstractDQ_eleme* nameOfMeasure, measureIdentification, measureDescription, evaluation evaluationMethodDescription, evaluationProcedure, and dateTime elem required result elements. The *AbstractDQ_element* /result is either a a DQ_QuantitativeResult, each of which has required and optional su of this report metadata should follow recommendations in NAP.
Data quality lineage (O) dataQualityInfo/-DQ_DataQuality/lineage	C-O	Not applicable to services. USGIN recommended practice is descri
Data quality lineage statement (O) dataQualityInfo/-DQ_DataQuality/lineage/-LI_Lineage/statement	C-C	If lineage information is included, a lineage/LI_Lineage/statement is specifies a "General explanation of the data producer's knowledg (NAP). USGIN recommended practice is described in section 4.19.
Data quality lineage source (O) dataQualityInfo/-DQ_DataQuality/lineage/-LI_Lineage/source	C-O	Each source/LI_Source element describes a source data resource th processStep. NAP provision is that LI_Source/description is manda LI_Source/sourceCitation and LI_Source/sourceExtent are not provi LI_Source/description includes the source medium name from the C MD_MediumNameCode, followed by <;><blank space> and a free "dvd; source satellite image." If the source is part of a processing chain, the LI_Source/processSte provides "Information about an event related to the creation proces (INCITS 453). This is interpreted to mean that the link from a sourc a process step for which the described source is an output. USGIN is described in section 4.19.
Data quality lineage process	C-O	An event in the development of the dataset. Each step requires a fr may have a free text rationale, dateTime stamp when process was

step (O) dataQualityInfo/-DQ_DataQuality/lineage/-LI_Lineage/processStep		CI_ResponsibleParty elements identifying parties involved in the proc many source/LI_Source associations to identify data that is input into practice recommended for USGIN is that source association from a inputs to a process, and processStep associations from a source ele resource to a process step that produced it. See USGIN recommen described in section 4.19.
[role] Portrayal catalog information (O) portrayalCatalogueInfo	O-O	portrayalCatalogueInfo/MD_PortrayalCatalogReference/portrayalCatal n element identifying a catalogue that contains symbols and rules t portrayal catalog is a collection of defined symbols used to depict, t a map. No documentation in ISO19115 about how this is supposed defines the structure of a Portrayal Catalogue. No USGIN recomme been adopted.
[role] Metadata constraint information (O) metadataConstraints	O-O	This element specifies use constraints for access to the metadata r for accessing the describe resource are in resourceConstraint/MD_Co MD_DatasetIdentification or MD_ServiceIdentification. Follow NAP f access constraints. NAP provision is that metadataConstraints/MD_Constraints/useLimitat MD_Constraints is used to specify metadataConstraints. When one of MD_LegalConstraints or MD_SecurityConstraints is used, useLimitation MD_LegalConstraints are specified by MD_RestrictionCode. ISO cod {copyright, patent, patentPending, trademark, license, intellectualPropertyRights, restricted, otherRestricti *Codelists* for discussion of encoding of codelist values. otherCons element required if accessConstraints or useConstraints is set to "oth example: "Data only to be used for the purposes for which they wer MD_SecurtyConstraints has various optional free text values, and a re MD_SecurityConstraints/classification from ISO MD_ClassificationC restricted, confidential, secret, topSecret}. See 4.17.3 of encoding of codelist values.
[role] Application schema information (O) applicationSchemaInfo	O-O	Information about the information schema of the resource applicationSchemaInfo/MD_ApplicationSchemaInformation element has name/CI_Citation, schemaLanguage free text, and constraintLanguage fr MD_ApplicationSchemaInformation element also allows inclusion of an document as ASCII, or a binary graphicsFile or softwareDevelopment applicationSchemaInfo elements may be used for different presenta schema, or for different kinds of schema (e.g. physical, logical, con
[role] Metadata maintenance information (O) metadataMaintenance	O-O	This element provides information about the maintenance schedule metadata record. Only one MD_MaintenanceInformation element may required MD_MaintenanceFrequencyCode. The ISO codelist is {co weekly, fortnightly, monthly, quarterly, biannually, a irregular, notPlanned, unknown}. See 4.17.3 *Codelists* for dis codelist values. Additional metadataMaintenance/MD_MaintenanceInformation/mai added to metadata records to document updates to the the content catalogs should use metadata maintenance notes to document any harvested metadata records; ideally a link to the original metadata provided.
[role] Series information (O) series	X-X	The MD_Metadata/series element that appears in the ISO19139 sche implement the metadata application model in ISO19115:2003 Figu class diagram defining the classes of geographic information to wh

		Not Used by USGIN.
[role] Described resource (O) `describes`	X-X	The `MD_Metadata/describes` element that appears in the ISO19139 s implement the metadata application model in ISO19115:2003 Figu class diagram defining the classes of geographic information to wh Not used by USGIN.
[role] Property type description (O) `propertyType`	X-X	The `MD_Metadata/propertyType` element that appears in the ISO1913 implement the metadata application model in ISO19115:2003 Figu class diagram defining the classes of geographic information to wh Not used by USGIN.
[role] Feature type description (O) `featureType`	X-X	Although an `MD_Metadata/featureType` element that appears in the IS appears to implement the metadata application model in ISO19115 a UML class diagram defining the classes of geographic informatio applies. Not used by USGIN.
[role] Feature attributes (O) `featureAttribute`	X-X	Although an `MD_Metadata/featureAttribute` element that appears in t appears to implement the metadata application model in ISO19115 a UML class diagram defining the classes of geographic informatio applies. Not used by USGIN.

3.2 Dataset Identification properties (MD_DataIdentification)

The difference between metadata for services, and metadata for other resources is in the `identificationInfo` part of the ISO19139 xml schema. This section documents use of `MD_DataIdentification` for metadata describing resources of interest in the geoscience information network that are not services. Service metadata utilizes the `SV_ServiceIdentification` element to provide a description and identification of a service (see 3.3 Service identification elements).

Note that the USGIN approach to metadata is predicated on the idea that the users first interest is finding information about some particular topic; how that information is accessed or acquired is a secondary consideration. Thus metadata records are about information resources (mostly datasets and other data products or publications) at the FRBR work level [FRBR, 2004]. Information on accessing data resources through service interfaces is considered a kind of distribution of the resource. The MD_DataIdentification part of a metadata record describes the information resource, and multiple `distributionInformation` elements specify access to that data through one or more services, or as various kinds of packaged, downloadable files. Given that a metadata record is about exactly one 'work', each metadata record has exactly one `MD_DataIdentification` element. Metadata records about services are targeted for the rather unusual situation in which a user is primarily interested in a specific kind of service and the details of its operation, typically this will be for web processing or brokering services. In practice, service protocols typically specify some kind of service specific self-description document (e.g. WSDL, WADL, OGC GetCapabilities, OpenSearch Description) that client software is more likely to understand than an ISO SV_ServiceIdentification element.

Table 3. Dataset Identification properties (MD_DataIdentification)

ISO 19115 (M/C/O) xPath from MD_Metadata	**NAP-USGIN M/C/O**	**Comments on MD_DataIdentification**
Resource citation (M) `identificationInfo/-MD_DataIdentification/-citation/CI_Citation`	M-M	CI_Citation cardinality exactly one required. The citation attribute pr identifies the intellectual origin of the content in the described resou a citation in a scientific journal. Required content for a `CI_Citation` el and `responsibleParty`. NOTE: if multiple identificationInfo sections a catalogs may ignore all but the first instance.
Resource title (M) `identificationInfo/-MD_DataIdentification/-citation/-CI_Citation/title`	M-M	A meaningful title that is unique within the scope of the catalog shou each resource that is described by a metadata record. USGIN recor that inform the human reader about the dataset's content as well as
Resource reference date (M) `identificationInfo/-MD_DataIdentification/-citation/CI_Citation/date/-CI_Date/date/`	M-M	Best practice is to include at least the date of publication or creation date of the resource reported in the citation corresponds to the most resource's content. `CI_Date` content includes a `date` and `dateType`. Da uses xs:dateTime data type, The dateTime data type is specified in "YYYY-MM-DDThh:mm:ss". `timezoneOffset` is optional (http://www.w3.org/TR/xmlschema11-2). Example date encoding: 2000-12-12T12:00:00+13:00, 2006-10-01T month or day is not known, encode as '01', for example '2006-01-01 the event used for the temporal aspect of the resource, and must be {creation, publication, revision}, which is a subset of the ISO DateTy date is distinct from the `dateStamp` for the metadata record (date and

		metadata update), or the EX_Extent/temporalElement that specifies the the resource content is applicable.
Unique resource identifier (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/-identifier/MD_Identifier	C-C	For USGIN, if the Citation has an identifier that is different from the described resource (MD_Metadata/dataSetURI), it must be included here value should be an identifier for the cited resource, without any assu http protocol. The identifier may be resolvable to a URL, if a protoco identifier scheme that is resolvable (e.g. http, doi…), but this is not n document, and should not be assumed when processing metadata doc The USGIN profile requires the use of MD_Identifier element to iden RS_Identifier must not be used. If additional codespace and version with the identifier, it should be encoded as MD_Identifier/authority/C alternateTitle and MD_Identifier/authority/CI_Citation/edition
Resource responsible party (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/-citedResponsibleParty	M-M	USGIN requires that the citation include at least one CI_ResponsibleP count of (individualName + organisationName + positionName) mus responsible party should identify the agent responsible for the intelle described resource. In this context, the CI_ResponsibleParty/role/CI_RoleCode@codeListValue is from {originat principalInvestigator, processor, author}, a subset of the ISO CI_Rol 4.17.3 *Codelists* for discussion of encoding of codelist values. For m content, the responsible party is what would normally be considered Best practice is to include point of contact information for the resour MD_DataIdentification/pointOfContact/CI_ResponsibleParty. Guidance would be helpful for consistency, but has not been developed as yet editor, contributor} View compiler and editor as equivalent roles, sin include compiler. Use the ISO19115-1 codelist.
Resource presentation form (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/-presentationForm	O-O	The form in which the original cited resource is available. Note that t original source of intellectual content in the described resource, and be different from the format for distribution described in the metadat be specified if there is a difference between the cited resource prese the distribution format(s) listed in the distributionInfo/MD_Distributi metadata record. This information is mostly for documentation; for access, the distribution formats are more relevant. presentationForm uses CodeList = CI_PresentationFormCode, with I {documentDigital, documentHardcopy, imageDigital, image mapDigital, mapHardcopy, modelDigital, modelHardcopy, p profileHardcopy, tableDigital, tableHardcopy, videoDigit audioDigital}. See section 4.17.3 *Codelists* for details on codelis
Resource series (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/series	O-O	Information about the (publication) series or collection of which the r NAP rule (name + issueIdentification) > 0 should be followed.
Resource other citation details (O) identificationInfo/-MD_DataIdentification/-citation/CI_Citation/-otherCitationDetails	O-O	A complete recommended bibliographic citation should be specified
Resource collective title (O) identificationInfo/-MD_DataIde	O-C	Title of the combined resource that the cited resource is part of, for e resource may be a paper in an anthology, in which case the antholo

ntification/-citation/CI_Citation/-collectiveTitle		collective title. Required if the cited resource is part of such a collect
Resource abstract (M) identificationInfo/-MD_DataIdentification/abstract	M-M	A free text summary of the content, significance, purpose, scope, et Exactly one value. Content from metadata harvested in other format CSDGM…) that does not map into ISO metadata elements can be i the abstract.
Resource purpose (O) identificationInfo/-MD_DataIdentification/purpose	O-O	"Summary of the intentions for which the dataset was developed. Pu objectives for creating the dataset and what the dataset is to suppor
Resource status (O) identificationInfo/-MD_DataIdentification/status	O-O	Value is from MD_ProgressCode codelist. ISO values are {complet historicalArchive, deprecated, onGoing, planned, requir underdevelopment}. For USGIN discovery metadata, values shoul {completed, onGoing, deprecated}. See section 4.17.3 *Codelist* usage.
Resource point of contact (O) identificationInfo/-MD_DataIdentification/-pointOfContact	O-C	Contact information for the agent who is currently the steward for th information is mandatory for physical resources such as core, cuttin manuscripts. Count of (individualName + organisationName + positi The CI_ResponsibleParty/role/CI_RoleCode is from CI_RoleCode code for physical resource point of contact are {custodian, owner, poi point of contact role codes may apply for other resources. See sec for details on codelist usage.
Resource maintenance (O) identificationInfo/-MD_DataIdentification/-resourceMaintenance	O-O	This element provides information about the maintenance schedule resource (or some subset/part of the resource specified by the scop description) described by the metadata record. 0 to many MD_Mainter elements may be included. Different MD_MaintenanceInformation elem have different napMD_ScopeCode or MD_ScopeDescription. Usage of poorly described, and no actual examples of usage could be found; allow identification of a set of attribute or features (by name?), or fea attribute instances (identified how?), or a dataset, to which the main applies. Use MD_MaintenanceFrequencyCode codelist. ISO values daily, weekly, fortnightly, monthly, quarterly, biannua asNeeded, irregular, notPlanned, unknown}. See section 4.17 on codelist usage.
Graphic overview of resource (O) identificationInfo/-MD_DataIdentification/graphicOverview	O-O	Highly recommended to include a URL providing a web-accessible v of the resource if it is applicable to the described resource, particula datasets that may be represented by maps. If MD_BrowseGraphic is inc MD_BrowseGraphic/filename character string is mandatory. USGIN Rec is to provide a complete URL as a gco:characterString value for the fileType/CharacterString should be a registered MIME type or stand abbreviation (e.g. napMD_FileFormatCode). This is a repeatable ele may present different resolutions, or different parts of resource. Nam overview should provide sufficient information for user to distinguish
Resource format (O) identificationInfo/-MD_DataIdentification/resourceFormat	X-X	This element is not used by NAP or USGIN; this information is enco MD_Metadata/distributionInfo/MD_Distribution/ in USGIN metadata (*MD_Distribution and MD_Distributor*).
Resource keywords (O) identificationInfo/-MD_DataIdentification/-	O-O	Keywords should be provided to facilitate the discovery of metadata the user. Keywords should be grouped according to Keyword Type MD_KeywordTypeCode {discipline, place, stratum, tempor

descriptiveKeywords/MD_Keyword		thesaurus in which they are defined (if applicable). See section 4.1 details on codelist usage. USGIN requires that MD_Keyword/keyword contain a CharacterString (s USGIN best practice is to include keywords in English.
Condition applying to access and use of resource (O) identificationInfo/-MD_DataIdentification/-resourceConstraints/	O-O	Restrictions on the access and use of a resource or metadata. Follo specification of resourceConstraints. This attribute provides informati to the described resource itself. In some situations, the metadata ma learn of the existence of a resource that they may not actually be ab further clearance. Constraints may be represented by MD_Constraint, MD_SecurityConstraint.
Aggregation information (O) identificationInfo/-MD_DataIdentification/-aggregationInfo/-MD_AggregateInformation	O-O	This element includes either a citation for or identifier of an associat with the type of association between the datasets, and optionally the the dataset. MD_AggregateInformation requires either aggregateDataSetName/CI_Cita aggregateDataSetIdentifier/MD_Identifier. MD_AggregateInformation/ mandatory, from DS_AssociationTypeCode. ISO codelist includes { largerWorkCitation, partOfSeamlessDatabase, source, ster 4.17.3 *Codelists* for details on codelist usage. USGIN recommended a dereferenceable URI for that resource in aggregateDataSetIdentifi *[note this text updated 2013-10-25 to reflect the information we're ac* related resources that do not have a metadata record, aggregateData may be used; this element is optional if aggregateDataSetIdentifier h For USGIN profile, this property, rather than MD_Metadata/parentIdent to indicate relationships between described resources.
Spatial Representation Type (O) MD_DataIdentification/spatialRepresentationType/	O-O	value from MD_SpatialRepresentationTypeCode list. ISO codelist in grid, textTable, tin, stereoModel, video}. See section 4.17 on codelist usage.
Resource spatial resolution (O) MD_DataIdentification/-spatialResolution/-MD_resolution/equivalentScale/MD_RepresentativeFraction/-denominator	C-C	If the spatial representation type code is vector, grid, textTabl equivalentScale/../denominator MUST be provided to express spatia arbitrary convention proposed to facilitate interoperablity). If a equiva is available, that should be supplied as well. The resolution distance smallest length between two resolvable points in the dataset. To cal given a resolution distance, recommended practice is to divide the r meters by 0.0005. This assumes that the smallest distance resolvab human usage is 0.5 mm.
Resource language (O) identificationInfo/-MD_DataIdentification/language	M-O	Language for content of described resource. Default value is 'eng'. I applicable to the described resource, a nilReason='notApplicable' at included on an empty language element. Multiple instances of this e the linguistic content of the resource is available in multiple languag Use the three-letter language code from the ISO 639-2/T three lette lowercase. ISO 639 codelists are available at http://www.loc.gov/standards/iso639-2/php/code_list.php.
Topic category identificationInfo/-MD_DataIdentification/-topicCategory	C-C	A topicCategory code must be provided when hierarchyLevel is set "dataset series". Codes are from MD_TopicCategoryCode, the ISO {farming, biota, boundaries, climatologyMeterologyAtmos elevation, environment, geoscientificInformation, healt imageryBaseMapsEarthCover, intelligenceMilitary, inland oceans, planningCadastre, society, structure, transport utilitiesCommunication}. See section 4.17.3 *Codelists* for deta

		Most USGIN resources will have MD_TopicCategoryCode = "geoscientificInformation", which is the default value for this pr topic categorization should be done using keywords.
Resource content extent identificationInfo/-MD_DataIdentification/extent/-EX_Extent	C-C	Defines the spatial (horizontal and vertical) and temporal region to w the resource applies. For USGIN, a geographicElement/geographicBou provided for any resource with content related to some geographic l geoscience resources, temporal extent expressed using named time geologic time scale should be specified as thematic keywords. If the not related to a geographic area, the place keyword 'non-geographic in the descriptiveKeywords element. In some situations, geographic indicated using location keywords, but this is strongly discouraged b location-based searches using geographic coordinates will not find s cases, geocoding tools from services like Google can often be used approximate bounding box.
Resource content extent description identificationInfo/-MD_DataIdentification/extent/-EX_Extent/description	C-O	Free text that describes the spatial and temporal extent of the datas geographic place names are used to express the geographic extent specifies that these should be encoded using keyword with keyword Geographic names may be duplicated in the EX_Extent/description.
Resource content extent bounding box identificationInfo/-MD_DataIdentification/extent/-EX_Extent/geographicElement/-EX_GeographicBoundingBox	O-C	USGIN profile requires that if an EX_Extent/geographicElement is supp geographic bounding box with bounding latitude and longitude expre Geodesic System WGS 84 decimal degrees. The corner coordinates for the geographic bounding box must not c because this may result in fatal errors with some CSW implementati must thus be represented as tiny rectangles. USGIN recommended the actual point location in the lower left corner of the rectangle.
Resource content extent geographic description identificationInfo/-MD_DataIdentification/extent/-EX_Extent/geographicElement/-EX_GeographicDescription	C-X	Not used by USGIN profile. In USGIN metadata this information sh using keywords for which the MD_KeywordTypeCode = 'place'; the has the same content as EX_GeographicDescription/authority/CI_Cita is the same as the EX_GeographicDescription/code.
Resource content extent bounding polygon identificationInfo/-MD_DataIdentification/extent/-EX_Extent/geographicElement/-EX_BoundingPolygon	C-O	To improve interoperability, USGIN mandates the use of Geographi instead of bounding polygon, if a bounding polygon is included it ma harvesting catalogs. If only a bounding polygon is provided, a mini must be calculated and inserted in the EX_GeographicBoundingBox record is harvested into a USGIN catalog.
Resource temporal extent (O) identificationInfo/-MD_DataIdentification/extent/-EX_Extent/temporalElement/-EX_TemporalExtent/extent/-TimePeriod	O-O	Property contains information about temporal extent to which resour many geoscience resources, this would be the geologic time period resource applies. USGIN mandates use of TimePeriod for all temporal ex extent for beginPosition@frame and endPosition@frame attributes are #ISO time extents, USGIN requires the values for beginPosition@frame and endP populated using numeric time coordinates in Ma, measured positive increas at 1950 CE (see Temporal extents). The default frame attribute value for coordinates is "urn:cgi:trs:CGI:StandardGeologicTimeMa" ISO 8601 Default Example:

		<gml:TimePeriod gml:id="IdModern"> <gml:name>Y2kX</gml:name> <gml:beginPosition frame="#ISO-8601">2010-01-00T00:00:00</gml:beginPosition> <gml:endPosition frame="#ISO-8601">2010-12-31T24:00:00</gml:endPosition> </gml:TimePeriod> Geologic Time Example: <gml:TimePeriod gml:id="IdJurassic"> <gml:name>Jurassic</gml:name> <gml:beginPosition frame="urn:cgi:trs:CGI:StandardGeologicTime </gml:beginPosition> <gml:endPosition frame="urn:cgi:trs:CGI:StandardGeologicTimeMa ">135 </gml:endPosition> </gml:TimePeriod>
Resource spatio-temporal extent (O) identificationInfo/-MD_DataIdentification/extent/-EX_Extent/temporalElement/-EX_SpatialTemporalExtent/	O-X	Not used. Although use of EX_SpatialTemporalExtent is allowed by IS USGIN mandates encoding space time location with EX_TemporalExte EX_GeographicBoundingBox.
Resource vertical extent (O) identificationInfo/-MD_DataIdentification/extent/-EX_Extent/verticalElement/-EX_VerticalExtent	O-O	Vertical extent is used to describe the elevation location of resource Common geoscience examples are resources related to vertical loc depth in a water body. The borehole trace is the vertical coordinate (CRS) within which described resources are located using coordinat distance from the collar (or ground level, or Kelly bushing) of the bo EX_VerticalExtent has minimumValue, maximumValue that are real numbe For interoperability, USGIN mandates that if vertical extent is specifi the verticalCRS xlink:href attribute value must be "http://www.spatialreference.org/ref/epsg/5714/" (replaces URN "urn:ogc:def:crs:EPSG::5714", this value may appear in older metac VerticalCRS with origin at World mean sea level (MSL), with elevati positive in meters(http://www.epsg-registry.org/). If vertical extent is the verticalCRS xlink:href attribute value must be an identifier for a w that can be dereferenced to get information on the borehole or soun

3.3 Service identification elements (SV_ServiceIdentification)

Table 4. Service Identification properties (SV_ServiceIdentification). See discussion of the use of MD_DataIdentification and SV_ServiceIdentification preceding Table 3.

ISO 19115 and 19119 (M/C/O) xPath from MD_Metadata	NAP-USGIN M/C/O	Comments on SV_ServiceIdentification
Resource service citation (M) `identificationInfo[1]/-SV_ServiceIdentification/-citation/CI_Citation`	M-M	The citation attribute provides information for citing the described serv scientific citation purposes, a citation for the intellectual content of the presented by the service would be found in the `MD_DataIdentification/` for datasets identified in the `operatesOn` section of `SV_ServiceIdentifica` defined by Webster as "an act of quoting". For USGIN purposes, this information to identify the intellectual origin or authority for the content resource, along the lines of a citation in a scientific journal. The purpo the service is to identify a particular service instance as a unique entit for a `CI_Citation` element are `title`, `date`, and `responsibleParty`.
Resource title (M) `identificationInfo[1]/-SV_ServiceIdentification/-citation/CI_Citation/title`	M-M	USGIN recommends that the title in a service identification citation sh the particular service instance, and inform the human reader about th function, and context.
Resource reference date (M) `identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/date/-CI_Date/date/`	M-M	The citation date for a service may indicate the creation date, when th became operational, the publication date, when the service first becar revision date, which specifies the date of most recent update. If the s online, a notAvailable or superseded date may be specified. The `Date` to specify the signficance of a reported date. This date is distinct from metadata record, or the `EX_Extent/temporalElement` that specifies the ti the resource content is applicable. The data type for the `date` element is xs:date, defined thus "date uses date/timeSevenPropertyModel, with *·hour·*, *·minute·*, and *·second·* rec *·timezoneOffset·* remains ·optional" (http://www.w3.org/TR/xmlschem encoding: 2000-12-12+13:00, 2006-10-01. If the month or day is not k '01', for example '2006-01-01'. ISO CI_DateTypeCode names that ap `{creation, publication, revision}`. See section 4.17.3 *Codelists* usage.
Unique resource identifier (O) `identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-identifier/MD_Identifier`	C-O	For USGIN, because the Citation is for the service, this identifier should be i `MD_Metadata/dataSetURI`, and is therefore optional. For USGIN purposes, this element content value is only an identifier for the for accessing the service. The USGIN profile requires the use of MD_Ide identify resources. RS_Identifier may substitute for MD_Identifier in th but the USGIN profile requires use of MD_Identifer. If additional codes content is associated with the identifier, it should be encoded as `MD_Identifier/authority/CI_Citation/alternateTitle` and `MD_Identifier/authority/CI_Citation/edition`
Resource responsible party (O)	M-M	USGIN requires at least one `CI_ResponsibleParty` with a rule that count organisationName + positionName) > 0. For a service, the point of co questions or reporting problems should be in

identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-citedResponsibleParty		SV_ServiceIdentification/pointOfContact/CI_ResponsibleParty. The serv responsible party should identify the parties responsible for creating (i publishing the service. ISO Role code names applicable to a service c {originator, principalInvestigator, processor, author, pu codelist values {resourceProvider, custodian, owner} should b SV_ServiceIdentification/pointOfContact element. See section 4.17.3 on codelist usage.
Resource presentation form (O) identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-presentationForm	O-O	The form in which the service is available, which in the case of a servi service implementation described by the metadata record, so the infor generally very useful. This profile specifies no convention for its use. N to the original source of intellectual content in the datasets operated o should be in MD_DataIdentification/citation/CI_Citation elements that srv:operatesOn associations. presentationForm uses the CI_PresentationFormCode codelist; ISO co applicable to a service citation include {documentDigital, imageD mapDigital, modelDigital, profileDigital, tableDigital, audioDigital}. NAP adds {multimediaDigital, diagramDigital *Codelists* for details on codelist usage.
Resource series (O) identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/series	O-O	Information about the series or collection of which the cited service is (name + issueIdentification) > 0. At this point there is not much preced services into a formal series, so in general this element is probably no services. This profile specifies no convention for its use.
Resource other citation details (O) identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-otherCitationDetails	O-O	Free text information useful to identify and cite the described service i specifies no convention for its use.
Resource collective title (O) identificationInfo/-SV_ServiceIdentification/-citation/CI_Citation/-collectiveTitle	O-O	Free text title of a "combined resource of which the service is a part." not much precedent for aggregating services into a collections, so in g probably not applicable to services. This profile specifies no conventio
Resource abstract (M) identificationInfo/-SV_ServiceIdentification/-abstract	M-M	A free text summary of the content, significance, purpose, scope, etc. described by this metadata. Exactly one value.
Resource purpose (O) identificationInfo/-SV_ServiceIdentification/-purpose	O-O	Text summary of the intentions for which the service was developed, i for creating the service and use cases it is designed to support. One v
Resource status (O) identificationInfo/-SV_ServiceIdentification/-status	M-M	Value is from MD_ProgressCode codelist. ISO Code names applicabl {completed, obsolete, onGoing, planned, required, underDev is synonymous with deprecated for this profile. See section 4.17.3 *Co* codelist usage.
Resource point of contact (O) identificationInfo/-SV_Servic	O-**O**^{R}	pointOfContact/CI_ResponsibleParty element for service metadata shou for a point of contact to report problems with the service. Element is o recommended! USGIN rule that count of (individualName + organisati

eIdentification/-pointOfContact		positionName) > 0. The CI_ResponsibleParty/role/CI_RoleCode@codeList CI_RoleCode; applicable name for the point of contact party are from {resourceProvider, custodian, owner}. Due to interoperability pr identifiers different from ISO identifiers for the same codelist elements use of ISO codelists. See section 4.17.3 *Codelists* for details on codel
Resource maintenance (O) identificationInfo/-SV_ServiceIdentification/-resourceMaintenance	O-O	This element provides information about the maintenance schedule or described by the metadata record. For a service, only one MD_Maintena elements may be included; for which the MD_ScopeDescription MD_Sc 'service'. If MD_MaintenanceInformation is present, then maintenanceAndUp mandatatory, populated by a MantenanceFrequencyCode; ISO name {continual, daily, weekly, fortnightly, monthly, quarterl annually, asNeeded, irregular, notPlanned, unknown}. See se for details on codelist usage. USGIN recommends following the NAP- practice that when SV_ServiceIdentification/status is set to "onGoing," e MD_MaintenanceInformation/dateOfNextUpdate or MD_MaintenanceInformati MaintenanceFrequency must be provided. Maintenance information specific to particular data the service presen in the dataset metadata for coupleResources associated with the serv
Graphic overview of resource (O) identificationInfo/-SV_ServiceIdentification/-graphicOverview	O-**O**[R]	Highly recommended to include a small image visual representation o provided by a map or image service. For geographic feature or data s overview might show the geographic distribution of available data. If M included, MD_BrowseGraphic/filename character string is mandatory. US practice is to provide a complete URL as a gco:characterString value property. Use napMD_FileFormatCode code values (http://www.fgdc.g register/codelists.html#IC_115) in fileType/CharacterString. See secti for details on encoding of the file format code, which is special becaus extension to the ISO base specification. Repeatable element; multiple values may present different resolutions resource. Names associated with overview should provide sufficient in distinguish these.
Resource format (O) identificationInfo/-SV_ServiceIdentification/-resourceFormat	O-X	The format of service response documents varies at the operation lev operation, different output formats may be requested. A listing of all po without bindings to the operations that respond with that format is not
Resource keywords (O) identificationInfo/-SV_ServiceIdentification/-descriptiveKeywords/MD_Keyword	O-O	Best Practice for USGIN profile metadata is to supply keywords to fac metadata records relevant to the user. **USGIN Keywords**: USGIN keyword vocabularies are in development this profile may include required keyword vocabularies. **Other Keywords**: Keyword Type - allowed ISO values from MD_Keyw {discipline, place, stratum, temporal, theme}. See section 4 details on codelist usage. MD_Keyword/keyword MUST contain a CharacterString element (see sect best practice is to include keywords in English.
Resource specific usage (O) identificationInfo/-SV_ServiceIdentification/-resourceSpec	O-X	Property not used by USGIN. Content may be ignored by clients.

ificUsage/		
Condition applying to access and use of resource (O) identificationInfo/-SV_ServiceIdentification/-resourceConstraints/	O-O	Information specifying restrictions on the access to and use of the des Constraints may be represented by MD_Constraint, MD_LegalConstraint, MD_SecurityConstraint. The attribute MD_Constraint/useLimitation is ma MD_LegalConstraint or MD_SecurityConstraint is provided. ISO19119 d as SV_ServiceIdentification/restrictions. Follow NAP specification th tification/resourceConstraints is to be used, and SV_ServiceIdentific not to be used. The metadataConstraints may allow a user to learn of th resource that they may not actually be able to access without further c
Aggregation information (O) identificationInfo/-SV_ServiceIdentification/-aggregationInfo/-MD_AggregateInformation	O-O	This element includes either a citation for or identifier of an associated along with the type of association, and optionally the activity that prod MD_AggregateInformation requires either aggregateDataSetName/CI_Citati SetIdentifier/MD_Identifier. associationType is mandatory, from DS_A ISO code names deemed applicable to service aggregation by this pr {crossReference, largerWorkCitation, source, }. See section details on codelist usage. Aggregation might be used to associate me layers with metadata for a service that provides a collection of layers (to cross reference services that can be chained, or to identify a resou to a service. If the related resource has an associated metadata record, USGIN rec is to include the identifier for that metadata record in aggregateDataSetIdentifier/MD_Identifier. For related resources that d metadata record, aggregateDataSetName/CI_Citation may be used; this aggregateDataSetIdentifier has a value. For USGIN profile, this property, rather than MD_Metadata/parentIdentif to indicate relationships between described resources. No testing has this element, or examples found for actual usage.
Resource service type (M) identificationInfo/-SV_ServiceIdentification/-serviceType	M-M	Exactly one value required. USGIN mandates use of a LocalName value (http://schemas.opengis.net/iso/19139/20060504/srv/serviceMetadata localName or ScopedName). There is not as yet a standard registry o identifiers that can serve as an authority for serviceTypes. An interim and identifiers is included in section 7.1 ServiceType (with the ad hoc 'http://resources.usgin.org/registry/serviceType201001'). Valid value are {WMS, WFS, WCS, CSW, …} Example: <srv:serviceType> <gco:LocalName codeSpace= "http://resources.usgin.org/registry/serviceType201001">WMS</gc </srv:serviceType>
Resource service type version (O) identificationInfo/-SV_ServiceIdentification/-serviceTypeVersion	O-C	Multiple serviceTypeVersion tags may not be implemented in some ha applications - USGIN recommends a reverse chronological order for s Constraint: if various versions are available, it is mandatory to list vers supported. Default is oldest version of service.
Resource service access properties (O) identificationInfo/-SV_ServiceIdentification/-accessProper	O-O	Optional MD_StandardOrderProcess element to provide information on the a service which include: fees, available date and time, ordering instructions, tu instructions and turnaround are not applicable to web services.

ties		
Resource service restrictions (O) identificationInfo/-SV_ServiceIdentification/-restrictions	O-X	Not used by USGIN; use resourceConstraints as per NAP.
Keywords (O) identificationInfo/-SV_ServiceIdentification/-keywords	O-X	Not used by USGIN; use descriptiveKeywords as per NAP
Resource service content extent (O) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent	C-C	Defines the spatial (horizontal and vertical) and temporal region to wh by the service apply. USGIN requires. Best Practice for USGIN is to in bounding box that specifies the extent to which resource content appl related to some geographic location. For geoscience resources, the te be expressed using time ordinal eras from a geologic time scale if the some particular geologic time. USGIN specifies count(description + geographicElement + temporalElem
Resource service content extent description ( ) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/description	C-C	Free text that describes the spatial and temporal extent of the dataset that description is mandatory if a geographicElement or temporalElement geographic place names are used to express the geographic extent th encoded using keywords with keyword type code = 'place'. Geograph duplicated in the EX_Extent/description.
Resource service content extent bounding box ( ) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/-geographicElement/-EX_GeographicBoundingBox	O-C	USGIN profile requires that if an EX_Extent/geographicElement is supplie geographic bounding box with bounding latitude and longitude expres decimal degrees. The corner coordinates for the geographic bounding box must not coi because this may result in fatal errors with some CSW implementation must thus be represented as tiny rectangles. USGIN recommended p actual point location in the lower left corner of the rectangle.
Resource service content extent geographic description ( ) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/geographic-Element/EX_Geographic-Description	C-X	Not used by USGIN profile, use keyword with type code = 'place'. Fo this information should be encoded using keywords, for which the MD = 'place'; the thesaurus/CI_Citation has the same content as EX_Geogr authority/CI_Citation, and the keyword is the same as the EX_Geograph
Resource service content extent bounding polygon ( ) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/-geographicElement/-EX_BoundingPolygon	C-X	To improve interoperability, USGIN mandates use of Geographic Bou polygons may be present, but may be ignored by harvesters.
Resource service temporal extent (O)	O-O	Information specifying the temporal extent to which resource is applica geoscience resources, this would be the geologic time period(s) to wh applies. Although the ISO19139 xml schema allows temporal extents to be

identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/temporalElement/EX_TemporalExtent/-extent/TimePeriod		ordered eras, TimePeriod MUST be used for temporal extent in USGIN p to make metadata interoperable. Values for beginPosition@frame and endPos populated if a temporal extent is provided. The default frame property value standard calendar date and time. For geologic time extents, USGIN requires beginPosition@frame and endPosition@frame to be populated using numeric measured positive increasing older with an origin at 1950 CE (see *Tempora* frame attribute value for geologic time coordinates is "urn:cgi:trs:CGI:Standa See section 4.21, below.
Resource service spatio-temporal extent (O) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/-temporalElement/-EX_SpatialTemporalExtent/	O-X	Although use of EX_SpatialTemporalExtent is allowed by ISO19139 and practice is to encode space time location with EX_TemporalExtent and E Box. Other optional extent elements may be included, but they may be implementations processing the metadata document.
Resource service vertical extent (O) identificationInfo/-SV_ServiceIdentification/-extent/EX_Extent/-verticalElement/-EX_VerticalExtent	O-O	Vertical extent is used to describe the elevation location of resources described service instance, if applicable. Common geoscience examp related to vertical location in a borehole or depth in a water body. The vertical coordinate reference system (CRS) within which described re using coordinates measured in linear distance from the collar (or grou bushing) of the borehole. EX_VerticalExtent has minimumValue, maximumValue that are real numbers For interoperability, USGIN mandates that if vertical extent is specifie verticalCRS xlink:href attribute value must be "http://www.spatialreference.org/ref/epsg/5714/"; this is the VerticalCR mean sea level (MSL), with elevations measured up positive in meters (http://www.epsg-registry.org/). If vertical extent is specified using dep xlink:href attribute value must be an identifier for a well or depth sound dereferenced to get information on the borehole or sounding geometry other information, reported depths should be assumed to be true verti the mean elevation of the geographic extent reported for the resource other information, reported depths should be assumed to be true verti the mean elevation of the geographic extent reported for the resource
Coupled Resource ( ) identificationInfo/-SV_ServiceIdentification/-coupledResource	O-O	This element correlates operations (identified by operationName) with d identifier). For logical consistenty, and SV_coupledResource/identifier equal to MD_DataIdentification/citation/CI_Citation/identifier/MD_Ide dataset that is the target of a SV_ServiceIdentification/operatesOn elem inline MD_DataIdentification/citation../code element, or a @uuidref att is necessary to implement the many-to-many relationship between da operations in a single service.
Coupled Resource operation name (M) identificationInfo/-SV_ServiceIdentification/-coupledResource/-SV_CoupledResource/-operationName	M-C	Mandatory if operations are specific to particular coupled resources. S service operation: GetMap, GetFeature, etc. There is no internal chec record that the given operation name is valid.
Coupled Resource identifier (M)	M-C	Mandatory in service metadata record if coupling type is tight or mixed tightly coupled dataset. Equal to MD_DataIdentification/citation/CI_Ci

identificationInfo/-SV_ServiceIdentification/-coupledResource/-SV_CoupledResource/identifier		identifier/MD_Identifier/code for a dataset that is the target of a SV_Se operatesOn element (either in an inline MD_DataIdentification/citation. @uuidref attribute).
Coupled Resource scoped name (X) identificationInfo/-SV_ServiceIdentification/-coupledResource/-SV_CoupledResource/ScopedName	X-O	OGC 07-045 application profile for ISO metadata using CSW 2.0.2 ex SV_CoupledResource with a ScopedName, defined as a scoped iden in the context of the given service instance (e.g. layer name or feature necessary for users to generate service requests (like GetMap or Get ISO service metadata.
Service coupling type (M) identificationInfo/-SV_ServiceIdentification/-couplingType	M-M	Type of coupling between service and associated data, if applicable. C by USGIN: • loose - Service instance is not associated with a specific datas datasetcollection. Loosely coupled services may have an asso types through the service type definition; no MD_DataIdentifica described. • mixed - Service instance is associated with a specific dataset but this can also be used with external data (i.e. data that is no operatesOn association). MD_DataIdentification describes the instance. • tight - Service instance only operates with specific data; MD_D element MUST be provided for the coupled data.
Service operations (M) identificationInfo/-SV_ServiceIdentification/-containsOperations	M-M	This element describes operations performed by the service, but the I includes insufficient detail to completely describe all parameters neces connection to a service. Widely used xml formats exist to describe ser including OGC getCapabilities.xml and W3C Web Service Description srv:containsOperations is a required element in a valid SV_ServiceIdent instance. At least one service operation MUST be described with the operationDescription/CharacterString = 'serviceDescription'. The identificationInfo/SV_ServiceIdentification/containsOperations/SV_Op connectpoint link for this operation MUST retrieve the service-specific document (e.g. WSDL, GetCapabilities, WADL).
Service operation name (M) identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/-operationName	M-M	Operation name following convention of the described service; manda associated with SV_OperationMetadata element with the operationDes String = 'serviceDescription'.
Service operation distributed computing platforms (M) identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/DCP	M-M	mandatory instance is associated with SV_OperationMetadata eleme operationDescription/CharacterString = 'serviceDescription'. USGIN p provision for use of this element.
Service operation description (O)	O-M	mandatory instance is associated with SV_OperationMetadata eleme operationDescription/CharacterString = 'serviceDescription'.

identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/-operationDescription		
Service operation invocation name (O) identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/-invocationName	O-O	service instance specific string to use in requests to invoke an operati
Service operation online resource (M) identificationInfo/-SV_ServiceIdentification/-containsOperations/-SV_OperationMetadata/-connectpoint	M-C	For loosely coupled services, operations and endpoints need to be de information endpoint (CI_OnlineResource) that returns a service self des MUST always be provded. The CI_OnlineResource/linkage URL should request if a single http GET request is sufficient. If the request require message package, or does not use standard web protocols, the CI_OnlineResource/description MUST provide instructions for executing
Service operates on (O) identificationInfo/-SV_ServiceIdentification/-operatesOn	O-C	"Provides information on the datasets that the service operates on" (IS If SV_ServiceIdentification/-couplingType is 'tight' or 'mixed', then ope must include a valid MD_DataIdentification element inline or by @uuidre explicitly links to an existing dataset metadata record that describes th The value of SV_ServiceIdentification/operatesOn@uuidref or SV_ServiceIdentification/operatesOn/MD_DataIden-tification/citation/CI_Citation/identifier/MD_Identifier/code must c the SV_ServiceIdentification/coupledResource/MD_CoupledResource/ident metadata record for the coupled dataset is a separate gmd:MD_Meta service described in the service metadata record should be identified the dataset. Explicitly linked reference example: <srv:operatesOn uuidref="13ce1e84-c887-4fd8-b888-8d021b1fa4c2" xlink:href=http://resources.azgs.org/geonetwork/srv/en/metada xlink:title="azgs:azgeochron"/>

4 Usage notes for Metadata Elements

This section presents additional information and discussion to supplement that in Table 2, Table 3, and Table 4.

4.1 Metadata file identifier

`MD_Metadata/fileIdentifier` is unique identifier for the metadata file. Some metadata profiles suggest that the metadata file identifier ID should be the same as the ID for the described resource. In the USGIN scheme, the metadata record is considered a separate resource, with a distinct identifier, from the resource it describes. The described resource identifier is the Unique resource identifier (`DatasetURI`, 4.8, below).

4.2 Metadata hierarchy

The ISO19115 specification (especially Annex H) discusses the use of metadata hierarchy, in which a resource may inherit metadata properties from parent metadata records in the hierarchy. For example a `dataset` in a dataset series might inherit all of the metadata content from the parent dataset series metadata record, except for dataset-specific data quality metadata. The linkage would be made through `MD_Metadata/parentIdentifier`. This kind of nesting, effectively database normalization, makes sense in a data management environment, but in a discovery environment, it present problems. Many queries would require joins with the parent metadata record, and software clients would be required to navigate the parent links to acquire 'inherited' properties from 'parent' records. For catalog service purposes, USGIN mandates that in metadata records returned by services, all inherited properties in such a hierarchy should be included explicitly in the metadata document, as opposed to implicitly through the `parentIdentifier` link. Internal document links may be used where allowed by the xml schema for identified elements repeated in a single response document. Resolvable xlink:href links to external resource are technically possible, but require that client software is xlink-aware, which is not always the case.

4.3 Metadata Contact vs. Resource Citation vs. Resource Contact

There are various locations to store contact information within an ISO 19139 metadata record. Here is a summary of the required contact properties and their significance as it pertains to the USGIN Profile.

- `MD_Metadata/contact/CI_ResponsibleParty` or "metadata point of contact" describes how to contact the party responsible for the **metadata** record to allow users to report errors, updates to metadata etc. The mandatory `CI_RoleCode` is set to "`pointOfContact`".
- `MD_Metadata/identificationInfo/[MD_DataIdentification || SV_ServiceIdentification]/-citation/CI_Citation/citedResponsibleParty/CI_ResponsiblerParty` provides information to identify the **intellectual origin** of the content in the described resource. This is straight forward when citing library resources (books, journals, etc.) but less clear when defining the intellectual origin of, for example, physical samples. The mandatory `CI_RoleCode` is set to one of the ISO codelist values {`custodian, owner, distributor, originator, pointOfContact, principalInvestigator, publisher, author`}.

- `MD_Metadata/identificationInfo/[MD_DataIdentification || SV_ServiceIdentification]/-pointOfContact/CI_ResponsibleParty` or "resource point of contact" contains information on who to contact to **access** the described resource. The mandatory `CI_RoleCode` is set to one of the ISO codelist values {`resourceProvider, custodian, owner, user, distributor, originator, pointOfContact, principalInvestigator, processor, publisher, author`}.

Additional contact information in the distribution section of the metadata provides point of contact for individual distribution processes to report problems with access to resources via different distributions.

4.4 Resource Title

Resource titles should provide sufficient information to distinguish the resource from other similar resources. They are not required to be globally unique, but bear in mind that users will be presented only with the resource title in typical brief search response documents. It is thus a disservice to have significant duplication of title strings or uninformative titles.

4.5 Resource Abstract

Ideally the resource abstract provides a succinct summary of the content of the resource, the purpose for which it was originally created, some indication of important quality parameters to help evaluate fitness for other purposes, any significant constraints on use of the resource, and a list of distribution options. After reading the abstract a user should have a good idea of whether the described resource contains the information they need and is accessible in a fashion that they can use.

4.6 Resource Type

The ISO 19115 `MD_Metadata/hierarchyLevel` property provides a high level categorization of resource types. The European INSPIRE Implementing Rules (MD_IR_and_ISO_20090218) proscribes the code list for the first hierarchyLevel xml element in an MD_Metadata document to be one of {`dataset, service, series`}, or the metadata set will be considered out of scope for the directive. Thus, metadata meant to be utilized by INSPIRE catalogs must follow this rule. The full ISO MD_ScopeCode list has a wider (and more useful) variety of resource categories; one or more hierarchyLevel elements using these codes could follow the first one with an INSPIRE-valid code in the first element to maintain INSPIRE compliance.

Table 1 in this document includes a more geoscience-domain-specific list of resource types, and values from this list should be used in one or more `hierarchyLevelName` elements. To enable resource-category-type searches to find narrower subcategories without complex query processing, `hierarchyLevelName` elements for the resource type and all broader/more general resource type categories should be included. The hierarchical categorization of the resources is encoded with the most specific category first, and progressively broader categories listed subsequently. Thus, harvesters that only take the first `hierarchyLevelName` element will get the most specific value. For example, if the resource is a photograph:

```
<gmd:hierarchyLevelName>
  <gco:CharacterString>Photograph</gco:CharacterString>
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
  <gco:CharacterString>StillImage</gco:CharacterString>
```

```
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
  <gco:CharacterString>Image</gco:CharacterString>
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
  <gco:CharacterString>Document</gco:CharacterString>
</gmd:hierarchyLevelName>
```

Note that the distinction of resource type and format is not always clear. Table 1 attempts to define resource types that are not specifically bound to a particular format, but are defined based on the kind of content. Format is interpreted as relating to specific approaches to encoding content and committing it to some sort of media.

4.7 Resource Locator

URL's for online access to resources are encoded in USGIN ISO 19139 metadata documents in the element `MD_Distribution/transferOptions/MD_DigitalTransferOptions/online/CI_Online-Resource`. Consistent use of this rule eliminates ambiguity on where to locate the URL to access a resource. Conventions for use of the `CI_OnlineResource` subelements `protocol`, `applicationProfile`, `name`, `description`, and `function` to enable metadata clients to reliably access referenced resources are discussed in section 4.15 CI_OnlineResource.

4.8 Unique Resource Identifier

The `MD_Metadata/DataSetURI` property should be a globally unique identifier for the described resource. The protocol used for this identifier is not specified by the USGIN Profile, but if it does not have a know resolution service, the capabilities document for a CSW service providing the metadata should have at least a text explanation of how to resolve URI's used by the service. Protocols with available resolvers include http (use the WWW DNS system) and doi (http://dx.doi.org/). Some authorities using urn: protocols are also implementing or have resolver services in place.

4.9 Browse Graphics

Links to browse graphics are useful to assist users evaluate resources that have a graphical aspect. The `MD_BrowseGraphic` element includes a fileType attribute that is important to inform client software accessing and displaying the graphic. For the USGIN profile, use standard MIME types in the `gmd:fileType/gco:CharacterString`. See the official Internet Assigned Numbers Authority (IANA) Mime Media Type registry at http://www.iana.org/assignments/media-types for a list of registered MIME types and their scope.

```
<gmd:MD_BrowseGraphic>
  <gmd:fileName>
        <gco:CharacterString>http://publicdocs.mnr.gov.on.ca/View.asp?-
        Document_ID=9632&Attachment_ID=18204</gco:CharacterString>
  </gmd:fileName>
  <gmd:fileDescription>
      <gco:CharacterString>Base Map from OMNR</gco:CharacterString>
  </gmd:fileDescription>
```

```
  <gmd:fileType>                <!-- This is a MIME Type -->
      <gco:CharacterString>image/jpg</gco:CharacterString>
  </gmd:fileType>
</gmd:MD_BrowseGraphic>
```

Code example 1. Encoding url, display name and file type for browse graphic.

4.10 Resolution and equivalentScale

For spatial datasets, some indication of the resolution of the data is very useful for evaluating fitness for use. From a data perspective, resolution in a spatial dataset specifies the the smallest resolvable distance between two points in the Earth for that dataset. For a grid or coverage, this would be the average distance between sample points. From data portrayal perspective, an `equivalentScale` is reported, representing the scale at which the portrayal was intended to be viewed when it was constructed. If the spatial representation type code is `vector, grid, textTable` or `tin`, an `equivalentScale/../denominator` MUST be provided to express spatial resolution. This is an arbitrary convention adopted to facilitate interoperablity. If a `equivalentScale/../distance` is available, that should be supplied as well. To calculate `equivalentScale` given a resolution distance, recommended practice is to divide the resolution distance in meters by 0.0005. This assumes that the smallest visually resolvable distance in a map display for human usage is 0.5 mm.

4.11 Resource Language

USGIN metadata is assumed to use American English and metadata documents should be returned in that language. The ISO 19139 XML includes elements and provisions to implement multiple languages in a single metadata package using `PT_Text` and `LocalizedCharacterString`, but in order to avoid complexity with USGIN recommended practice is to implement metadata search for different languages as different services. Each service would provide `CharacterStrings` in the same, single language, specified by the `MD_Metadata/language` element in returned metadata records.

4.12 Encoding of Vertical Extents

`EX_VerticalExtent` has `minimumValue`, `maximumValue` that are real numbers, and a `verticalCRS`. The vertical coordinate reference system (CRS) specifies how the bounds of the vertical extent are georeferenced. The USGIN profile accepts two approaches. Elevation is reported with reference to Earth mean sea level, measured positive upward. On the other hand, subsurface or subaqueous data are typically referenced relative to a borehole trace or sounding, measured positive downward from a local datum at the borehole or sounding orgin. The origin is typically a borehole collar at the ground surface, the Kelly bushing or drill floor on a drilling rig, or a ship deck., For to facilitate interoperability.

For interoperability, USGIN mandates that both elevation and depth vertical extent bounds MUST be specified in meters. In addition, if vertical extent is specified as an elevation, the `verticalCRS` xlink:href attribute value MUST be "http://www.spatialreference.org/ref/epsg/5714/"; this is the VerticalCRS with origin at World mean sea level (MSL), with elevations measured up positive in meters (http://www.epsg-registry.org/). If vertical extent is specified using depth, the `verticalCRS` xlink:href attribute value must be an identifier for a well or depth sounding that can be dereferenced to get information on the borehole or sounding geometry.

4.13 Content information: Entities and Attributes

One of the major deficiencies of the ISO 19115/19139 specifications is the absence of model elements to describe the entities and attributes in standard tabular datasets. Description of dataset schema using the ISO Technical Committee 211 (TC211) framework requires use of ISO19110. A standard XML schema for implementing this specification has recently become available, and entity-attribute metadata for dataset should be documented using the schemas located at http://www.isotc211.org/2005/gfc/gfc.xsd. The ISO19139 XML implementation of ISO19115 does not allow inclusion of the feature catalog inside the metadata record, but provides a citation element in the `gmd:contentInfo/gmd:MD_FeatureCatalogueDescription` element. The USGIN profile provision is to link to the feature catalog using an xlink:href attribute on the `gmd:featureCatalogueCitation`, identified as such by an accompanying xlink:title attribute with the value "Link to ISO19110 feature catalog". Code Example 2 shows recommended encoding. Section 8.4, ISO19110 Feature Catalog example, provides an example encoding of entity-attribute information for a feature dataset.

```
<gmd:contentInfo>
  <gmd:MD_FeatureCatalogueDescription>
    <gmd:includedWithDataset>
     <!-- true if a FeatureCatalog document is packaged with the dataset -->
       <gco:Boolean>false</gco:Boolean>
    </gmd:includedWithDataset>
   <!-- a dataset may contain multiple feature types. For an MS Access type database, each
   table    could be considered a featue type -->
        <gmd:featureTypes>
   <!-- the codespace is the namespace URI, should dereference to obtain XML schema (if
   applicable)-->
          <gco:LocalName codeSpace="
   http://xmlns.geosciml.org/geosciml-portrayal/2.0">GeologicContact_2</gco:LocalName>
        </gmd:featureTypes>
   <!-- follow pattern used by Stanford library http://opengeoportal.org/archives/category/-
   uncategorized); href is to an ISO19110 feature catalog,
   with featuretype identified by fragment ID -->
        <gmd:featureCatalogueCitation xlink:href="
   http://schemas.usgin.org/featurecatalog/12.xml#geologiccontact"
          xlink:title="Link to ISO19110 feature catalog"/>
   </gmd:MD_FeatureCatalogueDescription>
  </gmd:contentInfo>
```

Code Example 2. Encoding of feature catalog link for entity attribute description of datasets. See 8.4 ISO19110 Feature Catalog example for and example instance document.

4.14 Use of MD_Distribution and MD_Distributor

The ISO19115 model provides two possible paths for specifying information about how a user can access the resource. The `MD_Distribution` element may have 0 to many `distributionFormat`, `distributor`, and `transferOptions` child elements (see Figure 1). On the other hand, each of the `distributor` child elements may have 0 to many `distributorFormat` and `distributorTransferOption` elements. Several major existing applications that consume ISO19139 xml metadata files are configured to expect format and transfer option information to be at the `MD_Distribution/-distributionFormat` and `MD_Distribution/transferOptions` path. This works fine as long as there are not different format or transfer options from different distributors, or different transfer options for different formats. In these cases, a binding between distributor, format, and transfer options

necessitates use of the MD_Distribution/distributor/MD_Distributor path to distributorFormat and distributorTransferOptions (and distributionOrderProcess) information that works together. In order to accommodate both existing applications that utilize content in the MD_Distribution/-distributionFormat and MD_Distribution/transferOptions elements, and situations that require binding between distributor, order process, format, and transfer options, the USGIN profile mandates that if multiple MD_Distribution/distributionFormat or MD_Distribution/transferOptions elements are included in a document, all formats must be available via all the specified transfer options, and the content of these elements should be included in line. If multiple MD_Distribution/distributor elements are present, without child MD_Distributor/distributorFormat or MD_Distributor/distributorTransferOptions elements, then all formats and transfer options are available from all distributors.
To specify different bindings between distributor, order process, format, and transfer options, a separate MD_Distribution/distributor/MD_Distributor instance is included for each binding. At least one MD_Distribution/distributionFormat and one MD_Distribution/transferOptions element MUST be included for applications that expect content in these elements, and the format and transfer options specified by these elements should apply to the first distributor/MD_Distributor element. Repeated CI_ResponsibleParty, MD_StandardOrderProcess, MD_Format or MD_DigitalTransferOption elements in the distributor/MD_Distributor elements should be specified by reference (xlink:href to gml:id of first occurrence of the element within the document). The implication is that the distributionOrderProcess/MD_StandardOrderProcess, distributorFormat/MD_Format, and distributorTransferOptions/MD_DigitalTransferOptions child elements of a single MD_Distributor are all compatible with each other. Following this convention will produce schema valid metadata, and the distributor elements after the first can be ignored by applications expecting only a single distributor element.

In summary, for the USGIN profile, each distributor/MD_Distributor is a binding between one or more transfer options and the distributor formats that are available through that/those transfer options (MD_DigitalTransferOptions/onLine/CI_OnlineResource in particular). If different formats are available from the same distributor, but have different transfer options, these should be represented as different distributor/MD_Distributor instances.

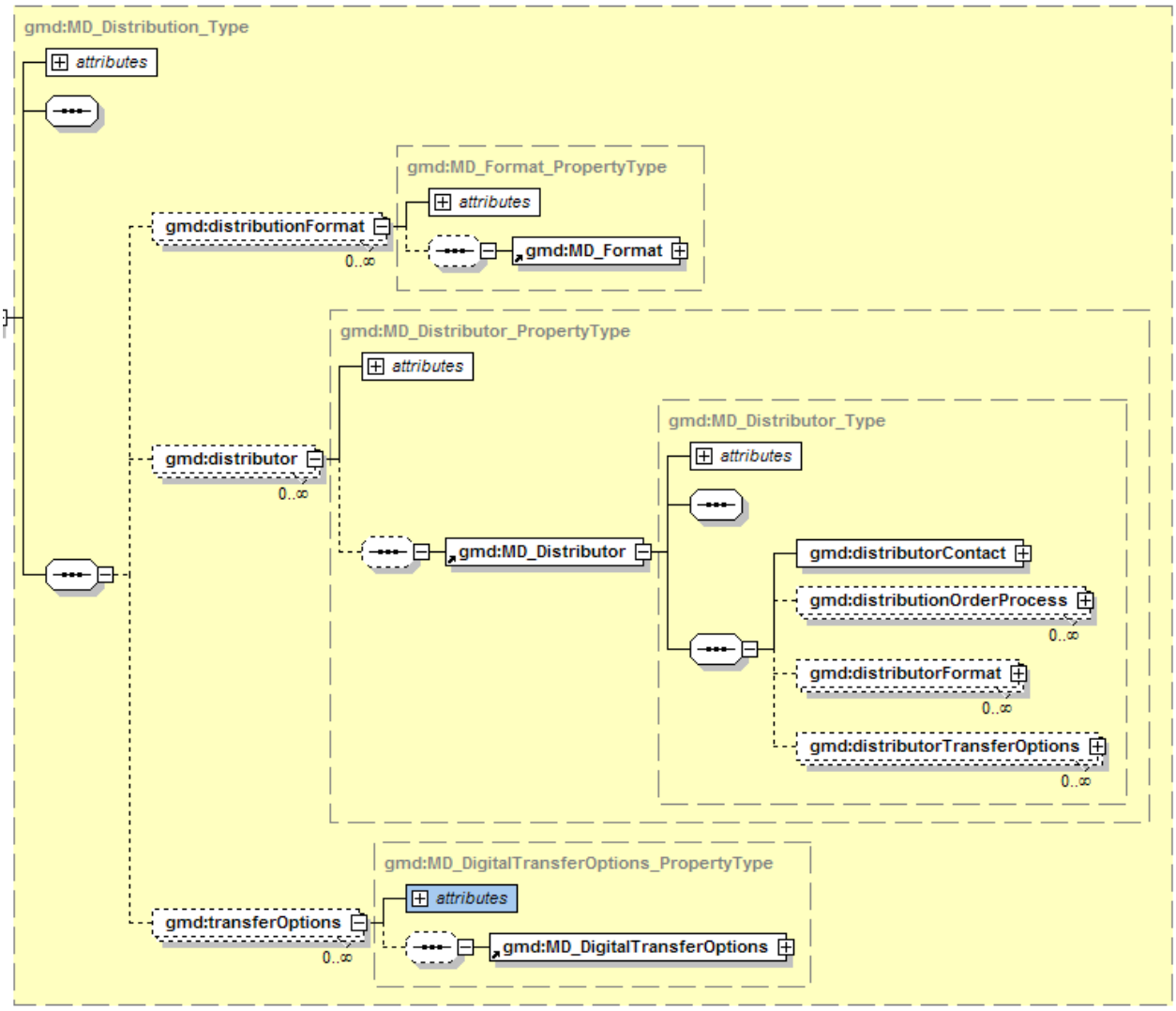


Figure 1. gmd:MD_Distribution_Type diagram

4.15 Distribution Format

If the resource distribution is a physical object, like a book, rock sample, paper document, the `distribution/../MD_Format/name` MUST be a term from the USGIN distribution format codelist (see Table 6). Note that format is partially orthogonal from resource type (Table 1). A document may be available in various digital (pdf, tiff, doc, txt) or non-digital (book, loose sheets) formats.

4.15.1 Digital resources

The format vocabulary needs to be designed to work in the framework of the `distribution/../-MD_DigitalTransferOptions`, which provides protocol, applicationProfile, name, and function subelements for online resources (CI_OnlineResource), and `MD_MediumNameCode` and `MD_Medium-FormatCode` for offline electronic resources (`MD_Medium`). For digital resources it provides terms to specify file-format information that does not have any other obvious home. Examples in INCITS 453, INSPIRE 19115/19, and ANZLIC 2007 populate `MD_Format/name` with values like 'ESRI ARC/INFO Coverage', 'ESRI shapefile', 'ESRI ARC/INFO Export e00', and 'MapInfo MID/MIF' all pertain to digital resources. If a MIME format (http://www.iana.org/assignments/media-types/)

is defined for a digital file format, the MIME media-type code should be used. If no appropriate MIME type is registered with IANA, USGIN mandates that the distribution format for digital resources should specify the file format using a pattern that includes vendor, application name, and file extension.
Pattern for digital resources: `[vendor:applicationName]/fileExtension`. The vendor and application names may not be applicable, and could be omitted, but the '/' and file extension should always be present. If the format consists of a single file, the file extension is a three letter file-type abbreviation assigned by the vendor. If the format consists of a package of files (e.g. an ArcGIS file geodatabase), the file extension is a name that in most cases should be obvious from vendor usage. The accompanying `MD_Format/version` value should indicate the version of application software if the format is specific to some version.
Service metadata describes a service as a resource, in which case distribution is effectively equivalent to the set of operations accessible at the service instance endpoint. The ISO 19119 service metadata model does not provide a clear way to associate individual operation responses with possible output formats for those responses. If all output formats are applicable to all service requests, they can be listed in `distributioninfo` as a collection of `distributionFormat/MD_Format` elements, but this is not a general solution. Another approach (e.g. GeoNetwork OpenSource, v. 2.4) to documenting output formats is to put format information in a collection of `srv:connectPoint/CI_OnlineResource/protocol` elements, with `connectPoint` elements for each format available on each operation. The USGIN profile convention is that operation metadata is best conveyed to metadata consumers by providing a link to a service-specific description file (getCapabilities or WSDL, see 4.22 Operation metadata), and this MUST be provided in a service metadata record by a `SV_OperationMetadata` element with an `operationDescription` value of 'serviceDescription'. The output formats offered by the service should be listed in a `distributionInfo` element as a collection of `distributionFormat/MD_Format` elements if all formats are applicable to all service requests, or if the mapping between requests and formats is obvious. Encoding of the format name should use whatever convention is used by the service to specify that output format in requests made to the service.
Table 5. Example format strings for digital files. These are to be used only if an appropriate MIME type is not defined.

ESRI:ARCINFO/Coverage
/shapefile
ESRI:ARCINFO/e00
PitneyBowes:MapInfo/mid
ESRI:ArcGIS/mdb
ESRI:ArcGIS/fileGeodatabase
Microsoft:Access/mdb

4.15.2 Non digital resources

The `MD_Format` element is the only format information for resources that do not have digital transfer options, and USGIN proposes Table 6 as a vocabulary for use to specify format of non-digital resources. Although this codelist could be implemented as a schema extension, for the time being we propose to use it as a controlled vocabulary specified by profile and practice, rather than schema. Use of such controlled vocabulary can be indicated by using xsi:type on the gco:characterString element to make the type gml:CodeType, which then requires a codeSpace

attribute. The distribution format Identifier from Table 6 should be used as the element value. Example encoding:

```
<gmd:MD_Format>
  <gmd:name>
      <gco:CharacterString xsi:type=”gml:CodeType”
    codeSpace="http://resources.usgin.org/registry/distributionFormatNames201001">
    sample:core</gco:CharacterString>
  </gmd:name>
  <gmd:version nilReason="notapplicable"/>
<gmd:MD_Format>
```

Table 6. USGIN Distribution formats for non digital resources. URI for this codelist is http://resources.usgin.org/registry/distributionFormatNames201001

Identifier	Name	Parent format	Scope
physicalArtifact	Physical artifact		described resource is a physical object
sample	Sample	physicalArtifact	Use for uncategorized sample.
sample:core	Core	sample	Cylindrical rock sample extracted from Earth with a coring drill
sample:cuttings	Cuttings	sample	Small rock fragments recovered from drilling process as sample of material being drilled
sample:fluid	Fluid	sample	Sample of a fluid
sample:handSample	Hand sample	sample	Single piece or pieces of material.
hardCopy	page-size document	physicalArtifact	A physical copy of a document on paper, film, or other similar material.
hardCopy:book	Book	hardcopy	Manuscript printed on paper, bound into a single volume
hardCopy:manuscript	Manuscript	hardCopy	Other printed or written representation on physical media, usually paper or mylar, includes unbound books, index cards, loose notes, file folders of papers
hardCopy:printedImage	Printed image	hardCopy	Image on paper or other opaque or semi-opaque media.

printedImage:paperMap	oversize sheet	printedImage	Map image on a single sheet
hardCopy:filmImage	photographic film	hardCopy	Image on film, viewed by passing light through the film. Includes single still images and collections of connected images for a movie.
audioRecording:tape	magnetic tape	physicalArti-fact	use for sound resources that are recorded on magnetic tape.
audioRecording:otherMedia	vinyl disk, wax cylinders	physicalArtifact	

4.16 CI_OnlineResource

The CI_OnlineResource element is used in a variety of contexts in ISO19115 content (see Table 7). This complex-content-element provides a linkage to some online resource related to the metadata or the described context resource. It is thus analogous to atom:link (IETF RFC-4287) or the link element described in IETF RFC-5988. For links that locate online documents accessible using standard browser and file type resolution technology, the link can be as simple as a single URL element that retrieves a representation of the resource. There are many other kinds of related resources that the onlineResource element may point to, including web services that provide access to a dataset resource, metadata for related or source resources, specifications for standards or extensions to standards.

Table 7. Contexts for CI_OnlineResource elements

Context	usage
contactInfo/onlineResource	Link to online resource to assist contacting a referenced individual or organization.
distributionInfo/transferOptions	link to a representation of the resource that is the subject of the metadata record
metadataExtension	link to document describing specifications for elements and attributes extending the base standard.

For the USGIN, it is desired that such links are accompanied by sufficient description that the linked resource can be accessed and provided to the metadata user automatically, with little or no operator intervention other than clicking on a link in a presentation of the metadata. A content specification for such machine-actionable links is discussed in a separate USGIN discussion paper. This approach essentially considers the metadata as a sort of hypermedia in which the CI_OnlineResource elements define the 'affordances', or actions that are available to a client.

The CI_OnlineResource element contains 6 child elements. The only required content is a URL that will access the online resource. All other properties are optional. The `protocol`, `applicationProfile`, `name` and `description` properties are all free text with loosely defined semantics as indicated by the element names.

4.16.1 CI_OnlineResource protocol

Since the base level web locator scheme (ftp, http) is indicated by the prefix of the URL (IETF RFC 1738), USGIN metadata uses the `protocol` attribute to specify a higher-level protocol in the network stack, e.g. a serviceType for data accessed through remote-procedure-call-on-HTTP type services.

4.16.2 CI_OnlineResource applicationProfile

The use of `applicationProfile` for file-based resources accessed via URL is discussed in Table 2. For links to documents for which the service type (encoded in the `CI_OnlineResource/protocol` property) and base protocol prefix in the `CI_OnlineResource/link` URL (e.g. 'http:') do not provide sufficient information to guide client software, the `applicationProfile` property is used to indicate a profile on the serviceType or some variation in document encoding or content conventions. For example WFS services may offer different features in different namespaces or encoding schemes, a catalog may offer different metadata encoding, or a resource-oriented service may offer representations using different encoding schemes. The same scheme may be used with different conventions, for instance different profiles for the use of ISO19139 or csw:record XML metadata encoding. RDF representations may be offered in XML, Turtle, or N3 encoding. Although these variations may be deducible using content negotiation or by accessing and parsing a service description document, much simple client logic is possible if the information is provided up front with the link. The actual string value that should be provided to specify an application profile should be defined by the agent originating the profile.

4.16.3 CI_OnlineResource name and description

In order to provide further clarity and guidance for the utilization of linked resources USGIN recommends use of the `CI_OnlineResource/name` property as indicated in Table 13. In order to identify the linkage element that locates the service description document (see 4.22 Operation metadata), USGIN mandates using `CI_OnlineResource/name` = "serviceDescription" in the `CI_OnlineResource` element with the linkage to the service description. The assumption is that any client that can connect to the service will know what to expect as the service description provided by that service.

The `description` property may be used to provide information about the online resource, and more usefully, to provide an explanation of how the other content of the `CI_OnlineResource` element is to be used to access the resource. The `linkage/description` should provide any necessary additional information in a text narrative as well. In some cases the protocol, applicationProfile and name properties may be insufficient to enable machine access to the resource through the provided link. The `CI_OnlineResource/description` element may include key values pairs to provide additional necessary information. USGIN recommendataion is to encode these in a 'parameter' object using JSON syntax. The parameters value is a list of key:value pairs enclosed in curly brackets ('{"key":"value", "key1":"value1"…}'). The keys should be the exact string that is required for the data access request parameter. For example, a dataset distributed through a particular layer in a multi-layer WMS:

```
<gmd:description>
  <gco:CharacterString>
    Whatever descriptive text you want.
    parameters:{"layers":"gtp_datagap_well_data_collection",
                "styles":"wellsDataSLD"}
  </gco:CharacterString>
```

```
</gmd:description>
```

The above example also shows how to encode multiple parameters.

In the case of a dataset distributed through a particular feature type in a multi-feature WFS:

```
<gmd:description>
  <gco:CharacterString>
    Whatever descriptive text you want.
    parameters:{"typeName":"BoreholeLithInterval2.0"}
  </gco:CharacterString>
</gmd:description>
```

In some cases, a WFS may offer multiple feature types based on the same base dataset, for example a geologic map service that offers a separate contacts and faults feature types from the same geologic lines dataset. In this case, the typeName values will be a comma-delimited list:

```
<gmd:description>
  <gco:CharacterString>
    Whatever descriptive text you want.
    parameters:{"typeName":"contacts,faults"}
  </gco:CharacterString>
</gmd:description>
```

4.16.4 CI_OnlineResource function property

The `function` property is populated from a codelist (see Table 8) and is used to indicate the expected action that actuating the link will trigger. (see section 4.17.3 *Codelists*).

Table 8. OnlineFunctionCode values from NAP (INCITS 453) and ISO 19115. '(ISO)' after the text in column 2 indicates function codes from ISO19115 (2006).

OnLineFunctionCode	USGIN profile usage
browsing	`CI_OnlineResource/linkage` is a valid URL for a web application that enables user to explore and seek information about the resource from a Web browser
download	`CI_OnlineResource/linkage` is a valid URL that will initiate transfer of a file containing the described resource to the local system. (ISO)
emailService	`CI_OnlineResource/linkage` is a valid mailto: link that will initiate an e-mail message to the correct party to request access to the described resource. (NAP)
information	`CI_OnlineResource/linkage` is a valid URL that will access a resource that provides information about the resource content, for example an explanatory web page, a downloadable document with narrative text describing the resource, or an XML encoded metadata document. (ISO)
offlineAccess	`CI_OnlineResource/linkage` is a valid URL that will access a web page providing instructions for requesting the resource from the provider. (ISO)

order	CI_OnlineResource/linkage is a valid URL that will access a web page to initiate an ordering process for obtaining the resource. (ISO)
search	CI_OnlineResource/linkage is a valid URL that will access a search interface for seeking out specific information content contained by resource, e.g. the metadata describes a database, and this linkage accesses a search interface to search the database (ISO)
upload	CI_OnlineResource/linkage is a valid URL for a web interface to transfer data from a local storage device or system to be included in the described resource. (NAP)
webService	CI_OnlineResource/linkage is a URL that accesses a standard web service. If the CI_OnlineResource/name is 'ServiceDescription' then the link is to a machine-readable standard service description document as defined by the service specification. Example description document may be a Web Services Description Language (WSDL) file or OGC getCapabilities file. (NAP)

4.16.5 Scoped name in service distributions

A distribution option for a dataset might be through a Web Map Service that contains one or more layers portraying the described dataset, along with layers portraying other datesets. Likewise, a dataset may be distributed as a particular feature type in a Web Feature Service that offers multiple features. Analogous internal knowledge may be required for various other service protocols. To completely specify access to the correct data representation requires knowing the layer name, feature type name, or some other information in addition to the base service connection information provided by the existing properties. USGIN recommended practice for this common situation is to include a parameters section in the linkage/description (see 4.16.3 CI_OnlineResource name and description, above).

4.17 Responsible parties and logos

Metadata should include a URL that locates a thumbnail logo for organizations related to the metadata origination, the organization hosting the catalog that returned the metadata, the organization that originated the data, and the organization hosting online services that provide access to the data. The standard place to put URL's in ISO19139 metadata is in the CI_Contact/onlineResource/CI_OnlineResource/linkage attribute. For URL's that indicate icon thumbnails, the CI_OnlineResource/name should be 'icon'.

The metadata originator information should be in a MD_Metadata/contact/CI_ResponsibleParty element with role code 'originator' to identify the original source of the metadata record, for which the CI_Contact/../CI_OnlineResource/linkage is a URL that points to an Icon for the metadata originator. This Icon will be displayed in search results to credit the metadata originator. Metadata harvesters should harvest and maintain this information so that the origin of metadata records can be credited.

The organization hosting the catalog that returned the metadata record should specified in a MD_Metadata/contact/CI_ResponsibleParty element with role code 'distributor', for which the CI_Contact/../CI_OnlineResource/linkage is a URL that points to an Icon for the metadata server hosting organization. This information need not be harvested, because it will be replaced by information describing the harvesting catalog service.

The organization that originated the data is specified by MD_Metadata/identificationInfo/MD_Data-

Identification/citation/../CI_ResponsibleParty with RoleCode ='originator', and /CI_OnlineResource/name='icon'. This will distinguish the citation responsible party element containing the icon linkage from CI_ResponsibleParty elements with RoleCode='author' or 'editor', which would provide an online linkage directly to the responsible party as specified by CI_OnlineResource protocol, applicationProfile, name, function, and description elements.

The organization hosting a service providing online access to described data is specified by MD_Metadata/distributionInfo/MD_Distribution/distributor/MD_Distributor/distributorContact/-CI_ResponsibleParty with RoleCode ='resourceProvider' or 'distributor', and ../CI_Online-Resource/name='icon'. Because the cardinality of distributorContact responsible party and online resources is 1, only one linkage can be provided for a distributor, and the metadata author must decide whether that will be a link to an icon, or a link to a web site or other resource related to the distributor.

```
<gmd:contact>
  <gmd:CI_ResponsibleParty>
    <gmd:organisationName>
      <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
    </gmd:organisationName>
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:onlineResource>
          <gmd:CI_OnlineResource>
            <gmd:linkage>
              <gmd:URL>http://www.azgs.az.gov/logo/metadata/azgs.png</gmd:URL>
            </gmd:linkage>
            <gmd:name>
              <gco:CharacterString>icon</gco:CharacterString>
            </gmd:name>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
      </gmd:CI_Contact>
    </gmd:contactInfo>
    <gmd:role>
      <gmd:CI_RoleCode
     codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_RoleCod
     e"
        codeListValue="originator">originator</gmd:CI_RoleCode>
    </gmd:role>
  </gmd:CI_ResponsibleParty>
</gmd:contact>
```

4.18 Extensions to CharacterString

4.18.1 Web extensions

ISO 19139 defines several extensions to gco:CharacterString in the gmx namespace. These are defined as members of an xml substitution group for gco:CharacterString, and include gmx:Anchor, gmx:FileName, and gmx:MimeFileType. gmx:Anchor is used for URL's linking to online web resources, and include a URI attribute associated with the character string that is the human-readable label for the link. gmx:FileName adds a filename URI attribute that specifies a

machine-readable absolute path to the location of the file, the human readable file name specified by the character string. `gmx:MimeFileType` adds a MIME type/subtype attribute to a character string that specifies a human readable file type. The gmx namespace is not imported into other ISO19139 schema in the normative schema. In order to create schema-valid documents that use these extensions, explicit namespace-declaration must be made to the gmx schema in instance documents. For the USGIN profile, these extension classes are not recommended, and may be ignored by client software.

4.18.2 Language localization

ISO 19139 defines an extension to `gco:CharacterString` that allows substitution by `PT_FreeText` or `LocalisedCharacterString`. `LocalisedCharacterString` adds a `locale/PT_Locale` property to the `CharacterSTring` element that can specify the language, country, and character encoding for the string. `PT_FreeText` allow substitution of a collection of `LocalisedCharacterString` elements for any `CharacterString`, each localized to a different language/country.

These various possibilities create potential to break interoperability. To avoid this problem, and to avoid complexity with `PT_Text` and `LocalizedCharacterString`, USGIN profile practice is to implement services for different languages as different catalog service instances, each of which serves CharacterStrings in a single language specified by the `MD_Metadata/language` element.

4.18.3 Codelists

ISO 19139 defines a "CodeListValue_Type" XML Class Type with three attributes:

`codeList`: contains a URL that references a codeList definition within a registry or a codelist catalogue. As currently used in the metadata services we have studied, the codeList is not used to identify a vocabulary; rather it provides a locator (functionally equivalent to xlink:href) for an online resource, typically a web page or xml file, that contains a listing of the codelist with the code values and scope notes. Different services provide different URL's, possibly linking to different kinds of resources (e.g. web page or xml file), for the same codelist. Thus, in practice, the values in this attribute can not be used for automated determination of the code list in use in a metadata document.

`codeListValue`: carries the value from the 'name' column of the codelist tables in ISO 19115, Annex B. This value serves to identify the codelist meaning in the context of the codelist catalogue (or registry) located by the `codeList` attribute. The codelist catalogue is expected to contain an explicit name and definition of the value in the default language of the metadata, as well as alternate expressions in different code spaces, some of them corresponding to the different locales supported by the metadata.

The `codeSpace` attribute is an optional identifier (URI); when present it refers to an alternative expression of the codelist value definition. In the example in ISO19139, section 8.5.5.1 (p. 30), the codeSpace URI for the domain code is the string "domainCode", and the value from the domainCode column in a codelist definition table in ISO 19115, Annex B is included as the value of the xml CodeList element in this case. This appears to be an attempt to allow alternate codeListValue strings to be associated with the codelist value concept for language localization or use of alternate identifiers.

Codelist elements in the ISO19139 XML schema are assigned to type CodeListValue_Type, and also included in a substitution group for `gco:CharacterString`. These codeList elements are thus substitutable for elements typed `gco:CharacterString`. Consequently, any CodeList instance

is an XML element that takes a string value (corresponding to the element value for a gco:CharacterString), and has three XML attributes defined by the CodeListValue_Type XML Class Type. A corresponding XML Class Property Type is defined for each of these CodeList elements, and this property type is used to restrict the values in XML `CharacterString` attributes to the code list.

The ISO specification uses an unfortunate choice of name for the 'codeListValue' attribute that is defined to be a identifier, apparently with the intention that it is a language-neutral concept identifier that might be associated with various language-localized labels for the concept. NAP CodeList registries (http://www.fgdc.gov/nap/metadata/register) contrast with the codelists defined in the tables in ISO 19115 Annex B in that the identifier (the 'name' column the ISO19115 Annex B tables) is an integer identifier with the prefix 'RI_'. This would appear to correspond functionally to the 'domainCode' values in the ISO19115 Annex B tables, which ISO19139 indicates should be the codeListValue when the `codeSpace`="domainCode".
NAP and INSPIRE usage is consistent with the ISO19139 definition of codeListValue as an identifier, with the name or label for the codeList concept included as the value of the CodeListValue attribute. The 'name' column in ISO 19115, Annex B tables, which is described as the content for the codeListValue by ISO19139, contains English words that are the same as the labels one would use in English. In the CT_CodeListCatalogues in the ISO publicly available standards registry for ISO 19139 (http://standards.iso.org/ittf/PubliclyAvailableStandards/-ISO_19139_Schemas/resources/), the CodeListDictionary/codeEntry/CodeDefinition elements only include gml:description and gml:identifier elements, but no gml:name elements. Based on this ISO guidance, USGIN profile mandates that controlled vocabulary terms from codelists are encoded thus:

Example 1:

```
<gmd:CI_DateTypeCode codeList=" http://standards.iso.org/ittf/PubliclyAvailableStandards/-
    ISO_19139_Schemas/resources/#CI_DateTypeCode"
    codeListValue="creation"/>
```

Example 2

```
<gmd:MD_CharacterSetCode
    codeList="
    http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_DateTypeCode"
    codeListValue="utf8"/>
```

Or

```
<gmd:MD_CharacterSetCode
    codeList="
    http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_DateTypeCode"
    codeListValue="utf8">UTF-8</gmd:MD_CharacterSetCode>
```

The alternate version of example 2 includes label string for the codelist concept as the element value string. Most applications appear to use only the codeListValue string for querying and content inspection. In order to indicate that standard ISO codelists are being used, the `codeList` attribute value MUST be " http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml ", which is the web location for the authoritative codelist dictionary XML document. Including the fragment identifier name for the particular codelist is informative only; this is the terminal token of the URL after the '#' character, and will be ignored by browsers.

The use of other codelists is implemented by providing a `codeList` attribute that points to the codelist actually used. The following example shows use of a DateTypeCode from the code list

in the North American Profile:

```
<gmd:CI_DateTypeCode
        codeList="http://www.fgdc.gov/nap/metadata/register/registerItemClasses.html#IC_87"
        codeListValue="superseded">superseded</gmd:CI_DateTypeCode>
```

Note that the ISO codelists use the name value for the `codeListValue` attribute. This is consistent with ISO codelist practice but is in contrast to the NAP codeList example encoding in Appendix E of NAP profile document (INCITS 453, 2009). In these examples the codeListValue is the 'Item identifier' from the NAP registry specified by the codeList, with the prefix 'RI_' added.

The ISO codelists are in much wider use at this time than the NAP codelists (as far as we can tell from surveying existing services), but we recognize that some of the terms added in the NAP codelists may be required for metadata describing some of the resources in the USGIN scope (Table 1). Table 8 summarizes differences between the ISO and NAP codelists. The recommended practice is to use ISO codelists wherever possible, encoded as in the examples above. NAP codes may be used where necessary, with the appropriate codeList URL: " http://www.fgdc.gov/nap/metadata/register/registerItemClasses.html" and optional codelist-specific document fragment token (e.g. #IC_87)
If a new codelist is created to restrict text in an ISO element whose type is simply CharacterString (e.g. HierarchyLevelName), then `characterString` element values are encoded by soft-typing the element that takes the character string using the xsi:type attribute. The following example uses the FileFormatCodeList, which is the only code list vocabulary added to the collection of codelists defined by ISO 19115 by the North American Profile.

```
<gmd:fileType xsi:type="gml:CodeType"
        codeList="http://www.fgdc.gov/nap/metadata/register/registerItemClasses.html#IC_115"
        codeListValue="jpg">
  <gco:CharacterString>jpg</gco:CharacterString>
</gmd:fileType>
```

As a convention for using controlled vocabularies on `gco:characterString` elements without the overhead of a defining a new namespace and xml schema, USGIN proposes that use of a controlled vocabulary be indicated by adding the attribute `xsi:type="gml:CodeType"` on the `gco:CharacterString` element. The type gml:CodeType requires a codeList attribute (see 4.14.2 Non digital resources and 7.2 Linkage name conventions), which MUST be the URI for the vocabulary used, with the implication that the `gco:CharacterString` element value and the codeListValue attribute will then be an identifier from that vocabulary. This essentially turns the CharacterString into a GML scoped name or gco:LocalName element.
Table 9. Codelist crosswalk between ISO, NAP and USGIN.

Codelist (ISO / NAP)	Coded Values/Names	Comments
`CI_DateTypeCode`	`creation, publication, revision`	ISO 19115 (B.5.2)
`napCI_DateTypeCode`	`…, notAvailable, inForce, adopted, deprecated, superseded`	NAP expansion
`CI_OnLineFunction-Code`	`download, information, offlineAccess, order, search`	ISO 19115 (B.5.3)
`napCI_OnLineFunction-Code`	`…, upload, webService, emailService, browsing, fileAccess, webMapService`	NAP expansion
`CI_PresentationForm-Code` `napCI_PresentationFo`	`documentDigital, documentHardcopy, imageDigital, imageHardcopy, mapDigital, mapHardcopy, modelDigital, modelHardcopy, profileDigital,`	ISO 19115 (B.5.4)

rm-Code	profileHardcopy, tableDigital, tableHardcopy, videoDigital, videoHardcopy, audioDigital	
	…, audioHardcopy, multimediaDigital, multimediaHardcopy, diagramDigital, diagramHardcopy	NAP expansion
CI_RoleCode napCI_RoleCode	resourceProvider, custodian, owner, user, distributor, originator, pointOfContact, principalInvestigator, processor, publisher, author	ISO 19115 (B.5.5)
	…, collaborator, editor, mediator, rightsHolder	NAP expansion
DQ_EvaluationMethod-TypeCode napDQ_Evaluation-MethodTypeCode	directInternal, directExternal, indirect	ISO 19115 (B.5.6)
DS_AssociationType-Code napDS_Association-TypeCode	crossReference, largerWorkCitation, partOfSeamlessDatabase, source, stereoMate	ISO 19115 (B.5.7)
	…, isComposedOf	NAP expansion
DS_InitiativeType-Code napDS_Initiative-TypeCode	campaign, collection, exercise, experiment, investigation, mission, sensor, operation, platform, process, program, project, study, task, trial	ISO 19115 (B.5.8)
MD_CellGeometryCode	point, area	ISO 19115 (B.5.9)
napMD_CellGeometry-Code	…, voxel	NAP expansion
MD_CharacterSetCode napMD_CharacterSet-Code	ucs2, ucs4, utf7, utf8, utf16, 8859part1, 8859part2, 8859part3, 8859part4, 8859part5, 8859part6, 8859part7, 8859part8, 8859part9, 8859part10, 8859part11, 8859part13, 8859part14, 8859part15, 8859part16, jis, shiftJIS, eucJP, usAscii, ebcdic, eucKR, big5, GB2312	ISO 19115 (B.5.10)
MD_Classification-Code	unclassified, restricted, confidential, secret, topSecret	ISO 19115 (B.5.11)
napMD_ClassificationCode	…, sensitive, forOfficialUseOnly	NAP expansion
MD_CoverageContent-TypeCode napMD_Coverage-ContentTypeCode	image, thematicClassification, physicalMeasurement	ISO 19115 (B.5.12)
MD_DataTypeCode not used by NAP	class, codelist, enumeration, codelistElement, abstractClass,	ISO 19115 (B.5.13) – The

and USGIN	aggregateClass, specifiedClass, datatypeClass, interfaceClass, unionClass, metaClass, typeClass, characterString, integer, association	MD_MetadataExtension Information element and its codelists are not used by NAP and USGIN.
MD_DimensionName-TypeCode napMD_DimensionName-TypeCode	row, column, vertical, track, crossTrack, line, sample, time	ISO 19115 (B.5.14)
MD_GeometricObject-TypeCode napMD_Geometric-ObjectTypeCode	complex, composite, curve, point, solid, surface	ISO 19115 (B.5.15)
MD_ImagingCondition-Code napMD_Imaging-ConditionCode	blurredImage, cloud, degradingObliquitty, fog, heavySmokeOrDust, night, rain, semiDarkness, shadow, snow, terrainMasking	ISO 19115 (B.5.16)
MD_KeywordTypeCode napMD_KeywordType-Code	discipline, place, stratum, temporal, theme	ISO 19115 (B.5.17)
	…, product, subTopicCategory	NAP expansion
MD_Maintenance-FrequencyCode napMD_Maintenance-FrequencyCode	continual, daily, weekly, fortnightly, monthly, quarterly, biannually, annually, asNeeded, irregular, notPlanned, unknown	ISO 19115 (B.5.18)
	…, semimonthly	NAP expansion
MD_MediumFormatCode napMD_MediumFormat-Code	cpio, tar, highSierra, iso9660, iso9660RockRidge, iso9660AppleHFS	ISO 19115 (B.5.19)
	…, UDF	NAP expansion
MD_MediumNameCode napMD_MediumNameCode	cdRom, dvd, dvdRom, 3halfinchFloppy, 5quarterInchFloppy, 7trackTape, 9trackTape, 3480Cartridge, 3490Cartridge, 3580Cartridge, 4mmCartridgeTape, 8mmCartridgeTape, digitalLinearTape, onLine, satellite, telephoneLink, hardcopy, hardcopyDiazoPolyester08, hardcopyCardMicrofilm, hardcopyMicrofilm240, hardcopyMicrofilm35, hardcopyMicrofilm70, hardcopyMicrofilmGeneral, hardcopyMicrofilmMicrofiche, hardcopyNegativePhoto, hardcopyPaper	ISO 19115 (B.5.20)
	…, hardcopyDiazo, hardcopyPhoto, hardcopyTracedPaper, hardDisk, USBFlashDrive,	NAP expansion

	1quarterInchCartridgeTape	
MD_ObligationCode not used by NAP and USGIN	mandatory, optional, conditional	ISO 19115 (B.5.21) - The MD_MetadataExtensionInformation element and its codelists are not used by NAP and USGIN.
MD_PixelOrientation-Code napMD_Pixel-OrientationCode	center, lowerLeft, lowerRight, upperRight, upperLeft	ISO 19115 (B.5.22)
MD_ProgressCode napMD_ProgressCode	completed, historicalArchive, obsolete, onGoing, planned, required, underDevelopment	ISO 19115 (B.5.23)
	…, proposed	NAP expansion
MD_RestrictionCode napMD_Restriction-Code	copyright, patent, patentPending, trademark, license, intellectualPropertyRights, restricted, otherRestrictions	ISO 19115 (B.5.24)
	…, licenseUnrestricted, licenseEndUser, licenseDistributor, privacy, statutory, confidential, sensitivity	NAP expansion
MD_ScopeCode napMD_ScopeCode	attribute, attributeType, collectionHardware, collectionSession, dataset, series, nonGeographicDataset, dimensionGroup, feature, featureType, propertyType, fieldSession, software, service, model, tile	ISO 19115 (B.5.25)
MD_Spatial-RepresentationTypeCode napMD_Spatial-RepresentationTypeCode	vector, grid, textTable, tin, stereoModel, video	ISO 19115 (B.5.26)
MD_TopicCategoryCode napMD_TopicCategory-Code	farming, biota, boundaries, climatologyMeterologyAtmosphere, economy, elevation, environment, geoscientificInformation, health, imageryBaseMapsEarthCover, intelligenceMilitary, inlandWater, location, oceans, planningCadastre, society, structure, transportation, utilitiesCommunication	ISO 19115 (B.5.27)
MD_TopologyLevelCode napMD_TopologyLevel-Code	geometryOnly, topology1D, planarGraph, fullPlanarGraph, surfaceGraph, fullSurfaceGraph, topology3D, fullTopology3D, abstract	ISO 19115 (B.5.28)

SV_CouplingType napSV_CouplingType	loose, mixed, tight	ISO 19119 (Amendment 1; C.2.8)
SV_Parameter-Direction napSV_Parameter-Direction	in, out, in/out	ISO 19119 (Amendment 1; C.2.9
LanguageCode	see http://www.loc.gov/standards/iso639-2/php/code_list.php	no complete NAP or ISO registry found
not used by ISO nap_DCPList	XML, CORBA, JAVA, COM, SQL, WebServices	NAP specific codelist – not used by USGIN due to poorly defined semantics and use.
not used by ISO napMD_FileFormatCode	bil, bmp, bsq, bzip2, cdr, cgm, cover, csv, dbf, dgn, doc, dwg, dxf, e00, ecw, eps, ers, gdb, geotiff, gif, gml, grid, gzip, html, jpg, mdb, mif, pbm, pdf, png, ps, rtf, sdc, shp, sid, svg, tab, tar, tiff, txt, xhtml, xls, xml, xwd, zip, wpd	NAP specific codelist – not formally used by USGIN, but these character strings should are to be used to populate fileType elements.

4.19 Geographic bounding box

USGIN profile requires that if an `EX_Extent/geographicElement` is supplied, at least one geographic bounding box with bounding latitude and longitude expressed using WGS 84 decimal degrees MUST be provided.

The corner coordinates for the geographic bounding box must not coincide in one point, because this may result in fatal errors with some CSW implementations. Point locations must thus be represented as tiny rectangles. USGIN recommended practice is to place the actual point location in the lower left corner of the rectangle.

4.20 Data Quality

4.20.1 Simple quality statement

For resource evaluation purposes, to answer a user's question 'is the described resource good enough for my purposes?', a free text statement is generally sufficient in place of a detailed xml document. For example, this might be a statement something like "*These data were compiled by the authors from field sheets and notes by scanning paper copies, georeferencing and digitizing on screen. Station locations are based on Garmin 12 GPS readings, except locally where they have been adjusted for consistency with the base map. Original GPS coordinates are reported in the station table. These data have been reviewed for completeness of description and internal consistency, but have not been independently field checked.*" This is the

only kind of quality information available for many resources; it is neither a quantitative measure nor technically a conformance result. A simple qualityStatement would suffice and provide significant value. There is no ISO19115 element for such a quality statement.

To implement this simple case, the proposed solution is that the first report conformance explanation in each DQ_DataQuality element instance would contain a free text discussion of data quality considerations for the indicated scope for that element. Thus, text at the XPath `DQ_DataQuality//report[1]//result[1]//explanation/CharacterString` will be interpreted as a self-explanatory quality statement. The use of any specific data quality element (e.g. DQ_QuantitativeAttributeAccuracy, DQ_RelativeInternalPositionalAccuracy..) to contain this explanation is arbitrary and may be ignored by client applications.

4.21 Lineage

Lineage in data quality section is a collection of processing steps that describe the provenance of a resource. Each step has some input resources, identified by source citations associated with the process step. The `LI_ProcessStep` element does not directly identify its output resource, so in a lineage that involves a chain of steps with intermediate resources, the sourceStep association from `LI_Source` links a resource to a processing step that it is output from.

A `LI_Lineage` element that contains a `source` but no `processStep` elements is redundant; this source information should be represented with the `MD_DataIdentification/citation/CI_Citation` element.

A GIS dataset originally digitized from a published geologic map, put online, obtained by an online download, and reprojected would report one `processStep` (reprojection) with `source/LI_Source` that has a `CI_Citation` for the downloaded data. This `LI_Source` would have a `sourceStep` pointing to an `LI_ProcessStep` for the original digital conversion from the paper map, and the `LI_ProcessStep/source/LI_Source` would contain the citation for the original paper map.

In order to enable XPath queries for any of the sources or processSteps in a processing chain, all related `LI_Source` and `LI_ProcessStep` elements should be directly nested within the `LI_Lineage` element, and the `processStep/source` and `LI_Source/sourceStep` associations should be by internal document references (see Code Example 3).

Code Example 3: A complex processing and source history using LI_Lineage.

```
<?xml version="1.0" encoding="UTF-8"?>
<LI_Lineage
  xmlns="http://www.isotc211.org/2005/gmd"
  xmlns:gco="http://www.isotc211.org/2005/gco"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.isotc211.org/2005/gmd
 http://schemas.opengis.net/iso/19139/20070417/gmd/dataQuality.xsd">
  <statement>
    <CharacterString>The digital data described by this metadata was originally compiled
        digitally from two published maps; this digital dataset was then reprojected to produce
        the described resource.</CharacterString>
 </statement>
 <processStep>
```

```
      <LI_ProcessStep id="1">
        <description>
          <CharacterString>digital compilation of 2 Maps</CharacterString>
        </description>
        <source xlink:href="#10"/>
        <source xlink:href="#20"/>
      </LI_ProcessStep>
    </processStep>
    <processStep>
      <LI_ProcessStep id="2">
        <description>
          <CharacterString>digital map compilation reprojected, should have some way to specify
            projection parameters?, output is LI_Source id=70 </CharacterString>
        </description>
        <source xlink:href="#40"/>
      </LI_ProcessStep>
    </processStep>
    <source>
      <LI_Source id="40">
        <description>
          <CharacterString>a digital compilation of 2 maps, output of processStep ID=1, input into
            reprojection process</CharacterString>
        </description>
        <sourceStep xlink:href="1"/>
      </LI_Source>
    </source>
    <source>
      <LI_Source id="10">
        <description>
          <LocalisedCharacterString>ultimate
            source--somepublishedmap</LocalisedCharacterString>
        </description>
  <!--no source processing recorded for production of paper map so no sourceStep-->
      </LI_Source>
    </source>
    <source>
      <LI_Source id="20">
        <description>
          <LocalisedCharacterString>another published map</LocalisedCharacterString>
        </description>
      </LI_Source>
    </source>
    <source>
      <LI_Source id="70">
        <description>
          <LocalisedCharacterString>a reprojected version of the digital
              compilation</LocalisedCharacterString>
        </description>
        <sourceStep xlink:href="2"/>
      </LI_Source>
    </source>
```

```
</LI_Lineage>
```

An `LI_Lineage` may be constructed that involves a number of resources and processing steps, and this lineage may be referenced by metadata for all the resources involved in the processing. The `LI_Lineage/source/LI_Source/sourceCitation/CI_Citation/identifier/MD_Identifier` is a reference to the `MD_Metadata/fileIdentifier` for the metadata for each resource in the chain. This approach allows the metadata record to record relationships through process steps between resources.

4.22 Temporal extents

Resource temporal extent (`identificationInfo/MD_DataIdentification/extent/EX_Extent/-temporalElement/EX_TemporalExtent/extent/TimePeriod`) is used to specify the temporal interval to which the content of a resource applies. Default reference frame for time is calendar date and time encoded using ISO-8601:

```
<gml:TimePeriod gml:id="Id2010">
    <!-- Note that gml:TimePeriod requires a gml:id string that must be unique
      in the scope of the instance document, and must have an alphabet
      character ('a-z', 'A-Z') as its first character -->
  <!-- USGIN requires the beginPosition and endPosition's frame property to be defined.
  The default value is #ISO-8601. -->
    <gml:beginPosition frame="#ISO-8601">2010-01-00T00:00:00Z </gml:beginPosition>
    <gml:endPosition frame="#ISO-8601">2010-12-31T24:00:00Z</gml:endPosition>
</gml:TimePeriod>
```

<gml:endPosition indeterminatePosition="now"/> is the correct way to represent "Present" in ISO or GML as one of the boundaries of a timePeriod.

The ISO 19139 xml schema allows temporal location to be quantified by a gml:TimeInstant or gml:TimePeriod element. In order to promote interoperability, the USGIN profile mandates use of gml:TimePeriod for specifying temporal extent for a resource.

For geologic time extents, the time coordinates for the beginPosition and endPosition should be expressed numerically in Ma. This convention allows search for resources pertinent to intervals of geologic time using simple numeric comparisons instead of the complex hierarchical concept expansions that would be necessary to use named eras from a stratigraphic time scale. Encoding example:

```
<EX_TemporalExtent>
  <extent>
    <gml:TimePeriod gml:id="y34096">
      <gml:beginPosition
        frame="urn:CGI:TemporalCRS:cgi:standardGeologyMa">220</gml:beginPosition>
      <gml:endPosition
     frame="urn:CGI:TemporalCRS:cgi:standardGeologyMa">140 </gml:endPosition>
    </gml:TimePeriod>
  </extent>
</EX_TemporalExtent>
```

The frame for the beginPosition and endPosition is a URI for standard geologic time, measured positive getting older, with an origin at 1950 CE, in units of millions of years.

4.23 Operation metadata

See the discussion on the use of service metadata in the USGIN profile above, preceding Table 3. The srv namespace elements based on ISO 19119 provide insufficient information and guidelines to access data through services without apriori knowledge of the service protocol. This is due in part to weakly defined semantics and use cases for the elements that are there (DCP, applicationProfile, protocol, MD_Format, serviceType, operationName vs. invocationName, connectPoint), and partly due to incomplete content model (where to put allowed outputFormat parameter values or supported query operations for CSW or WMS, what is syntax for delimiting parameters, are http request headers used). The ISO 19119 model for service metadata does not include a mechanism to specify valid values for operation parameters. For instance, OGC WMS and CSW services both support an output format parameter, and OGC capabilities documents provide a listing of the supported output formats, but where do these go in ISO19139 xml documents? Does the described service support http POST or GET method? This information is necessary in order to compose valid service requests.

Based on these considerations, the USGIN profile mandates that a service metadata record MUST include at least one service operation with the `operationDescription/CharacterString` = 'serviceDescription'. The `identificationInfo/SV_ServiceIdentification/containsOperations/-SV_OperationMetadata/connectpoint` link for this operation MUST retrieve the service-specific self description document (e.g. WSDL, GetCapabilities, WADL). The `CI_OnlineResource` in this connectPoint elment must have `CI_OnlineResource/name` = "serviceDescription" (from the table in section 7.2 Linkage name conventions).

WSDL and getCapabilities were designed to describe service operation, and it seems counterproductive to invent another scheme to do the same thing. Service metadata for processing services that are not coupled to particular input datasets should provide `srv:SV_OperationMetadata` to describe the operations offered by the service to inform human users. Because of the difficulty in creating usable abstract model that accounts for any and all possible services, it makes more sense to allow service description documents specific to different service frameworks to describe all the details in a machine-actionable format.

5 Abbreviations

CSW	Metadata Catalog for the Web. Also abbreviated as CS-W and CS/W
GeoSciML	Geoscience Markup Language
GML	Geographic Markup Language
GUID	Global Unique Identifier
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
UML	Unified Modeling Language
URI	Universal Resource Identifier
USGIN	U.S. Geoscience Information Network
WCS	Web coverage Service
WFS	Web Feature Service
XML	eXtensible Markup Language
XSD	XML Schema Definition
XSL	eXtensible Stylesheet Language
XSLT	XSL Transformations
XLink	XML Linking Language

6 References

6.1 Cited literature

[ANZLIC, 2007] ANZLIC Metadata Profile Guidelines, Version 1.0: Turner, ACT, ANZLIC - the Spatial Information Council, ISBN: 978-0-646-46940-9, 372 p., available at http://www.osdm.gov.au/Metadata/ANZLIC+metadata+resources/default.aspx (accessed 2010-11-18).

[Dublin Core] 2008-01-14 Dublin core Metadata Element Set, Version 1.1: Dublin Core Metadata Initiave, accessed at http://dublincore.org/documents/dces/.

[INSPIRE ISO19115/119] Drafting Team Metadata and European Commision Joint Research Centre, 2009-02-18, INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119,v. 1.1: European Commission Joint Research Centre, MD_IR_and_ISO_20090218, available at http://inspire.jrc.ec.europa.eu/index.cfm/pageid/101 (accessed 2010-11-18).

Franklin, Michael, Halevy, Alon, and Maier, David, 2005, From databases to dataspaces: a new abstraction for information management: ACM SIGMOD Record, V. 34, No. 4, ISSN:0163-5808.

[FRBR] Tillett, Barbara, 2004-03-02, What is FRBR?: Washington, DC, Library of Congress Cataloging Distribution Service, 8 pages. (originally published in Technicalities v. 25, no. 5, 2003), available at http://www.loc.gov/cds/downloads/FRBR.PDF (accessed 2013-06-12).

Heaney, Michael, 2000-01-14, An analytical model of Collections and their Catalogues: UK Office for Library and Information Networking, Third Issue, available at http://www.ukoln.ac.uk/metadata/-rslp/model/amcc-v31.pdf (accessed 2013-03-18).

Richard, S. M., and Grunberg, Wolfgang, editors, 2010, Metadata Recommendataions for Geoscience Resources: U.S. Geoscience Information Network Best Practices Document, Doc ID gin2010-11, v. 1.0.3, available at http://lab.usgin.org/profiles/doc/metadata-content-recommendations (accessed 2010-11-18).

7 Codelists

7.1 ServiceType

INSPIRE metadata Implementing Rules (*OJ L 326, 4.12.2008)* section D3 mandate the use of the value domain listed in Table 11 to categorize spatial data service types. These values are better suited for CI_OnlineFunctionCode used to specify CI_OnlineResource/online/Function. The USGIN team interprets the ISO scope notes to allow more useful content for service type, specifying an actual service specification like OGC WMS. USGIN draft ServiceType vocabulary is reported in Table 12.

Table 11. INSPIRE SPATIAL DATA SERVICE TYPE (for information only, not used by USGIN)

Type	Description
discovery	Discovery Service
view	View Service
download	Download Service
transformation	Transformation Service
invoke	Invoke Spatial Data Service
other	Other Services

Table 12. USGIN service type vocabulary. This is an interim listing of serviceTypes. The code list URI for this registry is http://resources.usgin.org/registry/serviceType201001. See also http://dclite4g.xmlns.com/ws.rdf for another rdf-based vocabulary. These should probably be merged.[NOTE 2014-03-03. Actual USGIN metadata is not using these precise identifiers. An activity is underway at https://github.com/OSGeo/Cat-Interop to develop a community vocabulary of service types.

Identifier	Name	Description
WMS	OGC Web Map service	provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases. A WMS request defines the geographic layer(s) and area of interest to be processed. The response to the request is one or more geo-registered map images (returned as JPEG, PNG, etc) that can be displayed in a browser application. The interface also supports the ability to specify whether the returned images should be transparent so that layers from multiple servers can be combined or not. (http://www.opengeospatial.org/standards/wms)
WFS	OGC Web Feature service	WFS offers direct fine-grained access to geographic information at the feature and feature property level; allows clients to retrieve or modify specific data items. That data can then be used for a wide variety of purposes, including purposes other than their producers' intended ones. http://www.opengeospatial.org/standards/wfs
WCS	OGC Web coverage	interface and operations that enable interoperable access to geospatial "coverages" [http://www.opengeospatial.org/ogc/-

	service	glossary/c]. The term "grid coverages" typically refers to content such as satellite images, digital aerial photos, digital elevation data, and other phenomena represented by values at each measurement point.
CSW	OGC Web catalog service	publish and search collections of descriptive information (metadata) about geospatial data, services and related resources. Providers of resources use catalogues to register metadata that conform to the provider's choice of an information model; such models include descriptions of spatial references and thematic information. (http://www.opengeospatial.org/standards/cat)
SOS	OGC Sensor observation service	manage deployed sensors and retrieve sensor data, specifically "observation" data. E.g. in-situ sensors (e.g., water monitoring) or dynamic sensors (e.g., satellite imaging). (http://www.opengeospatial.org/standards/sos)
WPS	OGC Web Processing service	rules for standardizing how inputs and outputs (requests and responses) for geospatial processing services, such as polygon overlay, are exchanged on the web to link services; defines how a client can request the execution of a process, and how the output from the process is handled, and an interface to register geospatial processes and for client to discovery and bind to those processes. (http://www.opengeospatial.org/standards/wps)
SPS	OGC Sensor planning service	interfaces for exchanging information about the capabilities of a sensor and how to task the sensor; designed to support queries to determine the feasibility of a sensor planning request, to submit such a request, to inquire about the status of such a request, to update or cancel such a request, and to request information about other OGC Web services that provide access to the data collected by the requested task.
OpenDAP	Open source data access protocol	(http://opendap.org/) a protocol for requesting and transporting data across the web. DAP 2.0 uses HTTP to frame the requests and responses. For details on DAP, see Data Access Protocol (DAP), version 2 which is a technical description of the Data Access Protocol.
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting	provides an application-independent interoperability framework based on metadata harvesting.

Example usage:

```
<srv:serviceType>
  <gco:LocalName
codeSpace="http://resources.usgin.org/registry/serviceType201001">WMS</gco:Local
Name>
</srv:serviceType>
```

7.2 Linkage name conventions

The cardinality of the `online` element in DigitalTransferOptions is 0..*. In order to distinguish the nature of various linkages that might be provided, above and beyond function, protocol, and

applicationProfile, USGIN profile mandates use of the following names to associate with links to identify important linkages.

Table 13. USGIN Names to identify special linkage URL's for CI_Online Resource. CodeList URI = http://resources.usgin.org/registry/linkageName201001

Identifier	Name (eng)	Usage
icon	Icon	linkage url is link to a thumbnail icon. Icon pixel height and width range?
serviceDescription	web service description	linkage url is link to getCapabilities or WSDL that describes a service using a formal syntax such that computer programs can automate connection to the service. OnlineFunctionCode for CI_OnlineResource should be 'Information'
baseURL	web service endpoint	Base url for service. Assumes that ServiceType specifies a well know service type such that requests can be constructed without significant additional information. OnlineFunctionCode for CI_Online-Resource should be 'webService'
serviceClient	web application	URL is linkage to a web application that allows the user to access the service. OnlineFunctionCode for CI_OnlineResource should be 'Browsing'
webpage	access information	URL locates a web page with instructions for accessing the service. This provides the user with information to implement a connection to the service, but does not enable automated service access. OnlineFunctionCode for CI_Online-Resource should be 'Information'
~~serviceMetadata~~	~~Service Metadata~~	~~linkage URL is link to a complete service metadata record. Use for distribution linkage to a service for which more information that can be provided by the CI_OnlineResource element is necessary to successfully automate access to the context resource through the service. OnlineFunctionCode for CI_Online-Resource should be 'Information'~~
download		

email request		
offline access		

Example usage:

```
<gmd:CI_OnlineResource>
  <gmd:linkage>
    <gmd:URL>http://75.101.143.247:8080/gsvr/wms?SERVICE=WMS&amp;REQUEST=getCa
      pabilities</gmd:URL>
  </gmd:linkage>
  <gmd:protocol>
    <gco:CharacterString>OGC:WMS</gco:CharacterString>
  </gmd:protocol>
  <gmd:name>
    <gco:CharacterString xsi:type=”gml:CodeType”
      codeSpace="http://resources.usgin.org/registry/linkageName201001">
      serviceDescription</gco:CharacterString>
  </gmd:name>
</gmd:CI_OnlineResource>
```

Use of such controlled vocabulary can be indicated by using xsi:type on the gco:characterString element to make the type gml:CodeType, which then requires a codeSpace attribute. The distribution format Identifier from Table 6 should be used as the element value. For compatibility with systems that can not process this encoding, the code identifier should be included as the element value as well as the codeListValue.

8 Examples

8.1 USGIN ISO 19139 Minimum Dataset Metadata

In the following listing, text in green is comments; XML elements are in blue, XML attributes are in black, and attribute values are in purple.

This example metadata record describes itself.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--*********************************************************************************************
*** Minimum example of a ISO 19139 Geospatial Dataset Metadata
*** based on the USGIN v1.2 Profile
*** by USGIN Standards and Protocols Drafting Team
*** U.S. Geoscience Information System (USGIN) - http://lab.usgin.org
*** Contributors: Wolfgang Grunberg, Stephen M Richard
*** 2013-06-13
***
*** DISCLAIMER: this is not an authoritative metadata example but an aide to get started.
*** Refer to the USGIN profile document for more specific and complete guidelines.
***
*** Validated against http://www.isotc211.org/2005/gmd (ISO 19115, CSW 2.0.2 AP ISO 1.0).
*** Follows the USGIN ISO 19139 Dataset Metadata Profile v1.2.
***
*** NOTES:
*** - Codelists:
*** Most ISO metadata profiles and applications use ISO codelists or codelists that use ISO's
                codelist names.
***
*** - Language code:
*** Use ISO <ISO639-2/T three letter language code - lower case> formatting.
***
*** KEY: (NAP-USGIN) - M/C/O/X (Mandatory, Conditional, Optional, Not Used)
***
*********************************************************************************************-->
        <!-- USGIN ISO 19139 geospatial dataset metadata record -->
<gmd:MD_Metadata
                    xmlns:gmd="http://www.isotc211.org/2005/gmd"
                     xmlns:gco="http://www.isotc211.org/2005/gco"
                     xmlns:gml="http://www.opengis.net/gml"
                     xmlns:xlink="http://www.w3.org/1999/xlink"
                     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                     xsi:schemaLocation="http://www.isotc211.org/2005/gmd
            http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd">
        <!-- (M-M) Metadata file identifier - A unique File Identifier (GUID) - USGIN recommends
                using a valid Universally Unique Identifier (UUID) -->
  <gmd:fileIdentifier>
    <gco:CharacterString>08fb00c8-0882-4bf7-b07f-fd37050c5efc</gco:CharacterString>
```

```
</gmd:fileIdentifier>
          <!-- (M-M) Metadata language - ISO <ISO639-2/T three letter language code - lower
      case formatting. -->
<gmd:language>
   <gco:CharacterString>eng</gco:CharacterString>
</gmd:language>

        <!-- (M-M) Metadata character set   - USGIN requires that a character set code is
              defined to facilitate CSW servers (deegree, GeoNetwork, etc.).   -->
<gmd:characterSet>
      <!-- MD_CharacterSetCode names: {ucs2, ucs4, utf7, utf8, utf16, 8859part1, 8859part2,
   8859part3, 8859part4, 8859part5, 8859part6, 8859part7, 8859part8, 8859part9, 8859part10,
   8859part11, 8859part13, 8859part14, 8859part15, 8859part16, jis, shiftJIS, eucJP, usAscii,
   ebcdic, eucKR, big5, GB2312} -->
    <gmd:MD_CharacterSetCode
   codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Charact
   erSetCode" codeListValue="utf8">UTF-8</gmd:MD_CharacterSetCode>
</gmd:characterSet>
         <!-- (M-M) Resource type - Define if this record is a: dataset (default), service, feature,
      software, etc. -->
<gmd:hierarchyLevel>
         <!-- MD_ScopeCode code names: {attribute, attributeType,
   collectionHardware,collectionSession, dataset, series, nonGeographicDataset,
   dimensionGroup, feature, featureType, propertyType, fieldSession, software, service, model,
   tile}. -->
   <gmd:MD_ScopeCode
     codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Scope
     Code" codeListValue="dataset">dataset</gmd:MD_ScopeCode>
</gmd:hierarchyLevel>
         <!-- (O-M) Resource hierarchy level name - USGIN makes this property mandatory to
              identify the USGIN resource type (see USGIN Profile, "Resources of Interest").
              Default CharacterString is “Dataset” Encode hierarchy by including
              hierarchyLevelName elements for all broader resource categories. E.g. default
              should also include a hierarchyLevelName=”Collection” element. For services
              USGIN hierarchyLevelName/CharacterString is “Service”. -->
<gmd:hierarchyLevelName>
   <gco:CharacterString>Dataset</gco:CharacterString>
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
   <gco:CharacterString>Collection</gco:CharacterString>
</gmd:hierarchyLevelName>
         <!-- (M-M) Metadata point of contact - Point of contact for the metadata record, e.g. for
              users to report errors, updates to metadata, etc. -->
<gmd:contact>
   <gmd:CI_ResponsibleParty>
    <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
    <!-- only one is mandatory, this example has organisationName -->
      <gmd:organisationName>
        <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
      </gmd:organisationName>
```

```
      <!-- Metadata contact information, either a phone number or e-mail address is required.
      This example has an e-mail address. ISO19115 makes contact role mandatory as well.
                -->
     <gmd:contactInfo>
       <gmd:CI_Contact>
         <gmd:address>
           <gmd:CI_Address>
  <!-- (O-M) Metadata point of contact e-mail address to meet profile requirement -->
             <gmd:electronicMailAddress>
               <gco:CharacterString>metadata@azgs.az.gov</gco:CharacterString>
             </gmd:electronicMailAddress>
           </gmd:CI_Address>
         </gmd:address>
       </gmd:CI_Contact>
     </gmd:contactInfo>
  <!-- (M-M) ISO 19139 Mandatory: contact role -->
     <gmd:role>
  <!-- CI_RoleCode names: {resourceProvider, custodian, owner, user, distributor, originator,
   pointOfContact, principalInvestigator, processor, publisher, author}   -->
     <gmd:CI_RoleCode
       codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_RoleC
       ode" codeListValue="pointOfContact">point of contact</gmd:CI_RoleCode>
     </gmd:role>
   </gmd:CI_ResponsibleParty>
</gmd:contact>
       <!-- (M-M) Metadata date stamp - USGIN profile requires use of
    dateStamp/gco:DateTime (Note this contrasts with INSPIRE mandate to use
    dateStamp/gco:Date). This is the date and time when the metadata record was created or
    most recently updated (following NAP). -->
<gmd:dateStamp>
  <!-- Requires an extended ISO 8601 formatted combined UTC date and time string
   (YYYY-MM-DDTHH:MM:SS) Some validators require time zone as well; recommend using
   GMT indicated by 'Z' -->
   <gco:DateTime>2010-01-14T10:00:00Z</gco:DateTime>
</gmd:dateStamp>
     <!-- (M-M) metadata standard - USGIN profile conformant metadata is indicated by the
               string “ISO-USGIN" -->
<gmd:metadataStandardName>
   <gco:CharacterString>ISO-USGIN</gco:CharacterString>
</gmd:metadataStandardName>
     <!-- (O-M) USGIN profile version -->
<gmd:metadataStandardVersion>
   <gco:CharacterString>1.2</gco:CharacterString>
</gmd:metadataStandardVersion>
      <!-- (M-M) Resource identification information - At least one of MD_DataIdentification
               (dataset, dataset series) or SV_ServiceIdentification (service) is required.   -->
<gmd:identificationInfo>
    <!-- Resource Dataset or Dataset Series Identification -->
   <gmd:MD_DataIdentification>
      <!-- (M-M) Resource citation - Information to identify the intellectual origin of the content in
               the described resource, along the lines of a citation in a scientific journal. Required
```

```
            content for a CI_Citation element are title, date, and responsibleParty. If the
            subject of the metadata record is derived from other sources, such that it is not
            considered identical with those sources, the provenance should be documented in
            LI_Lineage//LI_Source elements. -->
    <gmd:citation>
      <gmd:CI_Citation>
     <!-- (M-M) Resource title - Use titles that inform the human reader about the dataset's
            content; titles should be unique in the context of the metadata catalog that
            contains the metadata record. -->
        <gmd:title>
          <gco:CharacterString>USGIN minimum metadata example XML
           file.</gco:CharacterString>
        </gmd:title>
     <!-- (M-M) Resource reference date - Best practice is to include at least the date of
            publication or creation of the resource. The date of the resource reported in the
            citation corresponds to the resource's last update version according to its update
            frequency. CI_Date content includes a date and dateType. Date for USGIN profile
            uses gco:DateTime. -->
        <gmd:date>
          <gmd:CI_Date>
            <gmd:date>
<!-- Requires an extended ISO 8601 UTC date and time string (YYYY-MM-DDTHH:MM:SS)
 Some validators require time zone as well; recommend GMT indicated by 'Z' -->
              <gco:DateTime>2010-01-14T09:30:47Z</gco:DateTime>
            </gmd:date>
<!-- dateType is mandatory per ISO19115 -->
            <gmd:dateType>
<!-- CI_DateTypeCode names: {creation, publication, revision}. -->
              <gmd:CI_DateTypeCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
                CI_DateTypeCode"codeListValue="publication">publication</gmd:CI_DateTyp
                eCode>
            </gmd:dateType>
          </gmd:CI_Date>
        </gmd:date>
      <!-- (M-M) Resource responsible party - The citation attribute provides information for
            citing the intellectual origin of the described resource, analogous to a citation in a
            scientific journal. The citedResponsibleParty should specify the individual(s) and
            organizations responsible for the creation or origination of the resource. For most
            intellectual content, the responsible party is what would normally be considered
            the author of a work.    -->
        <gmd:citedResponsibleParty>
          <gmd:CI_ResponsibleParty>
 <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
 <!-- only one is mandatory, this example has organisationName -->
            <gmd:organisationName>
              <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
            </gmd:organisationName>
        <!-- cited responsible party contact information, either a phone number or e-mail
            address is required. This example has a voice telephone number. ISO19115
            makes contact role mandatory as well. -->
```

```
            <gmd:contactInfo>
              <gmd:CI_Contact>
                <gmd:phone>
                  <gmd:CI_Telephone>
                    <gmd:voice>
                      <gco:CharacterString>520-770-3500</gco:CharacterString>
                    </gmd:voice>
                  </gmd:CI_Telephone>
                </gmd:phone>
              </gmd:CI_Contact>
            </gmd:contactInfo>
  <!-- (M-M) ISO 19139 Mandatory: contact role   -->
            <gmd:role>
 <!-- CI_RoleCode names: {resourceProvider, custodian, owner, user, distributor, originator,
  pointOfContact, principalInvestigator, processor, publisher, author} -->
                                                        <gmd:CI_RoleCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
                CI_RoleCode"codeListValue="pointOfContact">point of
                contact</gmd:CI_RoleCode>
            </gmd:role>
          </gmd:CI_ResponsibleParty>
        </gmd:citedResponsibleParty>
      </gmd:CI_Citation>
    </gmd:citation>

    <!-- (M-M) Resource Abstract - A free text summary of the content, significance, purpose,
            scope, etc. of the resource. Exactly one value. -->
    <gmd:abstract>
      <gco:CharacterString>Example for the minimum required elements in a USGIN dataset
        metadata record. Note that this example includes conditional minimum elements that
        may or may not apply to a specific resource and its metadata. Although the resource is
        non-geographic, and extent element is included for example
        purposes.</gco:CharacterString>
    </gmd:abstract>
    <!-- (O-C) Resource point of contact (access contact) - CI_ResponsibleParty element here
            would contain information for point of contact to access the resource. This
            information is mandatory for physical resources such as core, cuttings, samples,
            manuscripts. This example metadata record is not about a physical resource, so
            the element is empty (included here only to anchor this comment... -->
<gmd:pointOfContact/>
    <!-- (O-C) Descriptive Keywords - Only required if the described resource is
            non-geographic, in which case the single keyword 'non-geographic' is mandatory.
            This example instance is about itself, and example metadata record, and is thus
            non-geographic so the keyword is mandatory. Note that and example bounding
            box is included for information purposes as well. -->
    <gmd:descriptiveKeywords>
      <gmd:MD_Keywords>
        <gmd:keyword>
          <gco:CharacterString>non-geographic</gco:CharacterString>
        </gmd:keyword>
```

```
            <gmd:type>
          <!-- Keyword Type - allowed values from MD_KeywordTypeCode names: {discipline,
        place, stratum, temporal, theme} -->
                <gmd:MD_KeywordTypeCode
              codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Ke
              ywordTypeCode" codeListValue="place">place</gmd:MD_KeywordTypeCode>
               </gmd:type>
            </gmd:MD_Keywords>
          </gmd:descriptiveKeywords>
          <!-- (M-M) Resource language - Multiple instances of this element indicate that the
         linguistic content of the resource is available in multiple languages -->
          <gmd:language>
        <!-- ISO 639-2/T three-letter language code (http://www.loc.gov/standards/iso639-2/). -->
             <gco:CharacterString>eng</gco:CharacterString>
          </gmd:language>
          <!-- (C-C) Topic category - A topicCategory code must be provided when hierarchyLevel
                  is set to "dataset" (this example document) or "dataset series". Most USGIN
                  resources will have MD_TopicCategoryCode = "geoscientificInformation", which is
                  the default value for this profile.   More specific topic categorization should be
                  done using keywords. -->
          <gmd:topicCategory>
        <!-- MD_TopicCategoryCode names: {farming, biota, boundaries,
         climatologyMeterologyAtmosphere, economy, elevation, environment,
         geoscientificInformation, health, imageryBaseMapsEarthCover, intelligenceMilitary,
         inlandWater, location, oceans, planningCadastre, society, structure, transportation,
         utilitiesCommunication} -->

    <gmd:MD_TopicCategoryCode>geoscientificInformation</gmd:MD_TopicCategoryCode>
          </gmd:topicCategory>
          <!-- (C-C) Resource content extent - Defines the spatial (horizontal and vertical) and
                  temporal region to which the content of the resource applies. For USGIN, a
                  geographicElement/geographicBoundingBox must be provided for any resource
                  with content related to some geographic location. If the described resource is not
                  related to a geographic area, the place keyword 'non-geographic' should be
                  included in the descriptiveKeywords element. -->
          <!-- note that although this example record is about a non-geographic resource, bounding
                  boxes are anticipated for most resources so an example element is included, but it
                  is not required. -->
          <gmd:extent>
            <gmd:EX_Extent>
          <!-- (O-C) Resource content extent bounding box -USGIN profile requires that if an
                  EX_Extent/geographicElement is supplied, it include a geographic bounding box
                  with bounding latitude and longitude expressed using WGS 84 decimal degrees.
                  The corner coordinates for the geographic bounding box must not coincide in one
                  point, because this may result in fatal errors with some CSW implementations.
                  Point locations must thus be represented as tiny rectangles. USGIN recommended
                  practice is to place the actual point location in the lower left corner of the rectangle.
                  -->
              <gmd:geographicElement>
                <gmd:EX_GeographicBoundingBox>
                  <gmd:westBoundLongitude>
```

```
                <gco:Decimal>-109.911001</gco:Decimal>
              </gmd:westBoundLongitude>
              <gmd:eastBoundLongitude>
                <gco:Decimal>-109.910999</gco:Decimal>
              </gmd:eastBoundLongitude>
              <gmd:southBoundLatitude>
                <gco:Decimal>34.772899</gco:Decimal>
              </gmd:southBoundLatitude>
              <gmd:northBoundLatitude>
                <gco:Decimal>34.772901</gco:Decimal>
              </gmd:northBoundLatitude>
            </gmd:EX_GeographicBoundingBox>
          </gmd:geographicElement>
        </gmd:EX_Extent>
      </gmd:extent>
    </gmd:MD_DataIdentification>
  </gmd:identificationInfo>
      <!-- Distribution (O-M) This element provides information to inform users how to obtain or
           access the described resource, which is considered part of the minimum useful
           content that should be provided by a metadata record. USGIN profile specifies that
           at least one MD_Distribution/distributor and one MD_Distribution/transferOptions
           element are required, and that the specified distributor provides the specified
           transfer options. -->
  <gmd:distributionInfo>
      <gmd:MD_Distribution>
          <gmd:distributor>
              <gmd:MD_Distributor>
       <!-- (O-C) distributor point of contact - CI_ResponsibleParty element here would contain
           information on who to contact if there is a problem with accessing the resource
           following the instructions in this distribution element. This information is
           mandatory, requiring at least a name (one of individual, organization, or position)
           and a voice telephone number or e-mail address. This example includes a position
           name and e-mail address. -->
                  <gmd:distributorContact>
                      <gmd:CI_ResponsibleParty>
                          <gmd:positionName>
                            <gco:CharacterString>Web Master</gco:CharacterString>
                          </gmd:positionName>
                          <gmd:contactInfo>
                              <gmd:CI_Contact>
                                  <gmd:address>
                                      <gmd:CI_Address>
                                          <gmd:electronicMailAddress>

<gco:CharacterString>webmaster@usgin.org</gco:CharacterString>
                                          </gmd:electronicMailAddress>
                                      </gmd:CI_Address>
                                  </gmd:address>
                              </gmd:CI_Contact>
                          </gmd:contactInfo>
                          <gmd:role>
```

```
    <!-- CI_RoleCode names: pointOfContact is mandatory for at least one distributor
                responsible party -->
                        <gmd:CI_RoleCode
                         codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.x
                         ml#CI_RoleCode"codeListValue="pointOfContact">point of
                         contact</gmd:CI_RoleCode>
                                </gmd:role>
                            </gmd:CI_ResponsibleParty>
                        </gmd:distributorContact>
                    </gmd:MD_Distributor>
                </gmd:distributor>
                <gmd:transferOptions>
                    <gmd:MD_DigitalTransferOptions>
                        <gmd:onLine>
                            <gmd:CI_OnlineResource>
        <!-- The CI_OnlineResource/linkage/URL element must contain complete URL to access
                the resource -->
                                <gmd:linkage>
                         <gmd:URL>http://repository.usgin.org/uri_gin/usgin/dlio/337</gmd:URL>
                                </gmd:linkage>
                            </gmd:CI_OnlineResource>
                        </gmd:onLine>
                    </gmd:MD_DigitalTransferOptions>
                </gmd:transferOptions>
            </gmd:MD_Distribution>
        </gmd:distributionInfo>
    </gmd:MD_Metadata>
```

8.2 USGIN ISO 19139 Dataset Metadata

In the following listing, text in green is comments; XML elements are in blue, XML attributes are in red, and attribute values are in purple.

This example metadata record describes a scanned well log document.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--*****************************************************************************************
*** Example ISO 19139 Geospatial Dataset Metadata based on the USGIN v1.2 Profile
*** by USGIN Standards and Protocols Drafting Team
*** U.S. Geoscience Information System (USGIN) - http://lab.usgin.org
*** Contributors: Wolfgang Grunberg, Stephen M Richard
*** 2013-06-14
***
*** DISCLAIMER: this is not an authoritative metadata example but an aide to get started; refer
                to
*** the USGIN profile document for more specific and complete guidelines.
***
*** Validated against http://www.isotc211.org/2005/gmd (ISO 19115, CSW 2.0.2 AP ISO 1.0).
*** Follows the USGIN ISO 19139 Dataset Metadata Profile v1.2.
***
*** NOTES:
*** - Codelists:
*** Most ISO metadata profiles and applications use ISO codelists. The codeList attribute points
                to a Uniform Resource Locator (URL) that locates a resource describing the code
                list values. It can be a URN or a URL.
***
*** KEY: - M/C/O/X (Mandatory, Conditional, Optional, Not Used)
***
*****************************************************************************************-->
     <!-- USGIN ISO 19139 geospatial dataset metadata record -->
<gmd:MD_Metadata   xmlns="http://www.isotc211.org/2005/gmd"
                xmlns:gco="http://www.isotc211.org/2005/gco"
                xmlns:gmd="http://www.isotc211.org/2005/gmd"
                xmlns:gmi="http://www.isotc211.org/2005/gmi"
                xmlns:gml="http://www.opengis.net/gml"
                xmlns:gmx="http://www.isotc211.org/2005/gmx"
                xmlns:gts="http://www.isotc211.org/2005/gts"
                xmlns:srv="http://www.isotc211.org/2005/srv"
                xmlns:xlink="http://www.w3.org/1999/xlink"
                xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                xsi:schemaLocation="http://www.isotc211.org/2005/gmd
                http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd">
      <!-- (M) Unique Identifier for the metadata record - USGIN recommends using a valid
                Universally Unique Identifier (UUID) -->
  <gmd:fileIdentifier>
    <gco:CharacterString>00C02E67-F1ED-473DA240068CCB041A73</gco:CharacterString>
  </gmd:fileIdentifier>
    <!-- (M) Metadata language ISO639-2/T three letter language code - lower case. -->
```

```
<gmd:language>
  <gco:CharacterString>eng</gco:CharacterString>
</gmd:language>
                    <!-- (M) Metadata character set   - default is "utf8", USGIN requires that a
              character set code is defined to facilitate CSW servers (deegree, GeoNetwork,
              etc.). -->
<gmd:characterSet>
                    <!-- MD_CharacterSetCode names: {ucs2, ucs4, utf7, utf8, utf16, 8859part1,
              8859part2, 8859part3, 8859part4, 8859part5, 8859part6, 8859part7, 8859part8,
              8859part9, 8859part10, 8859part11, 8859part13, 8859part14, 8859part15,
              8859part16, jis, shiftJIS, eucJP, usAscii, ebcdic, eucKR, big5, GB2312}.   -->
  <gmd:MD_CharacterSetCode
  codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Charact
  erSetCode" codeListValue="utf8">UTF-8</gmd:MD_CharacterSetCode>
</gmd:characterSet>

                    <!-- (M-M) Resource type - Define if this record is a: dataset (default),
              service, feature, software, etc. -->
<gmd:hierarchyLevel>
                    <!-- MD_ScopeCode code names: {attribute, attributeType,
              collectionHardware, col-lectionSession, dataset, series, nonGeographicDataset,
              dimensionGroup, feature, fea-tureType, propertyType, fieldSession, software,
              service, model, tile}. -->
  <gmd:MD_ScopeCode
    codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Scope
    Code" codeListValue="dataset">dataset</gmd:MD_ScopeCode>
</gmd:hierarchyLevel>
                    <!-- (O-M) Resource hierarchy level name - USGIN makes this property
              mandatory to identify the USGIN resource type (see USGIN Profile, "Resources of
              Interest"). Default USGIN hierarchyLevelName CharacterString is “Dataset”.
              Encode hierarchy by including hierarchyLevelName elements for all broader
              resource categories. E.g. default should also include a
              hierarchyLevelName=”Collection” element. For services USGIN
              hierarchyLevelName.CharacterString is “Service” -->
<gmd:hierarchyLevelName>
  <gco:CharacterString>Dataset</gco:CharacterString>
</gmd:hierarchyLevelName>
<gmd:hierarchyLevelName>
  <gco:CharacterString>Collection</gco:CharacterString>
</gmd:hierarchyLevelName>
                    <!-- (M-M) Metadata point of contact - Point of contact use- to report errors,
              updates to metadata -->
<gmd:contact>
  <gmd:CI_ResponsibleParty>
                  <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
    <gmd:individualName>
      <gco:CharacterString>Stephen Richard</gco:CharacterString>
    </gmd:individualName>
    <gmd:organisationName>
      <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
    </gmd:organisationName>
```

```
<gmd:positionName>
  <gco:CharacterString>Metadata Czar</gco:CharacterString>
</gmd:positionName>
<gmd:contactInfo>
  <gmd:CI_Contact>
    <gmd:phone>
      <gmd:CI_Telephone>
        <gmd:voice>
          <gco:CharacterString>520.770.3500</gco:CharacterString>
        </gmd:voice>
        <gmd:facsimile>
          <gco:CharacterString>520.770.3505</gco:CharacterString>
        </gmd:facsimile>
      </gmd:CI_Telephone>
    </gmd:phone>
    <gmd:address>
      <gmd:CI_Address>
        <gmd:deliveryPoint>
          <gco:CharacterString>416 W. Congress St., Suite 100</gco:CharacterString>
        </gmd:deliveryPoint>
        <gmd:city>
          <gco:CharacterString>Tucson</gco:CharacterString>
        </gmd:city>
        <gmd:administrativeArea>
          <gco:CharacterString>Arizona</gco:CharacterString>
        </gmd:administrativeArea>
        <gmd:postalCode>
          <gco:CharacterString>85701-1381</gco:CharacterString>
        </gmd:postalCode>
        <gmd:country>
          <gco:CharacterString>USA</gco:CharacterString>
        </gmd:country>
          <!-- (O-M) Metadata point of contact e-mail address - mandatory if voice phone
        number not supplied -->
        <gmd:electronicMailAddress>
          <gco:CharacterString>metadata@azgs.az.gov</gco:CharacterString>
        </gmd:electronicMailAddress>
      </gmd:CI_Address>
    </gmd:address>
          <!-- (O-O) online resources - this is the online resource to contact the metadata
       person-->
    <gmd:onlineResource>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
          <gmd:URL>http://www.azgs.az.gov</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
          <gco:CharacterString>http</gco:CharacterString>
        </gmd:protocol>
        <gmd:description>
          <gco:CharacterString>Arizona Geological Survey Web
```

```
                Site</gco:CharacterString>
            </gmd:description>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
        <gmd:hoursOfService>
            <gco:CharacterString>8 AM to 5 PM Mountain Standard time (no daylight
            savings)</gco:CharacterString>
        </gmd:hoursOfService>
        <gmd:contactInstructions>
          <gco:CharacterString>Contact Steve Rauzi [Ste-ve.Rauzi@azgs.az.gov] or call Oil
            and Gas Commission Staff at Arizona Geological Survey,
            520-770-3500.</gco:CharacterString>
        </gmd:contactInstructions>
      </gmd:CI_Contact>
    </gmd:contactInfo>
              <!-- (M-M) ISO 19139 Mandatory: contact role -->
    <gmd:role>
            <!-- a responsible party with role=pointOfContact is manda-tory for metadata
          contact -->
      <gmd:CI_RoleCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_Rol
        eCode" codeListValue="pointOfContact">pointOfContact</gmd:CI_RoleCode>
    </gmd:role>
  </gmd:CI_ResponsibleParty>
</gmd:contact>
                  <!-- (X-O) Metadata should include a URL that locates a thumbnail logo for
          organizations related to the metadata origination, the organization hosting the
          catalog that returned the metadata, the organization that originated the data, and
          the organization hosting online services that provide access to the data. -->
<gmd:contact>
  <gmd:CI_ResponsibleParty>
    <gmd:organisationName>
      <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
    </gmd:organisationName>
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:address>
          <gmd:CI_Address>
            <gmd:electronicMailAddress gco:nilReason="missing">
              <gco:CharacterString>metadata@usgin.org</gco:CharacterString>
            </gmd:electronicMailAddress>
          </gmd:CI_Address>
        </gmd:address>
        <gmd:onlineResource>
          <gmd:CI_OnlineResource>
              <!-- Icon image file (e.g. tif, png, jpg, gif) for the metadata originator. This Icon
          will be displayed in search results to credit the metadata originator. -->
            <gmd:linkage>
              <gmd:URL>http://www.azgs.az.gov/logo/metadata/azgs.png</gmd:URL>
            </gmd:linkage>
              <!-- URL that locates an icon thumbnails, the CI_OnlineResource/name MUST
```

```
            be 'icon' -->
            <gmd:name>
              <gco:CharacterString>icon</gco:CharacterString>
            </gmd:name>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
      </gmd:CI_Contact>
    </gmd:contactInfo>
                <!-- none of the codelist roles are applicable -->
    <gmd:role gco:nilReason="inapplicable"/>
  </gmd:CI_ResponsibleParty>
</gmd:contact>
                <!-- (M-M) Metadata date stamp - USGIN profile requires use of
            dateS-tamp/gco:DateTime (Note this contrasts with INSPIRE mandate to use
            dateS-tamp/gco:Date). This is the date and time when the metadata record was
            created or up-dated (following NAP). -->
<gmd:dateStamp>
  <gco:DateTime>2009-11-17T10:00:00</gco:DateTime>
</gmd:dateStamp>
                <!-- (M-M) USGIN profile conformant metadata is indicated by using
            "ISO-USGIN" -->
<gmd:metadataStandardName>
  <gco:CharacterString>ISO-USGIN</gco:CharacterString>
</gmd:metadataStandardName>
                <!-- (O-M) USGIN profile version, mandatory in this profile -->
<gmd:metadataStandardVersion>
  <gco:CharacterString>1.2</gco:CharacterString>
</gmd:metadataStandardVersion>
                <!-- (O-C) Dataset Identifier - For USGIN, this is a string that uniquely
            identifies the de-scribed resource. If the resource has an identifier, it should be
            included here; if the re-source will be referenced from other metadata, it must have
            an identifier here. If the da-taset is coupled to a service, the value of the
            MD_Metadata/dataSetURI attribute is the unique resource identifier used by
            srv:coupledResource to link the service with the da-taset. For the USGIN profile,
            the MD_Distribution/transferOptions/MD_DigitalTransferOptions/
            online/CI_OnlineResource is used to specify URLs for access to the resource. -->
<gmd:dataSetURI>

    <gco:CharacterString>http://azgs.az.gov/resource/00C02E67-F1ED-473D-A240-068CC
    B041A73</gco:CharacterString>
</gmd:dataSetURI>
                <!-- gmd:local would be here, but USGIN profile convention is that each
            catalog service is specific to one language, rather that expecting clients to be able
            to process gmd:PT_Locale and gmd:LocalisedCharacterString. gmd:local
            elements may be ignored by USGIN clients -->
                <!-- (O-O) Resource spatial representation - Spatial representation
            information for the da-taset (resource). Best practice is to include metadata for
            spatial representation if the de-scribed resource is a georeferenced dataset. -->
                <!-- in this example, the described resource is a document that does not
            contain geolo-cated data items, so spatial representation is not applicable -->
<gmd:spatialRepresentationInfo gco:nilReason="inapplicable"/>
```

```
                <!-- (O-O) Resource's spatial reference system - Description of the spatial
            and/or tem-poral reference systems used in the dataset. If
            {/MD_SpatialRepresentationTypeCode = "vector") or
            (../MD_SpatialRepresentationTypeCode = "grid") or
            (../MD_SpatialRepresentationTypeCode = "tin") then count referenceSystemInfo
            >= 1) }not applicable to the document described by this record -->
<gmd:referenceSystemInfo gco:nilReason="inapplicable"/>
                <!-- ********************************************************************************* -->
                <!-- (M-M) Resource identification information - At least one of
            MD_DataIdentification (da-taset, dataset series) or SV_ServiceIdentification
            (service) is required.   -->
<gmd:identificationInfo>
                <!-- Resource Dataset or Dataset Series Identification. In practice, anything
            that is not a service is described using MD_DataIdentification -->
  <gmd:MD_DataIdentification>
                <!-- (M-M) Resource citation - Information to identify the intellectual origin of the
            content in the described resource, along the lines of a citation in a scientific
            journal. Required content for a CI_Citation element are title, date, and
            responsibleParty. If the subject of the metadata record is derived from other
            sources, such that it is not considered identical with those sources, the
            provenance should be documented in LI_Lineage//LI_Source elements. -->
    <gmd:citation>
      <gmd:CI_Citation>
                <!-- (M-M) Resource title - Use titles that inform the human reader about the
            da-taset's content; titles should be unique in the context of the metadata catalog
            that contains the metadata record. -->
        <gmd:title>
          <gco:CharacterString>
            Scanned Borehole Compensated Sonic Log for 0391, Kerr-McGee08 Navajo
          </gco:CharacterString>
        </gmd:title>
                <!-- (M-M) Resource reference date - Best practice is to include at least the
            date of publi-cation or creation of the resource. The date of the resource reported
            in the citation corre-sponds to the resource's last update version according to its
            update frequency. CI_Date content includes a date and dateType. Date for USGIN
            profile uses gco:DateTime. -->
        <gmd:date>
          <gmd:CI_Date>
            <gmd:date>
                <!-- Requires an extended ISO 8601 UTC date and time string (YYYY-MM
            DDTHH:MM:SS) Some validators require time zone as well; recommend GMT
            indicated by 'Z' -->
              <gco:DateTime>2001-12-17T09:30:47</gco:DateTime>
            </gmd:date>
            <gmd:dateType>
                <!-- dateType is mandatory per ISO19115 -->
              <!-- CI_DateTypeCode values in citation context: {creation, publication,
            revision}. -->
              <gmd:CI_DateTypeCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
                CI_DateTypeCode"
```

```
            codeListValue="publication">publication</gmd:CI_DateTypeCode>
      </gmd:dateType>
    </gmd:CI_Date>
  </gmd:date>
          <!-- (C-C) Unique resource identifier   - For USGIN purposes, this element
       content value should be considered an identifier for the citation, without any
       assumption that it will use http protocol. The identifier may be resolvable to a URL,
       if a protocol prefix specifies an identifier scheme that is resolvable (e.g. http,
       doi…), but this is not necessary for a valid document, and should not be assumed
       when processing metadata documents. IF the Ci-tation has an identifier that is
       different from the identifier for the described resource (MD_Metadata/dataSetURI),
       it must be included here. The USGIN profile requires use of MD_Identifer. If
       additional codespace and version content is associated with the identifier, it should
       be encoded as MD_Identifier/authority/ CI_Citation/ alternateTitle and
       MD_Identifier/ authority/ CI_Citation/ edition   -->
  <gmd:identifier>
    <gmd:MD_Identifier>
      <gmd:code>
          <!-- 13 digit ISBN example -->
        <gco:CharacterString>isbn:000-0-000-00000-0</gco:CharacterString>
      </gmd:code>
    </gmd:MD_Identifier>
  </gmd:identifier>
          <!-- (M-M) Resource responsible party - The citation attribute provides
       information for cit-ing the intellectual origin of the described resource, analogous to
       a citation in a scientific journal. The citedResponsibleParty should specify the
       individual(s) and organizations re-sponsible for the creation or origination of the
       resource. For most intellectual content, the responsible party is what would
       normally be considered the author of a work.   -->
          <!-- in the context of the well log described by this example, the originator
       should be considered the operator who ran the log (individualName), or the
       logging company that ran the log; the drilling operator migh also be considered as
       originator. Unfortuantely, in this case that information is missing, so the provided
       contact information is for the custodian, not the author or originator. -->
  <gmd:citedResponsibleParty>
    <gmd:CI_ResponsibleParty>
          <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
      <gmd:individualName>
        <gco:CharacterString>Steve Rauzi</gco:CharacterString>
      </gmd:individualName>
      <gmd:organisationName>
        <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
      </gmd:organisationName>
      <gmd:positionName>
        <gco:CharacterString>Oil and Gas Administrator</gco:CharacterString>
      </gmd:positionName>
          <!-- cited responsible party contact information, either a voice phone number or
       e-mail address is required. Everything else is optional -->
      <gmd:contactInfo>
        <gmd:CI_Contact>
          <gmd:phone>
```

```
            <gmd:CI_Telephone>
              <gmd:voice>
                <gco:CharacterString>520-770-3500</gco:CharacterString>
              </gmd:voice>
              <gmd:facsimile>
                <gco:CharacterString>520-770-3505</gco:CharacterString>
              </gmd:facsimile>
            </gmd:CI_Telephone>
          </gmd:phone>
          <gmd:address>
            <gmd:CI_Address>
              <gmd:deliveryPoint>
                <gco:CharacterString>416 W. Congress St., Suite
                  100</gco:CharacterString>
              </gmd:deliveryPoint>
              <gmd:city>
                <gco:CharacterString>Tucson</gco:CharacterString>
              </gmd:city>
              <gmd:administrativeArea>
                <gco:CharacterString>Arizona</gco:CharacterString>
              </gmd:administrativeArea>
              <gmd:postalCode>
                <gco:CharacterString>85701</gco:CharacterString>
              </gmd:postalCode>
              <gmd:country>
                <gco:CharacterString>USA</gco:CharacterString>
              </gmd:country>
              <gmd:electronicMailAddress>
                <gco:CharacterString>Steve.rauzi@azgs.az.gov</gco:CharacterString>
              </gmd:electronicMailAddress>
            </gmd:CI_Address>
          </gmd:address>
        </gmd:CI_Contact>
      </gmd:contactInfo>
        <!-- (M-M) ISO 19139 Mandatory: contact role   -->
      <gmd:role>
        <!-- CI_RoleCode values in this context:   {originator, principalInvestigator,
       processor, author, custodian} -->
        <gmd:CI_RoleCode
          codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
          CI_RoleCode" codeListValue="custodian">custodian</gmd:CI_RoleCode>
      </gmd:role>
    </gmd:CI_ResponsibleParty>
  </gmd:citedResponsibleParty>
        <!-- (O-C) Dataset Presentation Form - The form in which the original cited
       resource is available. Note that the citation is to the original source of intellectual
       content in the de-scribed resource, and its presentation may be different from the
       format for distribution de-scribed in the metadata. This element should be specified
       if there is a difference between the cited resource presentation format and the
       distribution format(s) listed in the distribu-tionInfo/MD_Distribution section of the
       metadata record. This information is mostly for documentation; for discovery and
```

```
        data access, the distribution formats are more relevant. -->
          <!-- in this example, the original log was acquired by a pen on a strip chart
        recorder... -->
  <gmd:presentationForm>
            <!-- CI_PresentationFormCode names: {documentDigital, documentHardcopy,
        im-ageDigital, image-Hardcopy, mapDigital, mapHardcopy, modelDigital,
        model-Hardcopy, profileDigital, profileHardcopy, tableDigital, tableHardcopy,
        videoDigital, videoHardcopy, audioDigital} -->
      <gmd:CI_PresentationFormCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_
        PresentationFormCode" codeListValue="documentHardcopy">hardcopy on
        pa-per</gmd:CI_PresentationFormCode>
  </gmd:presentationForm>
          <!-- (O-O) Resource series - Information about the series or collection of which
        the cited resource is a part. -->
    <gmd:series>
      <gmd:CI_Series>
        <gmd:name>
        <!-- Name of the publication series or aggregate dataset of which the referenced
        dataset is a part. -->
          <gco:CharacterString>AZGS well log library</gco:CharacterString>
        </gmd:name>
      </gmd:CI_Series>
    </gmd:series>
          <!-- (O-O) Resource other citation details; Recommended practice: a complete
        recommended bibliographic citation should be specified in this element. -->
    <gmd:otherCitationDetails>
      <gco:CharacterString>
        Scanned Borehole Compensated Sonic Log for 0391, Kerr-McGee08 Navajo,1967:
        Tucson, AZ, Arizona Geological Survey Well Log Li-brary, scanned log file, 1 pdf
        document.
      </gco:CharacterString>
    </gmd:otherCitationDetails>
          <!-- (O-C) Resource collective title - Title of the combined resource that the cited
        resource is part of, for example the cited resource may be a paper in an anthology,
        in which case the anthology title would be the collective title. Required if the cited
        resource is part of such a collective work. -->
    <gmd:collectiveTitle gco:nilReason="inapplicable"/>
  </gmd:CI_Citation>
</gmd:citation>
            <!-- (M-M) Resource Abstract - A free text summary of the content,
        significance, purpose, scope, etc. of the resource. Exactly one value. -->
<gmd:abstract>
  <gco:CharacterString>Digital files containing Tiff images of scanned logs. Scanned
    using Neutra scanner hardware. Well penetrates Co-conino, Hermosa, Mississippian,
    Devonian (McCraken, Aneth) and bottoms in Precambrian. Total depth 3647 feet. No
    other logs were run. Well abandoned in 1973. </gco:CharacterString>
</gmd:abstract>
            <!-- (O-O) Resource purpose - Summary of the intentions for which the
        dataset was developed. Purpose includes objectives for creating the dataset and
        what the dataset is to support. No USGIN conventions for use of this element. -->
```

```
<gmd:purpose/>
                <!-- (O-O) Resource Status -   -->
<gmd:status>
              <!-- MD_ProgressCode: For USGIN discovery metadata, values should be
          re-stricted to {completed, onGoing, deprecated}. Obsolete is synonymous with
          deprecated.   -->
  <gmd:MD_ProgressCode
     codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Pr
     ogressCode" codeListValue="completed">completed</gmd:MD_ProgressCode>
</gmd:status>
              <!-- (O-C) Resource point of contact - CI_ResponsibleParty element contains
          information for point of contact to access the resource. This information is
          mandatory for physical resources such as core, cuttings, samples, manuscripts -->
<gmd:pointOfContact>
  <gmd:CI_ResponsibleParty>
    <!-- (M-M) (individualName + organisationName + positionName) > 0 -->
    <gmd:individualName>
      <gco:CharacterString>Steve Rauzi</gco:CharacterString>
    </gmd:individualName>
    <gmd:organisationName>
      <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
    </gmd:organisationName>
    <gmd:positionName>
      <gco:CharacterString>Oil and Gas Administrator</gco:CharacterString>
    </gmd:positionName>
          <!-- (O-C) Contact Information - If a resource point of contact is required then
          (phone + deliveryPoint + electronicMailAddress) > 0. -->
    <gmd:contactInfo>
      <gmd:CI_Contact>
        <gmd:phone>
          <gmd:CI_Telephone>
            <gmd:voice>
              <gco:CharacterString>520-770-3500</gco:CharacterString>
            </gmd:voice>
            <gmd:facsimile>
              <gco:CharacterString>520-770-3505</gco:CharacterString>
            </gmd:facsimile>
          </gmd:CI_Telephone>
        </gmd:phone>
        <gmd:address>
          <gmd:CI_Address>
            <gmd:deliveryPoint>
              <gco:CharacterString>416 W. Congress St., Suite
                  100</gco:CharacterString>
            </gmd:deliveryPoint>
            <gmd:city>
              <gco:CharacterString>Tucson</gco:CharacterString>
            </gmd:city>
            <gmd:administrativeArea>
              <gco:CharacterString>Arizona</gco:CharacterString>
            </gmd:administrativeArea>
```

```
            <gmd:postalCode>
              <gco:CharacterString>85701</gco:CharacterString>
            </gmd:postalCode>
            <gmd:country>
              <gco:CharacterString>USA</gco:CharacterString>
            </gmd:country>
            <gmd:electronicMailAddress>
              <gco:CharacterString>Steve.rauzi@azgs.az.gov</gco:CharacterString>
            </gmd:electronicMailAddress>
          </gmd:CI_Address>
        </gmd:address>
      </gmd:CI_Contact>
    </gmd:contactInfo>
          <!-- (M-M) ISO 19139 Mandatory: contact role -->
    <gmd:role>
          <!-- CI_RoleCode names: ISO role codes for physical resource point of contact
        are {custodian, owner, pointOfContact}; other point of contact role codes may
        apply for other resources.     -->
                                    <gmd:CI_RoleCode
        codeList=http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_
        RoleCode codeListValue="pointOfContact">point of contact</gmd:CI_RoleCode>
    </gmd:role>
  </gmd:CI_ResponsibleParty>
</gmd:pointOfContact>

          <!-- (O-O) Resource Maintenance - This element provides information about
        the mainte-nance schedule or history of the resource (or some subset/part of the
        resource specified by the scope and scope description) described by the metadata
        record. 0 to many MD_MaintenanceInformation elements may be included.     -->
<gmd:resourceMaintenance>
  <gmd:MD_MaintenanceInformation>
    <gmd:maintenanceAndUpdateFrequency>
          <!-- MD_MaintenanceFrequencyCode names: {continual, daily, weekly,
        fort-nightly, monthly, quarterly, biannually, annually, asNeeded, irregular,
        not-Planned, un-known} -->
      <gmd:MD_MaintenanceFrequencyCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
        _MaintenanceFrequencyCode" codeListValue="asNeeded">as
        need-ed</gmd:MD_MaintenanceFrequencyCode>
    </gmd:maintenanceAndUpdateFrequency>
  </gmd:MD_MaintenanceInformation>
</gmd:resourceMaintenance>
          <!-- (O-O) Graphic overview of resource - USGIN best practice is to provide
        xlink:href URL to file if it is available online, as an attribute of the
        MD_BrowseGraphic element. If   MD_BrowseGraphic is included,
        MD_BrowseGraphic/filename character string is manda-tory. Recommended
        practice is to use the Anchor extension of CharacterString xml ele-ment from
        ISO19139, which provides a url as an attribute and a text string as a label for the
        link.   -->
<gmd:graphicOverview>
  <gmd:MD_BrowseGraphic>
```

```
            <!-- file name is required. USGIN Recommended practice is to provide a
          com-plete URL as the CharacterString value -->
      <gmd:fileName>
        <gco:CharacterString>

          http://azgs.az.gov/resource/00C02E67-F1ED-473D-A240-068CCB041A73/previe
          w.jpg
        </gco:CharacterString>
      </gmd:fileName>
      <gmd:fileDescription>
        <gco:CharacterString>preview of scanned well log</gco:CharacterString>
      </gmd:fileDescription>
            <!-- Recommended practice is to provide a registered MIME type -->
      <gmd:fileType>
        <gco:CharacterString>image/jpg</gco:CharacterString>
      </gmd:fileType>
    </gmd:MD_BrowseGraphic>
  </gmd:graphicOverview>
                  <!-- (X-X) Resource Format - This element is not used by NAP or USGIN;
          this information is encoded in MD_Metadata/distributionInfo/MD_Distribution/ in
          USGIN metadata. -->
  <gmd:resourceFormat/>
                <!-- (O-O) Resource keywords - Best Practice for USGIN profile metadata is
          to supply keywords to facilitate the discovery of metadata records relevant to the
          user. USGIN re-quires that MD_Keyword/keyword contain a CharacterString.
          USGIN best practice is to include keywords in English -->
  <!-- Theme keywords -->
  <gmd:descriptiveKeywords>
    <gmd:MD_Keywords>
      <gmd:keyword>
        <gco:CharacterString>Scanned Gamma Ray Neutron</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>NMAL</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>borehole</gco:CharacterString>
      </gmd:keyword>
            <!-- Keyword Type: allowed values from MD_KeywordTypeCode
          names:{discipline, place, stratum, temporal, theme}   -->
      <gmd:type>
                                              <gmd:MD_KeywordTypeCode
          codeList=”http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
          KeywordTypeCode”
          codeListValue="theme">theme</gmd:MD_KeywordTypeCode>
      </gmd:type>
    </gmd:MD_Keywords>
  </gmd:descriptiveKeywords>
  <!-- Temporal keywords -->
  <gmd:descriptiveKeywords>
    <gmd:MD_Keywords>
```

```
      <gmd:keyword>
        <gco:CharacterString>Frasian</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>Upper Devonian</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>Devonian</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>Paleozoic</gco:CharacterString>
      </gmd:keyword>
            <!-- Keyword Type - allowed values from MD_KeywordTypeCode names:
          {discipline, place, stratum, temporal, theme} -->
      <gmd:type>
        <gmd:MD_KeywordTypeCode
          codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
          _KeywordTypeCode"
          codeListValue="temporal">temporal</gmd:MD_KeywordTypeCode>
      </gmd:type>
    </gmd:MD_Keywords>
  </gmd:descriptiveKeywords>
                <!-- Place keywords; if a resource is not associated with a spatial location, the
          key-word 'non-geographic' should be included as a place keyoword -->
  <gmd:descriptiveKeywords>
    <gmd:MD_Keywords>
      <gmd:keyword>
        <gco:CharacterString>Arizona</gco:CharacterString>
      </gmd:keyword>
      <gmd:keyword>
        <gco:CharacterString>T41N R27E S22 NE NE</gco:CharacterString>
      </gmd:keyword>
      <gmd:type>
            <!-- Keyword Type - allowed values from MD_KeywordTypeCode names:
          {discipline, place, stratum, temporal, theme} -->
        <gmd:MD_KeywordTypeCode
          codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
          _KeywordTypeCode"
          codeListValue="place">place</gmd:MD_KeywordTypeCode>
      </gmd:type>
    </gmd:MD_Keywords>
  </gmd:descriptiveKeywords>
                <!-- (O-O) Condition applying to access and use of resource - Follow NAP for
          specifica-tion of resourceConstraints. This attribute provides information for access
          control to the described resource itself. In some situations, the metadata may allow
          a user to learn of the existence of a resource that they may not actually be able to
          access without further clearance. Constraints may be represented by
          MD_Constraint, MD_LegalConstraint, or MD_SecurityConstraint.
          Example here of license statement. -->

  <gmd:resourceConstraints>
```

```
        <gmd:MD_LegalConstraints>
          <gmd:useLimitation>
            <gco:CharacterString>License: CC-BY 4.0
(http://creativecommons.org/licenses/by/4.0/):
    Attribution — You must give appropriate credit, provide a link to the license, and indicate if
changes were made. You may do so in any reasonable manner, but not in any way that
suggests the licensor endorses you or your use.
    No additional restrictions — You may not apply legal terms or technological measures that
legally restrict others from doing anything the license permits.
    Notices:  You do not have to comply with the license for elements of the material in the
public domain or where your use is permitted by an applicable exception or limitation.
     No warranties are given. The license may not give you all of the permissions necessary for
your intended use. For example, other rights such as publicity, privacy, or moral rights may limit
how you use the material.
                    </gco:CharacterString>
          </gmd:useLimitation>
          <gmd:useConstraints>
            <gmd:MD_RestrictionCode
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resourc
es/Codelist/gmxCodelists.xml#CI_RestrictionCode" codeListValue="license"/>
          </gmd:useConstraints>
        </gmd:MD_LegalConstraints>
      </gmd:resourceConstraints>

                <!-- (O-O) Aggregation information - The citation for or name of an aggregate
            dataset, the type of aggregate dataset, and optionally the activity which produced
            the dataset. -->
      <gmd:aggregationInfo>
              <!-- MD_AggregateInformation requires either
            aggregateDataSetName/CI_Citation or aggregateDataSetIdentifier/MD_Identifier.
            -->
        <gmd:MD_AggregateInformation>
            <!-- Related dataset name -->
          <gmd:aggregateDataSetName>
            <gmd:CI_Citation>
              <gmd:title>
                <gco:CharacterString>Related Resource's Title</gco:CharacterString>
              </gmd:title>
              <gmd:date>
                <gmd:CI_Date>
                  <gmd:date>
                    <gco:DateTime>2001-12-17T09:30:47</gco:DateTime>
                  </gmd:date>
                  <gmd:dateType>
                    <gmd:CI_DateTypeCode
                        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.
                        xml#CI_DateTypeCode"
                        codeListValue="publication">publication</gmd:CI_DateTypeCode>
                  </gmd:dateType>
                </gmd:CI_Date>
              </gmd:date>
```

```
      </gmd:CI_Citation>
    </gmd:aggregateDataSetName>
          <!-- Data Set Identifier -->
    <gmd:aggregateDataSetIdentifier>
      <gmd:MD_Identifier>
        <gmd:code>

              <gco:CharacterString>00000000-0000-0000-0000-000000000000</gco:Char
            acterString>
        </gmd:code>
      </gmd:MD_Identifier>
    </gmd:aggregateDataSetIdentifier>
          <!-- (M-M) Association Type is mandatory.. -->
    <gmd:associationType>
          <!-- Use DS_AssociationTypeCode names: {crossReference, largerWorkCitation,
        partOfSeamlessDatabase, source, stereoMate}   -->
      <gmd:DS_AssociationTypeCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#DS
        _AssociationTypeCode" codeListValue="crossReference">cross
        reference</gmd:DS_AssociationTypeCode>
    </gmd:associationType>
  </gmd:MD_AggregateInformation>
</gmd:aggregationInfo>
            <!-- (M-M) Resource language - Multiple instances of this element indicate
        that the lin-guistic content of the resource is available in multiple languages -->
<gmd:language>
          <!-- use ISO639-2/T three letter language code in lower case. -->
  <gco:CharacterString>eng</gco:CharacterString>
</gmd:language>
            <!-- character encoding of resource content. Not applicable to scanned image
        -->
<gmd:characterSet gco:nilReason="inapplicable"/>
            <!-- (C-C) Topic category - A topicCategory code must be provided when
        hierarchyLevel is set to "dataset" (this example document) or "dataset series".
        Most USGIN resources will have MD_TopicCategoryCode =
        "geoscientificInformation", which is the default value for this profile.   More specific
        topic categorization should be done using keywords. -->
<gmd:topicCategory>
          <!-- MD_TopicCategoryCode names: {farming, biota, boundaries,
        climatologyMeterol-ogyAtmosphere, economy, elevation, environment,
        geoscientificInformation, health, im-ageryBaseMapsEarthCover,
        intelligenceMilitary, inlandWater, location, oceans, planning-Cadastre, society,
        structure, transportation, utilitiesCommunication} -->
<gmd:MD_TopicCategoryCode>geoscientificInformation</gmd:MD_TopicCategoryCode>
</gmd:topicCategory>
            <!-- (C-C) Resource content extent - Defines the spatial (horizontal and
        vertical) and tem-poral region to which the content of the resource applies. For
        USGIN, a ge-ographicElement/¬geographic-BoundingBox must be provided for
        any resource with con-tent related to some geographic location. If the described
        resource is not related to a geo-graphic area, the place keyword 'non-geographic'
        should be included in the descriptive-Keywords element. -->
```

```
<gmd:extent>
  <gmd:EX_Extent>
          <!-- (C-C) Resource Content extent description - Free text that describes the
        spatial and temporal extent of the dataset. Note that if geographic place names are
        used to express the geographic extent, USGIN profile specifies that these should
        be encoded using key-word with keyword type code = ‘place.’ Geographic names
        may be duplicated in EX_Extent/description.-->
    <gmd:description>
      <gco:CharacterString>Some spatial/temporal description.</gco:CharacterString>
    </gmd:description>
    <gmd:geographicElement>
         <!-- (O-C) Resource content extent bounding box -USGIN profile requires a
        geographic bounding box with bounding latitude and longitude expressed using
        WGS 84 decimal de-grees. The corner coordinates for the geographic bounding
        box must not coincide in one point, because this may result in fatal errors with
        some CSW implementations. Point loca-tions must thus be represented as tiny
        rectangles. USGIN recommended practice is to place the actual point location in
        the lower left corner of the rectangle. -->
      <gmd:EX_GeographicBoundingBox>
        <gmd:extentTypeCode>
          <gco:Boolean>1</gco:Boolean>
        </gmd:extentTypeCode>
        <gmd:westBoundLongitude>
          <gco:Decimal>-109.911001</gco:Decimal>
        </gmd:westBoundLongitude>
        <gmd:eastBoundLongitude>
          <gco:Decimal>-109.910999</gco:Decimal>
        </gmd:eastBoundLongitude>
        <gmd:southBoundLatitude>
          <gco:Decimal>34.772899</gco:Decimal>
        </gmd:southBoundLatitude>
        <gmd:northBoundLatitude>
          <gco:Decimal>34.772901</gco:Decimal>
        </gmd:northBoundLatitude>
      </gmd:EX_GeographicBoundingBox>
    </gmd:geographicElement>
  </gmd:EX_Extent>

            <!-- (C-X) Resource content extent geographic description - Not used by
        USGIN profile, use keyword with type code = ‘place’ (with thesaurus if necessary).
        -->
            <!-- (C-X) Resource content extent bounding polygon - Not used by USGIN
        profile. To improve interoperability, USGIN mandates the use of Geographic
        Bounding Box; if bounding polygons are available they may be included, but
        clients may ignore content in this element. -->
  <!--
    <gmd:geographicElement>
      <gmd:EX_BoundingPolygon/>
    </gmd:geographicElement>      -->
</gmd:extent>
              <!-- (O-O) Resource temporal extent -   -->
```

```
      <gmd:extent>
        <gmd:EX_Extent>
          <gmd:temporalElement>
            <gmd:EX_TemporalExtent>
              <gmd:extent>
              <!-- Geologic time frame example -->
                <gml:TimePeriod gml:id="IdJurassic">
                  <gml:name>Jurassic</gml:name>
              <!-- USGIN requires the beginPosition and endPosition's frame property to be
              defined. urn:cgi:trs:CGI:StandardGeologicTimeMa identifies the standard Geologic
              time scale frame. -->
                  <gml:beginPosition
                    frame="urn:cgi:trs:CGI:StandardGeologicTimeMa">203</gml:beginPosition>
                  <gml:endPosition frame="urn:cgi:trs:CGI:StandardGeologicTimeMa
                    ">135</gml:endPosition>
                </gml:TimePeriod>
              </gmd:extent>
            </gmd:EX_TemporalExtent>
          </gmd:temporalElement>
        </gmd:EX_Extent>
      </gmd:extent>
                  <!-- (O-X) Resource spatio-temporal extent - Not used. USGIN mandates
              encoding space time location with EX_TemporalExtent and
              EX_GeographicBoundingBox -->
                  <!-- (O-O) Resource vertical extent   -->
      <gmd:extent>
        <gmd:EX_Extent>
          <gmd:verticalElement>
            <gmd:EX_VerticalExtent>
              <gmd:minimumValue>
                <gco:Real>-100</gco:Real>
              </gmd:minimumValue>
              <gmd:maximumValue>
                <gco:Real>200</gco:Real>
              </gmd:maximumValue>
              <!-- Use EPSG register of geodetic parameters such as at
              http://www.epsg-registry.org/. The default VerticalCRS is World mean sea level
              (MSL): http://www.spatialreference.org/ref/epsg/5714/ (replaces
              urn:ogc:def:crs:EPSG::5714) -->
              <gmd:verticalCRS xlink:href="http://www.spatialreference.org/ref/epsg/5714/">
              </gmd:verticalCRS>
            </gmd:EX_VerticalExtent>
          </gmd:verticalElement>
        </gmd:EX_Extent>
      </gmd:extent>
    </gmd:MD_DataIdentification>
  </gmd:identificationInfo>
                  <!-- ================================================================ -->
                  <!-- (O-O) Content information - Characteristics describing the feature
              cataloguecatalog, coverage, or image data. Not applicable to this resource. -->
  <gmd:contentInfo gco:nilReason="inapplicable"/>
```

```
                <!-- (O-O) Resource distribution information - This element provides
            information to inform users how to obtain or access the described resource.
            NOTE: there are several ways el-ements can be nested within MD_Distribution,
            please read the MD_Distribution sections in the USGIN profile document.   -->
<gmd:distributionInfo>
  <gmd:MD_Distribution>
                <!-- (O-O)   Resource distribution format - Information on the format or
            physical manifesta-tion of the resource. If the resource is a physical resource, like
            a book, rock sample, paper document, the distributionFormat/MD_Format/name is
            mandatory, and must be from the USGIN distribution format codelist. Value is
            optional for this document resource. -->
    <gmd:distributionFormat>
      <gmd:MD_Format>
        <gmd:name>
          <gco:CharacterString>application/pdf</gco:CharacterString>
        </gmd:name>
        <gmd:version>
          <gco:CharacterString>8.0</gco:CharacterString>
        </gmd:version>
      </gmd:MD_Format>
    </gmd:distributionFormat>
                <!-- (O-C) Resource distributor information - USGIN differs from NAP in this
            case (but not with ISO19115) by allowing multiple distributors, and binding
            between distributors, transfer options, and formats.   -->
    <gmd:distributor>
              <!-- For USGIN profile, each distributor/MD_Distributor is a binding between
            one or more transfer options and the distributor formats that are available through
            that/those transfer options (MD_DigitalTransferOptions/onLine/CI_OnlineResource
            in particular). If different formats are available from the same distributor, or have
            different transfer options, these should be represented as different
            distributor/MD_Distributor instances. See the USGIN Profile section 'Use of
            MD_Distribution and MD_Distributor' for instructions on use of these elements. -->
      <gmd:MD_Distributor>
        <gmd:distributorContact>
            <!-- contact point for questions, comments, problems with accessing the resource
            through this distribution -->
          <gmd:CI_ResponsibleParty>
            <gmd:organisationName>
              <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
            </gmd:organisationName>
            <gmd:contactInfo>
              <gmd:CI_Contact>
                <gmd:address>
                  <gmd:CI_Address>
                    <gmd:electronicMailAddress>
                      <gco:CharacterString>metadata@usgin.org</gco:CharacterString>
                    </gmd:electronicMailAddress>
                  </gmd:CI_Address>
                </gmd:address>
              </gmd:CI_Contact>
            </gmd:contactInfo>
```

```
      <gmd:role>
      <!-- CI_RoleCode codes applicable in this context include: {resourceProvider,
      dis-tributor, pointOfContact} -->
        <gmd:CI_RoleCode
          codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
          CI_RoleCode" codeListValue="distributor">distributor</gmd:CI_RoleCode>
      </gmd:role>
    </gmd:CI_ResponsibleParty>
</gmd:distributorContact>
      <!-- (O-C) Resource distributor format - USGIN profile specifies that the
      distributionIn-fo/MD_Distribution/distributionFormat may be included in the
      document (its schema val-id...), but distribution format information must be
      duplicated in a distributionInfo/distributor/MD_Distributor/distributorFormat element
      or the content can be lost -->
<gmd:distributorFormat>
  <gmd:MD_Format>
    <!-- Use USGIN distribution format code values. See "Online resource format
    names" in USGIN Profile -->
    <gmd:name>
    <-- registered MIME type IETF RFC-3778 -->
      <gco:CharacterString>application/pdf</gco:CharacterString>
    </gmd:name>
    <gmd:version>
      <gco:CharacterString>8.0</gco:CharacterString>
    </gmd:version>
  </gmd:MD_Format>
</gmd:distributorFormat>
<gmd:distributorTransferOptions>
  <gmd:MD_DigitalTransferOptions>
    <gmd:onLine>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
                <!-- The linkage element should contain the complete URL to access
          the resource direct-ly. CI_OnlineResource requires a Linkage element that is
          a gmd:URL. -->
          <gmd:URL>http://azgs.az.gov/resource/00C02E67-F1ED-473D-A240-068CC
          B041A73/borehole_report.pdf</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
    <!-- The protocol element defines a valid internet protocol used to access the
    resource. Recommended best practice (NAP) is that the protocol mnemonic
    should be taken from the controlled list of Official Internet Protocol Standards at
    http://www.rfc-editor.org/rfcxx00.html or the Internet Assigned Numbers Authority
    (IANA) at http://www.iana.org/protocols. ‘ftp’ or ‘http’ are common values. -->
          <gco:CharacterString>http</gco:CharacterString>
        </gmd:protocol>
        <gmd:name>
    <!-- The CI_OnlineResource/name element may duplicate the file name if the URL
    is a link to a file, but it is recommended to provide a user-friendly label for the file
    for presenta-tion in a user interface.   -->
          <gco:CharacterString>borehole_report.pdf</gco:CharacterString>
```

```
            </gmd:name>
          <!-- The CI_OnlineResource/description element may include a [[CDATA] section
          with key values pairs to provide additional necessary information. USGIN
          recommendataion is to encode these using JSON syntax. -->
            <gmd:description>
              <gco:CharacterString>Downloadable PDF document</gco:CharacterString>
            </gmd:description>
          <!-- (O-C) Resource distributor online distribution function -
          CI_OnlineResource/function is required by USGIN to indicate how linkage is to be
          used. If the resource is accessible as a web service, the metadata for the service
          should be separate metadata record with the dataset(s) exposed through the
          service identified in the service metadata record as cou-pledResources. -->
            <gmd:function>
          <!-- CI_OnlineFunctionCode names: {download, information, offlineAccess, order,
          search} -->
              <gmd:CI_OnLineFunctionCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.
                xml#CI_OnLineFunctionCode"
                codeListValue="download">download</gmd:CI_OnLineFunctionCode>
            </gmd:function>
          </gmd:CI_OnlineResource>
        </gmd:onLine>
      </gmd:MD_DigitalTransferOptions>
    </gmd:distributorTransferOptions>
  </gmd:MD_Distributor>
</gmd:distributor>
          <!-- (C-C) Resource distribution transfer options - MD_DigitalTransferOptions
          provides in-formation on digital distribution of resource. See USGIN Profile 'Use of
          MD_Distribution and MD_Distributor' for instructions on use of this element.
          Details on encoding for MD_DigitalTransferOptions are above in the
          distributorTransferOptions elements description. -->
<gmd:transferOptions>
  <gmd:MD_DigitalTransferOptions>
    <gmd:onLine>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
        <!-- The linkage element should contain the complete URL to access the resource
        directly. CI_Online-Resource requires a Linkage element that is a gmd:URL. -->

          <gmd:URL>http://azgs.az.gov/resource/00C02E67-F1ED-473D-A240-068CC
          B041A73/borehole_report.pdf</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
        <!-- The protocol element defines a valid internet protocol used to access the
        resource. Recommended best practice (NAP) is that the protocol mnemonic
        should be taken from the controlled list of Official Internet Protocol Standards at
        http://www.rfc-editor.org/rfcxx00.html or the Internet Assigned Numbers Authority
        (IANA) at http://www.iana.org/protocols. 'ftp' or 'http' are common values. -->
          <gco:CharacterString>http</gco:CharacterString>
        </gmd:protocol>
        <!-- (C-C) Resource distributor online distribution application profile -
```

```
            applicationProfile is required if the CI_OnlineResource/linkage does not connect to
            a web page, and another software application is needed to use the indicated file
            resource. The applicationProfile character string should specify the software using
            the following recommended syntax: “vendor:application name/application version”,
            e.g. “Microsoft:Word/2007”, or “ESRI:ArcGIS/9.3” -->
            <gmd:applicationProfile>
              <gco:CharacterString>Adobe:Acrobat/8.0</gco:CharacterString>
            </gmd:applicationProfile>
            <gmd:name>
            <!-- The CI_OnlineResource/name element may duplicate the file name if the URL
            is a link to a file, but it is recommended to provide a user-friendly label for the file
            that could be presented in a user interface.   -->
              <gco:CharacterString>borehole_report.pdf</gco:CharacterString>
            </gmd:name>
            <gmd:description>
              <gco:CharacterString>Downloadable PDF document</gco:CharacterString>
            </gmd:description>
            <!-- (O-C) Resource distributor online distribution function -
            CI_OnlineResource/function is required by USGIN to indicate how linkage is to be
            used. If the resource is accessible as a web service, the metadata for the service
            should be separate metadata record with the dataset(s) exposed through the
            service identified in the service metadata record as cou-pledResources. -->
            <gmd:function>
            <!-- CI_OnlineFunctionCode names: {download, information, offlineAccess, order,
            search}   -->
              <gmd:CI_OnLineFunctionCode
                codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#
                CI_OnLineFunctionCode"
                codeListValue="download">download</gmd:CI_OnLineFunctionCode>
            </gmd:function>
          </gmd:CI_OnlineResource>
        </gmd:onLine>
      </gmd:MD_DigitalTransferOptions>
    </gmd:transferOptions>
  </gmd:MD_Distribution>
</gmd:distributionInfo>
<gmd:dataQualityInfo>
            <!-- In this profile, data quality information is optional, but if included, the
            mandatory content for data quality information is a free text explanation of data
            quality considerations. The
            DQ_DataQuality[1]//report[1]//explanation/CharacterString element should contain
            a free text discussion/description of data quality considerations for the indicated
            scope. The use of any specific data quality element to contain this explanation is
            arbitrary and should not be considered meaningful in this context.   -->
  <gmd:DQ_DataQuality>
    <gmd:scope>
      <gmd:DQ_Scope>
        <gmd:level>
          <!-- MD_ScopeCode values for discovery metadata: {collectionHardware,
          collec-tionSession, dataset, series, dimensionGroup, fieldSession, software,
          service, model, tile} -->
```

```
      <gmd:MD_ScopeCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD
        _ScopeCode" codeListValue="dataset">dataset</gmd:MD_ScopeCode>
    </gmd:level>
          <!-- (C-C) Data quality scope level description - Not used by USGIN -->
  </gmd:DQ_Scope>
</gmd:scope>
<report>
          <!-- This is a container for a free text discussion/description of data quality
        considerations for the indicated scope. The use of any specific data quality
        element (e.g. DQ_DomainConsistency here) to contain this explanation is arbitrary
        and should not be considered meaningful in this context.   -->
  <DQ_DomainConsistency>
    <result>
      <DQ_ConformanceResult>
        <specification gco:nilReason="notApplicable"/>
        <explanation>
          <gco:CharacterString>Quality Statement: Original paper quality was poor quality,
            parts of scan are difficult to read. Scan resolution is 300 dpi.
          </gco:CharacterString>
        </explanation>
        <pass gco:nilReason="inapplicable"/>
      </DQ_ConformanceResult>
    </result>
  </DQ_DomainConsistency>
</report>
          <!-- (C-C) Resource lineage -   count(lineage/LI_Lineage/source +
        line-age/LI_Lineage/sourceStep + lineage/LI_Lineage/statement ) >0   for spatial
        dataset and spatial dataset series. Not applicable to services. -->
<gmd:lineage>
  <gmd:LI_Lineage>
          <!-- (C-C) Data quality lineage statement - General explanation of the data
        producer's knowledge of the dataset lineage. This is probably the most useful thing
        to supply. -->
    <gmd:statement>
      <gco:CharacterString>
        This dataset is maintained by the Arizona Geological Survey. Paper logs were
        obtained from the permitee and are on file at the Arizona Geological Survey. Logs
        were scanned in 2005 by S. J. Rauzi. Digital files are archived at the Arizona
        Geological Survey.
      </gco:CharacterString>
    </gmd:statement>
          <!-- (C-C) Data quality lineage process step - An event in the development of the
        dataset. Best practice recommended for USGIN is that source association from a
        process step is to inputs to a process, and processStep associations from a
        source element link an output resource to a process step that produced it.   -->
    <gmd:processStep>
      <gmd:LI_ProcessStep>
        <gmd:description>
          <gco:CharacterString>A detailed description of the scanning process could be
            included here. </gco:CharacterString>
```

```
            </gmd:description>
          </gmd:LI_ProcessStep>
        </gmd:processStep>
            <!-- (C-C) Data quality lineage source - not applicable in this example   -->
        <gmd:source/>
      </gmd:LI_Lineage>
    </gmd:lineage>
  </gmd:DQ_DataQuality>
</gmd:dataQualityInfo>
                  <!-- (O-O) Metadata constraint information - This element specifies use
            constraints for access to the metadata record.   -->
<gmd:metadataConstraints>
                  <!-- Constraints -->
  <gmd:MD_Constraints>
                <!-- metadataConstraints/MD_Constraints/useLimitation is mandatory when
            MD_Constraints is used to specify metadataConstraints.   -->
    <gmd:useLimitation>
      <gco:CharacterString>fair use</gco:CharacterString>
    </gmd:useLimitation>
  </gmd:MD_Constraints>
</gmd:metadataConstraints>
                  <!-- Legal constraint -->
<gmd:metadataConstraints>
  <gmd:MD_LegalConstraints>
                <!-- When one of the subtypes MD_LegalConstraints or
            MD_SecurityConstraints is used, useLimitation is optional. -->
    <gmd:useLimitation>
      <gco:CharacterString>one</gco:CharacterString>
    </gmd:useLimitation>
    <gmd:accessConstraints>
                <!-- MD_RestrictionCode names: {copyright, patent, patentPending, trademark,
            license, intellectualPropertyRights, restricted, otherRestrictions} - NAP expands
            with {licenseUnrestricted, licenseEndUser, licenseDistributor, privacy, statutory,
            confidential, sensitivity}. -->
      <gmd:MD_RestrictionCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_R
        estrictionCode" codeListValue="otherRestrictions">other
        restrictions</gmd:MD_RestrictionCode>
    </gmd:accessConstraints>
    <gmd:useConstraints>
                <!-- MD_RestrictionCode names: {copyright, patent, patentPending, trademark,
            license, intellectualPropertyRights, restricted, otherRestrictions} -->
      <gmd:MD_RestrictionCode
        codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_R
        estrictionCode" codeListValue="otherRestrictions">other
        restrictions</gmd:MD_RestrictionCode>
    </gmd:useConstraints>
                <!-- (C-C) otherConstraints is a free text element required by NAP if
            accessConstraints or useConstraints is set to "otherRestrictions." -->
    <gmd:otherConstraints>
      <gco:CharacterString>
```

```
            Data only to be used for the purposes for which they were collected.
          </gco:CharacterString>
        </gmd:otherConstraints>
      </gmd:MD_LegalConstraints>
    </gmd:metadataConstraints>
    <gmd:metadataConstraints>
                <!-- Security constraints; in general these will be in applicable to USGIN resources
                -->
      <gmd:MD_SecurityConstraints>
        <gmd:classification>
                    <!-- MD_SecurtyConstraints has various optional free text values, and a
                required MD_SecurityConstraints/classification from MD_ClassificationCode
                names: {unclassified, restricted, confidential, secret, topSecret}   -->
          <gmd:MD_ClassificationCode
            codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_Cl
            assificationCode"
            codeListValue="unclassified">unclassified</gmd:MD_ClassificationCode>
        </gmd:classification>
      </gmd:MD_SecurityConstraints>
    </gmd:metadataConstraints>
                        <!-- (O-O) Metadata maintenance information - This element provides
                information about the maintenance schedule or history of the metadata record.
                -->
    <gmd:metadataMaintenance>
      <gmd:MD_MaintenanceInformation>
        <gmd:maintenanceAndUpdateFrequency>
                    <!-- Only one MD_MaintenanceInformation element may be included, with a
                required MD_MaintenanceFrequencyCode names: {continual, daily, weekly,
                fortnightly, monthly, quarterly, biannually, annually, asNeeded, irregular,
                not-Planned, unknown} -->
          <gmd:MD_MaintenanceFrequencyCode
            codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_M
            aintenanceFrequencyCode" codeListValue="asNeeded">as
            needed</gmd:MD_MaintenanceFrequencyCode>
        </gmd:maintenanceAndUpdateFrequency>
        <gmd:maintenanceNote>
          <gco:CharacterString>This metadata record has been processed by the
              iso-19115-to-usgin-19115-data XSLT to ensure that all mandatory content for
              USGIN profile has been added.2013-04-04T12:00:00</gco:CharacterString>
        </gmd:maintenanceNote>
      </gmd:MD_MaintenanceInformation>
    </gmd:metadataMaintenance>
  </gmd:MD_Metadata>
```

8.3 ISO19110 Feature Catalog example

Example encoding of entity attribute information using the xml implementation of ISO 19110. Comments in green provide some explanation.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<gfc:FC_FeatureCatalogue xmlns="http://www.isotc211.org/2005/gfc"
xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:gfc="http://www.isotc211.org/2005/gfc"
xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gml="http://www.opengis.net/gml/3.2"
xmlns:gmx="http://www.isotc211.org/2005/gmx" xmlns:gsr="http://www.isotc211.org/2005/gsr"
xmlns:gss="http://www.isotc211.org/2005/gss" xmlns:gts="http://www.isotc211.org/2005/gts"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" id="FC001"
xsi:schemaLocation="<http://www.isotc211.org/2005/gfc
http://www.isotc211.org/2005/gfc/gfc.xsd>" uuid="99ffdfd0-775a-11e0-a1f0-0800200c9a34">

    <!-- mandatory -->
    <gmx:name>
        <gco:CharacterString>Feature Catalogue for Provisional Processed CTD Data from the
Deepwater Horizon Incident Response in the Gulf of Mexico May 08, 2010 through November
15 ,2010
</gco:CharacterString>
    </gmx:name>
    <!-- mandatory 1..* -->
    <gmx:scope gco:nilReason="unknown"/>

    <!-- mandatory -->
    <gmx:versionNumber gco:nilReason="unknown"/>
    <!-- mandatory -->
    <gmx:versionDate gco:nilReason="unknown"/>
    <gmx:language>
        <gco:CharacterString>eng</gco:CharacterString>
    </gmx:language>
    <gmx:characterSet>
        <gmd:MD_CharacterSetCode
codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_CharacterSe
tCode" -"utf8">UTF-8</gmd:MD_CharacterSetCode>
    </gmx:characterSet>
    <!-- can put subCatalog here, but its unclear what the use case is -->
    <!-- mandatory -->
    <gfc:producer>
        <!-- could use xlink:href to resource contact -->
        <gmd:CI_ResponsibleParty>
            <gmd:individualName>
                <gco:CharacterString>No Name Was Given</gco:CharacterString>
            </gmd:individualName>
            <gmd:organisationName>
                <gco:CharacterString>Arizona Geological Survey</gco:CharacterString>
            </gmd:organisationName>
            <gmd:contactInfo>
                <gmd:CI_Contact>
```

```
                        <gmd:phone>
                            <gmd:CI_Telephone>
                                <gmd:voice>
                                    <gco:CharacterString>520-770-3500</gco:CharacterString>
                                </gmd:voice>
                            </gmd:CI_Telephone>
                        </gmd:phone>
                        <gmd:address>
                            <gmd:CI_Address>
                                <gmd:deliveryPoint>
                                    <gco:CharacterString>
                                    416 W. Congress St., Suite 100
                                </gco:CharacterString>
                                </gmd:deliveryPoint>
                                <gmd:city>
                                    <gco:CharacterString>Tucson</gco:CharacterString>
                                </gmd:city>
                                <gmd:administrativeArea>
                                    <gco:CharacterString>AZ</gco:CharacterString>
                                </gmd:administrativeArea>
                                <gmd:postalCode>
                                    <gco:CharacterString>85701</gco:CharacterString>
                                </gmd:postalCode>
                                <gmd:electronicMailAddress>
<gco:CharacterString>metadata@azgs.az.gov</gco:CharacterString>
                                </gmd:electronicMailAddress>
                            </gmd:CI_Address>
                        </gmd:address>
                    </gmd:CI_Contact>
                </gmd:contactInfo>
                <gmd:role>
                    <gmd:CI_RoleCode
codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_RoleCode"
codeListValue="curator">curator</gmd:CI_RoleCode>
                </gmd:role>
            </gmd:CI_ResponsibleParty>
        </gfc:producer>
        <!-- mandatory 1..* -->
        <gfc:featureType>
            <!-- the id attribute should be used as a fragment identifer in URLs for the feature
catalog citation in ISO 19115 metadata -->
            <gfc:FC_FeatureType id="fragmentIdentifer">
                <!-- mandatory -->
                <gfc:typeName>
                    <gco:LocalName>Name of the feature</gco:LocalName>
                </gfc:typeName>
                <gfc:definition>
                    <gco:CharacterString>CTD Header Information</gco:CharacterString>
                </gfc:definition>
                <gfc:code>
```

```
                    <gco:CharacterString>URI for feature type--handy if there is a
registry</gco:CharacterString>
                </gfc:code>
                <!-- mandatory. Indicates if the feature type is abstract or not. default is FALSE-->
                <gfc:isAbstract>
                    <gco:Boolean>false</gco:Boolean>
                </gfc:isAbstract>
 <!-- mandatory. pointer to containing catalog; UUID is defined in root element for this catalog-->
                <gfc:featureCatalogue uuidref="87ffdfd0-775a-11e0-a1f0-0800200c9a66"/>
                <gfc:carrierOfCharacteristics>
                    <gfc:FC_FeatureAttribute>
                        <!-- mandatory -->
                        <gfc:memberName>
                            <gco:LocalName>Name of attribute1</gco:LocalName>
                        </gfc:memberName>
                        <gfc:definition>
                            <gco:CharacterString>explanation of attribute
1</gco:CharacterString>
                        </gfc:definition>
                        <!-- mandatory -->
                        <gfc:cardinality>
                            <gco:Multiplicity>
                                <gco:range>
                                    <gco:MultiplicityRange>
                                        <gco:lower>
                                            <gco:Integer>0</gco:Integer>
                                        </gco:lower>
                                        <gco:upper>
   <!-- I guess this is how one would say as many as you want... -->
                                            <gco:UnlimitedInteger isInfinite="true">
                                                5</gco:UnlimitedInteger>
                                        </gco:upper>
                                    </gco:MultiplicityRange>
                                </gco:range>
                            </gco:Multiplicity>
                        </gfc:cardinality>
                        <!-- optional -->
                        <gfc:definitionReference>
                            <gfc:FC_DefinitionReference>
                                <gfc:definitionSource>
                                    <gfc:FC_DefinitionSource>
            <!-- most likely this will be the same as the citation for the catalog producer -->
                                        <gfc:source
                                            xlink:href="#UUID of citation for containing catalog"/>
                                    </gfc:FC_DefinitionSource>
                                </gfc:definitionSource>
                            </gfc:FC_DefinitionReference>
                        </gfc:definitionReference>
                    <!-- optional -->
                        <code>
                          <gco:CharacterString>
```

```
          URI for this attribute if it is registered somewhere
                            </gco:CharacterString>
                          </code>
                            </gfc:valueMeasurementUnit>
                    <!-- optional -->
                          <valueType>
                              <gco:TypeName>
                    <!-- its unclear how this scheme is supposed to be used for complex datatypes
-->
                                      <gco:Name>
                                          <gco:CharacterString>
          datatype name for quantifying this attribute suggest using xs:types if applicable,
but might be URI for a complex element
                                          </gco:CharacterString>
                                      </gco:Name>
                                  </gco:TypeName>
                          </valueType>
<!-- if valueType is an enumeration, there may be collection of FC_listedValues documenting
the choices    -->
                      </gfc:FC_FeatureAttribute>
                  </gfc:carrierOfCharacteristics>
                  <carrierOfCharacteristics>
                      <FC_FeatureAttribute>
                          <memberName>
                            <gco:LocalName
codeSpace="http://stategeothermaldata.org/uri-gin/aasg/xmlschema/activefault/1.2">FeatureU
RI</gco:LocalName>
                          </memberName>
                          <definition>
                            <gco:CharacterString>
                            unique identifier for this feature. Best practice is to define an http
                              URI that will dereference to a normative description of the feature
                            </gco:CharacterString>
                          </definition>
                          <cardinality>
                              <gco:Multiplicity>
                                  <gco:range>
                                      <gco:MultiplicityRange>
                                          <gco:lower>
                                              <gco:Integer>1</gco:Integer>
                                          </gco:lower>
                                          <gco:upper>
                                              <gco:UnlimitedInteger>1</gco:UnlimitedInteger>
                                          </gco:upper>
                                      </gco:MultiplicityRange>
                                  </gco:range>
                              </gco:Multiplicity>
                          </cardinality>
                          <valueType>
```

```
                                <gco:TypeName>
                                    <gco:Name>
                                      <gco:CharacterString>xs:string</gco:CharacterString>
                                    </gco:Name>
                                </gco:TypeName>
                            </valueType>
                        </FC_FeatureAttribute>
                    </carrierOfCharacteristics>
                    <!-- other attributes   -->
                </gfc:FC_FeatureType>
            </gfc:featureType>
            <!-- the catalog may document many feature types -->
</gfc:FC_FeatureCatalogue>
```

Extensible Markup Language (XML)

A Primer

1/3/2013

Table of Contents

Section 1: Introduction 2

Section 2: An XML Primer 2

Section 3: XML Attributes 4

Section 4: XML Namespaces and Prefixes 6

Section 5: XML Review 8

Section 1: Introduction

The United States Geoscience Information Network (USGIN) specifies the use of interchange formats to exchange and present data in an interoperable manner. One such interchange format is XML.

This tutorial is designed to familiarize the reader with the basics of XML, thereby permitting the reader to interact more easily with the ubiquitous XML documents used by USGIN.

Section 2: An XML Primer

XML stands for *Extensible Markup Language*, referring to the fact that XML is designed to be readily adapted for a variety of uses.

Because XML is a markup language, both the *content* of an XML document and the *structure* applied to that content are visible to the user. The visible structure of markup language documents is provided markup language elements.

A markup language *element* is demarcated by markup language *tags*. Opening tags *open* the element, and closing tags *close* the element. The *content* of a markup language element exists between the opening and closing tags of the element.

In a markup language document, elements appear like so:

```
<tag>content</tag>
```

For some, the above example may seem reminiscent of HTML. That is because HTML is a markup language that uses the same syntax as XML (HTML stands for "Hypertext Markup Language").

So XML and HTML are both markup languages. Both XML and HTML documents are structured by *elements*, and both XML and HTML elements are demarcated by *tags*.

Markup language elements are *hierarchical*: elements can be nested within elements. Once an element has been opened, everything between the opening tag and the closing tag, including elements, is part of the element demarcated by the surrounding tags.

For example, an XML document containing the contact information for a State Geothermal Data web service might appear as follows:

```
<ContactInfo>

    <Phone>
```

```
            <Facsimile>(987) 654-3210</Facsimile>
            <Voice>(123) 456-7890</Voice>

        </Phone>

    </ContactInfo>
```

The above example of nested markup language elements is analogous to folders within folders on a computer (Figure 1). But markup languages documents don't use folders; they use *elements*.

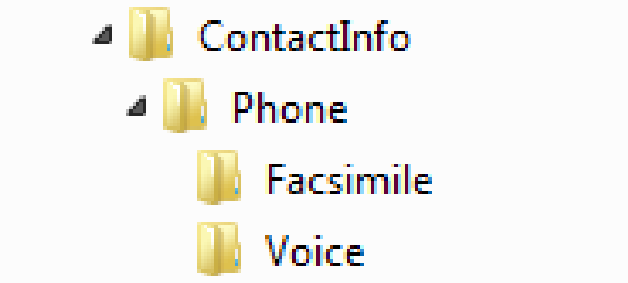


Figure 1: An example of nested folders within a computer filing system

All markup language documents begin with a *root element*: an element that contains all other elements in the document. In the case of HTML documents, this is the **HTML** element (`<html>`).

It is important to remember that, *unless otherwise specified*, elements *inherit* the characteristics of any elements within which they are nested; you will learn more about this in subsequent sections.

Markup language elements can close themselves. For example, the **horizontal rule** (`<hr>`) HTML element can be closed without a closing tag just by adding a **space** followed by a **forward slash (/)** to the opening tag of the element, like so:

```
<hr />
```

When interpreted by a browser, the **horizontal rule** tag produces a horizontal line, like so:

XML documents sometimes employ self-closing tags, usually when *attributes* modifying the element are more important than any content *within* the element.

Section 3: XML Attributes

Most XML documents tend to look a bit more complicated than previous examples suggest, owing to the presence of attributes.

Markup language attributes specify additional parameters that modify a given element; multiple attributes can modify a single element, and each attribute can specify multiple values.

In general, markup language attributes appear within the opening tag of a given element:

```
<tag attribute="value">content</tag>
```

Attributes can also be present within self-closing elements:

```
<tag attribute="value" />
```

As a practical example, the **paragraph** (`<p>`) element is used to designate a paragraph of text in an HTML document:

```
<p>The quick brown fox jumps over the lazy dog.</p>
```

When interpreted by a web browser, the sentence between the `<p>` tags of the **paragraph** element would appear as its own paragraph, like so:

The quick brown fox jumps over the lazy dog.

But the **paragraph** element can be modified by the **style** attribute, to produce all kinds of stylistic effects:

```
<p style="font-family:cursive; font-weight:bold;">The quick
brown fox jumps over the lazy dog.</p>
```

Here, the **style** attribute (highlighted in red), as applied to the **paragraph** element, specifies multiple values that will cause all content within this particular **paragraph** element to use the **cursive** font family; the text will also be **bolded**. When interpreted by a web browser, the above **paragraph** element would appear like so:

The quick brown fox jumps over the lazy dog.

Recall that elements can be nested; this means that any given element can include other elements.

For example: the **unordered list** (`<ul>`) and **list item** (`<li>`) elements are used to create lists in HTML documents. Once the **unordered list** element is opened, any **list item** elements within are interpreted by web browsers to create bullet points on a list, like so:

```
<ul>
     <li>Item 1</li>
     <li>Item 2</li>
     <li>Item 3</li>
</ul>
```

Interpreted by a web browser, this produces a bulleted list that would appear as follows:

- Item 1
- Item 2
- Item 3

As each of the three **list item** elements are nested within the **unordered list** element, anything that modifies the **unordered list** element *also* modifies the **list item** elements. For example:

```
<ul style="font-family:cursive; font-weight:bold;">
     <li>Item 1</li>
     <li>Item 2</li>
     <li>Item 3</li>
</ul>
```

Again, the **style** attribute has been used to modify the **unordered list** element (and everything nested within it), specifying **bolded** text using the **cursive** font family. This produces the following results

- **Item 1**
- **Item 2**
- **Item 3**

Note that attributes can also be used to further modify or even override the attributes of parent elements on an individual basis. For example, the **style** attribute could be used to make **Item 1** blue or Italicized, or it could be used to specify a different font family. These parameters would take precedence over those specified by the **style** attribute modifying the **unordered list** element.

XML elements can also be modified by attributes in the exact same way. For example:

```
<OperationsMetadata>
     <Operation name='GetCapabilities' />
```

```
</OperationsMetadata>
```

The **Operations Metadata** element (`<OperationsMetadata>`) contains metadata about the operations performed by a web service. Everything within this element, therefore, constitutes web service operation metadata.

In the above example, the **Operation** element (`<Operation>`) is modified by the **name** attribute (`name='GetCapabilities'`).

Here, the **name** attribute indicates that the name of one operation performed by the web service is **GetCapabilities**. Any content within the **Operation** element thereby constitutes metadata about the **GetCapabilities** operation. In this case, though, the **Operation** element is closed by a **forward slash** '/' and therefore contains no metadata.

Having made it this far, you are well on your way to understanding the basics of XML.

Section 4: XML Namespaces and Prefixes

Within XML documents, it sometimes becomes necessary to differentiate one element from another based on the context in which these elements are used.

For example, the `<table>` element is used differently between XML and HTML. Within an XML document in which both XML and HTML `<table>` elements are present, *namespaces* and *prefixes* can be used to differentiate XML table elements from HTML table elements.

Namespaces identify the context in which a given element is used, thereby differentiating that element from otherwise-identical elements.

Theoretically, *any* label can be used as a namespace, but XML namespaces should be URIs. Ideally, an XML namespace is an HTTP-capable URI that dereferences to a web-accessible description of the context identified by the namespace. Consequently, most XML namespaces in common use look like standard web addresses.

Going back to the above example, a namespace identifying HTML5 as the context for a given `<table>` element might appear as follows:

```
http://www.w3.org/TR/html5/
```

When entered into a web browser, this namespace dereferences to a description of HTML5.

Before a namespace can be used to differentiate a given element, the namespace must first be *declared* within the desired XML document. This is accomplished by means of the **XML namespace** attribute (`xmlns`); the namespace can be declared in two ways:

1. Declaring the namespace within the desired element, like so:

```
<table xmlns="http://www.w3.org/TR/html5/" />
```

2. Binding the namespace to an XML *prefix* that can be used as a proxy for the bound namespace within the element in which the binding took place (recall that nested elements inherit the characteristics of the parent elements in which they are nested)

Prefixes are short text strings that precede and differentiate markup language *tags*, *attributes*, and even the content of an element, like so:

```
<prefix:tag prefix:attribute="value">prefix:content</prefix:tag>
```

As indicated previously, prefixes serve as a proxy for a full namespace, but they can only do so within an element in which they have been bound to a namespace. In other words, a prefix cannot refer back to a bound namespace unless the prefix is used in an element in which the prefix has been bound to a namespace or which has inherited the binding from a parent element.

To give an example, let us assume that the `<table>` element we wish to differentiate via a bound namespace is nested within a `<body>` element. If we wished to declare a binding between a prefix and an appropriate namespace, we could do so either in the `<table>` element *or* within the `<body>` element within which the `<table>` element is nested, like so:

1. Binding takes place in the `<body>` element:

```
<body xmlns:html5="http://www.w3.org/TR/html5/">
     <html5:table>content</html5:table>

</body>
```

2. Binding takes place in the `<table>` element:

```
<html5:table xmlns:html5="http://www.w3.org/TR/html5/" />
```

In both examples, the **xmlns** attribute is used to bind the chosen prefix, html5, with the desired namespace, http://www.w3.org/TR/html5/; the prefix is then attached to the `<table>` element to invoke the namespace and thereby differentiate the `<table>` element from other `<table>` elements.

Typically, namespaces are bound to prefixes in parent elements and invoked via prefix where necessary in nested elements, except where it is absolutely necessary to bind a prefix to a namespace within the same element in which the namespace is invoked.

To give a more concrete example, color-coded in exactly the same manner as the previous example:

```
<wfs:WFS_Capabilities xmlns:wfs="http://www.opengis.net/wfs"
xmlns:ows="http://www.opengis.net/ows">
```

This is the root element (<WFS_Capabilities>) of the capabilities document that describes a National Geothermal Data System web feature service. Here, two namespaces are declared and bound to two different prefixes via the **xmlns** attribute:

- The **wfs** prefix is bound to the following namespace: http://www.opengis.net/wfs
- The **ows** prefix is bound to the following namespace: http://www.opengis.net/ows

Later on in the document, the following elements are nested within the root element in which the binding took place:

```
<ows:Abstract>Well header information for oil and gas wells
in Arizona.</ows:Abstract>

<wfs:Name>aasg:Wellheader</wfs:Name>
```

These prefixes indicate that the **Abstract** element is used in the context of an Open Geospatial Consortium OWS Common Implemetation Specification XML schema; likewise, the **Name** element is used in the context of an Open Geospatial Consortium web feature service. Information about the context of these elements can be found at the web location indicated by the namespaces to which the prefixes have been bound.

Section 5: XML Review

In this tutorial, we have established the following:

- XML is a markup language
- XML uses elements as basic structural components; elements are comprised by an opening tag, content, and a closing tag:

```
<tag>content</tag>
```

- Markup language elements can close themselves:

```
<tag />
```

- Elements can be *nested* within each other, much like folders on a computer:

```
<tag1>
     <tag2>

          <tag3 />
```

```
        </tag2>
    </tag1>
```

- Unless otherwise specified, nested elements inherit the characteristics of elements within which they are nested
- Markup language documents typically begin with a *root element* within which all other elements are nested
- Markup language elements can be modified by **attributes**; there can be multiple attributes per tag and multiple values per attribute:

```
<tag attribute="value" />
```

- Markup language elements can be differentiated from one another by *namespaces,* which identify a specific context in which a given element should be interpreted
- Namespaces must be declared using the **xmlns** attribute, either within a specific element or bound to a *prefix* that acts as a proxy for the full namespace
- Prefixes refer to and invoke a namespace that has been bound within a given element or the parents thereof; prefixes can modify tags, attributes, and content:

```
<prefix:tag prefix:attribute="value">prefix:content</prefix:tag>
```

With this background, it should now be possible to understand and interact with USGIN XML documents more easily.

This concludes the USGIN XML tutorial.

XML schema for Simple Feature WFS services in the National Geothermal Data System

January 23, 2015 SMR

In order to maximize the accessibility of WFS services deployed by the AASG geothermal data project, the first round of service deployments have been implemented as OpenGeospatial Consortium Simple Feature level 0 services. This document provides guidelines for construction of XML schema that implement geothermal domain feature and observation content models specified abstractly (independent of implementation approach) using lists of elements. AASG Geothermal data WFS Features consist of fields defined in a collection of Microsoft Excel content model workbooks accessible at http://www.stategeothermaldata.org/data_delivery/content_model_templates. These content models are in the process of being submitted for review and adoption for use by the wider DOE National Geothermal Data System and are presented as prototypes for US Geoscience Information Network (GIN) data services.

Implementation of the content models introduces some practical considerations that must be addressed to maximize their utility. The ESRI ArcGIS server that is being used by many data providers inserts an integer OBJECTID field as the first field in the feature content. This field should be populated with unique integers; although it can be omitted, ArcGIS WFS clients work better when an objectID field is included, so we are added this field in each feature class.

The feature geometry is included in an XML element named 'shape' that is always the last field in the XML. The geometry type for the shape is indicated in the FieldList entry for shape in the content model workbook.

The intention is that the XML schema is uniquely determined by the information listed in the FieldList worksheet in the Content model workbook. This includes the element name (which should be identical to the column heading in the template table or worksheet), a logical data type, and the xml data type that is used to implement the element in the schema. The field List includes rows for OBJECTID and shape elements, even though these are not (in general) populated in the template, but are supplied by the GIS into which the template data is loaded to deploy the WFS service.

This document provides a brief summary of provision in OGC specifications for conformance to the Simple Feature level 0 feature encoding using XML. There are two versions of the simple feature profile: 06-049r1 that uses GML 3.1.1, and 10-100r2 that uses GML 3.2.1.

Version 1 provisions

Abstracted from OGC 06-049r1 Geography Markup Language (GML) simple features profile, version 1.0

This profile prohibits use of gml:metaDataProperty elements for referencing metadata in instance documents

XML annotation required to specify the SF compliance level, placed at the top most nesting level of the schema such that the XPath expression:
`/xsd:schema/xsd:annotation/xsd:appinfo/gmlsf:ComplianceLevel`
can be used to determine the compliance level of the application schema

The following schema fragment shows how this annotation shall be declared in an application schema:

```
<xsd:annotation>
       <xsd:appinfo source=
"http://schemas.opengis.net/gml/3.1.1/profiles/gmlsfProfile/1.0.0/gmlsfLevels.xsd">
              <gmlsf:ComplianceLevel>0|1|2</gmlsf:ComplianceLevel>
              <gmlsf:GMLProfileSchema>
                     http://schemas.opengis.net/gml/3.1.1/profiles/gmlsfProfile/1.0.0/gmlsf[2].xsd
              </gmlsf:GMLProfileSchema>
       </xsd:appinfo>
</xsd:annotation>
```

The following XML fragment shows how to encode the GML application schema document's root element:

```
1 <xsd:schema
2 targetNamespace="target_name_space"
3 xmlns:prefix="target_name_space"
4 xmlns:xsd="http://www.w3.org/2001/XMLSchema"
5 xmlns:gml="http://www.opengis.net/gml"
6 xmlns:gmlsf="http://www.opengis.net/gmlsf"
7 elementFormDefault="qualified"
8 version="99.99.99">
```

In this profile, the gml:id attribute in the AbstractFeatureType is changed from optional in the GML specification [OGC 03-105r1] to mandatory in this profile.

In addition to the GML geometric property types defined in Table 6, this profile supports the following list of basic data types:
a) Integers (with optional specified value range and number of digits)
c) Measurements (that include a reference to units of measurement)
d) Character strings (with optional specified maximum length or fixed length)
e) Date
f) Boolean
g) Binary data
h) URI references to local or remote resources
i) References to other resources
j) Character strings whose values are from a code list
k) Reals (with optional specified number of digits, decimal precision and value range)

For compliance level SF-0, the attribute **minOccurs** can have the values 0 or 1 and **maxOccurs** can only have the value 1.

The following XML-Schema fragments show how to define property elements with content that is a measurement and includes a unit of measure in a GML application schema that complies with this profile:

```
1 <xsd:element name="propertyName" [minOccurs="0|1"] [maxOccurs="1"]>
2 <xsd:complexType>
3 <xsd:simpleContent>
4 <xsd:restriction base="gml:MeasureType">
5 <!-- … Optional facets described in sub-clause 8.4.4.2 … -->
6 </xsd:restriction>
7 </xsd:simpleContent>
8 </xsd:complexType>
9 </xsd:element>

SHORT FORM: <xsd:element name="propertyName" type="gml:MeasureType" [minOccurs="0|1"] [max-
Occurs="1"]/>
```

In all cases, geometry coordinates shall only be specified using the **gml:pos** for **gml:Point** or **gml:posList** elements for all other types

Simple feature v2 provisions:

Abstracted from OGC 10-100r2 Geography Markup Language (GML) simple features profile, version 2.0

Uses GML 3.2

Prohibits use of **gml:metaDataProperty** elements for referencing metadata in instance documents.

The maximum cardinality that can be declared in an application schema that complies with level SF-0 is 1. This means that the attribute **minOccurs** can have values of 0 or 1 and the attribute **maxOccurs** can only have the value of 1.

SF-0 does not allow the use of the **nillable** attribute when declaring elements in an application schema.

Application schemas that comply with level SF-0 can only declare, with the exception of geometric property types, feature properties that have simple or scalar values.

XML annotation required to specify the SF compliance level, placed at the top most nesting level of the schema such that the XPath expression:
`/xsd:schema/xsd:annotation/xsd:appinfo/gmlsf:ComplianceLevel`
can be used to determine the compliance level of the application schema

root element:

```
<xsd:schema
2   targetNamespace="target_name_space"
3   xmlns:prefix="target_name_space"
4   xmlns:xsd="http://www.w3.org/2001/XMLSchema"
5   xmlns:gml="http://www.opengis.net/gml/3.2"
6   xmlns:gmlsf="http://www.opengis.net/gmlsf/2.0"
7   elementFormDefault="qualified"
8   version="99.99.99">
```

As a GML application schema is revised, all first and second level revision changes shall require a new namespace identifier and thus a new value for the *target_name_space* placeholder

The following XML fragment imports the GML schema:

```
1   <xsd:import namespace="http://www.opengis.net/gml/3.2"
2    schemaLocation="http://schemas.opengis.net/gml/3.2.1/gml.xsd"/>
```

The following XML fragment imports the schema that declares the **ComplianceLevel** element used to declare the compliance level of an application schema:

```
1   <xsd:import namespace="http://www.opengis.net/gmlsf/2.0"
2        schemaLocation="http://…schema repository path…/gmlsfLevels.xsd"/>
```

An application schema compliant with this profile may import or include other application schemas based on this profile as long as the imported schemas have a compliance level declaration that is less than or equal to the importing schema's compliance level and are also compliant with this profile.

The following XML-Schema fragment shows how to define a feature type in a GML application schema that complies with level SF-0 of this profile:

```
1  <xsd:element name="FeatureTypeName"
2              type="[prefix:]FeatureTypeNameType"
3              substitutionGroup="gml:AbstractFeature"/>
4
5  <xsd:complexType name="FeatureTypeNameType">
6     <xsd:complexContent>
7        <xsd:extension base="gml:AbstractFeatureType">
8           <xsd:sequence>
9
10          <!-- … zero or more element declarations as described in 8.4.4 … -->
11
12          </xsd:sequence>
13       </xsd:extension>
14    </xsd:complexContent>
15 </xsd:complexType>
```

In addition to the GML geometric property types, this profile supports the following list of basic data types:

a) Integers (with optional specified value range and number of digits)
a) Measurements, that include a reference to units of measurement
b) Character strings (with optional specified maximum length or fixed length)
c) Date
d) Boolean
e) Binary data
f) URI references to local or remote resources
g) References to other resources
h) Character strings whose values are from a code list
i) Reals (with optional specified number of digits, decimal precision and value range)

The use of the attribute **nillable** is prohibited in compliance level SF-0 but is allowed for levels SF-1 (see 9.2.4) and SF-2.

Although XML-Schema [W3C XML-1] allows **minOccurs=maxOccurs=0**, this profile prohibits this combination.

The following XML Schema fragment defines the encoding of a property element whose content shall be from a list of values stored in one or more external dictionaries.

```
1 <xsd:element name="propertyName" [minOccurs="0|1"] [maxOccurs="1"]>
2    <xsd:complexType>
3       <xsd:simpleContent>
4          <xsd:restriction base="gml:CodeType">
5             <xsd:attribute name="codeSpace" type="xsd:anyURI" [use="optional"]
6                            default="…some uri…" | fixed="… some uri…"/>
7          </xsd:restriction>
8       <xsd:simpleContent>
9    </xsd:complexType>
10 </xsd:element>
```

default -- an instance document may point to an alternate value dictionary at runtime using the **codeSpace** attribute. **fixed** -- an instance document shall use values from the specified dictionary.

If the type is **xsd:decimal** then the facets *totalDigits* and *fractionDigits* listed in lines 5 and 6 shall be specified. If the type is **xsd:double**, then the facets *totalDigits* and *fractionDigits* listed in lines 5 and 6 shall not be specified.

Application schemas that comply with level SF-0 of this profile may freely use XML comments and annotations, using the **annotation** element, wherever they are legally allowed by XML-Schema and GML

NGDS Catalog Documentation

About

The National Geothermal Data System (NGDS) is a catalog of documents and datasets that provide information about geothermal resources located primarily within the United States (although information from other parts of the world is also included).

This complete and current catalog of available data, which is funded by the Department of Energy's Geothermal Technology Office, is designed to accelerate the development of U.S. geothermal resources, and can be used to:

- Determine geothermal potential
- Guide exploration and development
- Make data-driven policy decisions
- Minimize development risks
- Understand how geothermal activities affect your community and the environment
- Guide investments.

If you are new to this site, learn more about how you can get started using NGDS.

Below you will find more information about how the NGDS is typically used, who contributes data, and where it is hosted.

How is NGDS Typically Used?

The NGDS can be used in many ways, depending on your needs and interests. Generally, the NGDS is used by:

- Agencies, businesses, and researchers who wish to use the documents and datasets
- Individuals who want to contribute additional data
- Web developers who want to create custom applications that interact with NGDS data.

Learn more about typical uses and how you can get started using NGDS.

Who Contributes Data to NGDS?

NGDS data records are contributed by academic researchers, private industry, and state and federal agencies (including all fifty State Geological Surveys). Initially, five groups digitized their data and made their data records visible through the NGDS.

View a complete list of our NGDS data contributors on our Partners page or learn more about how to submit your data to the NGDS.

Where is NGDS Data Hosted?

Those who contribute data records to the NGDS host their data using their own computing resources according to the standards and protocols proposed by the U.S. Geoscience Information Network (USGIN).

The NGDS aggregates information about each of the documents and datasets, and makes these information records, which include direct access to most of the datasets, available to the public through the [[Map]] and [[Library]] interfaces in the NGDS.

This system allows NGDS contributors to retain complete control of their data, and allows the NGDS to be comprehensive yet scalable and decentralized.

If you're thinking of contributing documents or datasets to NGDS, learn more about the how to submit your data or read about our data models, exchange methods, and metadata.

What Data Formats are Available?

The documents and datasets you can find through the NGDS are delivered in the following ways:

Standardized, interoperable datasets that conform to interchange specifications are delivered through web services. These datasets are considered to be NGDS "Tier 3" structured data can be processed automatically by computer software with minimal user input. This type of data can be used by many different programs and in conjunction with any compatible datasets. For more information on data tiers and web services, please see the Data Models, Exchange Methods, and Metadata page.

- Documents and non-standardized datasets.

Although some of the data is available in non-standardized formats currently, the goal of the NGDS is to increasingly offer data that has been converted to standardized formats. To do this, we have adopted standards, protocols, and formats whenever possible.

What Standards Does the NGDS Meet?

NGDS metadata, which is used to describe each document or dataset, complies with International Organization for Standardization (ISO) standards. NGDS data and metadata can be transferred according to Open Geospatial Consortium (OGC) protocols such as Web Map Service (WMS), Web Feature Service (WFS), and Catalog Service for the Web (CSW).

Learn more about the NGDS data and metadata standards.

What Tools are Available to Access NGDS Data?

NGDS supports the popular ESRI Map Service protocol, so people may access the data through a variety of proprietary and free/open-source applications and software. See our Resources section for a listing of [[resources, models, and tools]] you can use to interact with NGDS data.

Getting Started

The National Geothermal Data System (NGDS) can be used in a variety of ways, but is typically used for:

- Finding U.S. geothermal documents and datasets
- Contributing new data
- Developing custom applications based on NGDS data

Below you'll find more information that will help you get started using the NGDS.

Finding and Using Data

The NGDS is commonly used by state and federal agencies, businesses interested in geothermal resource development, and individuals who are interested in knowing more about geothermal resources. Here's how you can use the NGDS to find and use our data:

- Use our Map to [[locate data records geographically]].
- Visit the Library to [[filter data records by keywords and other criteria]] such as category, author, or data type.
- The Resources section provides a list of [[other free and open-source applications]] that can also be used to interact with the NGDS data.
- The Help section is available to answer questions on the specifics of how to use the site Search, Map, or Library.

Contributing Your Data to NGDS

If you have information on geothermal resources, please consider making it available through the NGDS. Documents, datasets, or services that you catalog in the NGDS can be made openly accessible to the public, can remain privately accessible, or can be visible only to others within the contributing agency.

Here's how the process works:

- When you catalog your data with the NGDS, the data itself remains hosted on your own computing resources (or you can make arrangements for hosting with other NGDS nodes).
- A record of your data will be entered into the NGDS catalog through metadata records; these metadata records can then be accessed via the [[Map]] and [[Library]]. NGDS metadata describes a given data resource, any restrictions on its use, and how to obtain it.

Learn more about how to submit your data to the NGDS or read about our data models, exchange methods, and metadata.

Developing a Custom Application Using NGDS Data

NGDS is built on United States Geoscience Information Network (USGIN) standards, which allows Web developers to easily build innovative custom applications that use NGDS data. Some examples of third party applications might include data gap analysis, financial analysis, and other useful data manipulation.

Learn more about our [[data models, exchange methods, and metadata]] or [[find tools, forums, and developer communities]].

NGDS Exchange Methods

Metadata

All NGDS data records are organized and accessed using metadata. Metadata describes NGDS resources to support finding resources, evaluating them for a user's intended purpose, and obtaining the data. Each metadata record describes a resource that is considered an individual product with an author, editor or compiler, created with some particular purpose. The metadata record may describe a variety of 'distributions' or procedures for accessing the resource, such filling out an order form, clicking on a link, or linking to a web service with an appropriate client.

After hosting data using their own computing resources, NGDS data contributors submit one or more metadata records describing their data to an NGDS node. Metadata from NGDS nodes is harvested by the NGDS portal, which provides a single entry point to search the entire NGDS catalog (the catalog contains metadata describing information resources distributed across the various nodes in the network). Metadata records provide information necessary to obtain data through various distributions (file formats, web service protocols), allowing you to select the most appropriate way to access the software you are using. Metadata records in this catalog can be used to locate and access the data resources they describe.

Metadata records published by NGDS nodes are harvested by the central portal, which provides a web site for searching across all NGDS content, enabling users to discover NGDS data resources and learn how to access them. For more information about metadata, see the USGIN Metadata Tutorial.

Once the desired NGDS data resource is located using metadata that appears in the NGDS catalog, a user has several options for accessing that resource. Access options depend on the structure of the data resource. Typically, highly structured resources offer broader options for access and automated analysis.

Data Access Options

NGDS is uses the World Wide Web as a platform for accessing and exchanging data. Typical web-based data exchange methods are as follows:

- **File-based approaches**: Data is accessed in discrete digital packages called *files*. A file is located at specific URL (Web location). The content of a file can be structured to enable machine processing, but the file is only accessible as a unit.
- **Web application-based approaches**: Data is made available to users via Web-browser-based software. These applications typically provide a human user with functions to explore data with various kinds of visualizations, and in some cases to select particular data and download in files. These applications are typically tightly coupled to specific data sources.
- **Service-based approaches**: Data is available through *web services,* which provide interfaces for machine-to-machine interaction for finding, selecting and accessing data. Web services

require more sophisticated client and server software stacks and rigid quality control, but allow a greater level of automation in the data access and utilization process, fully taking advantage of interoperable interchange protocols and formats.

What Are Web Services?

A web service is a way to communicate data over the internet. Using web services allows NGDS data to be used in a variety of free-and-open-source applications, making it more freely available and easier to use.

NGDS purports to use web services with well-documented community specifications for service protocols and data interchange formats. Though web services require more sophisticated client and server software stacks and rigid quality control, web services have a number of advantages that meet the requirements of NGDS:

- Web services are web-accessible
- Data available as a web service can be consumed by web applications
- Web services are read-only and preserve data ownership
- A number of web service protocols are free-and-open-source, meaning that web services using these protocols can be hosted and accessed without proprietary software

Most NGDS web services provide functions to get data or metadata. The NGDS Node-in-a-Box (NIAB) software stack supports the creation and publication of Open Geospatial Consortium (OGC) compliant web services from data assembled in Excel spreadsheets.

Web Service Protocols in use with NGDS

- **WMS**: OGC Web Map Service (WMS). Provides raster images of georeferenced features for display in map views.
- **WFS**: OGC Web Feature Service (WFS). Provides structured data for georeferenced features including descriptive attributes.
- **ESRI Map Service**: Provides georeferenced rasters and features described by attributes. Requires ESRI ArcGIS server software.
- **CSW Applications**: OGC Catalog Service for the Web (CSW). Helps applications access metadata records used to locate NGDS information resources.
- **OPeNDAP**: Provides raster datasets as numeric array data.

As implemented in NGDS, these services provide the user read-only access to data published through the services. Transactional implementations allowing users to change content of the services are possible, but these are outside the scope of current NGDS specifications.

Interoperability

NGDS facilitates access to data conforming to the three aforementioned exchange methods through publication using interoperable protocols and formats. This enables greater automation of information

exchange, processing, and analysis, allowing human users to focus on understanding and learning from the data.

Data is categorized in one of three *interoperability tiers*; the higher the tier, the greater the level of interoperability. These tiers roughly conform to whether data can be accessed through a web service or must be downloaded or viewed as a file.

Tier 1: Unstructured data (text, images, or sound)

The simplest and most common access to resources is provided by simple Web links that result in a file download. Information contained in files can be accessed by users who have software that can recognize and open these files. This is the standard model for files accessible on the web, supported by HTTP servers and desktop web browser software.

Unstructured data requires user interpretation before it can be used for analysis. Users can utilize the information if they can understand the encoding and language, but the system provides no support for this understanding, and little or no automation is possible. Audio files must be transcribed; text files must be parsed and mined for data that is then broken down and structured in ways that can be processed by computers; images must be scanned, interpreted, and often georeferenced. Preparing Tier 1 data for analysis can be a painstaking and time-consuming process.

Tier 2: Structured data that does not conform to an NGDS schema

Tier 2 interoperability indicates that information content is structured (consistently organized) in a spreadsheet or database file such that it is amenable to computer processing; that said, Tier 2 data does not use a shared, documented interchange format. Data in this tier must be transformed by the data consumer on a case-by-case basis for integration with other datasets, requiring them to study each new data source to figure out how to extract the information they need. Obtaining data in a structured format is a step towards interoperability because once the format is understood, computer programs can be instructed to extract the desired information.

Tier 3: Structured data that conforms to an NGDS schema

Tier 3 data is structured data that conforms to an NGDS information exchange. Data that is published according to the exchange specification (content model, interchange format, service protocol) is interoperable with any other data published using that exchange. This is referred to as Tier 3 interoperability.

Creating Interoperable Data

Tier 3 data structured using a content model is among the most highly interoperable data found on the web. Tier 3 data can be accessed and analyzed automatically through a number of web services. NGDS uses content models based on USGIN-adopted standards and protocols to structure data to the Tier 3 format.

USGIN: Powering NGDS

The goal of USGIN is to facilitate the development of interoperable geoscience data-sharing networks that improve public access to geoscience information and aid scientific queries. For more information visit the USGIN website.

USGIN proposes:

- Protocols and standards.
- Profiles for the utilization of USGIN protocols and standards.
- Development tools to build distributed data-sharing networks using open-source software and existing World Wide Web infrastructure.

USGIN Architecture

To promote interoperability, USGIN adopts existing standards, protocols, and formats whenever possible. These include:

- Standards for metadata: ISO 19115, ISO 19119, and ISO 19139
- Protocols for data exchange:
  - Open Geospatial Consortium (OGC), Web Map Service (WMS), Web Feature Service (WFS), and Catalog Service for the Web (CSW)
  - ESRI Map Service
  - OPeNDAP
- GeoSciML-Portrayal schemas for geologic map data.

NGDS is based on protocols and standards adopted by USGIN. All NGDS content models are structured according to USGIN standards.

Publish Data

How to Submit Data to the NGDS

NGDS Node in a Box

The NGDS Node-in-a-Box (NIAB) is a software package designed to enable quick deployment of an NGDS node. The software installation includes a CKAN instance along with NGDS-specific extensions for conformance with NGDS metadata requirements, support for NGDS information exchanges, and publication of data through OGC services using a linked GeoServer instance. The NGDS NIAB CKAN extension provides users with a web-accessible user interface designed to help NGDS contributors submit and manage metadata records.

Installing the NIAB

Instructions on installing and configuring a NIAB instance are available with the following document:

http://ngds.github.io/install-and-run/NGDS_Installation_Documentation.htm

This document and other helpful information, including community conversation, are available at NGDS CKAN GitHub.

Finding an Existing NIAB

In addition to installing and configuring a specific Node-in-a-box instance, a registered user may contribute data using a public account on an existing Node-in-a-box if permission is granted from the Node-in-a-box owner. The data contributor is responsible for contacting the owner of the NIAB for permission. A listing of available NIABs, owners, and contact information is in development.

Overview: Becoming an NGDS Node

To become an NGDS node, one must place file-based information resources in a web-accessible location and then submit metadata describing these resources to the NGDS metadata catalog. Various workflows are available to create and submit metadata:

- Log in to an existing registered instance of the NGDS Node-in-a-Box (NIAB) and add metadata records through the provided interface.
- Set up and register an NGDS NIAB instance. This requires:
    1. Installing and configuring the NGDS NIAB software (see Installing the NIAB section above).
    2. Registering an NGDS NIAB instance with the NGDS portal (main aggregator) via email: metadata@usgin.org.
    3. Creating metadata records using the interface provided by your NGDS NIAB instance.
- Deploy some other metadata management tool that provides a CSW interface to NGDS-conformant ISO metadata for registered resources, and registering the catalog service with the NGDS central portal for harvesting. This will require contacting system adminstrators via email at metadata@usgin.org to harvest your CSW into the main aggregator at http://www.geothermaldata.org.
- Create NGDS-conformant ISO metadata records using any other tool and place them in a web-accessible folder that is registered with the NGDS portal for harvesting. Contact system adminstrators via email at metadata@usgin.org for more information.

How to Submit Your Metadata Records via a pre-installed NGDS Node-in-a-Box (NIAB)

Data providers can publish their data to NGDS via any NGDS Node-in-a-Box instance for which they have an account with data publication permission.

Connect to an NGDS NIAB using a web browser and click **Contribute** and populate the appropriate forms to submit a metadata record to the catalog. During this process, either upload a resource or specify the web location of an existing web-accessible data resource. Those providing data to a NIAB that is managed by another entity should make appropriate arrangements with the managers of the node

regarding future edits to that data; those who manage their own NIAB have complete control over the data submitted then harvested to the main aggregator for larger use in the system. The nodes (NIAB installations or other node types) will be harvested at intervals, so any edits made at a node will be reflected in the system with the next scheduled harvest.

Step 1: Create a Collection

The **Contribute** page leads to the **Create a Collection** form; as the name implies, populating this form will create a collection for a single or multiple metadata records. This is the first step in the file or resource upload process, which will create metadata in the system for that resource. The fields on this form are as follows:

- **Title**: Collection title
- **URL**: The URL at which this collection will be found. This URL is automatically generated when the **Title** field is populated, but it can be modified manually
- **Description**: A description of the collection and the metadata records stored in it
- **Tags**: Keywords that can be used to search for this collection and any metadata records stored in it
- **License**: The license under which any dataset(s) described by any metadata record(s) within this collection are published; among other things, the license determines the manner in which the dataset described by the metadata in this collection can be reused. Information on supported licenses can be found at the following locations:
    - Creative Commons
    - GNU Free Documentation License
    - Open Data Commons
    - UK Open Government License

When finished, click **Next: Add Data**.

Step 2: Add Data

Clicking **Next: Add Data** on the **Create Collection** form leads to the **Add Data** form. Populating this form creates a metadata record describing a specified data resource; this metadata record is added to the collection created in Step 1.

Four types of data may be upload. See Data Models, Exchange Methods, and Metadata page for helpful text on determining categories (tiers) of data which may be uploaded.

- **Structured Resource**: Tier 3 structured and standardized data (CSV files). Files submitted here must be “schema-valid”, which refers to the file as is conforming to the standards of a given information exhange schema. The blank Excel files for a given information exchange are found at http://schemas.usgin.org/models/. When data conforms to these specifications (field headings, data types, etc.) the file can be validated at http://schemas.usgin.org/validate/cm before up-loading to the node as a tier 3, structured dataset. After file upload, the option exists to publish the data as OGC web services (WFS and WMS), which have the greatest utility in the system. Once a “schema-valid” file is

uploaded, one must choose the appropriate information exchanges used from the **Content Model** drop-down lists.

- **Unstructured Resource**: Tier 1 and 2 data (text, images, structured (but not standardized) files like CSV or Excel, PDF files, etc.)
- **Offline Resource**: Metadata record would describe resources which do not exist online such as a core repository.
- **Data Service**: Register already published web services such as WMS, WFS, CSW, and OPeNDAP by providing the GetCapabilites or other URL in the **Resource** field.

When the appropriate fields have been populated:

- Click **Save & add another** to add another metadata record to the collection; this will repeat the steps described in Step 2.
- Click **Next: Additional Info** to finalize process by adding collection metadata.

Step 3: Collection Metadata

Clicking **Next: Additional Info** leads to the **Collection Metadata** form. Here, provide metadata that applies to the collection created in Step 1 and the metadata records within it created in Step 2.

The **Collection Metadata** form includes the following fields:

- **Visibility**: Determines whether or not this collection and the metadata record(s) within it are visible to the public
- **Authors**: The authors of the data resource(s) described by metadata record(s) in the collection
- **Maintainer**: The individual, agency, company, or organization responsible for the data resource(s) described by metadata record(s) in the collection
- **Spatial Location**: A rough approximation of the features in the data resource(s) described by metadata record(s) in the collection
- **Category**: A description of the type of data resource(s) described by metadata record(s) in the collection
- **URI**: A *uniform resource identifier* for the collection; see the USGIN URI Tutorial for more information
- **Quality**: A rough assessment of the quality of the data resource(s) described by metadata record(s) in the collection
- **Lineage**: The provenance of data resource(s) described by metadata record(s) in the collection
- **Status**: The workflow status of data resource(s) described by metadata record(s) in the collection
- **Publication Date**: The date on which data resource(s) described by metadata record(s) in the collection were published
- **Language**: The language in which the data resource(s) described by metadata record(s) in the collection are written

- **Dataset Extent**: The physical extent of features in the data resource(s) described by metadata record(s) in the collection; click the **Draw a rectangle button** ( ) to draw a box indicating the dataset extent

When done, click **Finish**. Doing so will finalize this collection and all metadata records within it; these records will be added to the catalog.

For **Structured Resources**, continue with the following one-click step to publish the resource as OGC WFS and WMS services:

Click **Publish as OGC.**

Curate your data contributions

Anyone can contribute data to NGDS by establishing web-accessible data resource(s) and then submitting metadata record(s) to a registered NGDS node. Metadata submitted to a registered node will be harvested by the NGDS portal, thereby becoming part of the NGDS catalog. Participants may also contribute data by installing and configuring a Node-in-a-box built on the CKAN platform. Users who choose to install the Node-in-a-box will independently self-host their own data, making it available to the NGDS network through common standards and protocols. NGDS offers numerous opportunities to developers who want to use this data to create specific applications. Please see the Develop with NGDS page for more information.

Those who wish to submit data to the NGDS should view the How to Submit Data to the NGDS page.

Those who wish to learn more about NGDS exchange methods, standards and protocols should view the Data Models, Exchange Methods, and Metadata page.

Retain Data Ownership

NGDS is designed from the ground up to allow contributors/nodes to maintain ownership and control of their data. It is the responsibility of the data provider to maintain the information they publish and look after its long-term stewardship.

Metadata records listed in the NGDS metadata catalog are publicly accessible, allowing anyone to learn of the existence of an NGDS information resource and the license under which the described resource can be viewed, distributed, and reused. Information on supported licenses can be found at the following locations:

- Creative Commons
- GNU Free Documentation License
- Open Data Commons
- UK Open Government License

Tools & Applications

Once you've used the National Geothermal Data System (NGDS) Map or& Library to locate the data records you are looking for, you can use the resources, models, and tools described here to interact with the data.

Here are the categories of tools and resources available:

- Geothermal exploration
- Financial
- Geographic Information Systems (GIS) and data applications

Have you developed a custom application to analyze NGDS data? If you have used NGDS data to develop a third party application that could also benefit others, pleaseshare it with us so we can include it here.

Geothermal Exploration

The following tools and resources can be used to facilitate your geothermal exploration.

- NREL Geothermal Prospector: A Web-accessible mapping tool developed for the geothermal power industry. This tool is designed to provide easy access to geothermal resource datasets and other data relevant to utility-scale geothermal power projects. Geothermal Prospector can load NGDS Web Map Services (WMS) and Web Feature Services (WFS) to facilitate exploration - e.g., Reservoir models.
- NGDS Data Explorer: A Web application designed to load and display NGDS WFS from a text or location search. When a WFS has been loaded, a user can then export attributes (standardized data) from this service as an HTML table, CSV file, or Microsoft Excel workbook. Please note that the user can load more than one WFS at a time.

Financial

The following tools, models, and resources can be used to predict and provide information about the costs of geothermal development.

- GeoFRAT: Geothermal Financial Risk Classification and Assessment Tool (GeoFRAT) is an NGDS application that helps users assess the resource potential and financial risks that may be associated with producing electricity from a given geothermal resource. When used in conjunction with other widgets and tools such as GETEM, GeoFRAT can help you conduct detailed risk analysis for a geothermal resource development project.
- Geothermal Electricity Technology Evaluation Model (GETEM): This tool models the estimated performance and costs of currently available U.S. geothermal power systems. GETEM can analyze and evaluate currently available technologies and estimate the cost of certain technologies five to 20 years in the future. The model is intended to help the

Geothermal Technologies Office (GTO) determine which proposed RD&D programs and projects offer the greatest prospective return on investment (ROI).

Geographic Information Systems (GIS) and Data Applications

These Web, desktop, and GIS software applications can be used to access, explore, visualize, and analyze NGDS data. Some GIS software is free or open source.

Free and Open-Source Desktop GIS Software

- gvSIG: Desktop GIS application designed for storing and analyzing data in compliance with Open Geospatial Consortium (OGC) standards.
- GrassGIS: Desktop GIS software suite used for geospatial data management and analysis, image processing, graphics and maps production, spatial modeling, and visualization. GrassGIS runs on Mac OSX, Linux, and Windows.
- QGIS: Desktop GIS software available under the GNU General Public License. QGIS is an official project of the Open Source Geospatial Foundation (OSGeo). QGIS runs on Linux, Unix, Mac OSX, Windows, and Android, and supports numerous vector, raster, and database formats and functionalities.
- uDig: Desktop GIS software available under the Eclipse Public License 1.0 (EPL-1.0) and BSD 3-Clause License. uDIG is a standalone application that provides a Java solution for desktop GIS data access, editing, and viewing. uDIG runs on Windows, Mac OSX, and Linux.Proprietary Desktop GIS Software.
- ArcGIS: Desktop GIS software designed to facilitate geoprocessing, geospatial data editing, and cartographic production.
- ArcGIS Explorer: ArcGIS Explorer Desktop is a free GIS viewer that gives you an easy way to explore, visualize, and share GIS information. This free software supports adding layers from ESRI's ArcGIS online catalog, KML files, and GIS Services.
- ArcGIS Online - ArcGIS Online is a cloud-based, collaborative content management system for maps, applications, data, and other geographic information. ArcGIS Online supports ArcGIS Server Web Services, OGC WMS, KML and CSV.

Data Applications

- GeoScene 3D: GeoScene3D is an application for 2D and 3D visualizing, interpreting, editing and publishing geological data, including drill or bore holes, terrain, and geologic units. This application is intended for 3D viewing, so there is not an extensive base map selection. Although additional steps are required to view Web map services in this application, it is WMS-aware.
- Geozilla: A free application that provides 3D terrain views with Web map service overlays.
- Google Earth: A free application that provides 3D terrain views with Web map service overlays.
- Worldwide Telescope: A free application that provides 3D terrain views. This application does not yet support WMS, but it can consume .CSV files to provide visualizations events as data points over time. These data points can be assigned different colors and sizes for

further visualization of data, though doing so can require extensive preparation. See more information on importing data.

State Geological Survey Web Map Applications

- NREL Geothermal Prospector - http://maps.nrel.gov/gt_prospector
- GETEM, DOE Geothermal Technologies Office - http://www1.eere.energy.gov/geothermal/getem.html
- GeoFrat, Energy and Geoscience Institute at the University of Utah - http://geofrat.egi.utah.edu/
- University of Nevada, Reno Interactive Geothermal Maps - http://www.nbmg.unr.edu/Geothermal/InteractiveMaps.html
- Washington State DNR Resources Map Application - https://fortress.wa.gov/dnr/geology/?Theme=geothermal
- Oregon DGMI Geothermal Data Map Application - http://www.oregongeology.org/sub/gtilo/index.htm
- West Virginia Geothermal Interactive Map - http://ims.wvgs.wvnet.edu/WVGeothermal/viewer.htm
- AASG - http://search.stategeothermaldata.org/
- SMU Geothermal Databases - http://geothermal.smu.edu/gtda/
- USGS Geothermal - http://energy.usgs.gov/OtherEnergy/Geothermal.aspx, http://viewer.nationalmap.gov/viewer/
- Geothermal Data Repository - http://gdr.openei.org/

Help

Here you will find additional information that can help you successfully use the following site features to access data through the National Geothermal Data System (NGDS):

- Search feature
- [[Map interface]]
- [[Library]]
- Managing Datasets
- [[Software for NGDS]]
- Frequently asked questions

Search Feature

This site's search feature provides access to a catalog of available geothermal resource data by indexing a database of metadata about those resources.

Metadata is basically "data about the data" – it is keywords and terms that the data contributors have used to describe each data resource. When you type a search word or phrase into the search box, the

NGDS compares it to the metadata records and returns the most appropriate records. Information on how to access a given data resource is contained in the metadata records returned in the search results.

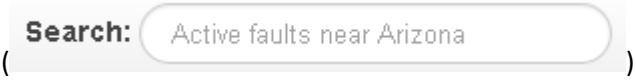
Map Interface

The [[map interface]] is typically used when you want to see a geographic distribution of data sites that may be of interest to you.

You can use the map to search the NGDS either by typing a keyword into the search box or by using your mouse to draw a "bounding box." A bounding box delineates geographic parameters on the map and the system retrieves the records that fall within those parameters.

Search results are displayed in list form and as colored circles on the map.

Map controls are as follows:

- **Search by Keyword** (

): Enter search terms and press the **Enter** key; the results will appear in list form beneath the search field as well as on the map.
- **Zoom In** ( ): Zooms in on the map.
- **Zoom Out** ( ): Zooms out.
- **Draw a Rectangle (or "Bounding Box")** ( ): Use this tool to search for results within a given area by drawing a rectangle on the map. Any record describing features within the bounding box you drew will appear as search results.
- **Zoom to Extent** ( ): Zoom to the original extent of the map.
- **Active Layer Menu** ( ): Mouseover for a toggle-able list of active layers in the map. This menu is populated as users choose to add available WMS resources to the map from their search queries.

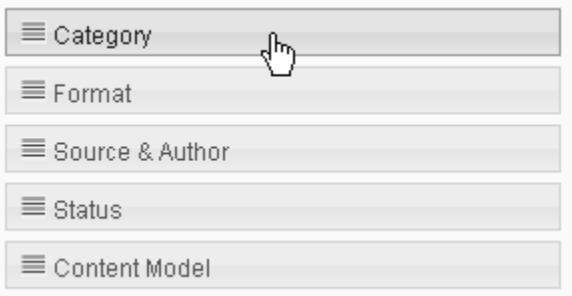
- **Search Results Menu** (): Toggle a menu of search results. This widget has some automatic capabilities; if a user makes a query by bounding box or keyword, the search results menu will automatically toggle itself.

Library

The [[Library interface]] allows you to use keywords and phrases to search the data records. Once you have results, you can narrow them to find the records you are looking for by choosing a set of criteria by which to filter your results.

Narrowing Your Search Results

1. Type a keyword or phrase into the search box to receive an initial set of results.
2. Choose one or more of the filter topics from the left side of the screen:



3. Within each topic, check the box by the criteria you wish to use as your filter(s):

   Selecting a check box will automatically update your search results to show only data that match the criteria you have selected. Note that you can use multiple criteria simultaneously to filter your results.

4. To remove a filter, select the name of the filter at the top of the screen:

You can filter your search results using the following topics:

- **Category**: Filter results by category and subcategory. For example, the category “drilling technology” includes a variety of subcategories related to drilling, such as the presence or absence of casing diagrams.
- **Format**: Filter results by format of the documents or datasets. Formats include data types and online resources, such as Web services.
- **Source & Author**: Filter results by data source and/or author.
- **Status**: Filter results by the status of the data within the context of the NGDS project. Data status tends to reflect the state of completion of that data. For example, typical NGDS documents or datasets will be either completed, ongoing, or deprecated.
- **Content Model**: Filter results by the content model to which the data conforms. All Tier 3 NGDS data conforms to a specific content model, which reflect properties of a given data set. A search for a specific content model would return results of data structured to conform to that content model.
- **Public/Private**: Filter results by the visibility of the data. When published, data contributors can designate whether or not a given document or dataset is publicly accessible or private. A setting of “private” renders the dataset invisible in public searches.

Managing Datasets

Delete a Dataset

In order to delete a dataset, click the **Edit** button on the page for the dataset. At the bottom of the page click **Delete.**

This does not permanently delete the dataset but does remove it from the Library. Since the dataset is not permanently deleted, the dataset's URL is not available for use by a new dataset.

To permanently delete a dataset go to the [[Node Admin]] page and purge the datasets. Now the dataset's URL is available for use again.

USGIN URIs

A Primer

11/21/2012

Table of Contents

Section 1: Introduction 2

Section 2: What is a URI? 2

Section 3: URNs and URLs 2

Section 4: USGIN URIs 3

Section 5: USGIN URI Syntax 4

Section 1: Introduction

Millions of database records and numerous datasets are shared according to USGIN specifications in a distributed, interoperable network. Organizing and maintaining uniqueness among these records and datasets requires a robust system of identifiers that is well-suited for the digital medium.

This document provides a general introduction to URIs and demonstrates URI applications within USGIN.

Section 2: What is a URI?

A Uniform Resource Identifier (URI) is a compact string of characters for identifying an abstract or physical resource (Berners-Lee ta al, 1998, rfc2396). Breaking this down into manageable chunks:

- URIs are *uniform:* they always use or conform to specific Internet Engineering Taskforce (IETF) URI syntax
- URIs identify a given resource.

Consequently: at their most basic level, URIs use a *consistent syntax* and *identify a resource*. Though multiple URIs can identify a given resource simultaneously, it is *imperative* that a single URI should not identify multiple different resources. Hence, each URI should be *globally unique.*

Generic IETF URI syntax may be found here; World Wide Web-specific IETF syntax may be found here.

Section 3: URNs and URLs

URIs come in two flavors: URNs and URLs. URNs and URLs are both URIs, and a URI can be a URN and a URL simultaneously.

- If a URI identifies a resource regardless of the location of that resource, then the URI is a **URN (Uniform Resource Name)**
- If a URI identifies a (web) location at which a bound resource may be found, then the URI is a **URL (Uniform Resource Locator)**

Though URIs (including URNs and URLs) must follow the appropriate IETF syntax to be considered URIs, most people are familiar with a number of URI analogs.

For example: Social Security numbers and car license plate numbers function as URNs, in that they use consistent syntax and identify resources without specifying that resource's location. Each individual Social Security number uses consistent syntax (NNN-NN-NNNN) and provides an identifier for a human resource (an individual person). Similarly, a car license plate number uses a combination of 6-7 letters and numbers and identifies a specific vehicular resources (a car). In both cases, the identifier does not indicate the location of the resource: a Social Security number does not indicate the location at which a

specific person may be found, nor does a license plate number indicate the location at which an individual car may be found.

Telephone numbers and street addresses are analogous to URLs, in that they follow a consistent syntax and identify a location at which a resource may be found. For example, street addresses use consistent syntax (NNNN [Street Name] [City] [State] [Zip Code] in the United States) and identify the location at which a resource (a shopping center, business, or office) may be found. Telephone numbers also function as URLs: they follow consistent syntax (NNN-NNN-NNNN) and identify a location (on the switchboard) at which a human resource (a person) may be found. In both cases, resources are associated with an identified location by means of a binding: a business resource is bound to a street address in the same way that a human resource is bound to a telephone number. If a business moves to another address or a person is no longer associated with a given phone number, then the binding is invalid.

All web addresses entered into your web browser are URLs. The web site http://www.google.com is a URL that identifies the location at which the Google search engine (a computing resource) may be found.

As with URIs, IETF-compliant syntax is available for both URNs and URLs.

Section 4: USGIN URIs

Like URLs, USGIN URIs can be entered into a web browser. Entering a USGIN URI into a web browser produces a *representation* of the resource identified by the URI. A representation is a symbolic proxy for something.

For example, a fault in the middle of the Mojave Desert can be represented by a symbol on a map, a digital photograph, or a text description of the fault. A USGIN URI identifying this fault would typically dereference to an XML representation of the fault (Figure 1). This XML document would provide structured information about the fault's attributes, including the fault's age, dip, and slip. These attributes can be used for analysis.

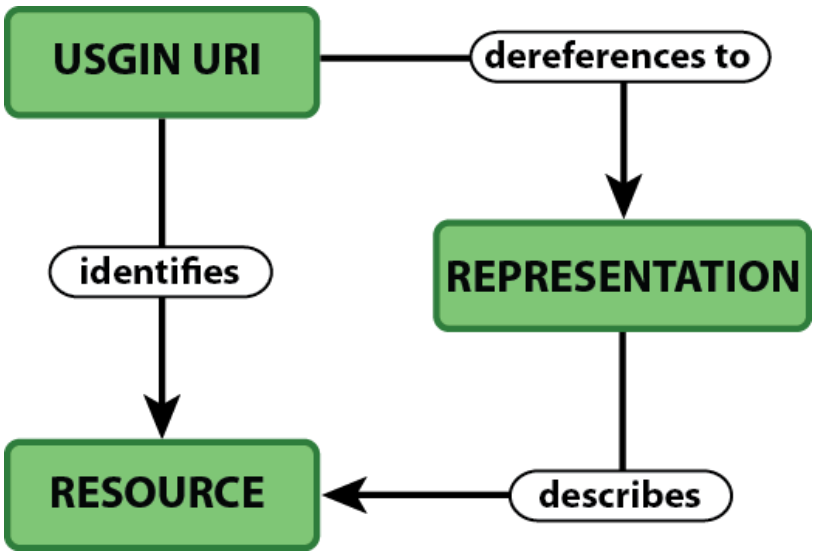


Figure 1: The relationship between USGIN URIs, XML documents, and the resource identified by the USGIN URI

However: even though they dereference to a web page when entered into a web browser, USGIN URIs are **not** URLs.

URLs identify a location at which a bound resource may be found. When entered into a web browser, a URL instructs the web browser to go to the indicated location and retrieve whatever resource is there.

By contrast: USGIN URIs are not URLs because the resource they identify is neither the representation found at the web location to which the USGIN URI dereferences, nor the web location at which said representation may be found

Returning to the example of a fault in the middle of the Mojave Desert: the *fault* is the resource identified by the USGIN URI. The USGIN URI does not identify the *representation* of said fault, nor the location at which said representation may be found. The USGIN URI identifies only the fault; the representation of the fault can be considered an intentional incidental.

Section 5: USGIN URI Syntax

As described above, USGIN URIs identify resources and use HTTP-compatible syntax to produce representations of said resources in web browsers. A USGIN URI can be broken down into a number of component tokens, which are represented in simplified form below:

http://host/uri-gin/authority/resource-type/resource-specific/

The tokens that comprise USGIN URIs are as follows:

- **http://** This token specifies the HTTP protocol for the URI.

- **host/** This token is a standard World-Wide Web domain name that is used to locate a particular server that has the necessary software to dereference the URI and return a representation of the resource identified by the URI. Currently, all USGIN URIs are dereferenced through the host located by:

    resources.usgin.org

- **uri-gin/** By USGIN convention, this token indicates that the URI is a USGIN URI. All URI components following this token are terminal components of the USGIN URI. The terminal components of the USGIN URI often reflect or incorporate existing resource identifiers that are not URIs, such as American Petroleum Institute (API) numbers. This is only practical, as existing resource identifiers are often sufficient for the purposes of USGIN - why reinvent the wheel?

- **authority/** This token identifies the local naming agent responsible for issuing and maintaining any resource identifiers that are present within the the terminal components of the USGIN URI. For example, if a USGIN URI identifies an oil well with an API number, and if the URI incorporates that API number into the URI, then the **authority/** token would reflect the authority responsible for issuing and maintaining API numbers: the American Petroleum Institute (or their agent in a particular state), abbreviated API. Examples of authority abbreviations in the context of USGIN URIs are as follows:

    - **American Petroleum Institute (API)**: http://host/uri-gin/api/
    - **United States Geological Survey (USGS)**: http://host/uri-gin/usgs/
    - **Arizona Geological Survey (AZGS)**: http://host/uri-gin/azgs/
    - **Department of the Interior (DOI)**: http://host/uri-gin/doi/
    - **Department of Energy (DOE)**: http://host/uri_gin/doe/

- **resource-type/** This token indicates the type of resource identified by the URI. Note that this token may itself be hierarchical, with resource subtypes delimited by the path separator character '/' Sample USGIN URI conventions can be found below:

    - Active or quaternary fault resource: http://host/uri-gin/authority/fault/
    - Borehole temperature resource: http://host/uri-gin/authority/bhtemp/
    - Thermal spring resource: http://host/uri-gin/authority/thermalspring/
    - Well header resource: http://host/uri-gin/authority/well/

- **resource-specific/** This token uniquely identifies a resource. Literally anything is possible within this component of the USGIN URI, although USGIN recommends constructing the string systematically, using existing resource identifiers and URL-safe characters. A list of URL-safe characters for resource-specific strings is as follows:

- 0-9
- a-z
- A-Z
- /
- - (dash)
- _ (underscore)

Consequently, a finished URI that follows USGIN specifications and identifies a specific car (vehicular resource) might appear as follows:

http://resources.cars.org/uri-gin/azdmv/vehicle/ABC123

Likewise, a finished URI that follows USGIN specifications and identifies a military aircraft (military resource) might appear as follows:

http://resources.airforce.gov/uri-gin/usaf/vehicle/aircraft/f22/1234567

This concludes the USGIN URI tutorial.

Archived USGIN Lab Discussions

Best Practices

Best Practices for USGIN Web Service Hosting

This blog is designed to encompass technical questions and troubleshooting for deploying and hosting USGIN services, particularly related to the AASG State Geothermal Data Project. Topics covered will included basic deployment procedures with WFS and WMS specifics, custom capablilites, handling large volume datasets, service and layers naming conventions, URIs, hosting online related files, metadata in the USGIN repository, and validating WFS and WMS services. The technical team at AZGS has encountered a need for such discussion as individual institutions and states begin hosting thier own USGIN AASG Geothermal Data Project services.

Convert File GDB to PostGIS Database

The goal of this post is to get our data from a File GDB to a format that GeoServer can use, specifically a PostGIS GDB. If we try and use a shapefile, the field names will be truncated and this will invalidate your schema that you have worked so hard to conform to the USGIN schema standards.

*Only works for 10.x and up geodatabases. You need to upgrade your gdb to 10.x if you are using an older geodatabase.

Geoserver App-Schema Set-Up: In Progress

Submitted by ccaudill on Fri, 02/01/2013 - 8:53am

One hurdle to using Geoserver to deploy AASG Geothermal Data services is the resolution of schemas. Creating additonal configuration files for Geoserver allows the use of "app-schemas" which will automatically download the indicated web-accessible schema (.xsd) and be referenced in the GetFeature request. This would solve a problem with service deployment using ArcGIS, where custom schema locations is necessary.

This blog will document the problems and break-throughs I've had so far in implementing Geoserver app-schemas. I have not yet been able to get schema resolution, but would like to make a post so that others can take into account my work thus far and chime in about other ideas or advances. This work is based on Tomcat running Geoserver, set up on an Oracle VM VirtualBox. The data for the services are pulled from a PostGIS database.

Service Deployment Using ArcGIS Server for WMS/WFS services

Once you have your data in the appropriate format, and the required fields populated, the next step is to create a service for the data. When it comes to AASG or OneGeology services, there are naming conventions and metadata rules.

If you are deploying a OneGeology WMS, please visit: http://onegeology.org/wmsCookbook/home.html for more detailed requirements. For AASG naming conventions, please visit this link

Web Service and Layer Naming Conventions

Submitted by ccaudill on Thu, 09/13/2012 - 11:06am

The AASG State Geothermal Data project provides OGC (Open Geospatial Consortium, opengeospatial.org/standards) compliant WMS and WFS services per the specifications outlined by the USGIN (see the USGIN site for more detials). AASG adopted the requirements and recommendations of USGIN for networking, data sharing, and enhanced interoperability. Such recommondations include naming conventions; the following outlines AASG-adopted naming conventions which require a specific format for naming Web Services and the Layers that contain the data within that service.

Creating WFS / WMS capabilites using ArcGIS service publishing wizard

Submitted by ccaudill on Mon, 08/27/2012 - 8:30am

When deploying a web service using ArcGIS publishing service, there are a few specifics that need to be considered when constructing the WFS and WMS Capabilites in the ArcGIS deployment wizard. This post is geared toward publishing a service for the AASG Geothermal Data Project, but the following practices are similar when deploying any USGIN-specification services using ArcGIS publishing products.

In the ArcGIS service publishing wizard or "Add New Service" wizard, you will be promted to choose Capabilities. Capabilites for both WMS and WFS are provided in AASG Geothermal Data Project services.

Validating your WFS and WMS services

Submitted by ccaudill on Fri, 08/17/2012 - 3:56pm

After a web service has been deployed, testing must be done to ensure that the service will work in the system, that is, be interoperable and available to end users. This group post will discuss using WFS GetFeature requests and WMS GetCapabilities requests for service validation. All OGC (Open Geospatial Consortium) services offer these requests which return an XML representation of the data. Some of the following information is taken from GeoSciML Portrayal Cookbook (available in full here http://usgin.org/sites/usgin.org/files/GeoSciML-PortrayalCookbook.pdf).

GetCapabilities Requests

If your service is deployed using ArcMap products, the request starts from your service rest page. Scroll to the "WFS" and "WMS" links on the service rest page. The URL given from these links is the Get Capabilities request, which consists of a service endpoint part, and the OGC request part:

ArcGIS-generated fields and AASG content model fields order and type

Submitted by ccaudill on Fri, 08/17/2012 - 10:02am

When deploying your service using ArcMap products, two fields are automatically generated. These are the OBJECTID and SHAPE fields. The XML schemas (http://schemas.usgin.org/schemas/) that are

associated with each AASG content model have a very specific ordering of the fields from that content model, and included in this specific order are the ArcMap-generated fields.

To acheive validation of your WFS service, and for your service's data to be interoperable and available within the system, the fields must be in the specific spelling, order, and data type that are assigned in the XML schema. These are usual errors we find when checking the WFS validation of hosted services.

Even though every content model has different fields, the ArcMap-generated fields are reliably in the same place in the field order. OBJECTID must come just **before** all the content model fields, and SHAPE must come directly **after** all the content model fields. (A "lax" function does allow Shape_length or Shape_Area or other fields to come after SHAPE field.)

Implementation of a Content Process "lax" function Element for AASG Schemas

Submitted by ccoleman on Fri, 03/23/2012 - 1:38pm

Discussion of using the "lax" function in Schema Documents

We've recently changed how we are structuring our XSD schemas by adding a "lax" process function element at the end of the sequence. The 'lax' function, or 'any Element' function, can be specified by the author of the schema document, and enables additional elements not specified by the schema to be included.

Data Interchange Document XML Schema Validation

Submitted by ccoleman on Thu, 03/22/2012 - 11:20am

Validation of XML data interchange documents with normative schemas using free, open-source software (XML Explorer)

One of the objectives of the US Geoscience Information Network (USGIN) is to establish a community of practice for developing, documenting, adopting, and using standard document encoding for data interchange. The technology for encoding schemes is evolving continuously, and current practice is to use XML encoding with data schema defined by XML schema.

Troubleshooting

Data Interchange Document XML Schema Validation

Submitted by ccoleman on Thu, 03/22/2012 - 11:20am

Validation of XML data interchange documents with normative schemas using free, open-source software (XML Explorer)

One of the objectives of the US Geoscience Information Network (USGIN) is to establish a community of practice for developing, documenting, adopting, and using standard document encoding for data

interchange. The technology for encoding schemes is evolving continuously, and current practice is to use XML encoding with data schema defined by XML schema.

How to Create a Styled Layer Descriptor (SLD) using Arc2Earth

Submitted by ccoleman on Thu, 03/22/2012 - 10:39am

If you decide to use GeoServer to deploy your Web Services, you will find that part of the process includes creating an SLD or Styled Layer Descriptor for portrayal of the data you intend to publish. A natural symbology for such layers is an ESRI layer file format. Using this format initially presented a problem, though; creating an SLD for data with a complicated symbology scheme, such as a geologic map, is not trivial using XML editing software-- even if you have a template to work from. Many times, conversion from RGB to Hexidecimal is required for each entry in your symbology layer. Even if one uses a built-in formula converter in Excel (or from various websites), a large amount of hand entry is still required.

Making Web Maps

Using WFS with MANY features in ArcGIS

Submitted by srichardAzgs on Fri, 05/10/2013 - 12:35pm

If you're using a WFS in ArcMap, it downloads and indexes ALL the data from the service. If you're connecting to a service that has ALOT of data (>?150,000 features), my experience is that you don't actually get all the data. One solution is to put a bounding box filter on the WFS connection to restrict the data request to the area you're actually interested in.

To do this, when you set up the WFS connection properties, and are in the 'Parameters' dialog (which you'll need to do to select the feature type from the service), put an XML bounding box filter expression in the 'XML Filter Expression'

here's a bounding box filter that works (WFS 1.0.0):

```
<ogc:Filter xmlns="http://www.opengis.net/ogc">
 <ogc:BBOX>
  <ogc:PropertyName>shape</ogc:PropertyName>
 <gml:Box srsName="EPSG:4326">
   <gml:coord>
    <gml:X>-109</gml:X>
    <gml:Y>37.00</gml:Y>
```

Using and Customizing Leaflet

Submitted by Ryan Clark on Wed, 03/14/2012 - 1:51pm

Leaflet is a simple javascript framework for building web-mapping applications. In thier words:

Leaflet is a modern, lightweight open-source JavaScript library for mobile-friendly interactive maps. It is developed by CloudMade to form the core of its next generation JavaScript API. Weighing just about 21kb of gzipped JS code, it still has all the features you will ever need for your web mapping needs, while providing a fast, smooth, pleasant user experience.

It is built from the ground up to work efficiently and smoothly on both desktop and mobile platforms like iOS and Android, utilizing cutting-edge technologies included in HTML5 and CSS3. The focus is on usability, performance, small size, A-grade browser support, flexibility and an easy-to-use API. The OOP-based code of the library is designed to be modular, extensible and very easy to understand.

When you're choosing a javascript framework to build your web-mapping application, you have a myriad of options. Why choose Leaflet? Well this is why I liked it:

ArcGIS Server and OGC Services

SDE-Trickery and Awesomeness: Data Transformations Behind-The-Scenes

Submitted by Ryan Clark on Thu, 06/02/2011 - 3:26pm

Imagine that you collected geographic data in one schema and you want to provide it in some other schema. For many of us, this only happens pretty much *all of the time*.

For example, maybe your organization has an enterprise database which stores all the information that you have about recent, active faults. Now you want to provide that data as a WFS service with fields that conform to the ActiveFaults template being used by the AASG Geothermal Data System. Perhaps you store all your geologic map data in such a database and you want to convert it into FeatureClasses that fit into NCGMP09-style geodatabases.

Here's a nice way to convert your data formats in such a way that you only have to do it once. You can continue to make changes to your data just like you always have, and there will always be a featureclass hanging out there in the new schema which conveys the most up-to-date information that you have. The pre-requisites are:

OGC Services don't work outside local network!

Submitted by Ryan Clark on Thu, 05/26/2011 - 5:35pm

A commonly overlooked step in the configuration of an ArcGIS Server is that you need to tell the server what its name is to the outside world. You see, on an internal network your computers might have names. As geologists we prefer names like "Eclogite", "Kyanite", "Harzburgite", etc... However, on the

internet computers generally have more complicated names. These names are things like "services.azgs.az.gov" or "www.google.com" or "lab.usgin.org".

By default, your ArcGIS Server knows what the computer's internal name is. However, once you want to share your services with people outside that local network, you need to tell the ArcGIS Server how those other people will find it. You do this by editing a file called "rest.config". If you installed IIS and ArcGIS Server in all the default locations, you'll find this file at C:\inetpub\wwwroot\ArcGIS\rest\rest.config.

Look through the file and find all the URLs in it. Make sure that instead of using the computer's internal name, those URLs use the computers external name.

For more information, here are a couple of links:

Dynamic Content as WFS

Submitted by Ryan Clark on Thu, 04/28/2011 - 3:57pm

There could be a number of reasons why you'd like to use a database view as a backend for a WFS. Perhaps your data is split across multiple tables and updated regularly. Perhaps the data is in a table without an explicit ESRI-style SHAPE field. That's the case in this example, where the geometry information is stored in a table as UTM coordinates.

Using Django for USGIN

Thought and ideas about using Django to accomplish USGIN-related... things.

Updated URI Redirection Engine

Submitted by Ryan Clark on Fri, 05/25/2012 - 11:44am

I've made a new version of the URI redirection engine that runs in Django (see the original post here). The tool allows you to provide a list of URI patterns that identify specific resources, and redirect traffic that matches that pattern to a URL where a representation of the resource exists. It allows for content-negotiation, and the new version also begins to build infrastructure by which various instances of the application can communicate with each other.

Each instance of the application is informed about the existence of various registers, which are essentially sets of URIs. The application may or may not know how to resolve URIs that are a part of any given register, but can send incoming requests to others servers capable of resolving that register's URIs.

PyGreSQL: This will be useful!

Submitted by Ryan Clark on Thu, 02/03/2011 - 6:49pm

PyGreSQL is a python library that allows you to write python code to interact with a PostgreSQL database. I can see some utility here if I start wanting to write Django methods that do things like create views of some of my tables.

May be useful in not too long! Don't forget about it!

Documentation: http://pygresql.org/pg.html

Installation:

sudo apt-get install python-pygresql

Postgres COPY for CSV importing

Submitted by Ryan Clark on Wed, 02/02/2011 - 2:55pm

Working on some integration between Geoserver, Django and NCGMP geodatabases, I often need to use PostgreSQL's COPY command to import data from CSV files into PostgreSQL. Here are some tips:

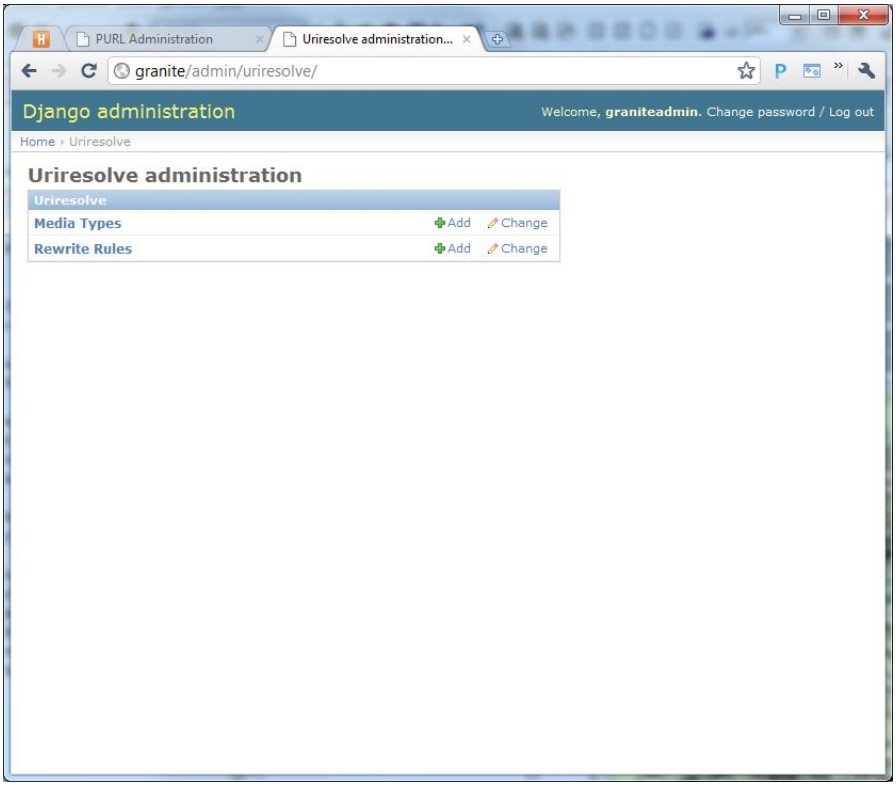
URI Redirection Engine (with pictures!)

Submitted by Ryan Clark on Mon, 01/24/2011 - 2:10pm

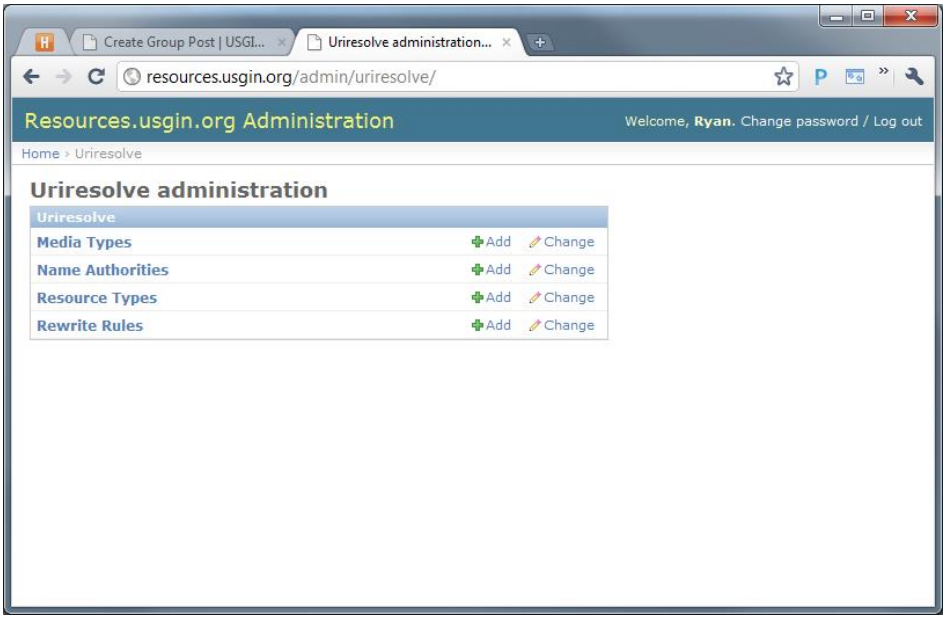
Here are a few screenshots to give a taste of what it is like to use our new Django app to manage a set of URI redirections:

Administration Overview:

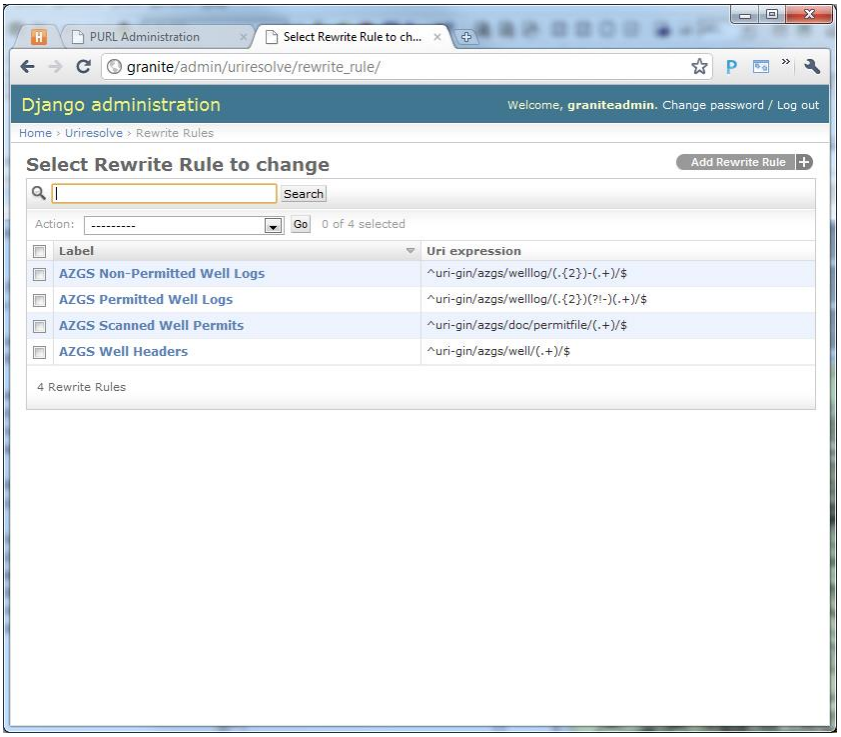
For any URI-Scheme: Basically, the Django app allows you to control a listing of MIME Media Types, and Rewrite rules.



USGIN URI-Scheme: The same basic structure, but the USGIN-specific app allows you to control a listing of Name Authorities and Resource Types, as described in [this post.]



Listing of Rewrite Rules: A nice overview of what rewrite rules you have in place, and a mechanism for searching for the one you might need to adjust.



URI dereferencing with Django / HTTP Content Negotiation

Submitted by Ryan Clark on Fri, 01/21/2011 - 9:46am

I've been building a Django app that essentially allows for the dereferencing of URIs that follow the USGIN Scheme. The engine also does some content-negotiation by reading HTTP GET accept-headers and responding with an appropriate file located somewhere on the internet. I had never even heard about content negotiation before embarking on this, and here are a few things that I learned:

Redirection Engine for USGIN-Compliant URI Resolution

Submitted by Ryan Clark on Fri, 01/07/2011 - 2:40pm

A common problem we run into is the unique identification of resources. This is a very broad statement because it represents a very broad problem. We have identification schemes for all kinds of things: people have names, web pages have URLs, books have ISBN numbers, etc. For the USGIN system to work smoothly, we need an identification scheme for resources that are part of the network. In USGIN we use Universal Resource Identifiers (URIs) which follow a specific format. For more information about the format, and for more background, see this post on lab.usgin.org.

Once we have this identification scheme built, it is important that the URIs can be dereferenced, or in other words that they are resolvable. We would like a user to be able to put the URI of a particular feature into the address bar of their web browser, push enter, and be presented with some sort of representation of their resource of interest. This is the niche that our Django-based redirection engine it trying to fill.

Django Setup

Submitted by Ryan Clark on Mon, 12/27/2010 - 11:03am

Introduction

I'm relatively new to Django, and maybe to servers in general, but setting up Django is a chore. Or at least figuring out how to set it up is. There is no installer. You have to have a pretty good understanding of how your HTTP server (IIS, Apache, whatever) works before you're going to be able to get anywhere. Then you'll have to know how to set up quite a few ancillary applications (e.g. Python, WSGI, FastCGI, Flup, MySQL, PostgreSQL). The details of which other applications you want to set up depend entirely on the evironment you're building - so that means it is a little different every time - and that means it is pretty hard to write one single "walkthrough" for how you should do it.

What I want to do here is detail my setup, and along the way make some suggestions about what seems to make things easier. So here goes:

Suggestion: Use Ubuntu (Most recent stable version, always.)

I say this because that apt system really makes installations a breeze. I fumbled for a while to try and setup WSGI (a module for Apache) to work on my Windows machine, to no avail. In Ubuntu it is so easy:

Geoportal on an Amazon Virtual Machine

PostgreSQL Trigger - make harvested/uploaded metadata records editable in Geoportal

Submitted by Ryan Clark on Mon, 12/27/2010 - 12:00pm

By default, metadata records that are harvested into the Geoportal are not editable. This includes any records in the Geoportal that were not created explicitly with interface provided by the web application ("Use dedicated editor to create metadata manually"). The SQL below creates a database trigger that allows harvested records to be edited.

Be aware that there is the potential for a lossy data translation when you do this. The harvested record is completely rewritten using the schema definitions that your Geoportal implements. If the original harvested record has any data that is not included in your schema implementation, that data will be lost after the original is edited. Another way to put that is, if you edit a harvested record, the edited version will only contain information that you see in the editor interface. If the original started with anything else in it, that will be lost.

Using jQuery for basic ISO Metadata record form validation

Submitted by averill.cate on Mon, 11/29/2010 - 11:50am

ESRI's Geoportal metadata record editing has some validity checking, but error checking does not occur until after a metadata record form is submitted. This can be a wasteful process because it involves having the client (the web browser) send a full request to the application server. If the submitted form has data validity errors then the whole form is transmitted back to the client for editing. One way to save the use of server resources (e.g., bandwidth) is by using jQuery to monitor information typed in html input boxes and check value validity before submitting the form. Since, ESRI's Geoportal comes with jQuery already enabled it is relatively to add validity checks. The first thing to do is create a custom javascript file (e.g. [geoportal_install_dir]/web/geoportal/catalog/js/custom.js). The contents of my script are:

```
$(document).ready(function()

{
  $('#mdEditor\\:identification_date').mask('9999-99-99T99:99:99');
  $('#mdEditor\\:general_datestamp').mask('9999-99-99');
```

Install Apache Directory Server

Submitted by Ryan Clark on Thu, 10/28/2010 - 3:38pm

ESRI's Geoportal wants an LDAP server, and suggests using Apache Directory Server.

I started a new VM using the image usgin2-atpms. Installation was as easy as:

wget http://apache.imghat.com//directory/apacheds/unstable/1.5/1.5.7/apacheds-1.5.7-i386.deb

sudo dpkg -i /home/ubuntu/apacheds-1.5.7-i386.deb

It was easy to connect to via SSH Tunnel to port 10389. The next step would be to fix some issues with the instance and its data-store, remove the extra MySQL installation, and then move on to configuration of the PostgreSQL database.

Drupal Development

Discuss ways to bend the Drupal content management system to your needs.

Creating a Format Agnostic Metadata Wizard

Submitted by wgrunberg on Fri, 05/28/2010 - 12:22pm

The USGIN Metadata Wizard is a metadata format agnostic web tool for creating, managing, and exporting new metadata records. This Drupal based service was conceived as a means to lowering the brier to metadata creation for those who are not able or willing to become proficient with metadata standards and informatics. The tool is based on a format agnostic metadata model which then can be exported into multiple metadata standards.

Creating ISO 19139 metadata through Drupal Views and Views Bonus Pack

Submitted by wgrunberg on Fri, 04/16/2010 - 2:00pm

Following is how we modified the Views Bonus Pack module for Drupal 6 to generate ISO 19139 XML metadata files for the USGIN Document Repository. This hack can be be used to relatively easily generate other XML metadata (FGDC, CSW records) or most complex XML. Ideally, this hack will evolve into a separate module. If you are interested or have questions, please participate in the Generate ISO 19139 metadata record XML files Views Bonus Pack issue thread.

Overview

The goal is to generate minimum ISO 19139 dataset metadata XML files (that conform to the USGIN profile) from the core and CCK fields used in the repository's Collection (ct_collection) and Document-like Information Object (DLIO, ct_dlio) content types.

Drupal Content Types used in the USGIN Document Repository

Submitted by wgrunberg on Fri, 04/16/2010 - 11:20am

Following are the two Drupal 6 content types that define the USGIN Document Repository (you may want to familiarize yourself with general Drupal concepts):

Collection (ct_collection)

The Collection content type acts as an arbitrary logical unit (or container) that groups the Document-like Information Object (DLIO) nodes which hold the actual data resource. At this time, Collections fulfill 3 objectives at once. Perhaps, they need to be separated at a later point.

Drupal Modules used in the USGIN Document Repository

Submitted by wgrunberg on Thu, 04/15/2010 - 4:28pm

Following is a break-down of the Drupal 6 modules used in the USGIN Document Repository site (you may want to familiarize yourself with general Drupal concepts):

Running Drupal on a Windows Server 2008 machine

Submitted by Ryan Clark on Wed, 04/14/2010 - 4:43pm

We recently struggled through figuring out how to make Drupal run on a Windows Server 2008 machine that was running IIS. Every Drupal site that we had dealt with prior to this was hosted on a Linux machine running Apache, MySQL and PHP, and so we needed to figure out what could be done differently to make it work on our in-house server, which was already running a production instance of ArcGIS Server under IIS7.

From the outset, we had two options:

Creating a Document Repository in Drupal

Submitted by wgrunberg on Wed, 04/14/2010 - 2:56pm

We just finished putting together a Drupal based Document Repository site (http://repository.usgin.org) as a testbed and as an alternative for other Repository systems such as DSpace. The goal was to fulfilled the following requirements in a ca. 2 week development period (200+ man hours):

GeoNetwork Configuration and Development

This group will focus upon GeoNetwork setup, configuration, and development. Any questions or comments are welcome!

Following are some resources that can help you learn about GeoNetwork:

- dev list
  - registration https://lists.sourceforge.net/lists/listinfo/geonetwork-devel
  - archive http://n2.nabble.com/GeoNetwork-developer-f2014756.html
- user list
  - registration https://lists.sourceforge.net/lists/listinfo/geonetwork-users
  - archive http://n2.nabble.com/GeoNetwork-users-f2013074.html
- SourceForge http://sourceforge.net/projects/geonetwork/
- Wiki http://trac.osgeo.org/geonetwork/
- Trac Tickets http://trac.osgeo.org/geonetwork/report
- USGIN Lab pages that contain "GeoNetwork" http://lab.usgin.org/search/node/GeoNetwork

Python Pylint/Pydev Code Checking in Eclipse

Submitted by lwolfe on Wed, 09/15/2010 - 6:16pm

Pylint is a good python code checking tool. Pydev is a good plugin for Eclipse for editing and running python programs. Much of this info is taken from Ryan Fraser's page . Wolfgang has an alternative description for installing pylint in the Komodo python IDE here

Windows 7 64-bit guide for GeoNetwork development within Eclipse

Submitted by wgrunberg on Wed, 07/28/2010 - 11:38am

This setup guide for downloading GeoNetwork using Subversion, building using Maven 2, developing, deploying and debugging on Tomcat and Eclipse on Windows 7 64-bit still uses the MckoiDB and should be updated to use MySQL.

Introduction

This document was developed for GeoNetwork's new Maven 2 based setup. See http://trac.osgeo.org/geonetwork/wiki/Maven for more information.

Note that several installation and configuration options may not be necessary for GeoNetwork development but can be necessary for other software development projects.

This document was compiled from experiments, notes, and web sites by the following contributors and sources (in alphabetical order):

Converting FGDC XML metadata records for WMS and WFS services to ISO 19139 via Python and GeoNetwork

Submitted by wgrunberg on Wed, 03/10/2010 - 1:04pm

This post describes a series of Python scripts to convert FGDC XML metadata for WMS and WFS services to ISO 19139 service metadata via GeoNetwork's OGC harvest service. Here are the steps:

1. Extract WMS and WFS GetCapabilities URLs from FGDC XML metadata and write them to a file.
2. Create a GeoNetwork Harvest Node from extracted GetCapabilities URLs and let GeoNetwork's OGC WMS/WFS harvester create ISO 19139 metadata from GetCapabilities response. Note: the resulting ISO 19139 metadata is only as good as the one in the GetCapabilities response. In addition, GetCapabilities responses do not include all fields necessary for minimum ISO 19139 metadata.
3. Copy some FGDC metadata entries (Title, Abstract, etc.) to newly created ISO 19139 metadata.

Our goal is to have working WMS and WFS metadata records in our CSW catalog that can be used to add the described services to analytical software such as ESRI's C-SW Client for ArcGIS Desktop. Following are the reasons why I am currently choosing this convoluted process instead of a direct FGDC to ISO 19139 metadata conversion:

Create a GeoNetwork OGC Harvest Node (WMS or WFS GetCapabilities to ISO 19139 metadata) through xml.harvesting.add request

Submitted by wgrunberg on Fri, 03/05/2010 - 11:05am

GeoNetwork offers a harvest service that, among other, follows a OGC WMS or WFS GetCapabilities URL and transforms the response into an ISO 19139 metadata record. Bear in mind that the resulting ISO metadata record is only as good as the GetCapabilities response and that the metadata can never be entirely ISO 19139 conformant without dummy values due to limitations of the OGC GetCapabilities schema.

GeoNetwork offers a handy user interface to add those "GeoNetwork Harvesting Nodes" and also exposes their functionality through a even handier XML service. See chapter 19.3 on "Harvesting Services" of the GeoNetwork opensource V 2.4 The Complete Manual. Following are my notes on creating an OGC harvesting node through GeoNetwork's harvesting service.

Tomcat configuration in Eclipse

Submitted by wgrunberg on Thu, 02/04/2010 - 1:35pm

Tomcat kept running out of memory when running it as a service from within Eclipse. Following is a solution:

Double-click on your Tomcat service under the Servers tab, and click on the "Open launch configuration" to open the Tomcat launch configuration window. On the Arguments tab, append the two following lines in the "VM arguments" form window:

-XX:MaxPermSize=128m
-Xmx1024m

How-to on running Tomcat within Eclipse:

- Eclipse Help: Creating a server
- IBM: Developing Web applications with Tomcat and Eclipse

Import MEF metadata archive into GeoNetwork through Python

Submitted by wgrunberg on Fri, 11/06/2009 - 1:53pm

This is a Python example script for importing GeoNetwork Metadata Exchange Format 1.1 (MEF) archives to GeoNetwork 2.4.2's mef.import service. The mef.import service requires a multipart/form-data POST through a modified MultipartPostHandler.py library which now supports Unicode (urllib2). It has been tested in Windows XP and Python 2.6.

Create GeoNetwork MEF files from ISO 19139 XML through Python

Submitted by wgrunberg on Fri, 11/06/2009 - 10:11am

This is an example Python script that showcases the creation of GeoNetwork Metadata Exchange Format 1.1 (MEF) archives from ISO 19139 metadata XML files.It has been tested in Windows XP and Python 2.6.

Metadata Interest Group

general forum for questions, comments, discussion about metadata

Discussion of Use of CI_Online Resource

Submitted by srichardAzgs on Fri, 02/21/2014 - 12:49pm

The CI_OnlineResource element is used in a variety of contexts in ISO19115 content. This complex-content-element provides a linkage to some online resource related to the metadata or the described context resource. It is thus analogous to atom:link (IETF RFC-4287) or the link element described in IETF RFC-5988. For links that locate online documents accessible using standard browser and file type resolution technology, the link can be as simple as a single URL element that retrieves a representation of the resource. There are many other kinds of related resources that the onlineResource element may point to, including web services that provide access to a dataset resource, metadata for related or source resources, specifications for standards or extensions to standards.

Lucene search syntax

Submitted by srichardAzgs on Mon, 02/03/2014 - 12:58pm

Useful guide to search term syntax for any applicaiton that's hitting a lucene index

http://lucene.apache.org/core/2_9_4/queryparsersyntax.html

JSON template for a 'complete' metadata record

Submitted by srichardAzgs on Mon, 12/12/2011 - 10:42pm

Incorporates concepts from ISO 19115, 19115-1, 19119, FGDC, 19110, and FRBR along with some additional stuff like recording the lineage of the metadata record, which is useful in a metadata harvesting federated system. Builds on the conceptual model presented in a previous discussion here.

This is an experiment to see what it would look like, trying to bring in just about any metadata property from those various specs... Lot's of comments in line.

The JSON is not perfect -- I'm a newbie.

Scenario for a community metadata editor

Submitted by srichardAzgs on Sat, 12/10/2011 - 12:19pm

Introduction

This is a scenario for a generic metadata edit tool, intended for brainstorming development of a community component that is transportable and reusable. Use of a standardized tool for editing metadata will reduce the cost of entry into the metadata game by reducing training and increasing

productivity, and should help improve the quality and consistency of metadata. The attached word document has the same text. Please add any other features or requirements that would be useful.

Here is a run down of some functionality.

The Scenario

Start: User opens the editor application.

Machine actionable links-- metadata, linked data, contexts...

Submitted by srichardAzgs on Mon, 10/03/2011 - 5:01pm

I spent some time researching this issue in the context of some other projects I've been following or involved in (OWS context Standards Working Group, energy industry ISO19115 metadata profile, and USGS CGI Web Application Integration Framework group). I would like to see this considered in a larger context to come up with a solution that can be used in these various contexts, because it's essentially the same problem.
Here' s the gist of the proposal for link attributes outlined in the attached doc:

href -- the link URL

type -- MIME type of response. Specifies encoding and optionally the native software application environment. We're sort of stuck with this because of existing usage. One big question is how much of serviceType, puprpose, and outputScheme it makes sense to conflate in the MIME type. I favor keeping the MIME type limited to low-level file formatting--to answer the question 'what software can read this file in a useful way' Understanding the content and intention of a document or the interface to a service should be specified by separate properties.

Googley Metadata

Submitted by Ryan Clark on Tue, 09/27/2011 - 10:19pm

We can all agree that search engines like Google have really defined the way that we search for information on the internet. So the question is, why aren't we using Google to search for our geologic datasets? Do we have to do all that formal, structured metadata? What the hell is CSW anyways?

As an experiment, I built a new site: metadata.usgin.org. This site is buit with Node.js, a really cool way to build a web server by running server-side javascript code. Its so easy. Here's my code. So what does it do?

It acts as a proxy server. You ask it for http://metadata.usgin.org/record/a386a4ba-e892-11e0-9e4a-0024e880c1d2, and it reads that file identifier out and issues a CSW GetRecordByID request to a CSW server. The response to that CSW request is an XML document, and nobody (including Google) really cares about that, so the server formats the record as a very simple HTML page.

Really Generic Functional Requirements for Metadata

Submitted by Ryan Clark on Thu, 09/08/2011 - 12:49pm

We write metadata so that we can search through it to find datasets we're interested in using for some purpose. I've recently started building a map-centric user-interface for doing this that has highlighted some simple functional requirements that metadata should meet in order to be helpful in accomplishing the task.

What I've done here is essentially listed the conceptual elements that seem neccessary in order to make a functional user-interface for searching. This has nothing to do with metadata syntax, just with metadata concepts.

Energy Industry Profile of ISO/DIS 19115-1 v1.0 Release Candidate is Available for Public Review

Submitted by srichardAzgs on Fri, 09/02/2011 - 7:51am

FYI-- this is a draft profile against the upcoming revision of ISO19115 for metadata, scoped to energy industry, but the scope overlaps requirements for USGIN (I've been very active in developing profile and looked after that). Please have a look at the docs if you're interested in metadata. Given the standards process time lines, it probably won't directly impact us until next year.

steve

Conceptual model for general purpose resource metadata

Submitted by srichardAzgs on Thu, 06/17/2010 - 11:29am

I have built a UML model (using Enterprise Architect 7.5 tool) of metadata content that is based on ISO19115/19119, but attempts to clean up some of the logical incoherence and inscrutable element labeling in those specifications, as well as adding hooks for more detailed information on specific resource types and resource lineage.

The attached files include a zip archive of the Enterprise Architect-generated html report on the UML model; unzip to a directory and start on the index.htm page.

The EA document (zipped into GenericCompleteMetadataConceptualModel.zip) is the actual UML model. You'll need EA 7.5 or later to open it. The project contains a Metadata package, within which the 'GenericMetadataModel' package contains the actual model. The other packages include UML models for related ISO specifications, elements of which have been incorporated (copied, renamed, modified -- not imported) into the generic model. They are included for reference.

The html documentation generated by Enterprise Architect can also be viewed at

http://usgin.org/GenericMetadataModelDoc/index.htm

Files:

Recommendations for metadata content to describe geoscience resources

Submitted by wgrunberg on Wed, 05/05/2010 - 3:31pm

Now that we have a better grasp of the IT and informatics challenges of a distributed and interoperable information network, we can think more again about its adoption - how to make it work and useful on the ground. Following was our rough chain of thought on metadata:

ISO 19139 service metadata record srv:DCPList codelist confusion

Submitted by wgrunberg on Tue, 11/24/2009 - 12:16pm

It appears that the GeoNetwork WMS harvest selects an incorrect codeListValue for a service metadata record's srv:DCPList (Distributed Computing Platform). Currently, it enters HTTP-GET and HTTP-POST. However, there is a bit of confusion within GeoNetwork and among standards and profiles. Specifically, one could interpret a conflict between DCP in a service's GetCapabilities (OGC) and a service's metadata record (ISO):

Useful site for CSW testing and resources

Submitted by srichardAzgs on Thu, 10/29/2009 - 10:29am

CSW requirements

appears to mostly be compiled by Armin Retterath, Kompetenz- und Geschäftsstelle GDI-RP

Has listing of active CSW servers, some requirements listings

http://www.mapbender.org/CSW_Requirements

ISO metadata editor tool comparisons

Submitted by srichardAzgs on Thu, 10/22/2009 - 4:41pm

The geospatial metadata community has benefitted greatly from the CSDGM Metadata Editor reviews created by Hugh Phillips and posted at: http://www.sco.wisc.edu/wisclinc/metatool/. As we make the transition toward implementation of an international metadata standard, the community requested that FGDC compile information about applications capable of creating ISO 19115 metadata. The information on this page was compiled to address that need.

Presentations and Posters

Post your posters and presentations related to USGIN topics.

Presentation for European Geosciences Union meeting, 2012-04-25

Submitted by srichardAzgs on Mon, 05/07/2012 - 9:55pm

Slides for talk at EGU, Vienna, April 2012, presented by SM Richard.

Presentation for NGDS Kickoff meeting 2009-11-03

Submitted by srichardAzgs on Mon, 05/07/2012 - 9:51pm

A rather large power point stack, for a one hour presenation on USGIN web services and their use for the National Geothermal Data System. Presented at the kick off meeting in Boise, ID, November 3, 2009 by SM Richard.

AGU Fall 2011 -- 2 posters: GeoSciML v3 release, and USGIN metadata

Submitted by srichardAzgs on Mon, 12/12/2011 - 10:54am

GeoSciML v3 'release' announcement poster

USGIN metadata poster

State Geothermal Data and the US Geoscience Information Network; SDSC Summer institute in Geoinformatics

Submitted by srichardAzgs on Mon, 12/12/2011 - 8:04am

Outline: Background; Approach: Metadata, Services, Interchange formats; Current status; Architecture and artifacts; Way forward

GSA national meeting 2011: USGIN Data integration framework

Submitted by srichardAzgs on Mon, 12/12/2011 - 7:54am

Progress 2010-2011, service inventory, promoting transition to open data and reusuable applications.

USGIN/NGDS Presentations at ESRI User Conference

Submitted by Ryan Clark on Thu, 07/14/2011 - 1:53pm

I gave two presentations at the ESRI User conference this year. One was geared mostly towards what we do with metadata in our systems, while the other was more of a broad overview of USGIN and a little bit of NGDS progress report.

Brief USGIN Overview: WRP GIS Committee Meeting, Mar. 8th, 2011

Submitted by Ryan Clark on Fri, 03/18/2011 - 11:52am

A brief overview of USGIN.

From WRP Meeting Notes:

III. Geoscience Information Network

Usgin Draft Specifications For Geoscience Resource Metadata And Identifiers

Submitted by srichardAzgs on Thu, 11/18/2010 - 5:11pm

Poster presented at Geological Society of America Meeting, Monday, Nov. 1, 2010 in Denver, Colorado. Overview of metadata recommendations, catalog system architecture, and proposed URI scheme.

NGDS & USIN Poster at DMT '10

Submitted by Ryan Clark on Wed, 05/26/2010 - 9:11am

This poster outlines the general goals of the National Geothermal Data System, describes its relationship to the US Geoscience Information Network, and shows some of the ways that the Arizona Geological Survey plans to make it easier for individual state surveys to contribute information to NGDS. It was presented in Sacramento by Ryan Clark at the Digital Mapping Techniques 2010 conference.

Presentation, USGS Menlo Park, 04/15/2010

Submitted by srichardAzgs on Fri, 04/16/2010 - 4:23pm

Powerpoint stack for presentation at USGS, Menlo Park, CA, 04/16/2010, by Steve Richard

USGIN posters at 2009 Geological Society of America (GSA) annual meeting

Submitted by wgrunberg on Wed, 10/21/2009 - 2:55pm

We presented the following posters at the 2009 GSA Annual Meeting in Portland, Oregon. Note, the PDF files are large.

XML ETL presentation at Data Preservation Workshop

Submitted by wgrunberg on Mon, 09/14/2009 - 6:30pm

Following is a presentation on XML Extract-Transform-Load (ETL) we did at the Geoscience Data Preservation Techniques Workshop, Indiana Geological Survey, Bloomington, IN on July 2009.

Although this presentation was geared towards National Geological and Geophysical Data Preservation Program (NGGDPP) metadata, the same tools can be used for other ETL needs.

ETL Debug Blog

Post to this group blog about your Extract-Transform-Load (ETL) efforts.

Helpful Regular Expression (regex) links

Submitted by wgrunberg on Wed, 05/12/2010 - 1:43pm

Following are helpful Regular Expression links

- Regular Expression HOWTO for Python
- Regular Expression Basic Syntax Reference
- Wikipedia: Regular Expression

Debugging Python scripts with Pylint

Submitted by wgrunberg on Wed, 03/17/2010 - 11:10am

I am currently using Pylint to check for bugs and formatting errors in my Python scripts. I used Brandon Corfman's nice Komodo Hacks: Integrating Pylint tutorial to set up Pylint and how to integrate it with the Komodo Edit IDE (see Easy setup of a Python 2.6 development environment on Windows about setting up Python).

Here is my **updated** **Pylint execution macro for Komodo Edit 5** which includes a verbose and an errors only reporting command.

Converting FGDC XML metadata records for WMS and WFS services to ISO 19139 via Python and GeoNetwork

Submitted by wgrunberg on Wed, 03/10/2010 - 1:04pm

This post describes a series of Python scripts to convert FGDC XML metadata for WMS and WFS services to ISO 19139 service metadata via GeoNetwork's OGC harvest service. Here are the steps:

1. Extract WMS and WFS GetCapabilities URLs from FGDC XML metadata and write them to a file.
2. Create a GeoNetwork Harvest Node from extracted GetCapabilities URLs and let GeoNetwork's OGC WMS/WFS harvester create ISO 19139 metadata from GetCapabilities response. Note: the resulting ISO 19139 metadata is only as good as the one in the GetCapabilities response. In addition, GetCapabilities responses do not include all fields necessary for minimum ISO 19139 metadata.
3. Copy some FGDC metadata entries (Title, Abstract, etc.) to newly created ISO 19139 metadata.

Our goal is to have working WMS and WFS metadata records in our CSW catalog that can be used to add the described services to analytical software such as ESRI's C-SW Client for ArcGIS Desktop. Following are the reasons why I am currently choosing this convoluted process instead of a direct FGDC to ISO 19139 metadata conversion:

Create a GeoNetwork OGC Harvest Node (WMS or WFS GetCapabilities to ISO 19139 metadata) through xml.harvesting.add request

Submitted by wgrunberg on Fri, 03/05/2010 - 11:05am

GeoNetwork offers a harvest service that, among other, follows a OGC WMS or WFS GetCapabilities URL and transforms the response into an ISO 19139 metadata record. Bear in mind that the resulting ISO metadata record is only as good as the GetCapabilities response and that the metadata can never be entirely ISO 19139 conformant without dummy values due to limitations of the OGC GetCapabilities schema.

GeoNetwork offers a handy user interface to add those "GeoNetwork Harvesting Nodes" and also exposes their functionality through a even handier XML service. See chapter 19.3 on "Harvesting Services" of the GeoNetwork opensource V 2.4 The Complete Manual. Following are my notes on creating an OGC harvesting node through GeoNetwork's harvesting service.

Authoritative fgdc-std-001-1998 (FGDC CSDGM) schema (XSD) and data definition document (DTD) URLs

Submitted by wgrunberg on Thu, 02/18/2010 - 3:51pm

Following are authoritative FGDC Content Standard for Digital Geospatial Metadata (CSDGM) XML schema document (XSD) and document type definition (DTD) URLs for FGDC-STD-001-1998. These documents are used to validate a FGDC XML metadata file.

- **http://www.fgdc.gov/metadata/fgdc-std-001-1998.dtd** - Usage:
  - <?xml version="1.0" encoding="ISO-8859-1"?>
    <!DOCTYPE metadata SYSTEM "http://www.fgdc.gov/metadata/fgdc-std-001-1998.dtd">
- **http://www.fgdc.gov/metadata/fgdc-std-001-1998.xsd** - Usage:
  - <metadata
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:noNamespaceSchemaLocation="http://www.fgdc.gov/metadata/fgdc-std-001-1998.xsd">

There are also annotated and modified FGDC-STD-001-1998 schema available for download:

FGDC XML schema woes

Submitted by wgrunberg on Wed, 02/10/2010 - 2:59pm

I am working on transforming FGDC XML metadata records to the USGIN 1.1 version of ISO 19139 metadata and having a hard time finding formal FGDC schema, name spaces, and schema locations. This is what I found so far:

XSLT transformations in Python through the Gnome libxml C parser

Submitted by wgrunberg on Mon, 11/09/2009 - 10:29am

- C
- ETL
- Python
- XSLT

This is an example script on how to do XSLT transformations in Python 2.6 through the Gnome libxml XML C parser and toolkit. The relative fast C library is available on multiple platforms but you will also need Python bindings for libxml2 and libxslt. There is a handy libxml2 and libxslt Python bindings installer for Windows which also includes the C libraries in DLL form.

Import MEF metadata archive into GeoNetwork through Python

Submitted by wgrunberg on Fri, 11/06/2009 - 1:53pm

This is a Python example script for importing GeoNetwork Metadata Exchange Format 1.1 (MEF) archives to GeoNetwork 2.4.2's mef.import service. The mef.import service requires a multipart/form-data POST through a modified MultipartPostHandler.py library which now supports Unicode (urllib2). It has been tested in Windows XP and Python 2.6.

Create GeoNetwork MEF files from ISO 19139 XML through Python

Submitted by wgrunberg on Fri, 11/06/2009 - 10:11am

This is an example Python script that showcases the creation of GeoNetwork Metadata Exchange Format 1.1 (MEF) archives from ISO 19139 metadata XML files.It has been tested in Windows XP and Python 2.6.

GeoNetwork authentication and CSW transactions through Python

Submitted by wgrunberg on Tue, 10/20/2009 - 10:51am

This is an example Python script that showcases GeoNetwork authentication, session handling, and CSW transactions.

XSLT to Transform WMS GetCapabilities response to CSW Insert transaction XML

Submitted by wgrunberg on Thu, 10/01/2009 - 10:13am

Tested with deegree-csw 2.3pre
Read the XSLT file for more information.

Attached is an example XSLT1 script to transform a WMS GetCapabilities 1.1.1 response to a CSW Insert transaction. The script is based on deegree's wms2iso19119.xsl (http://www.deegree.org/).
Note that currently it only supports WMS version 1.1.1 (<WMT_MS_Capabilities>) responses because it chokes on 1.3.0 (<WMS_Capabilities>) responses.

XML ETL presentation at Data Preservation Workshop

Submitted by wgrunberg on Mon, 09/14/2009 - 6:30pm

Following is a presentation on XML Extract-Transform-Load (ETL) we did at the Geoscience Data Preservation Techniques Workshop, Indiana Geological Survey, Bloomington, IN on July 2009.

Although this presentation was geared towards National Geological and Geophysical Data Preservation Program (NGGDPP) metadata, the same tools can be used for other ETL needs.

Building a GeoSciML WFS Server

I want to help make the process of building a GeoSciML WFS server easier. I'll be collecting links to relevant information, cookbooks, walkthroughs, etc. Hopefully this will chronicle my process of going from knowing practically nothing about how to implement a service to making one work.

Updated App-Schema Documentation

Submitted by Ryan Clark on Thu, 10/15/2009 - 1:51pm

The official documentation for the app-schema extension has been revamped! There is a lot more to look through here - absolutely a must-read for anyone interested in using GeoServer to serve GeoSciML features.

http://docs.geoserver.org/trunk/en/user/data/app-schema/index.html

A Very Useful Conversation

Submitted by Ryan Clark on Fri, 09/18/2009 - 3:22pm

I've been in the midst of a daily back-and-forth email discussion with the app-schema developers on the GeoServer-Users mailing list. Reading through the conversation may be helpful, as they've already resolved a number of issues I've had. Here's a link to the thread:

http://www.nabble.com/Geoserver-2.0-availability---included-app-schema-capabilities-td24624783.html

NCGMP to GeoSciML: MappedFeature

Submitted by Ryan Clark on Fri, 09/18/2009 - 3:20pm

I posted an example of the <typeMappings> section of a GeoServer/app-schema mapping file here: A Sample GeoServer Mapping file for NCGMP09 to GeoSciML. Checkout this other post to see where your mapping file should be, and how your datastore will refer to it: GeoServer Complex-Feature WFS Configuration: Directory Structure.

The <typeMappings> section is where you indicate what fields in your source data (NCGMP09) go into which fields in your output data (GeoSciML). Here I want to describe the different levels of complexity you can undergo when making a mapping file.

The Simplest Case: From one Field to Another

From NCGMP09 to GeoSciML (and back again)

Submitted by Ryan Clark on Fri, 09/18/2009 - 3:07pm

I'm almost done building a complete mapping file that will output GeoSciML MappedFeatures from polygon features in a database that follows the NCGMP09 schema. This schema, developed as part of the USGS National Cooperative Geologic Mapping Program, is a proposed standard database schema for

the publication of single digital geologic map. An ESRI Personal Geodatabase implementation of the NCGMP09 schema is available on the website linked above.

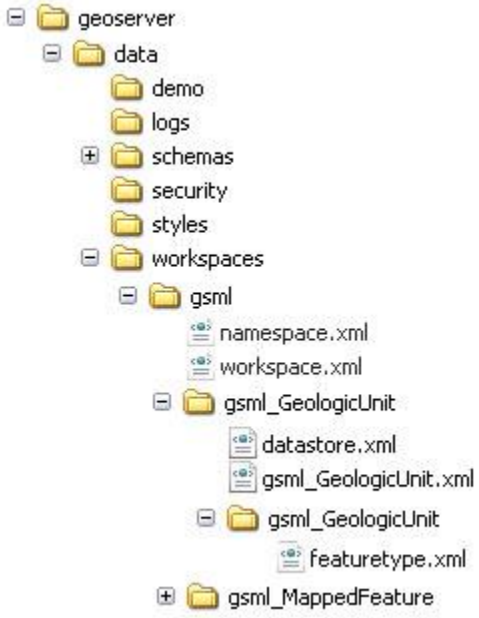
GeoServer Complex-Feature WFS Configuration: Data Connections in the Mapping File

Submitted by Ryan Clark on Fri, 09/18/2009 - 3:06pm

When you configure a complex feature service, you have to manually map your data connection parameters into an XML document called a mapping file. I found that the easiest way to make this connection is as follows:

GeoServer Complex-Feature WFS Configuration: Directory Structure

Submitted by Ryan Clark on Fri, 09/18/2009 - 2:57pm



At present, there is no user-interface component that helps you configure a complex-feature WFS in GeoServer. I'm sure that it is being worked on, but in the meantime, the configuration is done manually, by building the appropriate directory structure and putting XML configuration files in the right places.

- This is an example of the directory structure that has to be in place.
- Login to post comments
- Read more

It Works!

Submitted by Ryan Clark on Tue, 09/15/2009 - 10:58am

It works! Once you get the app-schema extension installed and the tutorial data in place, it works just fine. Make sure that in your GetFeature requests you specify the outputFormat as GML3, as it won't work at all in GML2 (which GeoServer defaults to if no outputFormat is specified).

Also, be aware that the app-schema extension is only supported at WFS version 1.1.0.

Now I can move on to beginning to map my data into GeoSciML features. I'll write some more walkthrough as I put it together.

GeoServer: Installing the App-Schema Extension

Submitted by Ryan Clark on Tue, 09/15/2009 - 10:56am

I've been spending the last couple of days working with GeoServer version 2.0-beta2, which introduces the ability to serve complex features. Serving complex features is a critical component of providing geologic data in GeoSciML format.

GeoServer accomplishes this through the use of what's called the "app-schema" extension. Unfortunately, there is not yet a user-interface for configuring the extension, and so you set up your complex features manually, by creating folders and configuration files in the right places that say the right things.

Complex Features in GeoServer

Submitted by Ryan Clark on Tue, 09/15/2009 - 10:55am

GeoServer has released a 2nd Beta of version 2.0. You can read about it on their blog.

This is exciting because they've introduced what they're calling the "app-schema extension". Essentially, it represents real support for complex features. This will open the door to GeoSciML delivery through the friendly GeoServer interface.

Groundwork: Basic software installations

Submitted by Ryan Clark on Mon, 09/14/2009 - 12:03pm

There have already been successful implementations of GeoSciML Web Feature services, using a few different approaches.

Getting Started

Submitted by Ryan Clark on Mon, 09/14/2009 - 11:57am

I'm starting the process of implementing a WFS that delivers GeoSciML data. I intend to make Group Posts here detailing what I'm trying, what works, what doesn't, in the hopes that it wil both help me keep track of what I've done, and also be of use to someone else attempting to do something similar. If I find things that are of particular use, I'll end up making more formal walkthroughs. I'm coming at this without having much background in open-source software, OGC services or XML, so I'll have a lot to learn.

CSW Debug Blog

Post to this group blog about your metadata Catalog Service for the Web (CSW) client and server implementation efforts and debugging sagas. I encourage occasional rants of the cathartic nature.

ISO 19139 service metadata record srv:DCPList codelist confusion

Submitted by wgrunberg on Tue, 11/24/2009 - 12:16pm

It appears that the GeoNetwork WMS harvest selects an incorrect codeListValue for a service metadata record's srv:DCPList (Distributed Computing Platform). Currently, it enters HTTP-GET and HTTP-POST. However, there is a bit of confusion within GeoNetwork and among standards and profiles. Specifically, one could interpret a conflict between DCP in a service's GetCapabilities (OGC) and a service's metadata record (ISO):

OGC Schema Issues in Relation to CSW

Submitted by Ryan Clark on Wed, 11/18/2009 - 11:38am

In order to create metadata for both static datasets and dynamic, online services and for use with CSW, the OGC created an xml schema that merges the schema for ISO19115 (dataset metadata) and ISO19119 (service metadata) (see secion D.1.5, page 105 in OGC 07-045). The way that was accomplished was by creating a schema located at http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd. The contents of that schema are quite simple -- it looks like this:

Connecting to GeoNetwork's MckoiDB through the Druid Database Manager

Submitted by wgrunberg on Fri, 11/13/2009 - 12:45pm

Here is a tutorial on how to connect to GeoNetwork's MckoiDB through the Druid Database Manager: http://geonetwork-opensource.org/documentation/how-to/druid_dbms

Note that the GeoNetwork 2.4.2 installer will not prompt you to provide a MckoiDB user name and password. Those are generated automatically and stored in

\geonetwork\web\geonetwork\WEB-INF\db\account.prop

Useful site for CSW testing and resources

Submitted by srichardAzgs on Thu, 10/29/2009 - 10:29am

CSW requirements

appears to mostly be compiled by Armin Retterath, Kompetenz- und Geschäftsstelle GDI-RP

Has listing of active CSW servers, some requirements listings

http://www.mapbender.org/CSW_Requirements

GeoNetwork 4.2.1 exploration notes v3

Submitted by wgrunberg on Thu, 10/08/2009 - 2:59pm

BUG: what I consider a bug
FEATURE: what I consider a missing feature
TODO: possibilities that need to be explored or decide on

GeoNetwork 4.2.1 exploration notes v2

Submitted by wgrunberg on Wed, 10/07/2009 - 9:22am

BUG: what I consider a bug
FEATURE: what I consider a missing feature
TODO: possibilities that need to be explored or decide on

GeoNetwork 4.2.1 exploration notes v1

Submitted by wgrunberg on Wed, 10/07/2009 - 9:08am

BUG: what I consider a bug
FEATURE: what I consider a missing feature
TODO: possibilities that need to be explored or decide on

CSW metadata editors we are exploring

Submitted by wgrunberg on Thu, 10/01/2009 - 3:20pm

Now that we know what kind of CSW services we can pull off, we need to pay more attention to creating and maintaining metadata. Right now, I just do a bunch of ETLs from our in-house data repositories (ArcSDE, Access, Excel, etc.) to CSW Insert Transaction XML files. That method is not user friendly and editing individual metadata records are a pain.

So, here are some CSW metadata editors I am looking at:

Ruminations on XML and large datasets

Submitted by wgrunberg on Wed, 09/16/2009 - 11:57am

While working on thoroughly understanding CSW & metadata related OGC and ISO standards, I can not but start questioning XML and its use to carry and format data. After all, USGIN's goal is not just to implement interoperable USGIN metadata and data profiles but also to encourage their widespread adoption. The following blog entry "How XML Threatens Big Data" by Michael Driscoll discusses some of the limitations XML poses to dealing with large datasets and ease of use.

Metadata about Metadata for a Service That Provides Metadata

Submitted by Ryan Clark on Wed, 09/16/2009 - 11:56am

... kind of sounds like overkill, but there's reason behind the confusion!

Imagine that you're trying to find some data. In the USGIN scheme, in order to find data you search through CSW services, which are catalogues full of ISO1939 metadata documents. Each of these documents points you to a dataset or service that is somewhere, hopefully online, for you to take a look at. These CSW services are like the card catalogues of the digital era.

List of interesting metadata implementation recommendations and profiles

Submitted by wgrunberg on Wed, 09/16/2009 - 11:55am

Following is a list of some metadata implementation recommendations and profiles I am looking at.

Configuring deegree-csw

Submitted by wgrunberg on Wed, 09/16/2009 - 11:54am

This is a list of our configuration changes too deegree-csw - specifically version 2.3-pre (missing link) with an PostgreSQL backend.

Follow instructions in deegree's documentation

See deegree_csw_2.0.2_documentation_en.pdf. Make sure the document is in a version corresponding to your version of deegree-csw (in this case "deegree Web Catalogue Service v.2.2").

Geospatial data and metadata protocols, services, standards, and profiles

Submitted by wgrunberg on Wed, 09/16/2009 - 11:50am

The beginnings of a list of geospatially enabled data and metadata protocols, services, standards, and profiles we are getting exposed to. Please, comment with corrections and additions.

What is the difference between CSW, CS-W, and CS/W?

Submitted by wgrunberg on Wed, 09/16/2009 - 11:49am

Well .. I couldn't find any difference between CSW, CS-W, and CS/W.

CSW Clients and Servers we are testing

Submitted by wgrunberg on Wed, 09/16/2009 - 11:47am

Following is a list of Web Service client, server, and hybrid applications we are aware of that are supposed to handle the CSW 2.0.2 protocol. Check out OGC's registry of products for more CSW implementations.

Public CSW services to test against

Submitted by wgrunberg on Wed, 09/16/2009 - 11:46am

Here are some CSW services I use to test CSW clients against

CatalogConnector CSW Client

Submitted by wgrunberg on Mon, 06/15/2009 - 12:17pm

USGIN CatalogConnector test client **http://catalog.usgin.org/catalogConnector/** (version 0.5)

This promising, stand-alone, Java-based, CSW web client is currently being developed by SDI Catalonia (IDEC) and USGIN at http://sourceforge.net/projects/catalogconnecto/

Appendix E: Establishment and Maintenance of Nodes and Hubs

1. Hub/Node selection
    a. Node Criteria and Participants (spreadsheet)……..2
    b. Node Specification Criteria……7
    c. Node Map Graphic……14
    d. Hub Deployment Timeline (spreadsheet) – submitted separately as Excel file through E-link 241.3 (report number- DOE-AZGS-2850-7). Also available online at https://github.com/ngds/documents/blob/master/AASG-State-Contributions/E-1-d-HubDeploymentTimeline.xlsx?raw=true.
2. Establishment Instructions
    a. NGDS Installation Instructions……15
    b. Reviewing, Deploying, Validating Instructions……48
    c. NGDS Software Requirements Summary (with comments)……92
    d. Geoserver Installation Instructions……150
3. Node Maintenance
    a. Data Replication Process Documentation: Hyper-V 2012 Replication on Workgroup Servers ……161
4. Hub Testing
    a. Python Script: http://resources.arcgis.com/en/help/main/10.1/index.html#/Example_Check_a_folder_for_stopped_services/0154000005tr000000/ - Submitted to OSTI
5. Data Quality Reporting
    a. Issue Tracker: https://github.com/ngds/metadata-issues - submitted to OSTI

NGDS Node Specification Criteria	Central Catalog Node	AASG Node	SMU Node	DOE-GDR (NREL)
1. The capabilities offered MUST play a role in making geothermal-relevant data assets accessible.	✓	✓	✓	✓
2. The capabilities offered MUST be publicly Web-accessible.	✓	✓	✓	✓
3. A node MUST offer at least one of the following capabilities:	✓	✓	✓	✓
a. Host a Web-accessible repository of geothermal-relevant data assets with metadata conforming to the USGIN ISO profile (USGIN Standards and Protocols Drafting Team, 2010-11) published through an NGDS catalog node.	✗	✓	✓	✓
b. Host a web-accessible folder that 1) contains NGDS-conformant metadata files and 2) is registered for harvesting by an NGDS catalog node.	✗	✗	✗	✗
c. Host NGDS-conformant web services (WMS, WFS, WCS, etc.) that are registered in an NGDS catalog.	✗	✓	✓	✗
d. Host an NGDS catalog node, which is a server operating a CSW 2.0.2 service that offers metadata conforming to the USGIN ISO profile.	✓	✓	✓	✓
Is it a Node?	✓	✓	✓	✓
Notes:			at present the SMU catalog is limited to metadata describing data services, not yet publications	

Above criteria are drawn from the NGDS Node Specification Criteria document at:
http://gdsdpwg.net/NGDSDesign/Shared%20Documents/NGDS%20Nodes/NodeSpecificationCriteria20130206.docx

NGDS Node Specification Criteria	Geothermal Prospector (NREL)	OIT GeoHeat Center Library	NBMG (UNR) Data Repository	AZ (hub)	IL (hub)
1. The capabilities offered MUST play a role in making geothermal-relevant data assets accessible.	✓	✓	✓	✓	✓
2. The capabilities offered MUST be publicly Web-accessible.	✓	✓	✓	✓	✓
3. A node MUST offer at least one of the following capabilities:	✗	✓	✓	✓	✓
a. Host a Web-accessible repository of geothermal-relevant data assets with metadata conforming to the USGIN ISO profile (USGIN Standards and Protocols Drafting Team, 2010-11) published through an NGDS catalog node.	✗	✓	✓	✓	✗
b. Host a web-accessible folder that 1) contains NGDS-conformant metadata files and 2) is registered for harvesting by an NGDS catalog node.	✗	✗	✗	✓	✓
c. Host NGDS-conformant web services (WMS, WFS, WCS, etc.) that are registered in an NGDS catalog.	✗	✗	✓	✓	✓
d. Host an NGDS catalog node, which is a server operating a CSW 2.0.2 service that offers metadata conforming to the USGIN ISO profile.	✗	✗	✗	✓	✗
Is it a Node?	✗	✓	✓	✓	✓
Notes:					

Above criteria are drawn from the NGDS Node Specification Criteria document at:
http://gdsdpwg.net/NGDSDesign/Shared%20Documents/NGDS%20Nodes/NodeSpecificationCriteria20130206.docx

As of 1/16/2014:

Some of these nodes are projecting to have more and more capabilities in the near future.

NGDS Node Specification Criteria	NV (hub)	KY (hub)	NH	KS	MN	WA	OR
1. The capabilities offered MUST play a role in making geothermal-relevant data assets accessible.	✓	✓	✓	✓	✓	✓	✓
2. The capabilities offered MUST be publicly Web-accessible.	✓	✓	✓	✓	✓	✓	✓
3. A node MUST offer at least one of the following capabilities:	✓	✓	✓	✓	✓	✓	✓
a. Host a Web-accessible repository of geothermal-relevant data assets with metadata conforming to the USGIN ISO profile (USGIN Standards and Protocols Drafting Team, 2010-11) published through an NGDS catalog node.	✗	✗	✗	✗	✗	✗	✗
b. Host a web-accessible folder that 1) contains NGDS-conformant metadata files and 2) is registered for harvesting by an NGDS catalog node.	✓	✓	✗	✗	✗	✗	✗
c. Host NGDS-conformant web services (WMS, WFS, WCS, etc.) that are registered in an NGDS catalog.	✓	✓	✓	✓	✓	✓	✓
d. Host an NGDS catalog node, which is a server operating a CSW 2.0.2 service that offers metadata conforming to the USGIN ISO profile.	✗	✗	✗	✗	✗	✓	✗
Is it a Node?	✓	✓	✓	✓	✓	✓	✓

Notes:

Above criteria are drawn from the NGDS Node Specification Criteria document at:
http://gdsdpwg.net/NGDSDesign/Shared%20Documents/NGDS%20Nodes/NodeSpecificationCriteria20130206.docx

NGDS Node Specification Criteria	WI	DE	AL	MT	HI	IN	MI
1. The capabilities offered MUST play a role in making geothermal-relevant data assets accessible.	✓	✓	✓	✓	✓	✓	✓
2. The capabilities offered MUST be publicly Web-accessible.	✓	✓	✓	✓	✓	✓	✓
3. A node MUST offer at least one of the following capabilities:	✓	✓	✓	✓	✓	✓	✓
a. Host a Web-accessible repository of geothermal-relevant data assets with metadata conforming to the USGIN ISO profile (USGIN Standards and Protocols Drafting Team, 2010-11) published through an NGDS catalog node.	✗	✗	✗	✗	✗	✗	✗
b. Host a web-accessible folder that 1) contains NGDS-conformant metadata files and 2) is registered for harvesting by an NGDS catalog node.	✗	✓	✗	✗	✗	✓	✗
c. Host NGDS-conformant web services (WMS, WFS, WCS, etc.) that are registered in an NGDS catalog.	✓	✓	✓	✓	✓	✓	✓
d. Host an NGDS catalog node, which is a server operating a CSW 2.0.2 service that offers metadata conforming to the USGIN ISO profile.	✓	✓	✓	✓	✓	✓	✓
Is it a Node?	✓	✓	✓	✓	✓	✓	✓

Notes:

Above criteria are drawn from the NGDS Node Specification Criteria document at:
http://gdsdpwg.net/NGDSDesign/Shared%20Documents/NGDS%20Nodes/NodeSpecificationCriteria20130206.docx

NGDS Node Specification Criteria	WA	ND	UT	AK	Node X	data value note
1. The capabilities offered MUST play a role in making geothermal-relevant data assets accessible.	✓	✓	✓	✓	✓	enter 1 or 0
2. The capabilities offered MUST be publicly Web-accessible.	✓	✓	✓	✓	✓	enter 1 or 0
3. A node MUST offer at least one of the following capabilities:	✓	✓	✓	✓	✗	calculated
a. Host a Web-accessible repository of geothermal-relevant data assets with metadata conforming to the USGIN ISO profile (USGIN Standards and Protocols Drafting Team, 2010-11) published through an NGDS catalog node.	✗	✗	✗	✗	✗	enter 1 or 0
b. Host a web-accessible folder that 1) contains NGDS-conformant metadata files and 2) is registered for harvesting by an NGDS catalog node.	✗	✗	✗	✗	✗	enter 1 or 0
c. Host NGDS-conformant web services (WMS, WFS, WCS, etc.) that are registered in an NGDS catalog.	✓	✗	✗	✗	✗	enter 1 or 0
d. Host an NGDS catalog node, which is a server operating a CSW 2.0.2 service that offers metadata conforming to the USGIN ISO profile.	✓	✓	✓	✓	✗	enter 1 or 0
Is it a Node?	✓	✓	✓	✓	✗	calculated
Notes:						

Above criteria are drawn from the NGDS Node Specification Criteria document at:
http://gdsdpwg.net/NGDSDesign/Shared%20Documents/NGDS%20Nodes/NodeSpecificationCriteria20130206.docx

How to be an NGDS node

NGDS Specification Version 1.0.
Last edit: 2014-11-14

History:

- 2013-05-17 Accepted by the NGDS GDSDP Workgroup.
- 2013-10-08 Minor revision to manifest file provisions made by SMR based on comments from Matt MacKenzie
- 2013-10-08 Post v1.0 to GitHub

Contents

Introduction 2
NGDS Node requirements 2
Details 3
Discussion 4
Functional NGDS node categories 5
Files-only node 5
Service node 5
Catalog Node 5
Data node 5
NGDS Node Self-Description Document 5
Standards Statements 6
Requirements 6
Status of this Memo 6
References 6

Introduction

This document proposes requirements that must be met in order to label an internet resource 'NGDS node'.

NGDS Node requirements

An NGDS node is a web-accessible server that hosts at least one of the functional capabilities enumerated below to play a role in making geothermal-relevant data assets accessible.

To be considered a node in the National Geothermal Data System, the following criteria must be met:

1. The capabilities offered MUST play a role in making geothermal-relevant data assets accessible.
2. The capabilities offered MUST be publicly Web-accessible.
3. A node MUST offer at least one of the following capabilities:
    a. Host a Web-accessible repository of geothermal-relevant data assets with metadata conforming to the USGIN ISO profile (USGIN Standards and Protocols Drafting Team, 2010-11) published through an NGDS catalog node.

b. Host a web-accessible folder that 1) contains NGDS-conformant metadata files and 2) is registered for harvesting by an NGDS catalog node.
c. Host NGDS-conformant web services (WMS, WFS, WCS, etc.) that are registered in an NGDS catalog.
d. Host an NGDS catalog node, which is a server operating a CSW 2.0.2 service that offers metadata conforming to the USGIN ISO profile.

NGDS nodes SHOULD self-identify by providing an **NGDS node self-description document**. The NGDS NODE SELF-DESCRIPTION DOCUMENT

Every NGDS node SHOULD offer an NGDS Node self-description document available via http GET at standard path. Recommended practice is to use the IETF "/.well-known/host-meta" path (see http://tools.ietf.org/html/rfc6415). Because many web-server configurations are set up to hide dot files (files or folders with names beginning with a '.') from directory browsing or even direct access, it is also recommended to a simple JSON-formatted manifest file in the root directory of the server. This file should contain the same information.

This document is a representation of the node resource. The purpose of this document is to make the node self-describing, both as a **Hypermedia** document for machine clients and as a web page for people to read. The document would be a publicly accessible HTML document, discoverable using standard web search technology (Google, Yahoo, Bing, etc.), with text description of the node, the node operator, node capabilities, expected lifetime, acknowledgements for support, etc. The template below defines an HTML profile for including information and links that a software agent could use to inspect the capabilities and access desired function. This might be encoded in HTML link elements in the head section, or using 'rel', 'class', and 'type' attributes on HTML anchors or other elements in the document. Some work needs to be done to define the vocabulary of 'rel' or 'class' attributes, or microdata properties

section (below) outlines proposed content that describes the node and enumerates the NGDS capabilities offered. The template is intended to be a hypermedia document that can be used by web applications to automate connection to node resources. If provided, the self-description document MUST be accessible at http://hostName.xxx/.well-known/host-meta, following the pattern specified by IETF RFC 6415 [http://tools.ietf.org/html/rfc6415].

Details

Hypermedia: Extension of text as an encoding scheme to any media stream that includes information and controls through which the user (or automaton) obtains choices and selects actions. (R. Fielding, 2008, http://roy.gbiv.com/untangled/2008/rest-apis-must-be-hypertext-driven (and discussion))

Metadata conforming to USGIN ISO profile: See "Use of ISO metadata specifications to describe geoscience information resources" (USGIN Standards and Protocols Drafting Team, 2010-11) for detailed guidance on this profile, and "Guidelines for implementation of USGIN metadata content recommendations in ISO 19139 XML metadata" (USGIN Standards and Protocols Drafting Team, 2012) for additional implementation details.

Metadata registered in an NGDS catalog node: an arrangement is in place so that the registering party creates metadata records that are harvested into the target catalog node. Metadata may be created through a forms interface directly to the catalog, harvested from a non-NGDS catalog end point (e.g OAI-PMH, open-search, ATOM etc.), harvested from a Web Accessible directory, or harvested by a file transfer/upload workflow

NGDS catalog node: a server operating a CSW 2.0.2 service that offers metadata conforming to the USGIN ISO profile that describes resources intended as part of the NGDS. Such nodes MUST be registered with the NGDS core aggregating node, either directly or indirectly through another node, such that its metadata may be harvested.

NGDS core aggregating node: The central system administration node, includes a catalog repository that harvests metadata from all NGDS catalog nodes. The NGDS core portal accesses this repository for system-wide searches.

NGDS conformant web service: a web service that implements an NGDS data exchange, defined by a service protocol, interchange format, and exchange-specific profile.

NGDS portal: a web application (NGDS client) that provides a user interface for searching the entire NGDS catalog.

NGDS home page: a web site that is the public face of the NGDS, providing access to the portal, information about the system, tutorials, governance information, and specifications.

NGDS roles:

- **originator**: agent that created the intellectual property (work)
- **contributor**: agent that prepares a resource for publication and makes it available
- **submitter**: agent that generates metadata to register the resource
- **owner**: agent that holds the intellectual property rights for a resource. Copyright holder, legal role.
- **provider**: agent that operates the service endpoint from which a resource is obtained
- **steward**: agent that is responsible for the maintenance/update of a resource

Discussion

A node is a server that provides NGDS resources, not a client that uses the offered resources. A server that hosts applications utilizing NGDS services is a vital part of the system, but is not considered a node. Such applications would be expected to be registered in the system by posting metadata to an NGDS catalog.

Applications enabling search across the entire NGDS catalog are an essential part of the system. At least one such search application will be recognized as the official portal to the system, accessed from the NGDS home page on the web.

Each node is under the stewardship of a data **provider**. Individual organizations may host more than one node. For example different servers may host OGC web services, a CSW service, and a repository containing

document resources, all under the stewardship of a single data provider. These would be considered three distinct nodes. A node may host data from more than one data **contributor**.

If a server is hosting resources that are described by metadata in an NGDS catalog, and those resources meet the 'geothermal-relevant' criteria, then the operator of that node MAY declare it to be an NGDS node, but doesn't have to.

Putting an NGDS self-description document on the server at the specified location constitutes a declaration that that server considers itself an NGDS node. One of the functions of the self-description document would be that a system registry of known NGDS nodes could have a web crawler application that searches for such self-description documents and starts a node-registration/approval/validation process when it finds new ones. This would enable a 'pull' model registration.

Functional NGDS node categories

Both data resources and metadata can be made accessible either packaged in files or through services that offer capabilities extending HTTP. Thus NGDS nodes can be categorized according to whether they present information in files or through web services. An individual node my offer both kinds of information access.

Files-only node

A publish-only node hosts web accessible files that are either information resources registered in an NGDS catalog, or NGDS-conformant metadata records in a web-accessible folder registered for harvest to an NGDS catalog.

Service node

A service node offers at least one NGDS service that is registered in an NGDS catalog, either through the metadata for datasets exposed by the service, or as a catalog service.

Catalog Node

A service node that offers a CSW service to search the metadata collection hosted by that node

Data node

A service node that offers at least one registered data service (e.g. WFS, WMS, WCS, OpenDAP)

NGDS Node Self-Description Document

Every NGDS node SHOULD offer an NGDS Node self-description document available via http GET at standard path. Recommended practice is to use the IETF "/.well-known/host-meta" path (see http://tools.ietf.org/html/rfc6415). Because many web-server configurations are set up to hide dot files (files or folders with names beginning with a '.') from directory browsing or even direct access, it is also recommended to a simple JSON-formatted manifest file in the root directory of the server. This file should contain the same information.

This document is a representation of the node resource. The purpose of this document is to make the node self-describing, both as a **Hypermedia** document for machine clients and as a web page for people to read. The document would be a publicly accessible HTML document, discoverable using standard web search technology (Google, Yahoo, Bing, etc.), with text description of the node, the node operator, node capabilities, expected lifetime, acknowledgements for support, etc. The template below defines an HTML profile for including information and links that a software agent could use to inspect the capabilities and access desired function. This might be encoded in HTML link elements in the head section, or using 'rel', 'class', and 'type' attributes on HTML anchors or other elements in the document. Some work needs to be done to define the vocabulary of 'rel' or 'class' attributes, or microdata properties

The content for the self-description page would include the following:

1. Node Name: [Name to identify this node]
2. Supporting organization: [name, contact point to report issues]
3. Description: [is there any topical focus, group of contributors, other special characteristics of capabilities offered]
4. Node type: term from controlled vocabulary to indicate type of node. One of {filesOnly, catalog, dataService, catalogAndDataService}
5. Resources:
    a. Data services: List of service types, links to end points, like ArcGIS server REST page?
    b. Metadata publishing profile: how exposed (WAF, CSW, other, including specific profiles used?)
6. How can content of this node?
7. URL for RSS feed for service announcements and information broadcasts.

See the accompanying example document. When the content is established, this could be converted to a template to facilitate creation of new documents for NGDS nodes.

Standards Statements

Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in IETF RFC-2119 [RFC2119].

Status of this Memo

This NGDS Specification licensed under the CC BY license. NGDS Specification Drafts are working documents of the National Geothermal Data System Technology Working Group. Note that other groups may also distribute working documents as NGDS Specification Drafts. Current NGDS Specification Drafts can be accessed in the NGDS/System-design gitHub repository.

NGDS Specification Drafts are draft documents valid until rejected or adopted by the NGDS GDSDPWG, and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use NGDS Specification Drafts as reference material or to cite them other than as "work in progress."

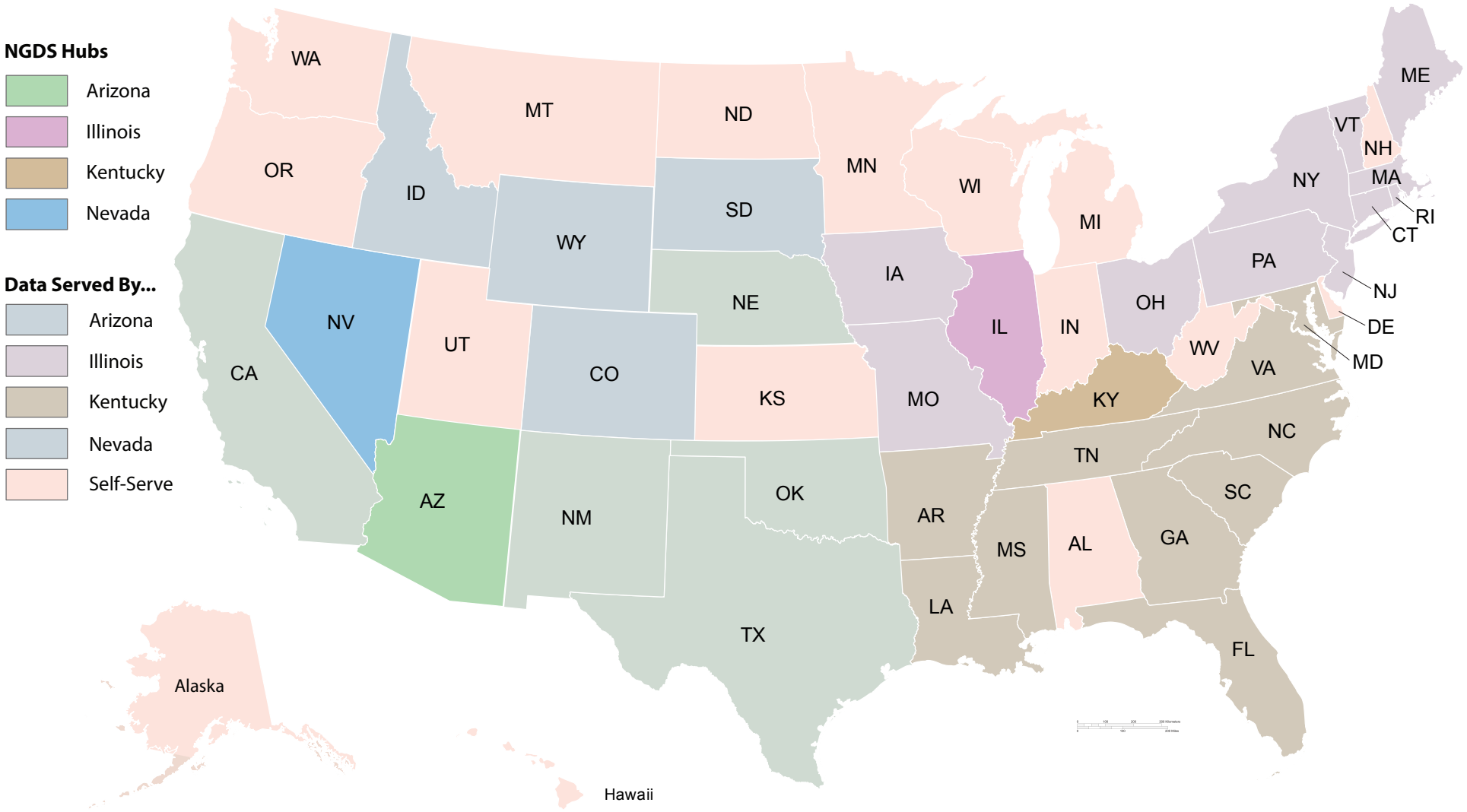
References

USGIN Standards and Protocols Drafting Team, 2010, Metadata Recommendations for Geoscience Resources: U.S. Geoscience Information Network Best Practices Document, Doc ID gin2010-11, v. 1.0.3, available at http://lab.usgin.org/profiles/doc/metadata-content-recommendations (accessed 2010-11-18).

USGIN Standards and Protocols Drafting Team, 2012, Guidelines for implementation of USGIN metadata content recommendations in ISO 19139 XML metadata: U.S. Geoscience Information Network Best Practices Document, available at http://repository.usgin.org/uri_gin/usgin/dlio/499.

USGIN Standards and Protocols Drafting Team, 2010-11, Use of ISO metadata specifications to describe geoscience information resources, Doc ID gin2010-009, available at http://repository.usgin.org/uri_gin/usgin/dlio/337.

NGDS Hub States



Updated January 15, 2015

Siemens AG

National Geothermal Data System
Software Installation Instructions Document
Version: 0.91

CONTENT OWNER: Christoph Kuhmuench

Version:	Date:	Author:	Details
0.1	Roberto Silva Filho	05/28/2013	Initial Draft Created
0.2	Monica McKenna	06/11/2013	Minor updates
0.3	Monica McKenna	06/24/2013	Combining comments from a few people
0.4	Monica McKenna	06/25/2013	Added appendix with summary of development.ini changes
0.5	Monica McKenna	07/22/2013	A little re-organization, more hints, and added gdal
0.6	Monica McKenna	07/24/2013	Updating with feedback
0.7	Christoph Kuhmuench	12/26/2013	Updating to latest installer.
0.8	Jordan Matti	2/7/2014	Many changes; tracked changes best viewed *without* showing format changes (**Review** pane -> **Show Markup** -> **Formatting**)
0.9	Christy Caudill	2/19/2014	Moved Window OS/ Oracle VM install to Appendix A.
0.91	Jordan Matti	2/20/2014	Minor changes (formatting, headers, comments)

All future revisions to this document shall be approved by the content owner prior to release.

1. Contents

1. Contents i

1. Preface 2

1. Purpose and Audience 2

2. Document Roadmap 2

3. System Scope and Background 2

1.3.1. What is the NGDS Software Stack? 3

2. Quick Installation Guide 3

3. Prerequisites 4

4. Install the NGDS Software Stack 5

1. Update Apt-Get 6

2. Install a Java Development Kit (JDK) 6

3. Install Git 6

4. Obtain the NGDS Software Stack Installation Files 7

5. Set Installation Parameters 7

6. Run the Installation Script 7

7. Final Steps 7

4.7.1. Start Apache Tomcat 7

4.7.2. Create a "Public" Organization for Your Node 8

5. Troubleshooting your NGDS Installation 8

Appendix A Installing a Virtual Machine on a Windows Operating System 10

1. Creating an Ubuntu Linux Virtual Machine using VirtualBox . 10

5.1.1. Download and install Oracle VM VirtualBox Manager 10

5.1.2. Create an Ubuntu Linux Virtual Machine 10

5.1.3. Configure your new Virtual Machine 13

5.1.4. Download an Ubuntu ISO image 15

5.1.5. Mount the Linux installation .ISO file in your virtual machine 15

5.1.6. Install Ubuntu Linux 12.04 16

5.1.7. Take a Snapshot 18

2. Accommodating a corporate firewall (OPTIONAL) 18

5.2.1. Install and Configure CNTLM 18

5.2.2. Configure your virtual machine environment to use CNTLM as its proxy 19

5.2.3. What to do if cntlm and proxy continue to cause issues 21

Appendix B The development.ini file 22

Development.ini Considerations 23

Appendix C Architectural and Deployment Diagrams 24

Domain Model 25

Notes on CKAN in Production Mode 25

Known Issues 26

List of Figures

Figure 1: NGDS is a grid of repositories ..... **Error! Bookmark not defined.**

Figure 2:Create a new linux virtual machine ..... **Error! Bookmark not defined.**

Figure 3:Set the VM memory to at least 1024MB ..... **Error! Bookmark not defined.**

Figure 4:Create a virtual HD. The disk size can be changed later **Error! Bookmark not defined.**

Figure 5:Specify the image type..... **Error! Bookmark not defined.**

Figure 6: Dynamically allocated disk space will allow the image to grow as new data is imported into NGDS ..... **Error! Bookmark not defined.**

Figure 7:Configure the hard drive size. We recommend large values, based on the amount of data the node should store. The minimum is 24GB. ..... **Error! Bookmark not defined.**

Figure 8:Oracle VM Linux machine configuration ..... **Error! Bookmark not defined.**

Figure 9:You may wish to enable the shared clipboard..... **Error! Bookmark not defined.**

Figure 10:9 Mounting ISO image in the Linux VM ..... **Error! Bookmark not defined.**

Figure 11: Linux Ubuntu installation screen ..... **Error! Bookmark not defined.**

Figure 12:Login for the new VM ..... **Error! Bookmark not defined.**

Figure 13: Installing guest additions ..... **Error! Bookmark not defined.**

Figure 14: NGDS Software Stack in Production Mode. ..... **Error! Bookmark not defined.**

Figure 15: NGDS System Context View ..... **Error! Bookmark not defined.**

Figure 16: NGDS Domain Model as a Class Diagram ..... **Error! Bookmark not defined.**

Figure 17: NGDS Domain Model of Resources as a Class Diagram 27

Figure 18: NGDS High-level Components 28

1. Preface

National Geothermal Data System (NGDS) was a Department of Energy-funded effort to facilitate public access to information about geothermal resources from public and private sources. NGDS data is available through a distributed, scalable network of data providers.

1. Purpose and Audience

This document is a step-by-step tutorial to help new developers and users setup both **Publisher** and **Aggregator** instances of the **NGDS Software Stack**.

This document is intended for a technical audience who need to understand the concepts and the reasoning of the installation process. Targeted audiences include:

- **NGDS System Administrators**
- Software Architects
- Software Developers
- DoE Monitors

This document purports to:

- Describe NGDS and how it works
- Identify and describe the basic components that are necessary to install the NGDS Software Stack
- Provide step-by-step installation instructions for the NGDS Software Stack
- Outlines the process of installing the NGDS Software Stack in **production** mode, as compared to **development** mode.

This fulfills one of the main goals of NGDS: to provide a basis for a sustainable open-source software project that is attractive for an open source team to maintain. With this documentation, future NGDS System Administrators will be able to quickly understand the system and become productive. This document accommodates the possibility that an alternative organization may take over architectural oversight of NGDS software in the future.

2. Document Roadmap

This document outlines the architecture of NGDS and is structured in the following way:

- Section 2: A quick installation guide for advanced users.
- Section 3: NGDS Software Stack prerequisites
- Section 4: NGDS Software Stack installation guide
- Section 5: NGDS Software Stack installation troubleshooting
- Appendix A: Installation guide for an Ubuntu Linux virtual machine in VirtualBox
- Appendix B: An overview of the **development.ini** file
- Appendix C: NGDS architecture and diagrams and notes

3. System Scope and Background

NGDS is a distributed data-sharing network. NGDS data providers host data using their own computing resources and submit metadata describing their data to web-accessible NGDS metadata repositories (referred to hereafter as **publisher** nodes).

Metadata submitted to registered **publisher** nodes is regularly harvested by NGDS **aggregator** nodes, which create web-accessible metadata catalogs. A metadata catalog can be used to find and access any data described by metadata records in the catalog.

NGDS **publisher** and **aggregator** nodes are provided by the **NGDS Software Stack**.

NGDS requirements are discussed in detail in the requirements document [REQ2.7]. The software Architecture is discussed in detail in the software architecture Document [ARCH1.0].

1.3.1. What is the NGDS Software Stack?

The **NGDS Software Stack** is a collection of applications designed to interact with NGDS data, metadata, and interchange formats.

When installed, the NGDS Software Stack allows the computer on which it is installed to become an NGDS node. There are two types of NGDS nodes:

- **Publisher** nodes: When installed on a server and configured to act as a **publisher** node, the NGDS Software Stack provides a web-accessible interface that can be used to submit and manage metadata records. Metadata that has been added to a **publisher** node that has been *registered* with an NGDS **aggregator** node will be harvested by the **aggregator** node at regular intervals.
- **Aggregator** nodes: When installed on a server and configured to act as an **aggregator** node, the NGDS Software Stack provides a web-accessible metadata catalog to which harvested metadata records are added. An NGDS **aggregator** node will harvest metadata from any registered catalog service (CSW) or any registered NGDS **publisher** node.

Note that the NGDS Software Stack can also be installed in two *modes*:

- **Production** mode: a stable release of the software stack
- **Development** mode: used by developers to create new versions of the software stack for subsequent release

2. Quick Installation Guide

This section provides an overview of the installation process for the NGDS Software Stack. This section should only be used by experienced system administrators and developers who have installed NGDS before.

To install the NGDS Software Stack:

1. Install Xubuntu Linux 12.04 64-bit LTS and create a user with super-user privileges and the following credentials
    a. Username: ngds
2. Login as: ngds
3. Open an Ubuntu Linux terminal and execute the following commands:

```
% sudo apt-get update
% sudo apt-get upgrade
```

4. Install the Oracle Java Development Kit (JDK)
    a. Add the Oracle JDK path to PATH
    b. Set the JDK_HOME environment variable
5. Install Git
6. Download the NGDS CKAN Extension by executing the following commands:

```
% cd ~
% mkdir tmp
% cd temp
% git clone https://github.com/ngds/ckanext-ngds.git
% cd ckanext-ngds/installation
% sudo ./install_ngds.sh
```

7. Start Apache Tomcat and execute the following commands:

```
% cd /opt
% sudo chown –R ngds.ngds ngds
% cd /opt/ngds/tomcat/bin
% ./catalina.sh run
```

If the installation was performed correctly, the web-accessible interface provided by the NGDS Software Stack can be reached at:

http://127.0.0.1/

Having navigated to the above address, perform the following:

1. Log in with the following credentials:
   - Username: admin
   - Password: admin
2. Navigate to the following URL:
   - http://127.0.0.1/organization
3. Add a new organization; the name of the organization should be:
   - Public

Note: **admin** is not a secure password. Though we will use it in examples for the remainder of this document, we recommend that you change the password to something more secure if you are concerned about security.

3. Prerequisites

Installing and configuring the individual components utilized by the NGDS Software Stack requires a physical or virtual computer with the following properties:

- Network access
- A properly configured Ubuntu Linux distribution 12.04 or higher operating system installed (example: Xubuntu 13.04 desktop-i386.iso)
- A user account with Super-User (Administrator) privileges
- At least 1024 megabytes of RAM; a physical computer that will be used to host a virtual machine should have sufficient RAM to allocate at least 1024 MB of RAM to a virtual machine

Appendix A of this document describes the steps necessary to create your own virtual machine and install Ubuntu Linux; Section 4 of this document describes the steps necessary to install the NGDS Software Stack as both a **publisher** and **aggregator** node.

4. Install the NGDS Software Stack

The NGDS Software Stack depends on a number of operating system components that must be installed on a computer before that computer can become an NGDS node. These components include:

- Java Development Kit (JDK)
- Git
- Apache SOLR
- PostgreSQL database
- PostgreSQL extensions for Geographical Information Systems (POSTGIS)
- Geoserver
- Apache Tomcat
- CKAN
- Python extensions
- gdal

To install these components, the computer on which they will be installed must have access to the Internet.

Many of these components will be installed automatically by the NGDS Software Stack installation script; the **Java Development Kit**, **Git**, **Apache SOLR**, and **Geoserver** must be installed manually.

Figure 14 provides a visual representation of the manner in which these components interact. Components near the top of the figure are *nested* within components near the bottom of the figure:

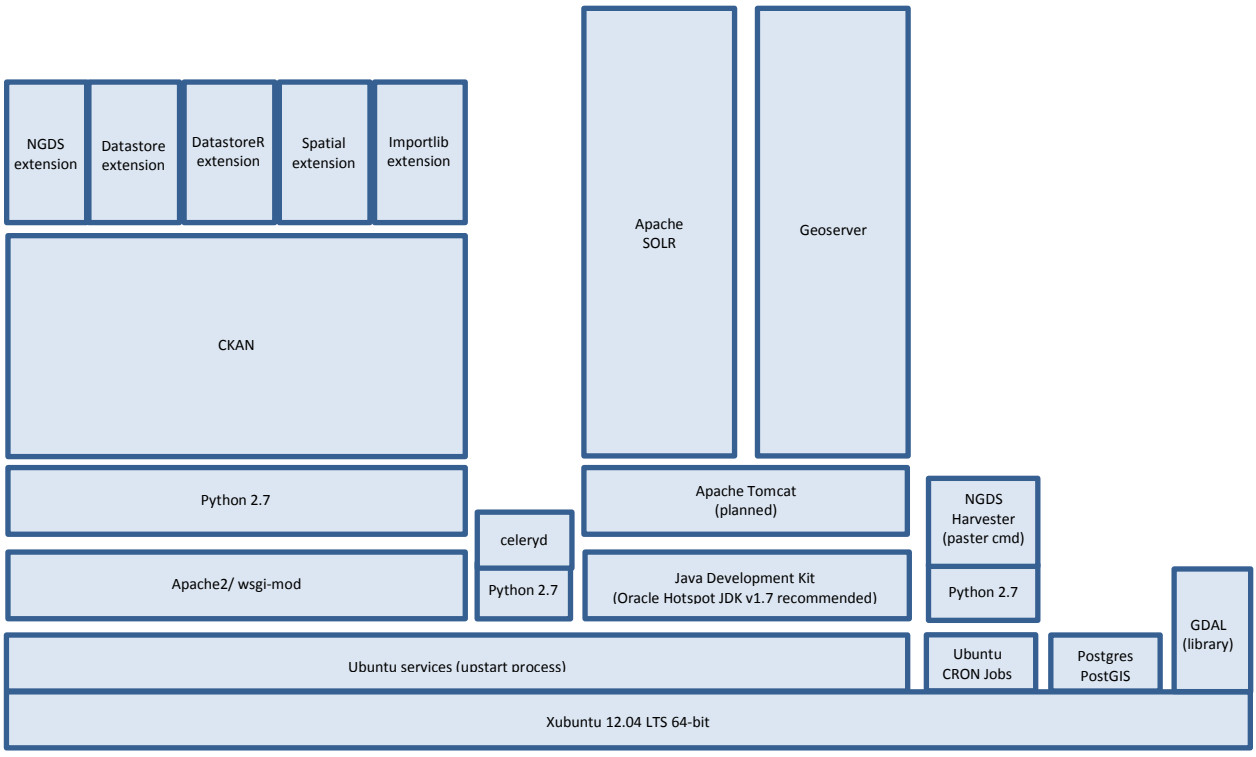


Figure 1: NGDS Software Stack in Production Mode

1. Update Apt-Get

Before installing the components listed in Section 4, you should verify that your **apt-get** program is working and up-to-date

To do so, open an Ubuntu Linux terminal and execute the following commands:

```
% sudo apt-get update
% sudo apt-get upgrade
% sudo reboot
% sudo apt-get upgrade
```

The **reboot** command is used because some upgrades require you to restart your machine before subsequent upgrades can be applied.

2. Install a Java Development Kit (JDK)

The first component of the NGDS Software Stack is a Java Development Kit (JDK) of your choice; the Oracle Hotspot JDK is recommended.

To install the Oracle Hotspot JDK, you will need to download an archive from the Oracle website:

http://www.oracle.com/technetwork/java/javase/downloads/index.html

Be sure to download the 64-bit version. At the time of writing of this document, the most up-to-date Oracle JDK 64-bit archive is:

jdk-7u45-linux-x64.tar.gz

After downloading the archive, extract it by opening an Ubuntu Linux terminal and executing the following commands:

```
% cd /opt
% sudo tar xvfz ~/Downloads/jdk-7u45-linux-x64.tar.gz
% sudo ln –s jdk-7u45-linux-x64 jdk
```

Navigate to the **etc** directory and use a text editor to edit the **environment** file there. Add the environment variable JAVA_HOME and modify the PATH variable so it appears as follows:

```
PATH=/opt/jdk/bin:<previous path value>
JAVA_HOME=/opt/jdk
```

To make sure that all changes are in effect reboot the system, open an Ubuntu Linux termal, and execute the following command:

```
% java –version
```

The command should return a version number of the installed JDK.

3. Install Git

To install the **git** program, make sure you are logged in as **ngds**.

Open an Ubuntu Linux terminal and execute the following commands:

```
% sudo su
% sudo apt-get install git git-core
```

4. Obtain the NGDS Software Stack Installation Files

To obtain the installation files for the NGDS Software Stack, open an Ubuntu Linux terminal and execute the following commands:

```
%  cd ~ngds
%  mkdir tmp
%  cd tmp
%  git clone https://github.com/ngds/ckanext-ngds.git
```

5. Set Installation Parameters

Before running the NGDS Software Stack installation script, you will need to specify any required installation paramaeters. To do so, navigate to the following directory:

ckanext-ngds/installation/

Use a text editor to edit the following file:

install-ngds.sh

The most important variables to specify are:

- APPS (line 53): Installation base directory. Default is: /opt/ngds
- DEPLOYMENT_MODE (line xx): Choose between *central* or *node* mode. Default is: *node*

Other variables such as the Apache Tomcat home directory can be configured as well.

Do not change anything beyond line 95, which reads, “DO NOT CHANGE ANYTHING BELOW THIS POINT”.

6. Run the Installation Script

Open an Ubuntu Linux terminal and execute the following commands:

```
%  cd ~ngds
%  cd tmp
%  sudo ./install_ngds.sh
%  cd /opt
%  sudo chown –R ngds.ngds ngds
```

This will install the NGDS Software Stack.

7. Final Steps

4.7.1. Start Apache Tomcat

After you have installed the NGDS Software Stack, you need to run the Apache Tomcat server software hosting Apache SOLR and Geoserver. To do so, open an Ubuntu Linux terminal and execute the following commands:

```
%  cd /opt/ngds/tomcat/bin
%  ./catalina.sh run
```

The web-accessible interface provided by your NGDS Software Stack can now be reached by navigating to the following IP address in a web browser of your choice:

http://127.0.0.1

4.7.2. Create a "Public" Organization for Your Node

If the installation was performed correctly, the web-accessible interface provided by the NGDS Software Stack can be reached at:

http://127.0.0.1/

Having navigated to the above address, perform the following:

4. Log in with the following credentials:
    - Username: admin
    - Password: admin
5. Navigate to the following URL:
    - http://127.0.0.1/organization
6. Add a new organization; the name of the organization should be:
    - Public

Note: **admin** is not a secure password. Though we will use it in examples for the remainder of this document, we recommend that you change the password to something more secure if you are concerned about security.

After these steps the system is configured to be used.

5. Troubleshooting your NGDS Installation

If the installation seems to stall out, check the output of the installation script to look for error messages

The most common errors are:

1) **Typos**: typos appearing in commands or paths can be very difficult to spot and can sometimes lead to unclear error messages. Check your text and paths carefully. Some scripts (such as BASH) are case-sensitive, so a lower-case or upper-case letter in the wrong place can cause problems.
2) **Permission Errors**: Permission errors occur when you try to perform an action without super-user capabilities. If you notice permission errors, use the **sudo** command ("super-user do") to open up permissions on the directories involved.
3) **Path Errors**: Some paths in the **development.ini** file terminate in a **slash** character (/) and some do not; the presence or absence of a **slash** character can impede an installation. Compare your **development.ini** file with the entries listed in Appendix A of this document to make sure your **development.ini** file has **slash** characters in the right places. See Appendix A for more information on the development.ini file.

After evaluating the output of the installation script and fixing any errors you find, re-run the installation script.

If the NGDS CKAN Extension website fails to load correctly after running the installation script without apparent errors, we recommend running the **paster** command through Eclipse.

Eclipse is debugging software that can, in some cases, locate errors in a given operation and allow you to fix them as they are detected. Typically, Eclipse identifies typos or path errors.

See the Eclipse wiki (http://wiki.eclipse.org/Main_Page) for more information.

Appendix A Installing a Virtual Machine on a Windows Operating System

If you will be installing the Ubuntu Linux operating system on a virtual machine, you will need to choose appropriate virtualization software. Virtual machines are supported by virtualization software that provides an abstract hardware representation emulating real host hardware. Virtualization allows the installation of a full operating system within a host OS.

In other words: a virtual machine is a computer that is created by a software application running on a host computer; a virtual machine therefore exists *entirely* within the memory of the host machine on which it is hosted. The obvious advantage here is that the resources of a single powerful host machine can be allocated to host many virtual machines for different purposes.

Though a virtual machine is not required for this project, this project requires a specific configuration of the Ubuntu Linux OS; a virtual environment is ideal for the installation of this configuration.

Currently, two free virtual environment managers are available: VMware Player, and Oracle VirtualBox. They can be downloaded on the links below:

- VMWare Player: http://www.vmware.com/products/player/
- Oracle VM VirtualBox: https://www.virtualbox.org/wiki/Downloads

This tutorial was developed using VirtualBox version 4.2.10 for Windows. Here, we install Linux Ubuntu 12.04 LTS from Canonical.

1. Creating an Ubuntu Linux Virtual Machine using VirtualBox

The steps in Section 3.2 of this document describe the installation of Ubuntu Linux (or Xubuntu) on a virtual machine supported by version 4.2.10 of Oracle VM VirtualBox. Newer versions of VirtualBox can be utilized.

5.1.1. Download and install Oracle VM VirtualBox Manager

Download and install version 4.2.10 of Oracle VM VirtualBox on a Windows computer of your choice. Doing so allows your Windows computer to support one or more virtual machines.

Download the software from: https://www.virtualbox.org/wiki/Downloads

Run the installer and follow the on-screen instructions to install VirtualBox.

5.1.2. Create an Ubuntu Linux Virtual Machine

Run the VirtualBox application installed in Section 2.2.1 and use it to create a virtual machine:

1. Run the VirtualBox application
2. Create a new virtual machine
3. Specify the following (Figure 1):
    a. **Name**: NGDS
    b. **Type**: Linux

c. **Version**: Ubuntu

4. Allocate at least 1024 MB of ram to your virtual machine (Figure 2)
5. Create a hard drive for your new virtual machine (Figure 3)
6. Specify the type of hard drive used by your virtual machine (Figure 4); the drive type you select determines the compatibility of the virtual hard disk you create with different virtualization software
7. Specify disk space allocation (Figure 5); dynamic allocation allows your virtualization software to allocate more hard drive space from the virtualization platform to this virtual hard drive as-needed
8. Allocate disk space to your virtual hard drive (Figure 6); this allocates a specified amount of hard drive space from the virtualization platform to the virtual machine

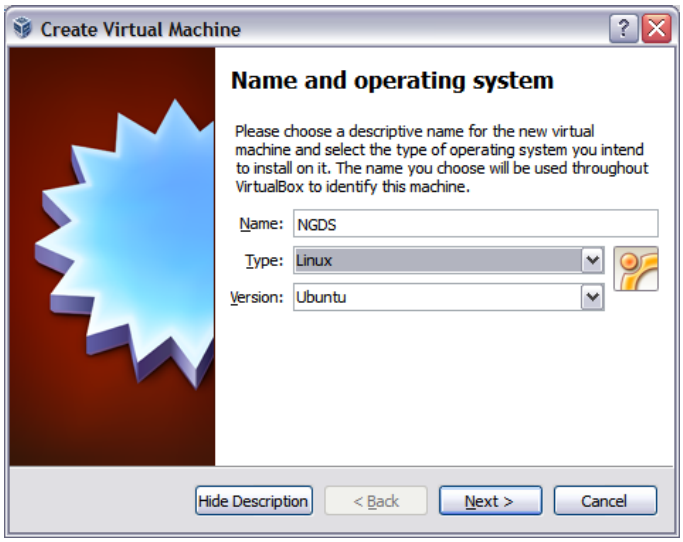


Figure 2: Create a new Linux virtual machine

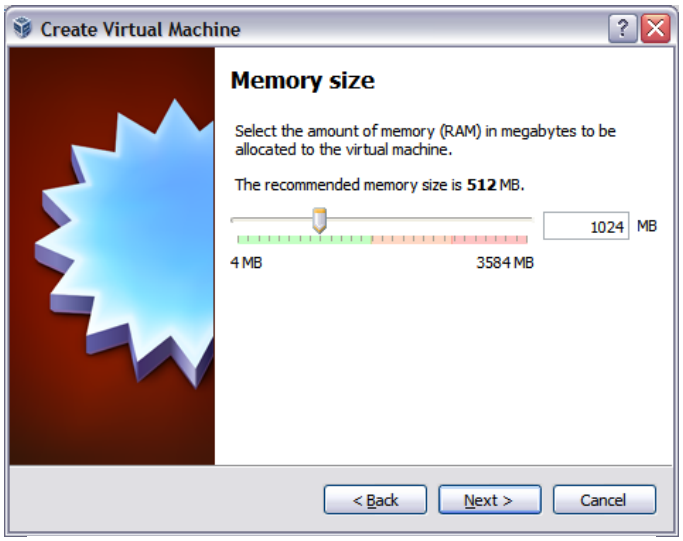


Figure 3: Allocate at least 1024 MB of memory to the VM

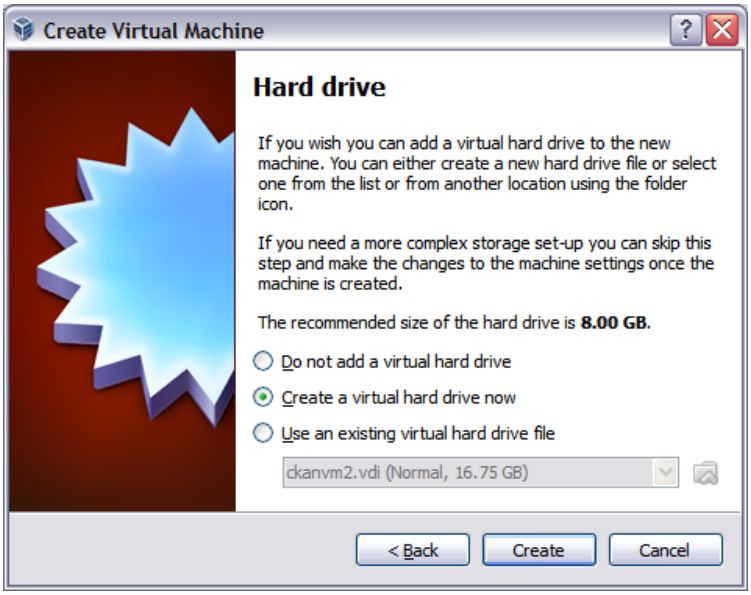


Figure 4: Create a virtual hard disk

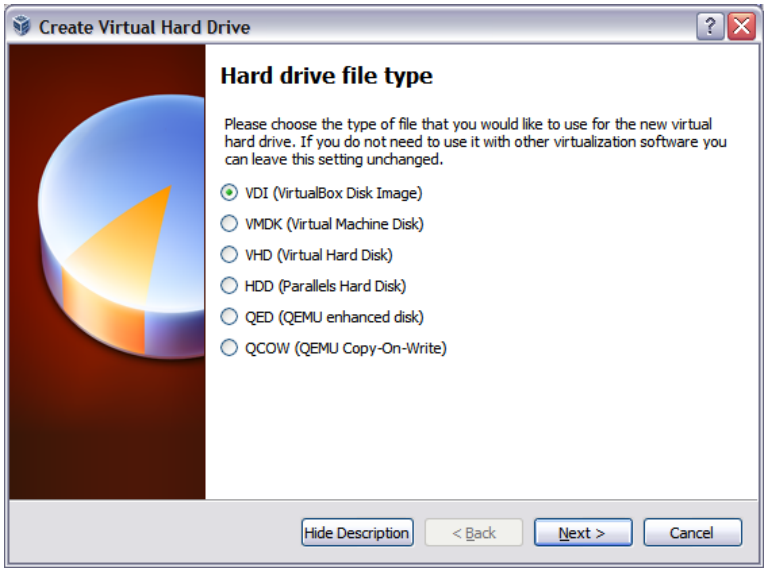


Figure 5: Specify the image type

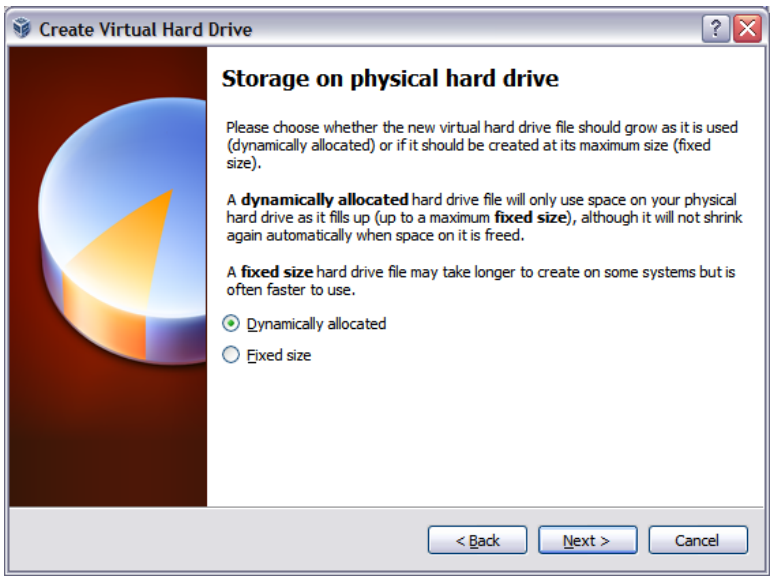


Figure 6: Specify storage allocation

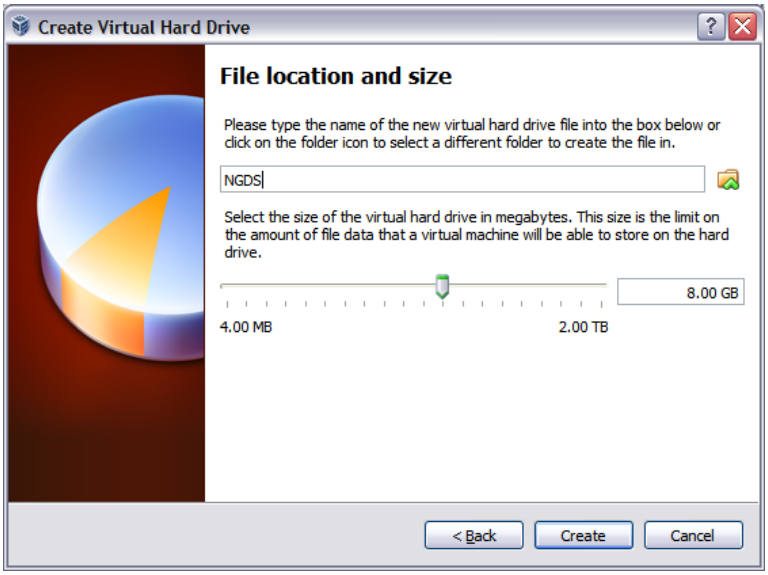


Figure 7: Configure virtual hard drive

5.1.3. Configure your new Virtual Machine

Open the **Oracle VM VirtualBox Manager** (Figure 7); select the virtual machine you created in Section 3.2.2 and click **Settings**.

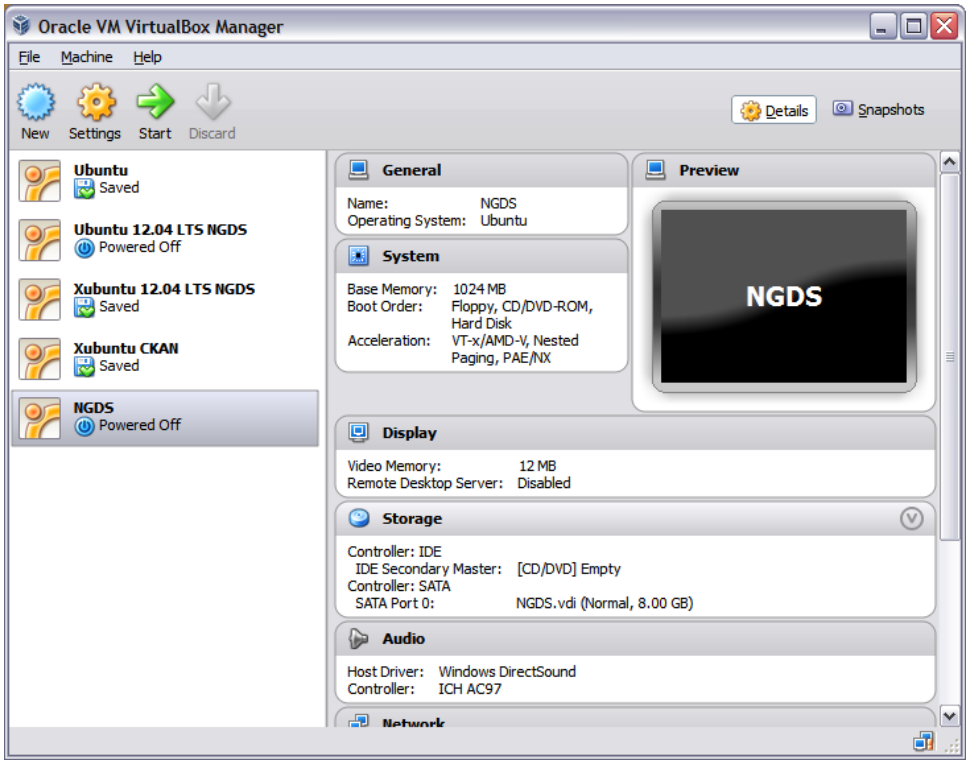


Figure 8: Configuring a virtual machine in VirtualBox

First, enable the **Shared Clipboard**:

1. Select **General** settings
2. Select the **Advanced** tab (Figure 8)
3. Click the **Shared Clipboard** dropdown menu
4. Select **Bidirectional**

This will enable anyone who connects to this virtual machine to copy and paste between the virtual machine and the computer used to connect to the virtual machine (including the computer on which the virtual machine is hosted).

Virtual machines, being virtual, are distinct from the computer that is used to connect to them and therefore do not necessarily share the same clipboard by default.

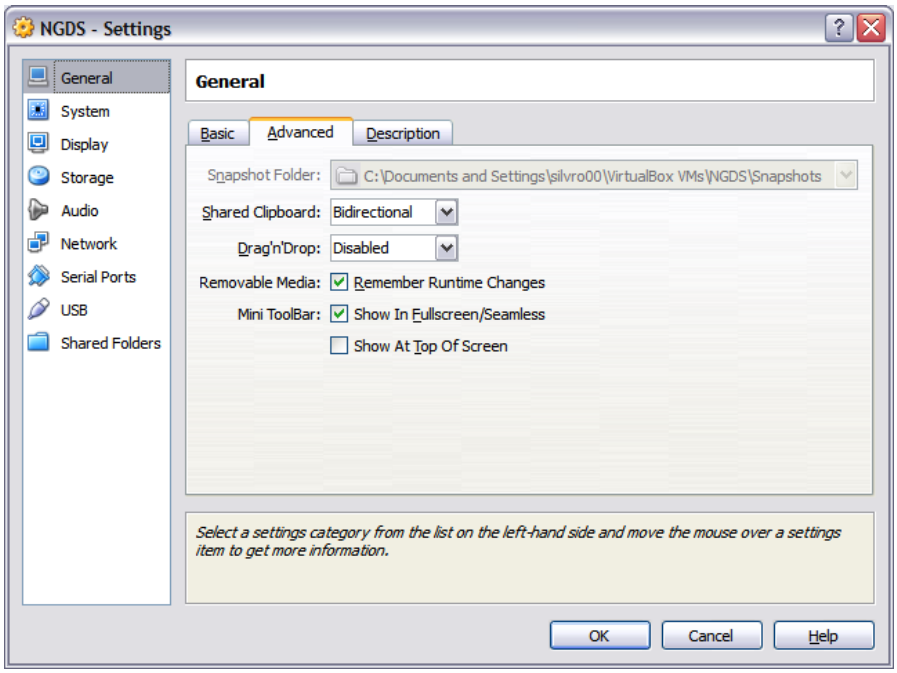


Figure 9: Enabling the shared clipboard

Now that your virtual machine is created and configured, you can install an existing Linux distribution.

5.1.4. Download an Ubuntu ISO image

An ISO image is a type of virtual CD (ISO stands for *International Standards Organization*; an ISO image is an image of an ISO-standard CD).

CD images are files that can be loaded and read by virtual CD drives. Virtual CD drives are software applications that emulate a CD-ROM drive in much the same way that an entire computer can be emulated by virtualization software.

To install Ubuntu on a virtual machine, you will need an ISO image of an Ubuntu installation file, available at: http://releases.ubuntu.com/12.04/

This tutorial utilizes the Long Term Service (LTS) version of Ubuntu, which features long-term support (3 years): http://releases.ubuntu.com/12.04/ubuntu-12.04-desktop-i386.iso

5.1.5. Mount the Linux installation .ISO file in your virtual machine

After downloading an ISO image but *before* starting it, mount it within the VirtualBox environment and use it to install the Ubuntu operating system on the virtual machine you created in Section 3.2.2:

1. In the **Oracle VM VirtualBox Manager** (Figure 8), select the virtual machine you created in Section 3.2.2 and click **Settings**
2. In the **Settings** window, click **Storage** (Figure 9).
3. In the **Storage** panel, click the **CD** icon next to the **CD/DVD Drive** dropdown menu.

4. Navigate to the ISO image you downloaded in Section 3.2.4; select the ISO file
5. In the **Storage** panel, click **OK** to mount the image

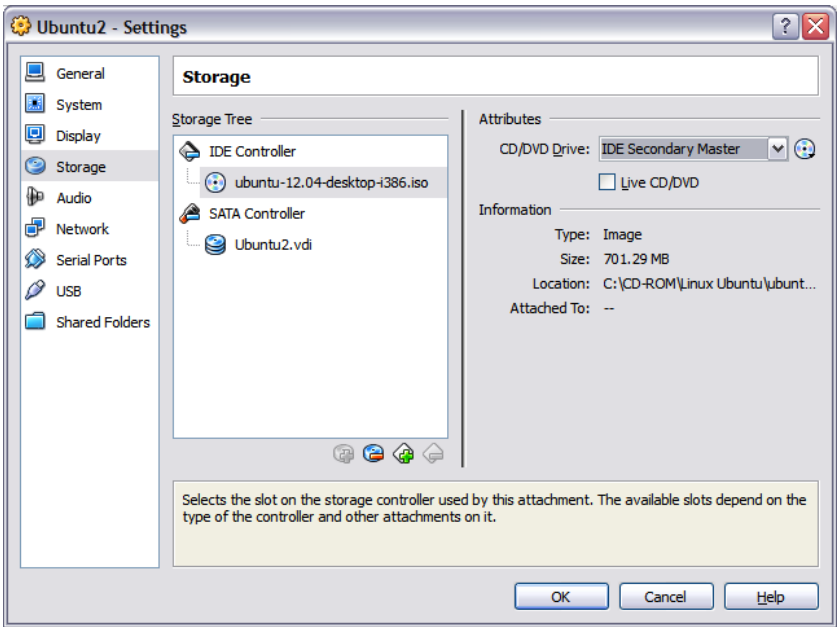


Figure 10: Mounting the Ubuntu ISO image in the VM

5.1.6. Install Ubuntu Linux 12.04

In the **Oracle VM VirtualBox Manager** (Figure 7), select the virtual machine you created in Section 3.2.2 and click **Run**. This will start your virtual machine.

When started, your virtual machine will prompt you to install the operating system loaded in the image in much the same manner as you would on a physical computer.

Click **Install Ubuntu** to begin; follow the on-screen instructions (Figure 11).

When you are prompted to do so, create a user with the username **ngds**; specify a password of your choice.

When the installation is complete, restart your virtual machine and log in using the username **ngds** and the password you specified during the Ubuntu installation process (Figure 11).

In addition to the above, it is recommended that you install the **Guest Additions** module.

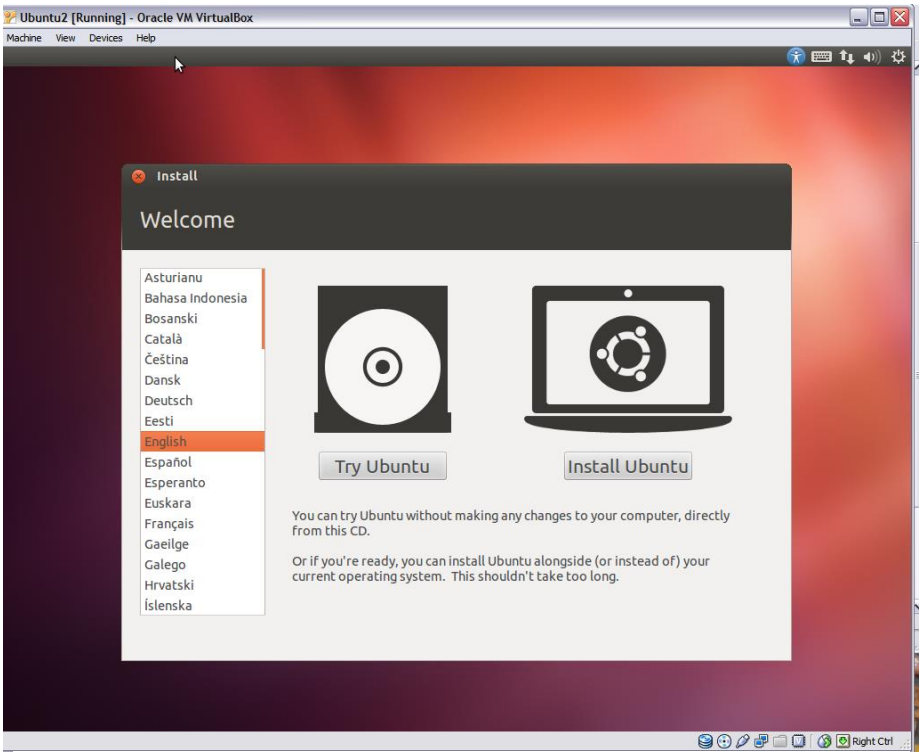


Figure 11: The Ubuntu Linux installation screen

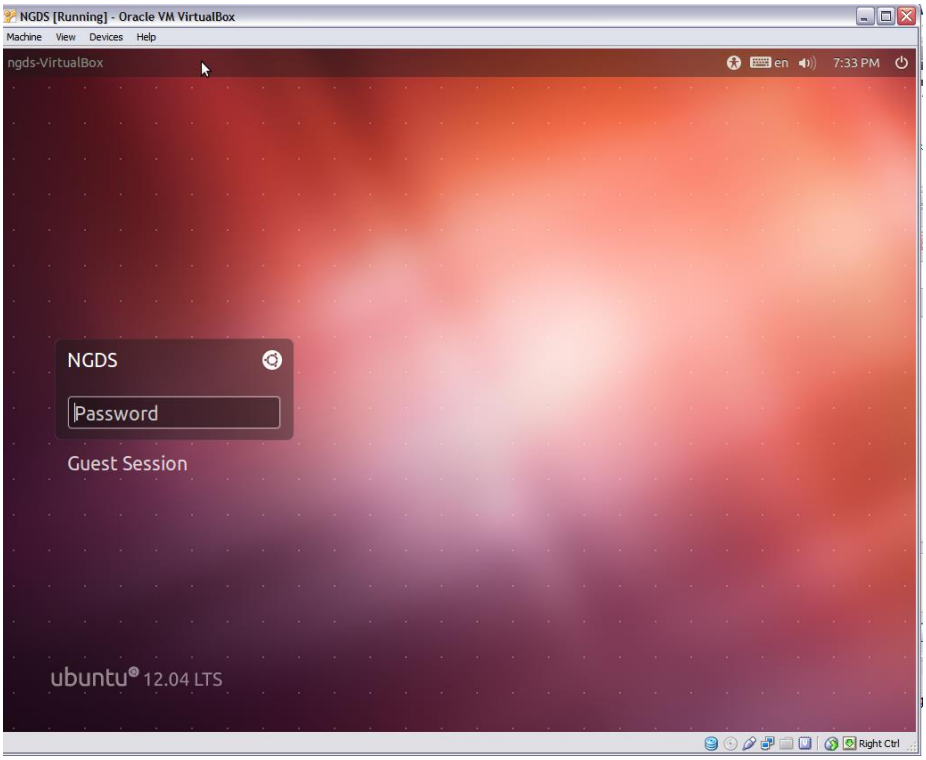


Figure 12: Logging on to Ubuntu Linux

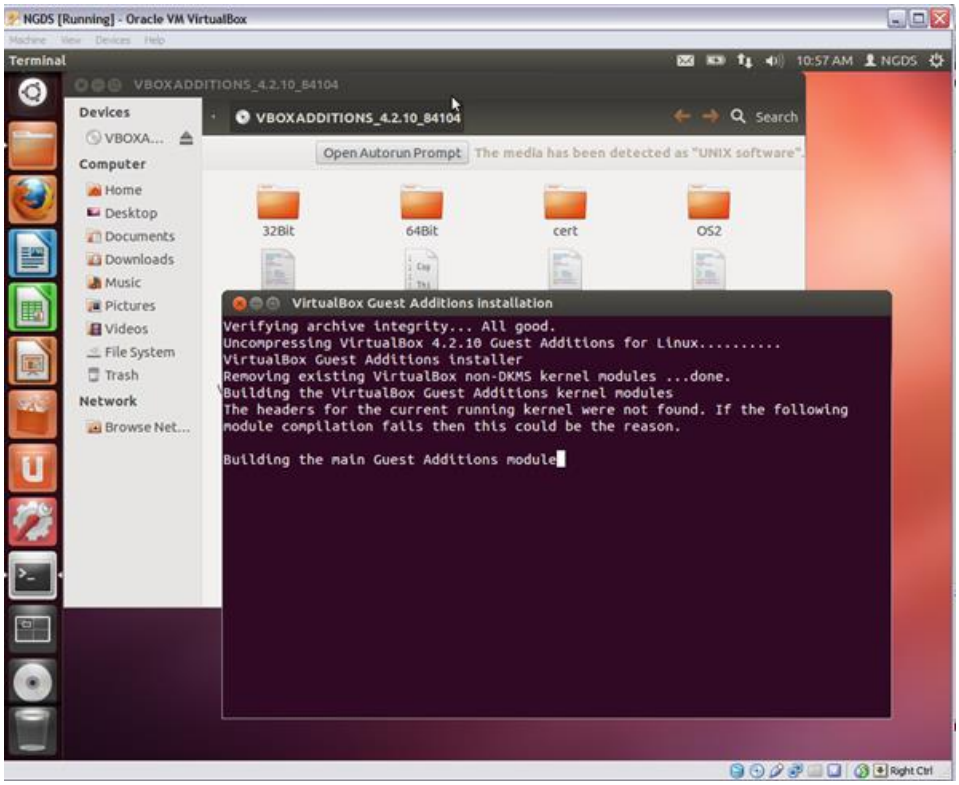


Figure 13: Installing the Guest Additions module for Ubuntu Linux

5.1.7. Take a Snapshot

Take a Snapshot of your virtual machine before continuing. A Snapshot is a record of the virtual machine that can be used to restore it to its condition at the time the Snapshot was taken. Snapshots are typically used as precautions against failure at a later date.

A Snapshot can be taken via the **VirtualBox Manager**.

2. Accommodating a corporate firewall (OPTIONAL)

If the computer you are using to host your virtual machine is behind a corporate firewall, your virtual machine may not have immediate Internet access. Internet connectivity is required in order to install NGDS Software Stackcomponents on your virtual machine (as will be discussed in Section 3).

5.2.1. Install and Configure CNTLM

CNTLM is a *proxy* that authenticates the user with a log-in and password, a typical requirement for corporate firewalls. If you are not behind a firewall that requires authentication, you can skip this step.

CNTLM is available at: http://cntlm.sourceforge.net/

After installing CNTLM on your host machine, use a text editor to modify the **cntlm.ini** file; here, specify the credentials your host machine uses to bypass your corporate firewall. An example appears in Table 1 below:

```
Username    yourcorporateproxyusernamehere
Domain      us008
Password    yourpasswordhere
# List of corporate proxies
Proxy       proxyfarm-us.3dns.netz.sbs.de:84
Proxy       129.73.8.72:8080
Proxy       129.73.11.208:3128
NoProxy     localhost, 127.0.0.*, 10.*, 192.168.*
# local port used by CMTLM
Listen      3128
```

In the example above, text strings preceded by a pound sign or hash symbol (#) are *comments* for the benefit of human operators; comments are not interpreted by any program reading the **cntlm.ini** file.

When configuring CNTLM, be sure to specify a localhost (**NoProxy**) entry with appropriate IP addresses and an appropriate port. The default CNTLM port is 3128. Asterisks (*) are *wildcard characters* which indicate the range of available possibilities for a given character – so 10.* can be 10.0, 10.1, or 10.2, all the way up to 10.9.

To use CNTLM, make sure CNTLM is running on your host machine whenever you run the virtual machine you created in Section 3.2. If CNTLM is not running on the host machine, your virtual machine will be unable to establish an Internet connection.

CNTLM can be executed by command prompt or set to run as a Windows service. Starting CNTLM from a command prompt is useful within a development environment because doing so allows you to manually restart CNTLM in response to freezes or crashes.

5.2.2. Configure your virtual machine environment to use CNTLM as its proxy

Log in to the virtual machine you created in Section 3.2; navigate to the **etc** directory and use a text editor to manually edit the **environment** file. Add the proxies specified in Section 3.3.1 to the **environment** file; an example appears below:

```
http_proxy=http://10.0.2.2:3128/
https_proxy=http://10.0.2.2:3128/
ftp_proxy=http://10.0.2.2:3128/
no_proxy="localhost,127.0.0.1,192.168.50.1,192.168.50.2"
HTTP_PROXY=http://10.0.2.2:3128/
HTTPS_PROXY=http://10.0.2.2:3128/
FTP_PROXY=http://10.0.2.2:3128/
NO_PROXY="localhost,127.0.0.1,192.168.50.1,192.168.50.2"
```

Alternatively, you can use the Ubuntu Network Configuration application to manually specify the desired proxies (Figure 13).

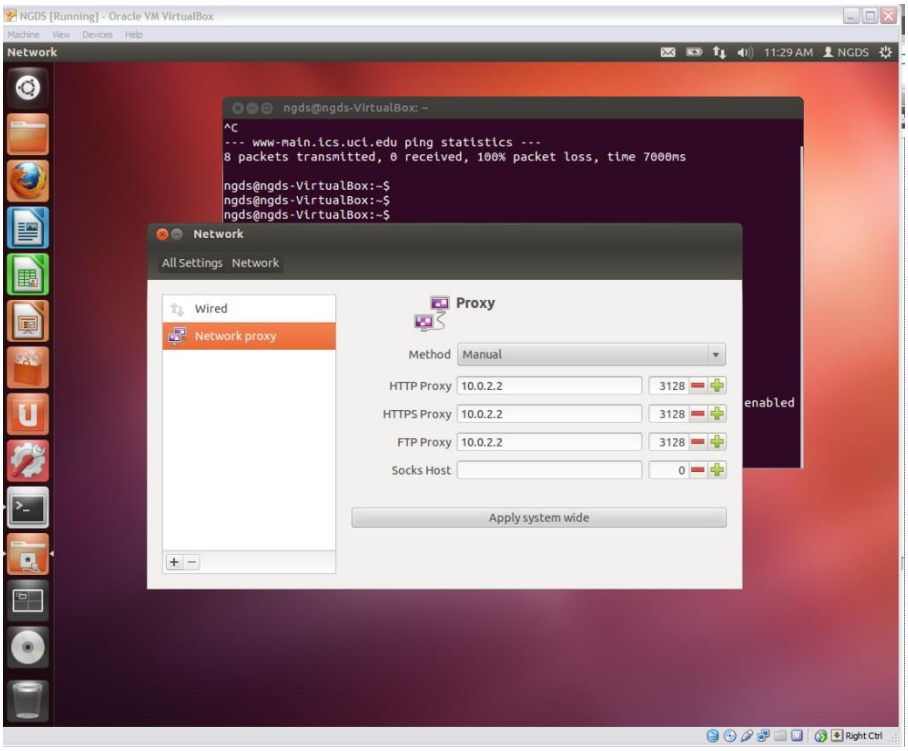


Figure 14: Configuring a proxy in Ubuntu Linux

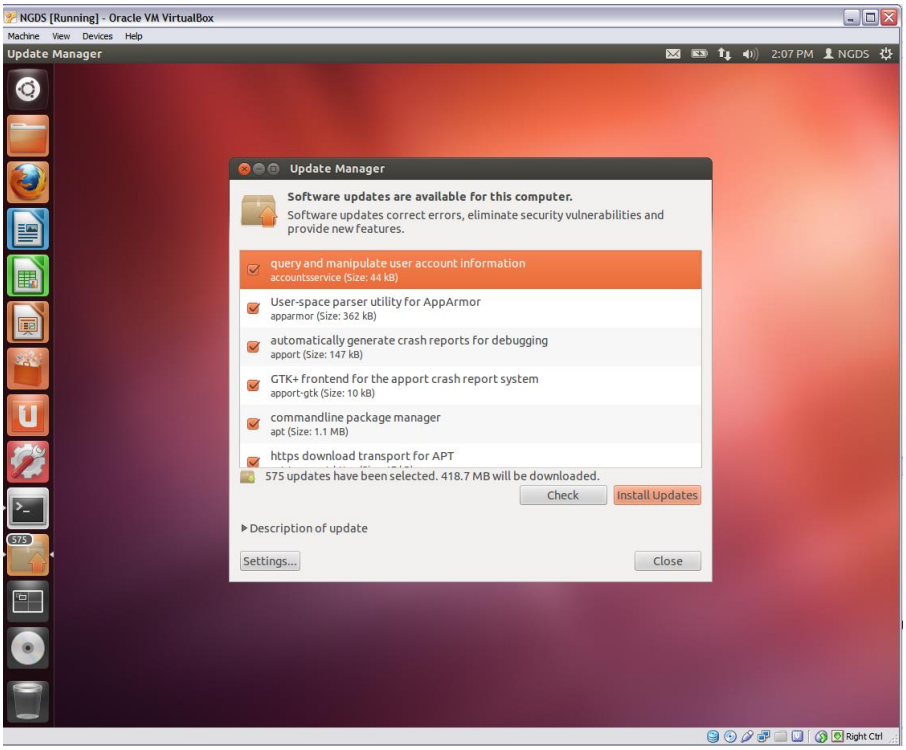


Figure 15: Installing updates in Ubuntu Linux

5.2.3. What to do if cntlm and proxy continue to cause issues

1) If possible, finish the install on a virtual machine connected to the Internet instead of a local intranet. If this is not possible, you will need to configure your virtual machine's settings in such a way that you are able to use the **apt get** command; negotiating an intranet may require installation of CNTLM within your virtual machine, as well.
2) If CNTLM causes issues after you have successfully installed the software, but then when you try to open the web sites locally hosted and CNTLM then causes issues, establish port forwarding within your virtual machine to forward the ports of interest (e.g. 5000, etc) to your physical machine, and browse the web sites on your physical machine. At least at CT RTC this solves the issues with the proxy.

Appendix B The development.ini file

The initial **development.ini** file is automatically generated by the installation script. The NGDS plug-ins list that is in this file needs to be modified. This list needs to be manually updated. The installation script will have prompted you to update this list of plugins and other values in the development.ini file.

Important note: **development.ini** is divided into various sections; section headers are demarcated by [brackets]. All CKAN parameters described in this document must be entered into the [app:main] section!

The code block below contains a summary of all changes that need to be made to the **development.ini** file during the NGDS CKAN Extension installation process.

```
sqlalchemy.url = postgresql://testuser:pass@localhost/testdb

ckan.datastore.write_url = postgresql://testuser:pass@localhost/test_datastore

ckan.datastore.read_url = postgresql://readonlyuser:pass@localhost/test_datastore


ngds.resources_dir=/home/yourname/pyenv/src/ckanext-ngds/ckanext/ngds/base/resources/


ckan.geoserver.url=http://localhost:8080/geoserver/rest

ckan.geoserver.workspace_name=NGDS

ckan.geoserver.workspace_URL=http://localhost:5000/ngds


ckan.site_id = ckan_instance

solr_url = http://localhost:8983/solr


ckan.storage.bucket=subdir

ofs.impl=pairtree

ofs.storage_dir=/home/yourname/storage/

ngds.deployment=node

ngds.default_group_name=public
```

```
  ngds.logo_text=CONTRIBUTING GEOTHERMAL DATA

  ngds.bulk_upload_dir=/home/yourname/upload

  ngds.facets_config=/home/yourname/pyenv/src/ckanext-ngds/facet-config.json



  ckan.plugins = stats json_preview recline_preview datastore datastorer ngdsui
ckan_harvester metadata geoserver csw
```

Development.ini Considerations

- Plugins must be set *incrementally* during installation; they cannot be added all at once. Instead, plugins must be added *after* their dependencies have been added
- You need to choose between the **datastorer** and **harvest** plugins depending on your CKAN installation:
    - When running NGDS in **node** mode (**ngds.deployment = node)**, you need the **datastorer** plugin
    - When running NGDS in **central** mode (**ngds.deployment = central)**, you need the **harvest** plugin
    - These two plugins *cannot* be installed in the same instance of the NGDS CKAN Extension!

Trailing slashes on paths must be identical to those specified in the code block above

Appendix C Architectural and Deployment Diagrams

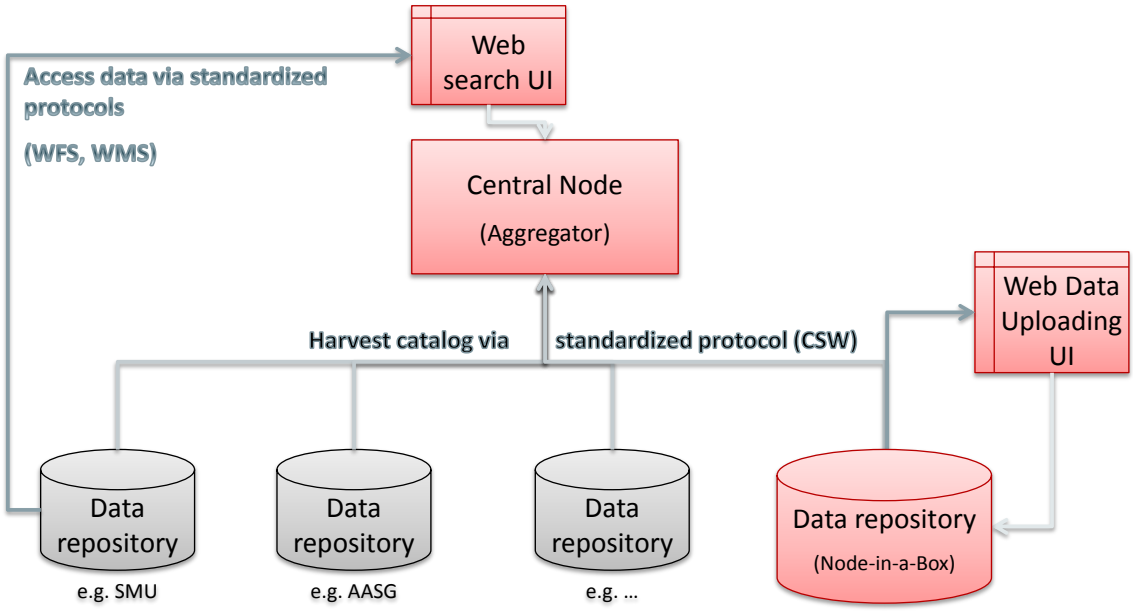


Figure 16: A diagram of NGDS

What is CKAN?

CKAN stands for **Comprehensive Knowledge Archive Network**.

CKAN is modular free-and-open-source data portal software. When properly installed on a server, CKAN provides a web-accessible interface by which users can submit and manage metadata records. The CKAN user interface also allows users to configure automated metadata harvesting from registered CKAN instances (an *instance* is a specific installation of the CKAN software); metadata harvested in this way is used to generate a web-accessible catalog. These traits are well-suited to the requirements of NGDS.

A CKAN *extension* is a user-generated modification of the CKAN software. The NGDS CKAN Extension is a CKAN extension designed to interact with NGDS data, metadata, and interchange formats.

Domain Model

The Domain Model of NGDS can be represented as a class diagram (Figure 18)

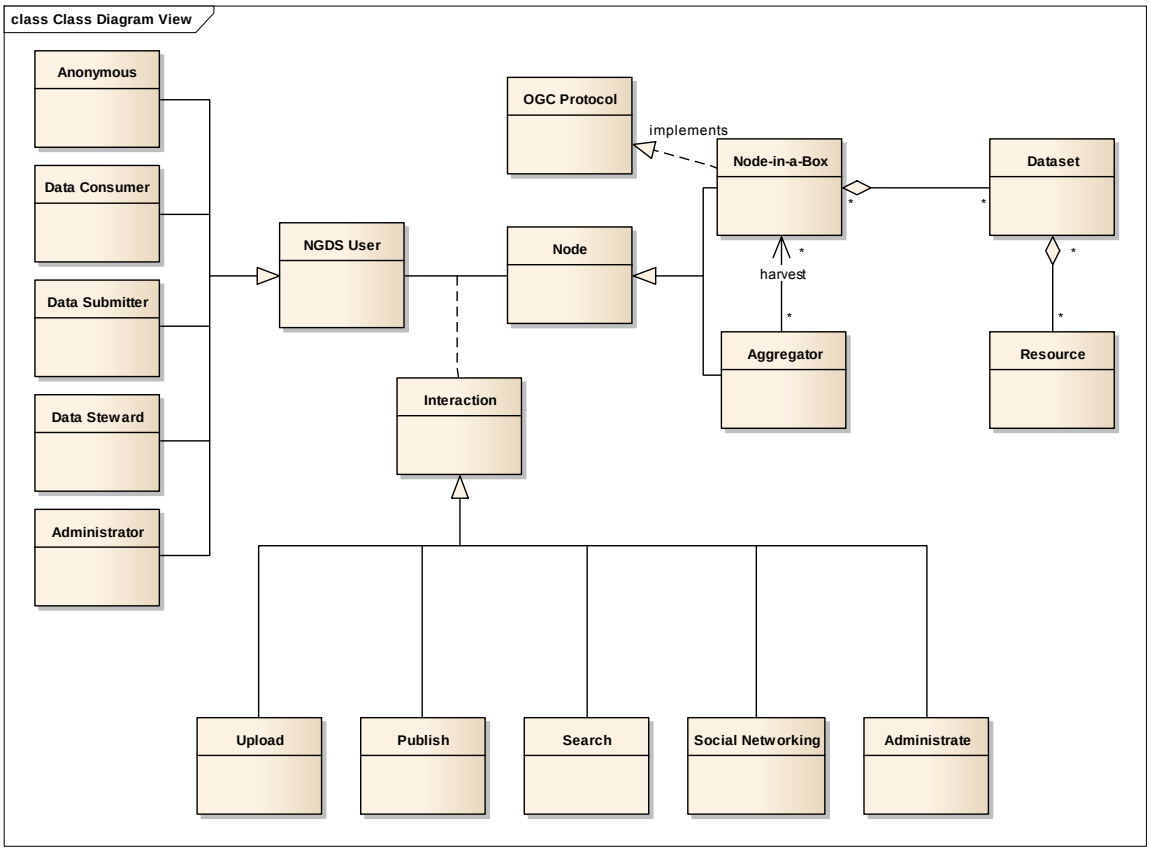


Figure 17: NGDS Domain Model as a Class Diagram

Notes on CKAN in Production Mode

When running CKAN in **production** mode, consider the following:

- CKAN runs as wsgi job behind apache2; you can control it with the following command:

  sudo service apache2 start|stop|restart|status

- The celeryd runs as a service; you can control it with the following command:

  sudo service ngds-celeryd start|stop|restart|status

- Tomcat currently has to be started manually but I will try to place it either behind apache2 or start it as a service with the following command:

  cd /opt/ngds/tomcat/bin; ./catalina.sh run

- If the system is configured to act as a central node a paster command that triggers harvesting has to be scheduled via cron. Adrian can provide the details as he knows more about harvesting than I do.
- Both SOLR and Geoserver are hosted by the same tomcat in order to reduce the amount of resources needed to run the system.

Comments on CKAN in Development Mode

CKAN should not be run behind apache2. Instead, we recommend the following commands:

```
% sudo service apache2 stop
% cd /opt/ngds/bin/default
% . ./bin/activate
% cd ckan
% ln -s /opt/ngds/etc/ckan/default/production.ini ./development.ini
% paster serve development.ini
```

Known Issues

- Configure the CATALINA_OPTS variable to provide more stackspace for tomcat (the default values are too low for production mode)
- Configure tomcat to either run behind apache2 or as a service (I will try my best to do that before end of next week)

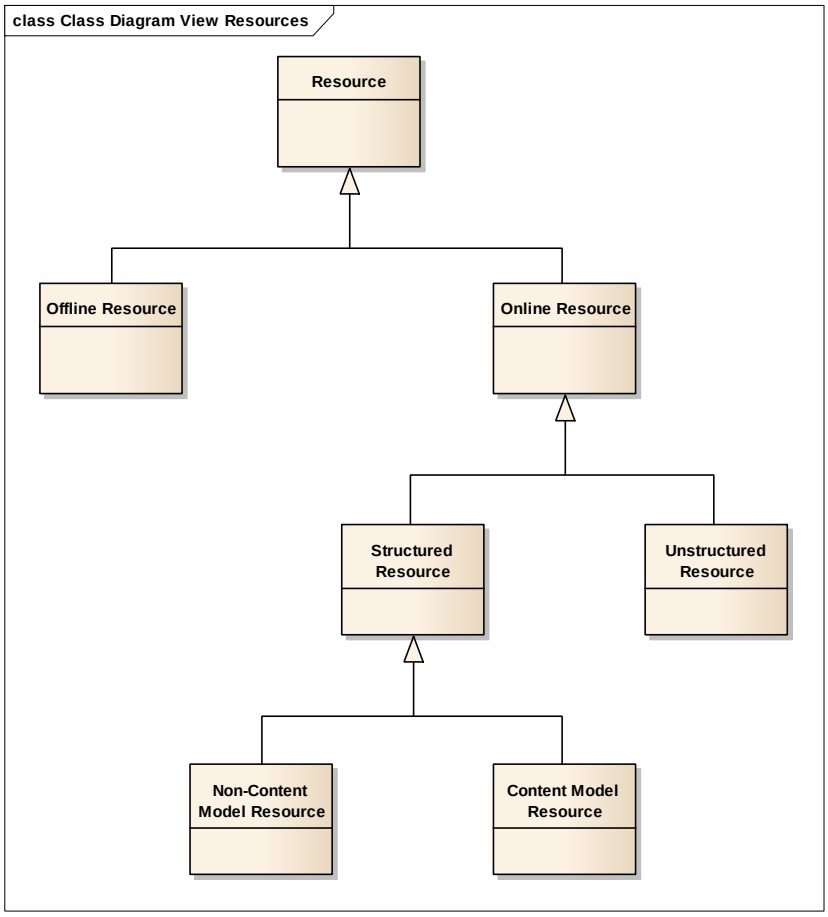


Figure 18: NGDS Domain Model of Resources as a Class Diagram

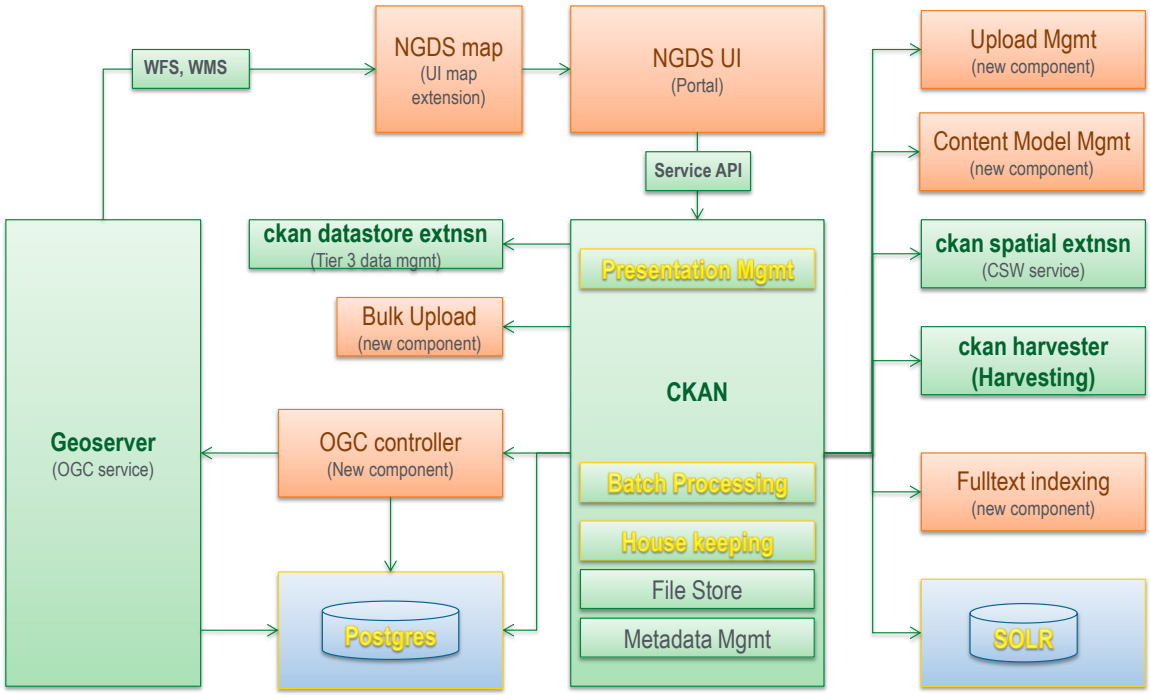


Figure 19: NGDS High-level Components

Reviewing, Deploying & Validating Services Submitted for the NGDS

Also accessible at https://docs.google.com/document/d/1H08dObu5pWze7g3sSap3uDcI4DwYzuIEMfcP0V7Xj9s/edit?pli=1#

Introduction

The National Geothermal Data System (NGDS) is an interoperable, distributed data system providing access to information resources related to geothermal energy. Data is added to the system from a network of data providers.

In order to keep the NGDS running smoothly, contributed data should meet the information exchange specifications for the system. The purpose of this document is to step through a suggested workflow for publishing NGDS data services that meet the system specifications.

Ingredients

Creating, editing and publishing data to NGDS requires three key ingredients: data, hardware and software.

Data: Input geothermal data into one of the United State Geoscience Information Network (USGIN) content models. These models are available for download from the USGIN models webpage: http://schemas.usgin.org/models/.

Hardware: Hardware requirements will be determined by software requirements and amount of traffic hitting the server.

Software: Software requirements can vary. Minimally, a data management system and software that can create web services are needed. For the purpose of this document, we will focus on the software needed to deploy web services using ArcGIS and OpenGeoSuite.

Generally, you will need:

- Data processing, editing and management software, for example: Excel, Access, ArcGIS Desktop and the Excel To NGDS Service Tool
    - This document will consider use of ArcMap for the workflow in using the 'Excel To NGDS Service Tool' for data processing. If not using ArcMap, Appendix G of this document describes manual review of a dataset. Use of OpenGeo Suite to create Web Services is described in Section 7.2.
- Enterprise database management system (pick one): Microsoft SQL Server, Oracle, PostgreSQL
- Geospatial data server (pick one): ArcGIS Server or Geoserver/OpenGeoSuite

Table of Contents

Section 1: Data QA/QC

Section 2: Create URIs

Section 3: Run the Excel to NGDS Service Tool

Section 4: Fix Errors Indicated by the Excel to NGDS Service Tool

Section 5: Finish Creating the Feature Class

- 5A. Finish Creating the Feature Class using Microsoft Access
- 5B. Finish Creating the Feature Class using ArcMap

Section 6: Load Data Into Production Database and Publish Services

- Section 6.1: Load Data and Publish Service Using ArcGIS
  - A. Load Data
  - B. Create an ArcMap Project
  - C. Publish services using ArcGIS Server
- Section 6.2: Load Data and Publish Service Using GeoServer
  - A. Load Data
  - B. Publish services using GeoServer

Section 7: Validating Services

- A. WMS Validation
- B. WFS Validation

Section 8: Create A New Resource in the Repository

Appendix A – Rewrite Rule Examples

Appendix B – Naming Conventions

Appendix C - Service Description Example

Appendix D – Online Resource Examples

Appendix E - Performing a Reprojection and Creating the Feature Class

Appendix F: Creating an ArcGIS (v.10) Service

Appendix G: Manually Reviewing the Data and Creating the NGDS Service

Section G1: Import Excel Data Into Access and Review Data Types

Section G1a: Excel to Geodatabase without Access

Section G2: Register Access Table with ArcGIS Geodatabase

Section G3: Create a Feature Class

Section 1: Data QA/QC

It is assumed for the workflow of this document that your data is already in one of the Excel NGDS content models. If not, download the appropriate content model from the USGIN model schemas site: http://schemas.usgin.org/models/, input your data and proceed. For more information on entering your data into the content models, visit the USGIN content models tutorial http://usgin.org/content/working-your-content-model.

1. **Field Names in First Row**. For use of the 'Excel to NGDS Service" tool to process your data (Section 3), ensure that the first row of your datasheet contains the field names.
2. **Check for Reasonable Data**. Ensure that data is of good quality and is as described in the 'Description' field of the FieldList tab in the Excel content model workbook.
3. **Create Meaningful Labels**. Label field should contain a description that is meaningful because the text in this field will be the label for the records when viewed in a map document. For example, a record in the Label field in the AZBoreholeTemperatures service might read "Temperature Observation for well api:412548976".

Section 2: Create URIs

Understanding URIs

URIs, or Uniform Resource Identifiers, dereference to an online resource. These online resources are representations of the service data, like an online file or XML representation of the data service. An implemented rewrite rule enables the return of these resources by inputting the URI into a web browser.
Project developers with NGDS use a rule based approach to rewrite http URIs to URLs, retrieving those useful representations of information resources. For NGDS data services, all URIs should follow a pattern:

http:/resources.usgin.org/uri-gin/[name authority]/[token]/[unique identifier]/

In the Borehole Temperature Observation content model, the ObservationURI for a single record might be:

http:/resources.usgin.org/uri-gin/azgs/bhtemp/0003/

where "azgs" is the **name authority**, "bhtemp" is the **token** and "0003" is the unique identifier. Find a list of assigned name authorities and tokens for NGDS project participants https://github.com/usgin/uriredirect/wiki/NGDS-Name-Authorities-and-Tokens. For all records in the Borehole Temperature dataset, the unique identifier should be the only part of the ObservationURI that changes.

NOTE: URIs must end with a '/' to work.

For a more in-depth description of URIs, with redirection diagrams, visit http://usgin.org/content/usgin-uris.

Creating URIs

1. In Excel, fill out URI columns/fields, as required. The URI will be unique for each record in the dataset, as well as for each unique resource type in the dataset. Resource types might include well headers, geologic unit descriptions, sources.
2. Insert temporary 'working' columns in worksheet, as needed. Remember to delete them later.
3. Create an Excel formula that combines the URI 'root' (http://resources.usgin.org/uri-gin/[name authority]/[token]/) and the unique identifier, per record. This process is the same for return of online files or XML-representations of the data service.

B2 f_x =CONCATENATE("http://resources.usgin.org/uri-gin/azgs/bhtemp/",A2,"/")

	A	B	C
1	TEMPID	TEMP_URI	ObservationURI
2	0001	http://resources.usgin.org/uri-gin/azgs/bhtemp/0001/	http://resources.usgin.org/uri-gin/azgs/bhtemp/0001/
3	0002	http://resources.usgin.org/uri-gin/azgs/bhtemp/0002/	http://resources.usgin.org/uri-gin/azgs/bhtemp/0002/
4	0003	http://resources.usgin.org/uri-gin/azgs/bhtemp/0003/	http://resources.usgin.org/uri-gin/azgs/bhtemp/0003/
5	0004	http://resources.usgin.org/uri-gin/azgs/bhtemp/0004/	http://resources.usgin.org/uri-gin/azgs/bhtemp/0004/
6	0005	http://resources.usgin.org/uri-gin/azgs/bhtemp/0005/	http://resources.usgin.org/uri-gin/azgs/bhtemp/0005/

4. For implementation of a URI redirection (making your URI function as a URL in a browser), email URIredirection@usgin.org.

Section 3: Run the Excel to NGDS Service Tool

In order to facilitate the process of reviewing the data to confirm that it conforms to the information exchange schemas, we have created an ArcGIS data processing tool. This tool can be downloaded from the USGIN GitHub Repository at https://github.com/usgin/ExcelToNGDSServiceTool. The inputted Excel file is validated against the appropriate schema (.xsd), then a Geodatabase (.mdb) is created that houses a Feature Class which ready for service deployment. If this tool will not be used and a manual review of the data is required, see Appendix G.

1. Go to https://github.com/usgin/ExcelToNGDSServiceTool
2. Click the **Zip button** to download the repository as a **Zip file**.
3. Save the **entire** unzipped folder 'ExcelToNGDSServiceTool-master' to C:\Users\[user]\Documents\ArcGIS\
4. Open ArcMap
5. Open the Catalog window (within ArcMap)
6. Navigate in Catalog to your Home directory, then choose the Documents\ArcGIS folder
7. Open the 'ExcelToNGDSServiceTool' folder saved in Step 3
8. Open the Toolbox 'ExcelToNGDSService'
9. Double-click the script 'Excel to Service'

a. Drag the Excel file reviewed in Sections 1 and 2 into the Excel file box, or browse to select it
b. Select the sheet containing the data (if multiple tabs exist in your Excel workbook)
c. Select the appropriate schema
d. Enter the Service Name, ideally in the format: State Prefix + Service (see Appendix B)
e. Select the appropriate layer name
f. The Validation Only button validates the Excel spreadsheet without creating the Feature Class; leave this button unchecked to create the Feature Class

As the Tool runs the Excel file is **Not Modified**. Data changes will only appear in the ArcGIS Feature Class that is created. If the Tool output states that the conversion was successful a geodatabase will have been created in the same folder as the original Excel file. Inside the geodatabase there will be a Feature Class for each of the layers of the service.

The following text outlines the checks completed by the tool. If you receive validation failures, proceed to Section 4. If the tool completes and validation is successful, proceed to Section 5. The Tool will compare the Excel file with the schema to check the following conditions:

- Data is checked that it conforms to UTF-8 and Windows 1252 (encoding used by server) - **Validation Fail** otherwise
    - *See https://en.wikipedia.org/wiki/Windows-1252 for a Win 1252 character table*
- Number of fields must be the same in the Excel file as in schema - **Validation Fail** otherwise
- Order of fields must be the same in the Excel file as in schema - **Validation Fail** otherwise
- Field names must be the same in the Excel file as in schema - **Validation Fail** otherwise

- If a field name contains the word “URI” (and the field is *not* MetadataURI or SourceURI or SourceCitationURI) data in that column is checked and modified to:
    - Remove whitespace and carriage returns from the URI
    - URI is checked that it contains “http://resources.usgin.org/uri-gin/” - **Validation Fail** otherwise
    - Checked that they contain at least 7 backslashes - **Validation Fail** otherwise
    - Checked that they end with a backslash and if not a backslash is added
    - If the field is the first URI field then this is the primary URI field and the URIs in this field are checked to make sure they are each unique - **Validation Fail** otherwise
    - Appropriate use of the name authority and token are **NOT** checked
- Cells containing more than 255 characters are given an appropriately long field length in the ArcGIS Feature Class created by the tool so all data is preserved
- Latitude and Longitude Coordinate fields (LatDegree, LatDegreeWGS84, LongDegree, and LongDegreeWGS84) are checked to assure that their values fall within -90 to 90 and -180 to 180 respectively - **Validation Fail** otherwise.
- Fields “MaximumRecordedTemperature," "MeasuredTemperature," "CorrectedTemperature" and "Temperature" must have values between 0 and 999 (if the placeholder -999 or -9999 is not being used) - **Validation Fail** otherwise
- The TemperatureUnits field is checked to make sure that the same units are used for every record, either F or C - **Validation Fail** otherwise
- The SRS field, if found, is checked that the coordinate system is either:
    - WGS 84 (if the cell contains the numbers 4326 for EPSG:4326 or 84 for WGS84)
    - NAD 27 (if the cell contains the numbers 4267 for EPSG:4267 or 27 for NAD27)
    - NAD 83 (if the cell contains the numbers 4269 for EPSG:4269 or 83 for NAD83)
    - Anything else in this column will be considered “Unknown” and WGS 84 will be used for the coordinate system, thus possibly requiring a manual reprojection at the end (see Appendix E if manual reprojection is necessary)
    - When creating the Feature Class, the Tool will apply the appropriate projection and reproject to WGS 84 as necessary, writing EPSG:4326 in the ‘SRS’ field.
- Fields indicated by the schema as required are given the following values if empty:
    - Text fields given the value “Missing”
    - Double fields given the value “-9999”
    - Date fields give the value "1/1/1900T00:00"
- Fields that are not empty are checked that their data is of the correct type. If the data is not of the correct type for the field then:
    - Text fields given the value “Missing” (if a required field) or left empty (if not a required field)
    - Double fields given the value “-9999” (if a required field) or left empty (if not a required field)
    - Date fields, if not recognized as a date field in Excel, are converted from the following formats to the type date: “1/28/1975T00:00” or “1/28/1975T00:00:00” or “1975-1-28T00:00:00” or “1975-1-28T00:00” or “1/28/1975” or “1975-1-28.” If conversion is not possible and the field is required there is a **Validation Fail** otherwise the field is left empty. Date fields are explicitly specified in newer schemas as “xs:date.” In older schemas they are simply specified as “xs:string.” In

this case the Excel to NGDS Service Tool assigns the date type to any fields specified as strings containing the word Date, except for the field "WellDrillDateType."

Section 4: Fix Errors Indicated by the Excel to NGDS Service Tool

1. **Fix any errors in the Excel file that caused a Validation Fail** - Notices given that the tool is changing a cell value do not need to be fixed manually. The change will appear in the final Feature Class attribute table.
2. **Schema Field Mismatch Errors** - Any errors with the fields of the Excel file not matching will likely necessitate downloading the schema (XSD) from http://schemas.usgin.org/models/ and modifying the Excel file in review to match the fields of the schema. Field names are found after '<xs:element name=' with the following exceptions, which must not be fields in the Excel file.

a. <xs:element name=[Layer Name] …
b. <xs:element name="OBJECTID" type="xs:int" minOccurs="0">
c. <xs:element name="Shape" type="gml:MultiCurvePropertyType">

3. **Date Issues** - Dates in Excel need to be either recognized by Excel as a date or as text in the format "1/28/1975T00:00" or "1/28/1975T00:00:00" or "1975-1-28T00:00:00" or "1975-1-28T00:00" or "1/28/1975" or "1975-1-28" (without the quotes).
4. **Spatial Reference System Warning** - If the tool is unable to determine which coordinate system the data is in, or reproject as necessary, a coordinate system of WGS 84 will be assumed and "EPSG:4326" will be written in the SRS column. A warning indicating that the spatial reference system will need to be **defined** and *then* a **reprojection** to WGS 84 will be required will be issued. To first define the projection follow the steps below and then see Appendix E for instructions on performing a reprojection.

. In ArcCatalog, open ArcToolbox window
a. Expand Data Management Tools, Projections and Transformations, Feature
b. Double-click on 'Define Projection'
c. Double-click on 'Project'
i. Add Feature Class that needs to have its projection defined as the 'Input Dataset'
ii. Select the 'Input Coordinate System' indicated by the **Original** Excel file (the SRS field in the new Feature Class has already been updated to say "EPSG:4326" but the coordinates were not changed).
iii. Follow the steps in Appendix E to reproject to WGS 84 (EPSG:4326)

5. Run the tool again until the Validation is successful.

Section 5: Finish Creating the Feature Class

5A. Finish Creating the Feature Class using Microsoft Access

If the Excel file validates, the ArcMap tool creates a Feature Class within a Personal geodatabase (.mdb). By opening the .mdb directly using Microsoft Access, data and fields can be easily manipulated. Microsoft provides a free runtime version of Access which can be installed on any Windows desktop, and run any Access database application (http://www.microsoft.com/en-us/download/details.aspx?id=4438).

The following steps need to be performed to finish creating the NDGS Feature Class:

1. Open the Access database (.mdb)
2. In Design View, **move the Shape** field to the end, so it is the last field. This is important for the eventual service to validate against the schema.
3. **RelatedResources** field: This field is meant to correlate other relevant services or online documents. If other services from the same state are running, find a column of matching data (ID or APINo, for example) to correlate with dataset. Joining by this field, populate the RelatedResources field with the main URI from the other service data.
4. If the service contains multiple layers then the following additional steps will need to be performed:

a. Open ArcCatalog and navigate to the geodatabase just created.

b. Feature classes will be made by the tool for every layer name. In ArcCatalog delete the Feature Classes not being used.

c. Close ArcCatalog and open the Access database (.mdb)

d. The fields in each Feature Class are created by the tool to be identical. In Design View use the schema to delete the fields that do not belong from each Feature Class.

e. Edit the URIs in each layer to be unique, if necessary. For example, AnalysisURI field in Aqueous Chemistry and Rock Chemistry layers will need to be unique.

5B. Finish Creating the Feature Class using ArcMap

If the Excel file validates, the ArcMap tool creates a Feature Class within a Personal geodatabase (.mdb). NGDS schemas require that the 'Shape' field be the last field in the finished web feature service, and hence the last field in the Feature Class. The ExceltoNGDSServices tool does not place the Shape field at the end, so this must be performed manually. As explained in Section 5A, this can be done with use of Microsoft Access. The following outlines how this can be done using ArcMap alone, where an XML document of the Feature Class is exported from the Personal geodatabase (.mdb) and imported into a File geodatabase (.gdb).

1. Open ArcCatalog
2. Right-click the Feature Class created in Section 3, and choose 'Export' ➔ 'XML Workspace Document'

a. In the dialog box, choose 'Schema Only'

b. Specify an output location and a file name

c. Click 'Next' ➔ 'Finish'

3. Open XML Workspace document in an editor (Notepad++ or XMLSpy, for example) and manually move the 'Shape' field to the end. From here you can also edit the field names, field sizes, move other things around if necessary. The following is an excerpt from an XML Workspace Document; the entire text represents the 'Shape' field and will need to be moved to the last field:

```
<Field xsi:type="esri:Field">
  <Name>Shape</Name>
  <Type>esriFieldTypeGeometry</Type>
```

```
 <IsNullable>false</IsNullable>
 <Length>0</Length>
 <Precision>0</Precision>
 <Scale>0</Scale>
 <Required>true</Required>
 <GeometryDef xsi:type="esri:GeometryDef">
   <AvgNumPoints>0</AvgNumPoints>
   <GeometryType>esriGeometryPoint</GeometryType>
   <HasM>false</HasM>
   <HasZ>false</HasZ>
   <SpatialReference xsi:type="esri:GeographicCoordinateSystem">
                <WKT>GEOGCS[&quot;GCS_WGS_1984&quot;,DATUM[&quot;D_WGS_1984&quot;,
      SPHEROI
                D[&quot;WGS_1984&quot;,6378137.0,298.257223563]],PRIMEM[&quot;Greenwich
      &quot;,0.
                0],UNIT[&quot;Degree&quot;,0.0174532925199433]]</WKT>
     <XOrigin>-400</XOrigin>
     <YOrigin>-400</YOrigin>
     <XYScale>999999999.99999988</XYScale>
     <ZOrigin>-100000</ZOrigin>
     <ZScale>10000</ZScale>
     <MOrigin>-100000</MOrigin>
     <MScale>10000</MScale>
     <XYTolerance>8.983152841195215e-009</XYTolerance>

     <ZTolerance>0.001</ZTolerance>
     <MTolerance>0.001</MTolerance>
     <HighPrecision>true</HighPrecision>
     <LeftLongitude>-180</LeftLongitude>
     <WKID>4326</WKID>
   </SpatialReference>
   <GridSize0>3.7252903001966386e-007</GridSize0>
 </GeometryDef>
 <ModelName>Shape</ModelName>
</Field>
```

4. Create a new, empty File Geodatabase by right-clicking in desired file location in ArcMap and choose ‘New’ ➔ ‘Create New File Geodatabase’

5. Right-click the new File Geodatabase and choose ‘Import’ ➔ ‘XML Workspace Document...’

a. In the dialog box, choose ‘Schema Only’ and navigate to the .xml file created in Step 2. This will create a new Feature Class from the edited XML

6. Right-click the new Feature Class and choose ‘Load’ ➔ ‘Load Data...’

. Input the original Feature Class created in Section 3

a. Map the fields from the original to the new Feature Class created in Step 4

From here, you are ready to create an ArcMap project (.mxd, see Section 6.1B) and create a web service using ArcServer (see Section 6). Section 8 is dedicated to creating metadata and making your new service available within the NGDS.

Section 6: Load Data Into Production Database and Publish Services

The final step is to load data into the production database and deploy services. This can be done in an ArcGIS environment or using open source tools. To deploy services, the data first needs to be loaded into a database that is connected to the internet. The data is then published to the internet using ArcGIS Server or Geoserver.

There are two loading and publishing scenarios presented in this document:

1. Section 6.1: Load data and publish services using ArcGIS
2. Section 6.2: Load data and publish services using OpenGeo Suite Tools

A distributed enterprise database system is recommended for staging and publishing web services. Schema-valid data tables are edited in a 'staging database' and pushed to a 'production database' for service deployment.

Section 6.1: Load Data and Publish Service Using ArcGIS

ArcCatalog will be used to load data from the personal geodatabase into the staging and production databases.

The staging database at the Arizona Geological Survey is currently a Microsoft SQL Server database. The production database is a PostGIS-enabled PostgreSQL database.

In this section, an ArcSDE connection will first be made to the Microsoft Server SQL database. After the data from the personal database is loaded into the SQL database, a replica of that dataset will be created in the PostGIS-enabled PostgreSQL database.

A. Load Data

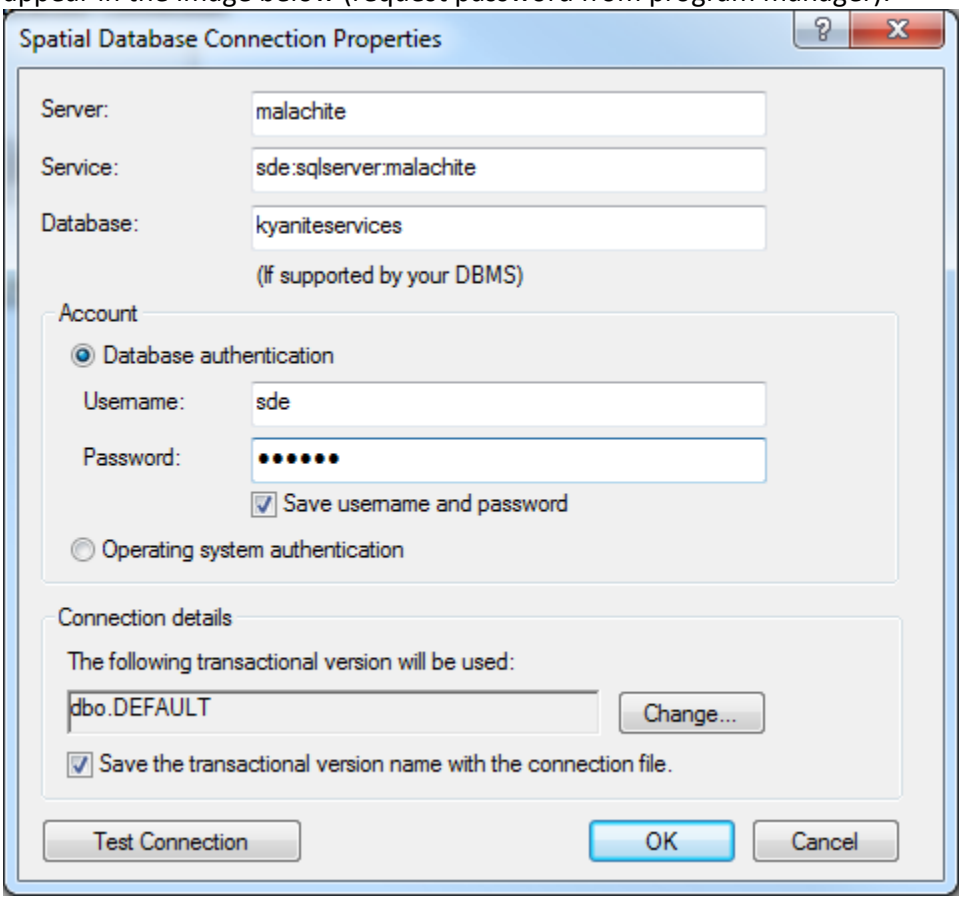
1. Open ArcCatalog.
2. Connect to personal geodatabase folder.

a. Click on Connect to Folder icon if you need to connect to the database folder.

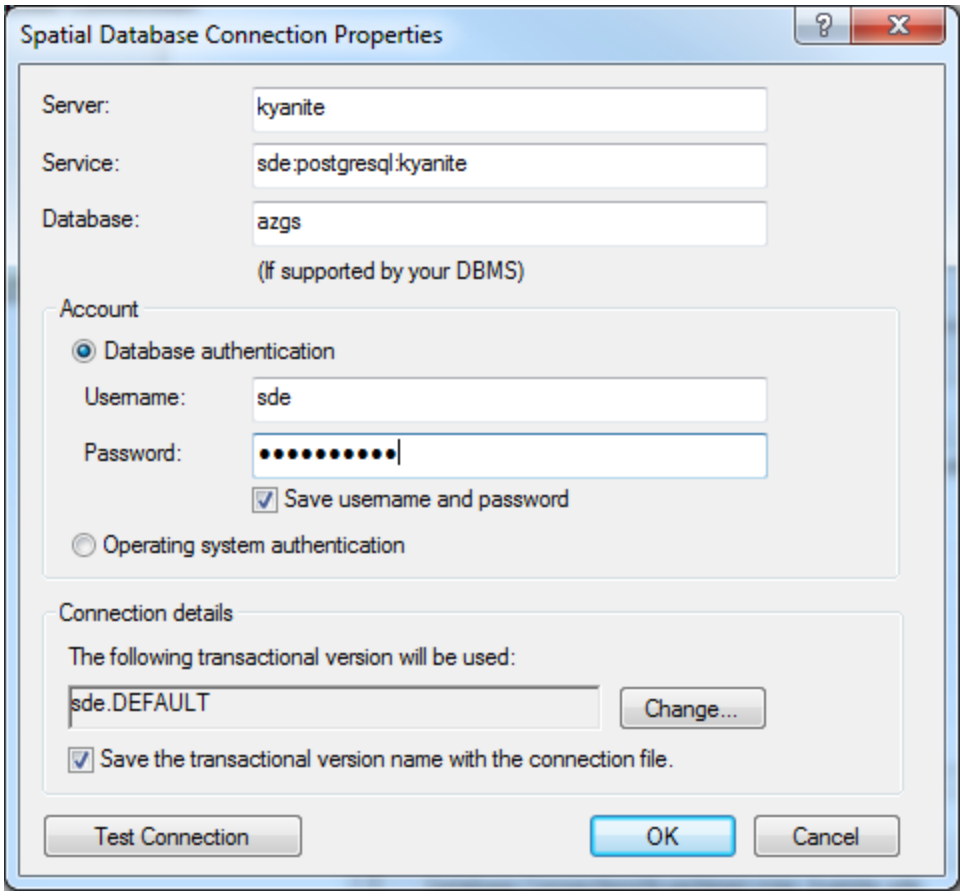
3. Connect to staging and production enterprise geodatabases (ArcSDE Connection):

. Double-click Add Spatial Database Connection

a. A new window opens, Spatial Database Connection Properties. Input the values as they appear in the image below (request password from program manager):



Staging geodatabase connection properties

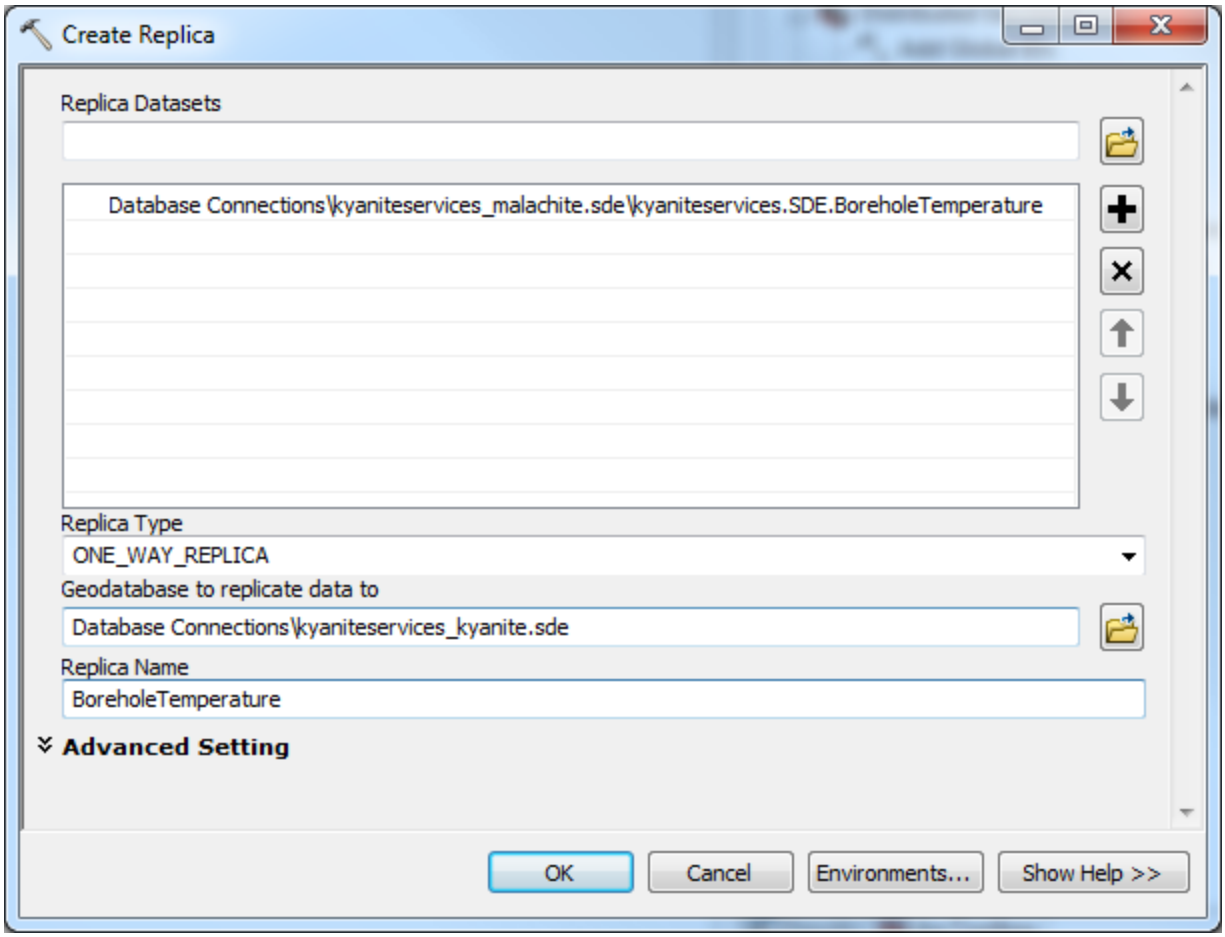


Production geodatabase connection properties

4. In ArcCatalog, copy Feature Class from geodatabase created in Section 5 and paste into staging database.
5. Right-click the Feature Class in staging database, select Add Global IDs.
6. Right-click the Feature Class in staging database, select Register as Versioned. **DO NOT** check the box next to Register the selected objects with the option to move edits to base.
7. Create Replica

a. Open ArcToolbox window
b. Expand Data Management Tools, Distributed Geodatabase
c. Double-click on Create Replica; parameter values should be

i. Replica Datasets: input Feature Class from staging database
ii. Replica Type: `ONE_WAY_REPLICA`
iii. Geodatabase to replicate data to: input production database
iv. Replica Name: should be the same name as the Feature Class name in the staging database
v. Click OK



Create Replica

The data is ready to add to ArcMap.

B. Create an ArcMap Project

1. Open a Blank ArcMap
2. Add Feature Class from production database
3. Click Add Data
4. Navigate to Feature Class
5. Symbolize layer

a. From the Feature Class, right click and select Properties. Under the Symbology tab, choose Import.

b. Choose a layer from the project schemas site http://schemas.usgin.org/models/ selecting the layer file that corresponds to the content model.

c. Select Value Field on which field in the dataset will be symbolized, if prompted. For example, BoreholeTemperatures services are symbolized by the "MeasuredTemperature" field; ActiveFaults services are symbolized by the "Symbol" field; SeismicHypocenters services are symbolized by the "Magnitude" field.

d. Click OK. Layer symbology has been applied.

e. Click OK again to close Layer Properties window.

6. Create metadata for the map document. The text will display on the arcgis services rest page, near the bottom under "Document Info".

. From the ArcMap File dropdown menu, select Map Document Properties

a. Title (example): "Arizona Borehole Temperature"

b. Summary (example): "Arizona Borehole Temperatures"

c. Author, just the organization (example): "Arizona Geological Survey"

d. Credits (example): "Data provided by the Arizona Geological Survey; service provided by the Arizona Geological Survey".

e. **NOTE:** If data is being served by another state, then you can specify that. For example, Illinois is serving Borehole Temperature Data for the State of Massachusetts. The credits in this case might read "Data provided by the Massachusetts Geological Survey; service provided by the Illinois State Geological Survey".

f. Tags (example): "geothermal, wells, geochemistry, water, temperatures, borehole, aasg, wms, wfs"

g. Check box next to "Store relative pathnames to data sources"

7. Create data frame metadata. Text will appear in the "Description" area on the arcgis services rest page.

. Right-click Data Frame, select Properties.

a. Click on General tab. Fill out the following information:

i. Name (example): "AZBoreholeTemperatures"

ii. Description (example): "This layer contains data describing borehole temperatures in the state of Arizona, obtained by the AASG for the Geothermal Data Project. The data table includes general information on the location of the borehole, temperature at some depth within the well, measurement procedure, location uncertainty and remarks. Citations are included in the dataset.

1. The HeaderURI for a particular borehole is the cross-referencing link (foreign key) used to associate the borehole with web based information on the well construction, pictures or other information, specific to one feature.
2. The data is displayed using a temperature gradient color scheme ramping from green (cool) to red (hot). Layer files containing the symbology for these data layers can be found in the USGIN repository at http://repository.usgin.org/uri_gin/usgin/dlio/206. The original data and metadata can be viewed or downloaded from the USGIN Repository at http://repository.usgin.org/uri_gin/usgin/dlio/318."

NOTE: At the end of the Description, include a link to the repository page (as shown above). For NGDS data, go to http://repository.stategeothermaldata.org/repository/browse/ and copy the link of the page for the appropriate dataset.

iii. Credits (example): "Data provided by the Arizona Geological Survey; service provided by the Arizona Geological Survey".

1. If data is being served by another state, then you can specify that. For example, Illinois is serving Borehole Temperature Data for the State of Massachusetts. The credits in this case might read "Data

provided by the Massachusetts Geological Survey; service provided by the Illinois State Geological Survey".

4. Create layer metadata. The text will appear in the Layer description on the arcgis services rest page)

iii. Right-click layer. Select Properties.

iv. Click on General Tab. Enter the following:

1. Layer Name (example): BoreholeTemperature
2. Description: Same as Data Frame
3. Credits: Same as Data Frame

1. Save the map document.

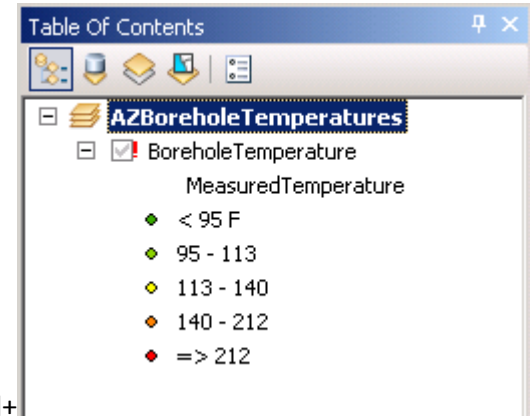
C. Publish services using ArcGIS Server

This document assumes that ArcGIS Server is installed and configured properly. Refer to ArcGIS Desktop 10.0 help documentation for more details about server configuration: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/What_is_ArcGIS_Server/00530000 0001000000/

1. Confirm that data from production database are connected to ArcMap layers

a. Open ArcMap.

b. If the links are broken (indicated by a little red exclamation point next to the layer), repair data sources. Refer to ArcGIS Desktop Help document on Repairing Broken Data Links.



P@ssword+

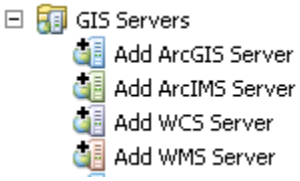
c. Save and Close ArcMap.

4. Publish Service

c. Open ArcCatalog.

d. Connect to GIS Server.

e. Expand GIS Servers



d. If needed, make the initial connection to ArcGIS Server.

i. Double-click Add ArcGIS Server

1. Select Manage GIS Services; click Next
2. Input Server URL and Host Name
3. Follow prompts to connect to server
4. Once connected, the connection should remain

e. To publish service (See Appendix F for screenshots)

i. Right-click on ArcGIS Server

ii. Select Add New Service …

1. Add GIS Service window pops up. Fill in fields:

a. Name: Map service name. See Appendix B for service names.

b. Type: Select Map Service

c. Description: Type description. See service example description in Appendix C.

2. Click Next. A new window pops up. Fill in parameter fields:

. Map Document: Navigate to and select ArcMap service document

a. Directory: Select output directory. This should most likely be the default c:\arcgisserver\arcgisoutput

b. Server Cache Directory: Select Cache directory. This should most likely be the default c:\arcgisserver\arcgiscache

c. Maximum Number of Records Returned by Server: default value (1000) okay

3. Click Next. A new window pops up. Select capabilities and operations.

. Check boxes next to:

a. Mapping

i. Map

ii. Query

iii. Data

b. WMS

. Fill out WMS service properties (metadata). This will be the contact information for the data provider.

i. Change the OnlineResource; see Appendix D for examples.

ii. Enable the “Use layer names from the map document”

c. WFS

. Fill out WFS service properties (metadata). This will be the contact information for the service provider.

i. Change the OnlineResource; see Appendix D for examples.

ii. Enter the correct namespace (found in first line of schema).

iii. Enter the prefix “aasg”.

4. Click Next. A new window pops up. Accept defaults for pooling.
5. Click Next. A new window pops up. Accept defaults for processes.
6. Click Next. A new window pops up. Select radio button to start service right now.
7. Click Finish

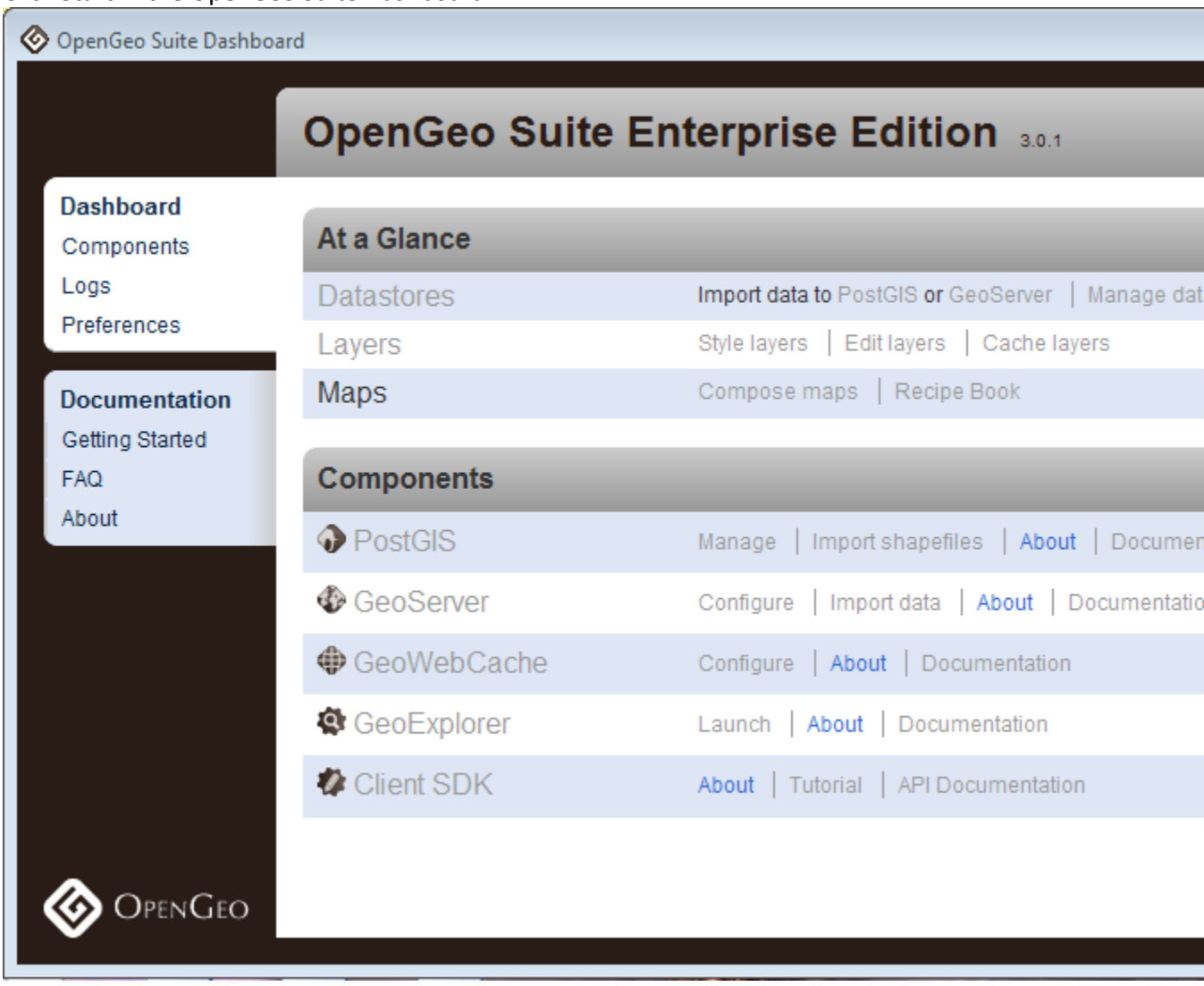
The service has been created and needs to be validated against the schema. Proceed to Section 8.

Section 6.2: Load Data and Publish Service Using GeoServer

GeoServer can be installed separately, or part of a software package known as OpenGeo Suite (http://opengeo.org/) which includes Postgres and PostGIS (database management systems) and OpenLayers. These are free and open-source software options. OpenGeo Suite is recommended and the following workflow describes loading a shapefile into PostGIS environment using this software. Using the PostGIS interface, a shapefile will be imported then converted to a table within the PostGIS database. This data table will be the source of data for the web service. For a more in-depth explanation of GeoServer set up and installation, see repository.stategeothermaldata.org/resources/documents/NGDS/GeoserverDocumentation.docx.

A. Load Data

1. Click Start in the OpenGeo Suite Dashboard



2. Under Components, click 'Import shapefiles' next the the PostGIS icon.
3. Click 'View connection details' to connect to database and folder for deployment.
4. Click Add File; navigate to schema-valid shapefile (shp) or database file (dbf).

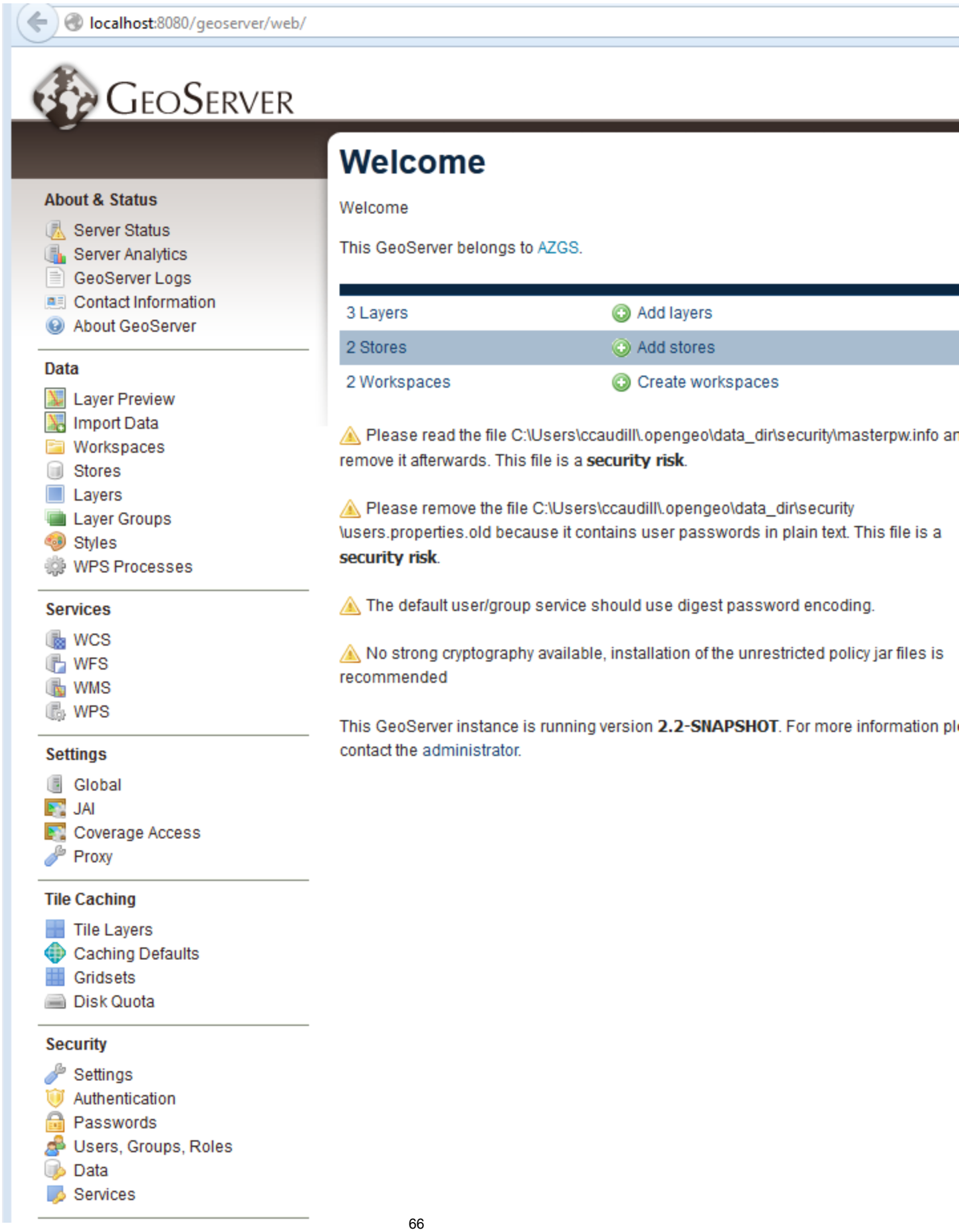
5. Highlight file which now appears in the field and click Import; If the shapefile has a 'geom' field and a defined SRID, the spatial index will be automatically created with the Shape to PostGIS Converter.

Proceed to section 10.2 to connect to this database and deploy the service using GeoServer.

B. Publish services using GeoServer

The following outlines how to use OpenGeo Suite on Windows platform to deploy WMS/WFS services using a connection to PostGIS database tables. From the GeoServer somepage, under About & Status, click Contact Information to enter the service provider information that will serve as your metadata contact information in service requests. For the following example, the GeoServer data directory was assigned as file location C:\Program Files (x86)\OpenGeo\OpenGeo Suite\webapps\geoserver\data.

1. Using the OpenGeo Suite Dashboard, click Start and choose GeoServer to open the GeoServer interface.


localhost:8080/geoserver/web/
GeoServer
About & Status
Server Status
Server Analytics
GeoServer Logs
Contact Information
About GeoServer
Data
Layer Preview
Import Data
Workspaces
Stores
Layers
Layer Groups
Styles
WPS Processes
Services
WCS
WFS
WMS
WPS
Settings
Global
JAI
Coverage Access
Proxy
Tile Caching
Tile Layers
Caching Defaults
Gridsets
Disk Quota
Security
Settings
Authentication
Passwords
Users, Groups, Roles
Data
Services
Welcome
Welcome
This GeoServer belongs to AZGS.
3 Layers
Add layers
2 Stores
Add stores
2 Workspaces
Create workspaces
Please read the file C:\Users\ccaudill\.opengeo\data_dir\security\masterpw.info an
remove it afterwards. This file is a security risk.
Please remove the file C:\Users\ccaudill\.opengeo\data_dir\security
\users.properties.old because it contains user passwords in plain text. This file is a
security risk.
The default user/group service should use digest password encoding.
No strong cryptography available, installation of the unrestricted policy jar files is
recommended
This GeoServer instance is running version 2.2-SNAPSHOT. For more information pl
contact the administrator.

2. Under Data, click Workspaces.

a. Click Add new workspace. The workspace specifies a namespace, so each data type will need its own separate workspace.

b. Choose a name to identify your workspace (i.e., ThermalSprings)

c. Fill in the namespace, given in the first line of an NGDS schema, or XSD, (http://schemas.usgin.org/models/)

d. Fill in the Contact information (metadata). This will serve as the contact information for the data provider.

e. Click Submit

3. Under Stores, click Add new Store. The Store will house the connection parameters to the data table in PostGIS database.

. If using PostGIS, choose this from the Vector Data Sources list. Choose the workspace you've just created from the drop-down list. The Data Source Name* will be the name of the data table as it appears exactly in your PostGIS database (i.e., azthermalsprings).

a. Click the box next to "Enabled".

b. Once connection parameters are filled in, click Save.

4. Under Layers, click Add a new resource.

. From the drop-down list, choose the workspace:Store you've just created.

a. If connected to PostGIS database, this should give all the available tables in the specified folder. Navigate through the list to the data table (i.e., azthermalsprings) and click Publish to the right of that data table name.

b. This takes you to Edit Layer.

c. In the Data tab, under Bounding Boxes, choose 'Compute from data' under 'Native Bounding Box; choose 'Compute from native bounds' under 'Lat/Lon Bounding Box'. This will pull locational bounding boxes from the data table.

d. Click Save at bottom.

5. If ArcMap was used to create the original data table, it is important to note that GeoServer will not correctly read the SRID. This must be manually indicated by creating a SQL view of your layer that will serve as your web service. Under Layers, click Add a new resource.

. Choose the workspace:Store you created previously.

a. When the available data tables pop up, there is an bolded option at the top labeled 'Configure new SQL view...' Click this.

b. In the Create new SQL view page, fill in the 'View Name' that will serve as the final layer name, as indicated by the appropriate schema.

c. In 'SQL statement' use a SQL statement that will call all the fields from the desired data table as such:

select * from azthermalsprings

Click Refresh.

a. When the fields appear, check objectid as the Identifier. For the shape field, type is indicated as Geometry. The SRID box must be filled in with the correct EPSG code (i.e., 4326). Click Save.

b. This takes you to Edit Layer.

c. In the Data tab, under Bounding Boxes, choose 'Compute from data' under 'Native Bounding Box; choose 'Compute from native bounds' under 'Lat/Lon Bounding Box'. This will pull locational bounding boxes from the data table.

d. In the publishing tab, under WMS Settings, a Style (or symbology) may be chosen.

i. To add a Style: under Styles from GeoServer home, click Add a new style.
ii. Choose a name and the workspace for this style (i.e., ThermalSprings)
iii. To use the SLDs for the NGDS project, scroll down to SLD file and click Browse. Browse to your downloaded SLD copies, avaialbe from http://repository.stategeothermaldata.org/repository/resource/50ec3aefb656b70647f32e38bc1b7479/
iv. Click Upload. Click Submit.

e. Click Save.

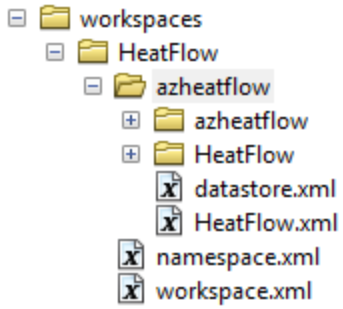
2. Perform a GetFeature request to confirm data return. Construct a GetFeature URL, where the typeName=workspace:layer. For example: http://localhost:8080/geoserver/wfs?service=wfs&version=1.3.0&request=GetFeature&typeName=ThermalSpring:ThermalSpring

7. Create app-schema configuration for the new service. Creating an additional configuration file (mapping file) for GeoServer allows the use of "app-schemas" which will automatically download the indicated web-accessible schema (.xsd) and be referenced in the GetFeature request. This solves a problem with service deployment using ArcGIS, where custom schema locations is necessary. (For more in-depth explanation of app-schema implementation and associated files, see http://repository.stategeothermaldata.org/resources/documents/GeoserverDocumentation.docx)

a. If necessary, download the app-schema extension from http://sourceforge.net/projects/geoserver/files/GeoServer%20Extensions/

i. Download the app-schema plugin zip file for the same version of GeoServer.
ii. Unzip the app-schema plugin zip file to obtain the jar files inside. Do not unzip the jar files.
iii. Place the jar files in the WEB-INF/lib directory of your GeoServer installation.
iv. Restart GeoServer to load the extension.

b. Create a mapping file and place in appropriate location. Example location: C:\Program Files (x86)\OpenGeo\OpenGeo Suite\webapps\geoserver\data\workspaces\HeatFlow\azheatflow



See an example working directory here: https://github.com/ccaudill/GeoserverWorkspace/tree/master/HeatFlow

Example mapping file: https://github.com/ccaudill/GeoserverWorkspace/blob/master/HeatFlow/azheatflow/HeatFlow.xml

c. Edit datastore.xml

Example datastore file: https://github.com/ccaudill/GeoserverWorkspace/blob/master/HeatFlow/azheatflow/datastore.xml

8. Perform a GetFeature request to confirm data return and use of schema. Construct a GetFeature URL, where the typeName=workspace:layer. For example: http://localhost:8080/geoserver/wfs?service=wfs&version=1.3.0&request=GetFeature&typeName=HeatFlow:HeatFlow

The service has been created and needs to be validated against the schema.

Section 7: Validating Services

WMS and WFS services must be validated to ensure the data will be interoperable within the system. For WMS, visual confirmation in a GIS program is sufficient for validation. For WFS, the GetFeature URL for the service will be validated against the schema (schema validation).

A. WMS Validation

Validate the WMS service using WMS Capabilities URL.

ArcMap

a. Navigate to the service rest page and click on 'WMS'
b. Copy the WMS GetCapabilites URL, just left of the '?'
c. Open ArcCatalog. Expand GIS Servers. Choose 'Add WMS Server'
d. Paste in the URL and click 'Get Layers' and 'OK'. Drag the added layer into ArcMap.
e. Ensure that the WMS draws with correct symbology. Also, choose the Identify tool and select a feature. Click on a URI to test the redirects.

uDig free and open-source desktop GIS, available here: http://udig.refractions.net/

a. Click on the "Layer" drop-down list and choose "Add"
b. Choosing data type to add, click "Web Map Server". Paste the WMS GetCapabilities URL into the box.
c. Click "Next". This should add the data directly to the map.
d. To display feature attributes, select the layer in the "Layers" box on the left. On the right, select "Info."
e. Click on a feature on the map and open the "Information" tab on the bottom to display attributes as listed in the service. Click on a URI to test the redirects.

B. WFS Validation

Validate the service against the schema using WFS GetFeature request URL.

USGIN WFS Validator tool: http://schemas.usgin.org/validate/wfs

a. Input the WFS GetCapabilites URL from the service and click Submit
b. Choose the FeatureType and Number of Features to validate
c. Choose the Content Model and Version to validate the service against. This will actually validate the GetFeature request against the schema (XSD). The version of the content model should be the latest version used; see http://schemas.usgin.org/models/ for the versions in use.

Free versions of XML Explorer and Notepad ++
available here: http://xmlexplorer.codeplex.com/, http://notepad-plus-plus.org/

a. Navigate to the service rest page and click on 'WFS'
b. Create a Get Feature Request: In the browser, change the URL by deleting "Capabilites" and replacing with "Feature" and adding to the end of the URL the layer name. This is shown below in bolded text:

http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/AZActiveFaults/MapServer/WFSServer?request=GetCapabilities&service=WFS

http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/AZActiveFaults/MapServer/WFSServer?request=GetFeature&service=WFS&TypeName=ActiveFault&MaxFeatures=2

c. Copy this URL. In XML Explorer "Open Url …". Paste in the Url. Once loaded save the file. Open in Notepad++.
d. From http://schemas.usgin.org/models/ save the schema (.xsd) to validate file against to a local file location.
e. In Notepad++:

i. In the first element, in the attribute "xsi:schemaLocation=", find the 2nd URL that appears (will end with the layer name) and delete. In it's place, write the name of the .xsd saved in step d above (i.e., ActiveFault1.1.xsd).
ii. Save the document as .xml and having the same name and in the same location file location as the .xsd saved in step d (i.e., ActiveFault1.1.xml). Close.

f. Open XML Explorer; Click Open and navigate to the xml file that was just edited in step e.
. Upon opening this .xml file, the program will look at the indicated schema to automatically validate the file.

1. Any returns with a red X will mean a change in the data, field type, or field heading is needed, and likely a re-import of the Feature Class into a staging or deployment database.
2. A return of no errors indicates that the .xml file (and hence the GetFeature request and the service) was successfully validated.

Section 8: Create A New Resource in the Repository

Creating a new resource in the NGDS repository that represents this data service exposes it to the catalog where it becomes searchable within the system. This also creates metadata for the service. This step is necessary to fulfill the Tier 3 USGIN requirements. For additional information on accessing the repository (and data delivery cycle in general), please visit http://usgin.org/sites/usgin.org/files/Data_Delivery_Cycle.pdf. For more information on metadata and how the repository and catalog fit into the project architecture, visit http://usgin.org/sites/usgin.org/files/sgd-repository_tutorial.pdf. The following is a walk-through of how to create a repository resource, and thus metadata, for your new data service.

1. Go to http://repository.stategeothermaldata.org/repository/ to Create a New Resource for the service, if it does not already exist.
2. Upload the an Excel file, database file, or shapefile that represents this service to the repository by clicking "Attach a new file" under the Manage File Attachments on screen left

when creating your new resource. After uploading, the file name may be changed at the bottom of the page under Links.

3. It is important to add the links to the service as this allows the service to be added to the searchable catalog interface at http://search.geothermaldata.org/.
    a. Scroll to the bottom of the page to the Links section. Click the Globe icon.
    b. Add links for ESRI, OGC:WFS, and OGC:WMS services copied from the REST page.
        i. Example ESRI link: http://mapserver.kgs.ku.edu/arcgis/rest/services/NGDS/KSBoreholeTemperatures/MapServer
        ii. Example OGC:WFS link: http://mapserver.kgs.ku.edu/arcgis/services/NGDS/KSBoreholeTemperatures/MapServer/WFSServer?request=GetCapabilities&service=WFS
        iii. Example OGC:WMS link: http://mapserver.kgs.ku.edu/arcgis/services/NGDS/KSBoreholeTemperatures/MapServer/WMSServer?request=GetCapabilities&service=WMS

2. Edit Author and Distributor information.
3. Add Keywords.
4. Add Description (must be greater than 50 characters).
5. Add to Collections on screen left. Collections are categorizations of metadata, where state, organization, and data type are specified.
6. Click Save.

Appendix A – Rewrite Rule Examples

Rewrites for return of online file location

In the 'Redirect to' field at http://resources.usgin.org/admin/, the root redirect URI http:/resources.usgin.org/uri-gin/ will replace the root of the online file location.

online file location example:
http://kgs.uky.edu/kgsweb/geothermal/ar/Logs/123.tiff
file (or unique record identifier): 123.tiff

name authority: args
resource type: logs_tiff
pattern: (.+)/
'Redirect to': http://kgs.uky.edu/kgsweb/geothermal/ar/Logs/$1
final URI in spreadsheet: http:/resources.usgin.org/uri-gin/args/logs_tiff/123.tiff/

where pattern (.+)/ and $1 in the 'Redirect to' field indicates the online file has a simple one-folder path. An example of redirecting to a file having more than one folder to redirect to is as follows:

online file location example:
http://kgs.uky.edu/kgsweb/geothermal/ar/Logs/OilandGas/documents/123.tiff
file (or unique record identifier): 123.tiff

name authority: args
resource type: logs_docs
pattern: (.+)/(.+)/(.+)/

'Redirect to': http://kgs.uky.edu/kgsweb/geothermal/ar/Logs/$1/$2/$3
final URI in spreadsheet: http:/resources.usgin.org/uri-gin/args/logs_docs/OilandGas/documents/123.tiff/

Rewrites for return of XML representations of data service
These will be entered into the 'Redirect to' field at http://resources.usgin.org/admin/

General pattern for rewrites are as follows. Replace the highlighted portions; see examples below.

[HostURL]/[ServiceName]/MapServer/WFSServer?request=GetFeature&service=WFS&TypeName=aasg:[LayerName]&FILTER=%3Cogc:Filter%3E%3Cogc:PropertyIsEqualTo%3E%3Cogc:PropertyName%3Eaasg:[FieldNameForURI]%3C/ogc:PropertyName%3E%3Cogc:Literal%3Ehttp://resources.usgin.org/uri-gin/[NameAuthority]/[Token]/$1/%3C/ogc:Literal%3E%3C/ogc:PropertyIsEqualTo%3E%3C/ogc:Filter%3E

Arizona-hosted example:

HostURL: http://services.azgs.az.gov/arcgis/services/aasggeothermal/
ServiceName: NMWellLogs
LayerName: WellLog
FieldNameForURI: LogURI
NameAuthority: nmbgmr
Token: welllog
Pattern: (.+)/
Media type: application/xml
Redirect to:
http://services.azgs.az.gov/arcgis/services/aasggeothermal/NMWellLogs/MapServer/WFSServer?request=GetFeature&service=WFS&TypeName=aasg:WellLog&FILTER%3Cogc:Filter%3E%3Cogc:PropertyIsEqualTo%3E%3Cogc:PropertyName%3Eaasg:LogURI%3C/ogc:PropertyName%3E%3Cogc:Literal%3Ehttp://resources.usgin.org/uri-gin/nmbgmr/welllog/$1/%3C/ogc:Literal%3E%3C/ogc:PropertyIsEqualTo%3E%3C/ogc:Filter%3E

Illinois-hosted examples:

http://geothermal.isgs.illinois.edu/arcgis/services/aasggeothermal/MEaqWellChemistry/MapServer/WFSServer?request=GetFeature&service=WFS&TypeName=aasg:$1&FILTER=%3Cogc:Filter%3E%3Cogc:PropertyIsEqualTo%3E%3Cogc:PropertyName%3Eaasg:AnalysisURI%3C/ogc:PropertyName%3E%3Cogc:Literal%3Ehttp://resources.usgin.org/uri-gin/megs/chemistry/aqchem/$1/$2/%3C/ogc:Literal%3E%3C/ogc:PropertyIsEqualTo%3E%3C/ogc:Filter%3E

http://geothermal.isgs.illinois.edu/arcgis/services/aasggeothermal/ILDrillStemTests/MapServer/WFSServer?request=GetFeature&service=WFS&TypeName=aasg:DrillStemTest&FILTER=%3Cogc:Filter%3E%3Cogc:PropertyIsEqualTo%3E%3Cogc:PropertyName%3Eaasg:ObservationURI%3C/ogc:PropertyName%3E%3Cogc:Literal%3Ehttp://resources.usgin.org/uri-gin/isgs/dst/$1/%3C/ogc:Literal%3E%3C/ogc:PropertyIsEqualTo%3E%3C/ogc:Filter%3E

Appendix B – Naming Conventions

Data Frame (Service Name) and Layer Names must be without special characters or spaces. The layer name must be as exactly specified 'element name' in schema. The service name must include the state's abbreviated name concatenated with the Service type (usually plural). The following is a list of service names and layer names approved for use in the NGDS/USGIN schema:

Service Names + State Prefix (eg AZActiveFaults)	Layer Name
State Prefix + ActiveFaults	ActiveFault
State Prefix + AqueousChemistry1_10	BaseMetals
State Prefix + AqueousChemistry1_10	CommonAnalytes
State Prefix + AqueousChemistry1_10	FreeGas
State Prefix + AqueousChemistry1_10	GasIsotopes
State Prefix + AqueousChemistry1_10	IsotopesDissolved
State Prefix + AqueousChemistry1_10	MajorDissolvedConstituents
State Prefix + AqueousChemistry1_10	MinorDissolvedConstituents
State Prefix + AqueousChemistry1_10	Nitrogen
State Prefix + AqueousChemistry1_10	SingleAnalyte
State Prefix + AqueousChemistry1_10	WaterDissolvedGas
State Prefix + AqueousChemistry1_10	WaterIsotopes
State Prefix + AqueousChemistry1_10	WaterQuality
State Prefix + BoreholeLithIntercepts	BoreholeLithIntercept
State Prefix + BoreholeLithIntervals	BoreholeLithInterval
State Prefix + BoreholeTemperatures	BoreholeTemperature
State Prefix + ContourLines	ContourLine
State Prefix + DirectUseSites	DirectUseSite
State Prefix + DrillStemTests	DrillStemTest
State Prefix + FluidFluxInjectionDisposal	FluidFluxInjection
State Prefix + GeothermalAreas	GeothermalArea
State Prefix + GravityStations	GravityStation
State Prefix + HeatFlow1_23	HeatFlow
State Prefix + HeatPumpFacilities	HeatPumpFacility
State Prefix + HydraulicProperties	HyraulicProperty
State Prefix + LASFileLocations	BoreholeTemperatureLASLog
State Prefix + PhysicalSamples	PhysicalSample
State Prefix + PowellCummingsGeothermometry	GasAnalysis
State Prefix + PowellCummingsGeothermometry	LiquidAnalysis
State Prefix + PowerPlantProduction	PlantProduction
State Prefix + PowerPlantFacilities	PowerPlantFacility
State Prefix + RadiogenicHeatProduction	RadiogenicHeatProduction
State Prefix + RockChemistry	Isotopes

State Prefix + RockChemistry	NobleGases
State Prefix + RockChemistry	RareEarths
State Prefix + RockChemistry	SingleAnalytes
State Prefix + RockChemistry	StableIsotopes
State Prefix + RockChemistry	TraceElements
State Prefix + RockChemistry	USeries
State Prefix + RockChemistry	Volatiles
State Prefix + RockChemistry	WRMajorElements
State Prefix + SeismicHypocenters	Hypocenter
State Prefix + TemperatureLogs	BoreholeTemperatureGeophysicalLog
State Prefix + ThermalConductivity	ThermalConductivity
State Prefix + ThermalSprings	ThermalSpring
State Prefix + VolcanicVents	VolcanicVent
State Prefix + WellFluidProduction	FluidProduction
State Prefix + WellHeaders	Wellheader
State Prefix + WellLogs	WellLog
State Prefix + WellTests	WellTest

Appendix C - Service Description Example

Service Description Example

This web map service (WMS) was published using ArcGIS Server v. 10.0 and is compliant with OGC (Open Geospatial Consortium) version 1.30 specifications. This service provides dynamic, spatially referenced geographic information using data collected for the National Geothermal Data System (http://www.geothermaldata.org/). In addition to the WMS capabilities, this service was designed to be interoperable with both WFS (Web Feature Services) as well as KML (Keyhole Markup Language). The WFS capabilities allow the client to query, make additions and/or modifications to an existing dataset. WFS can be utilized through the interoperability extension in ArcCatalog. For more information on using the ArcGIS data interoperability extension visit http://www.esri.com/software/arcgis/extensions/datainteroperability /common-questions.html. A KML service allows the client to view an image of the data in three dimensions, using free software available for download on the internet such as ArcGIS Explorer or Google Earth. For more information on OGC specifications, visit http://www.opengeospatial.org/standards.

Appendix D – Online Resource Examples

For AZGS-hosted services:
http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/OKBoreholeTemperatures/MapServer/WMSServer
http://services.azgs.az.gov/ArcGIS/services/aasggeothermal/OKBoreholeTemperatures/MapServer/WFSServer

For Illinois-hosted services:

http://geothermal.isgs.illinois.edu/arcgis/services/aasggeothermal/MEaqWellChemistry/MapServer/WMSServer
http://geothermal.isgs.illinois.edu/arcgis/services/aasggeothermal/MEaqWellChemistry/MapServer/WFSServer

For Kentucky-hosted services:
http://kgs.uky.edu/arcgis/services/aasggeothermal/LAWellheaders/MapServer/WMSServer
http://kgs.uky.edu/arcgis/services/aasggeothermal/LAWellheaders/MapServer/WFSServer

Appendix E - Performing a Reprojection and Creating the Feature Class

Coordinates of original data (North America) might reference NAD 83 or NAD 27 SRS. The specified coordinate system for NGDS is European Petroleum Survey Group (EPSG) 4326, which is the code equivalent to WGS 84. Therefore, to conform to the NGDS EPSG:4326 specification, the NAD 83 or NAD 27 data needs to be geographically transformed. Use the tools in ArcToolbox to perform the projection and transformation.

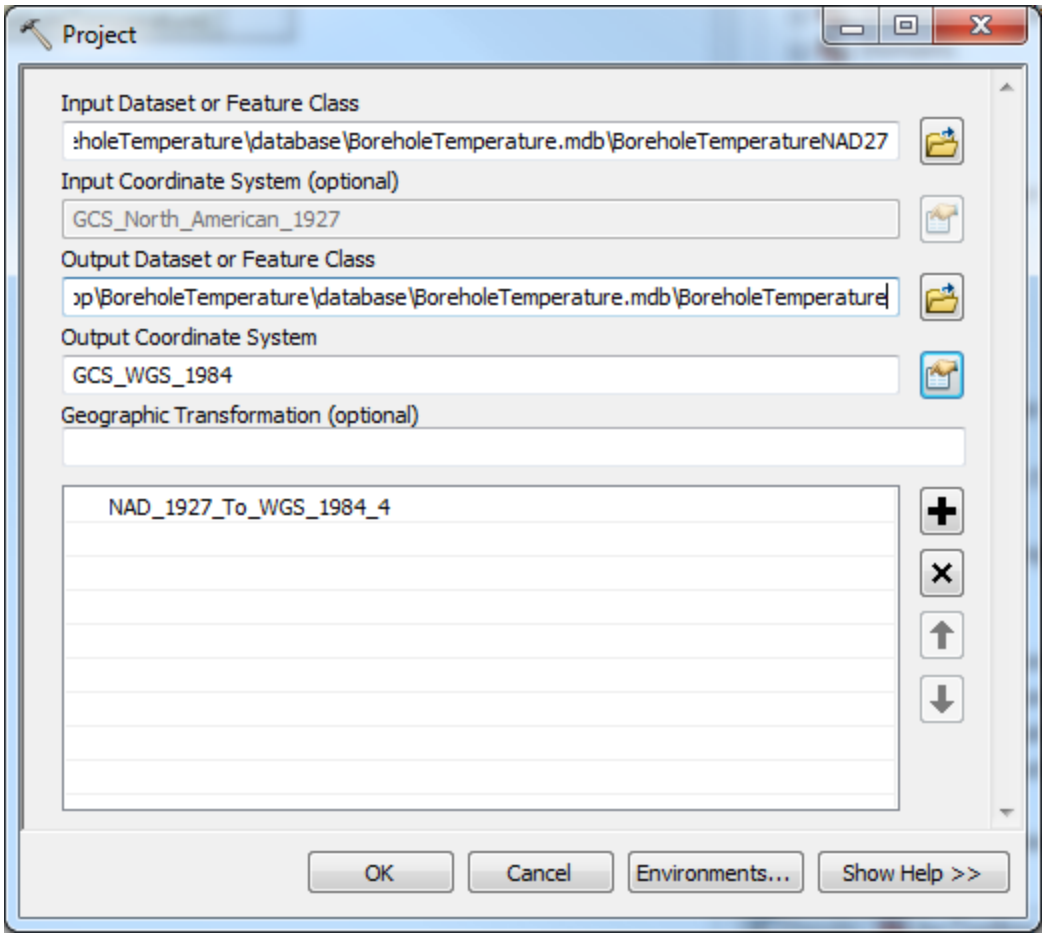
1. In ArcCatalog, open ArcToolbox window
2. Expand Data Management Tools, Projections and Transformations, Feature
3. Double-click on 'Project'

a. Add Feature Class that needs to be projected as the 'Input Dataset'
b. 'Input Coordinate System' value defaults to the coordinate system for the Feature Class.
c. Specify output location for the Feature Class to be the same database as the table. Name the Feature Class the layer name (Appendix B). Feature class name for Arizona Borehole Temperature example would be "AZBoreholeTemperatureWGS84".
d. Select Output Coordinate System.
e. Select Geographic Transformation. Common transformations for North America:
i. NAD 27 -> WGS 84: NAD_1927_To_WGS_1984_4
ii. NAD 83 -> WGS 84: NAD_1983_To_WGS_1984_1

4. Click OK.

For other transformations, please refer to the ArcGIS documentation that is included in the ArcGIS installation. This typically can be found in the following system directory:

C:\Program Files (x86)\ArcGIS\Desktop10.0\Documentation\geographic_transformations.pdf



SECTION 7: CALCULATE NEW COORDINATES

The last step in the data transformation process is calculating WGS84 coordinates and the EPSG.

1. In ArcCatalog, open ArcToolbox window
2. Expand Data Management Tools, Fields
3. Double-click on calculate field

a. Add Feature Class from Section 6 as the Input Table

b. Calculate Latitude

i. Select `LatDegree` for field name

ii. Type `!shape.firstpoint.Y!` in Expression field

iii. Select `Python_9.3` for Expression Type

iv. Click OK

c. Calculate Longitude

. Select `LongDegree` for field name

i. Type `!shape.firstpoint.X!` in Expression field

ii. Select `Python_9.3` for Expression Type

iii. Click OK

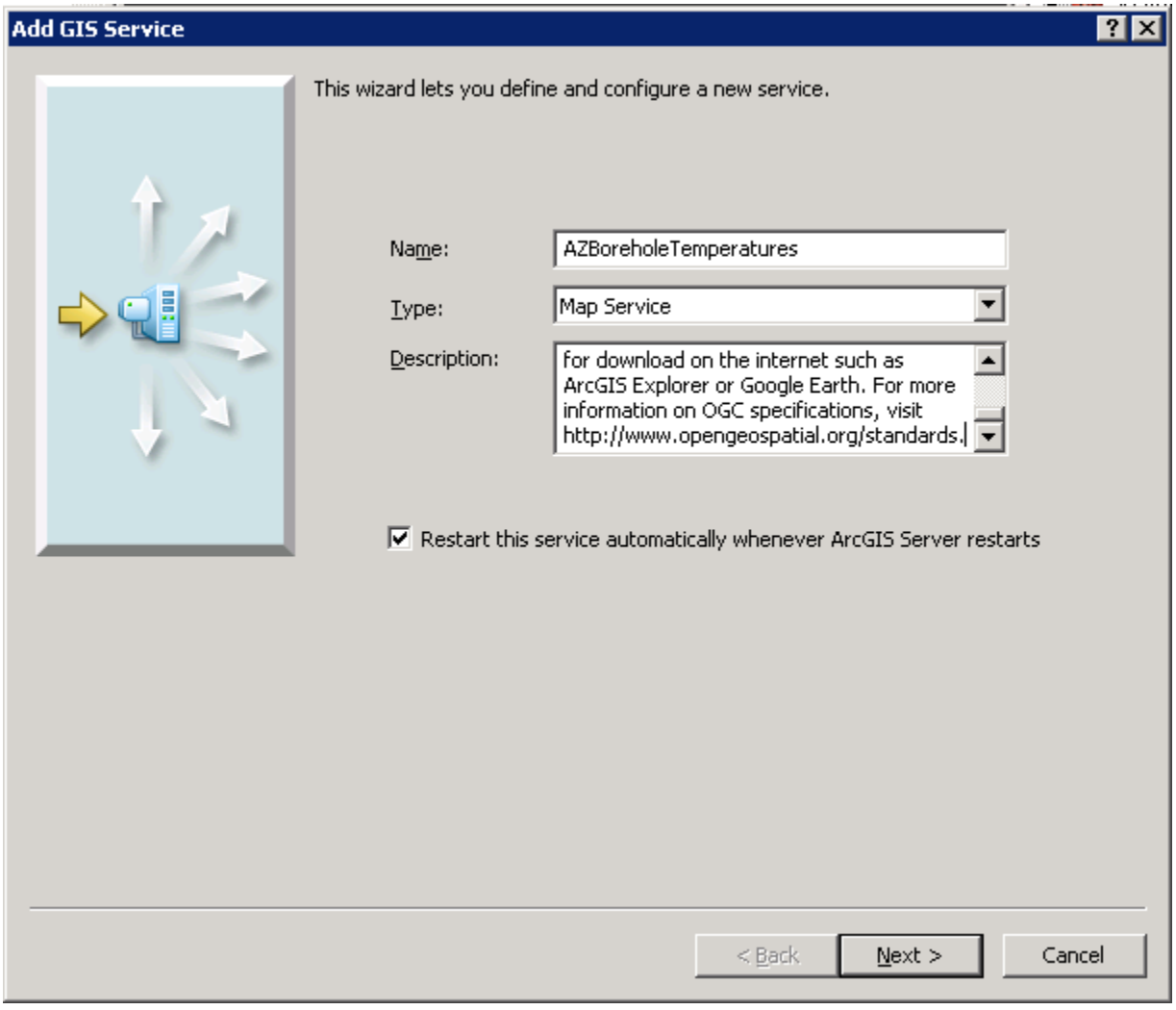
d. Double-click on calculate field, again

. Input same Feature Class

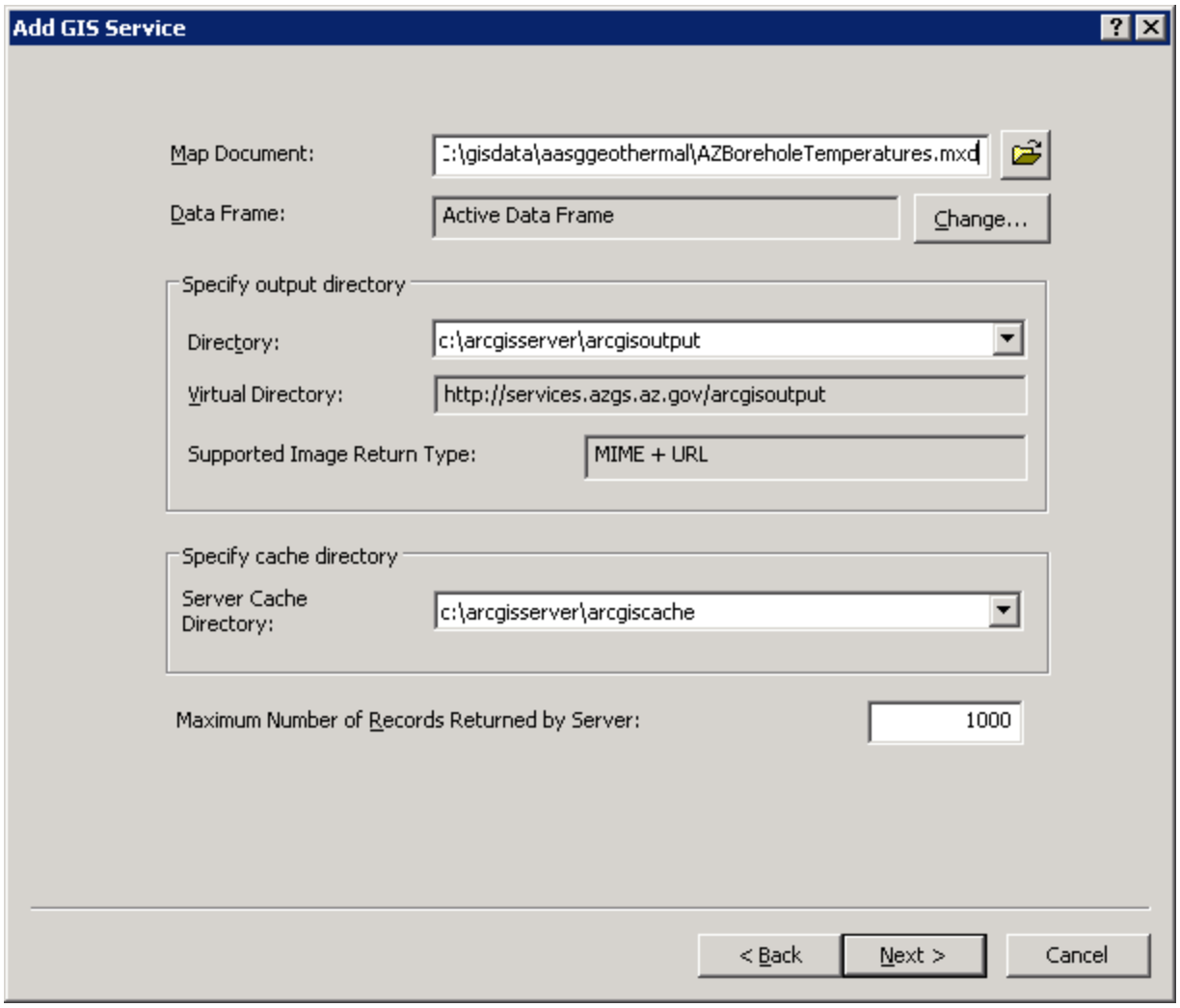
i. Select SRS for field name
ii. Assuming data has gone through data transformation process and is now in a WGS 84 projection, type "EPSG:4326" in Expression field (quotes included).
iii. Select VB for Expression Type
iv. Click OK

The data is now transformed into the NGDS specified spatial reference system and the SRS field is updated to read EPSG:4326.

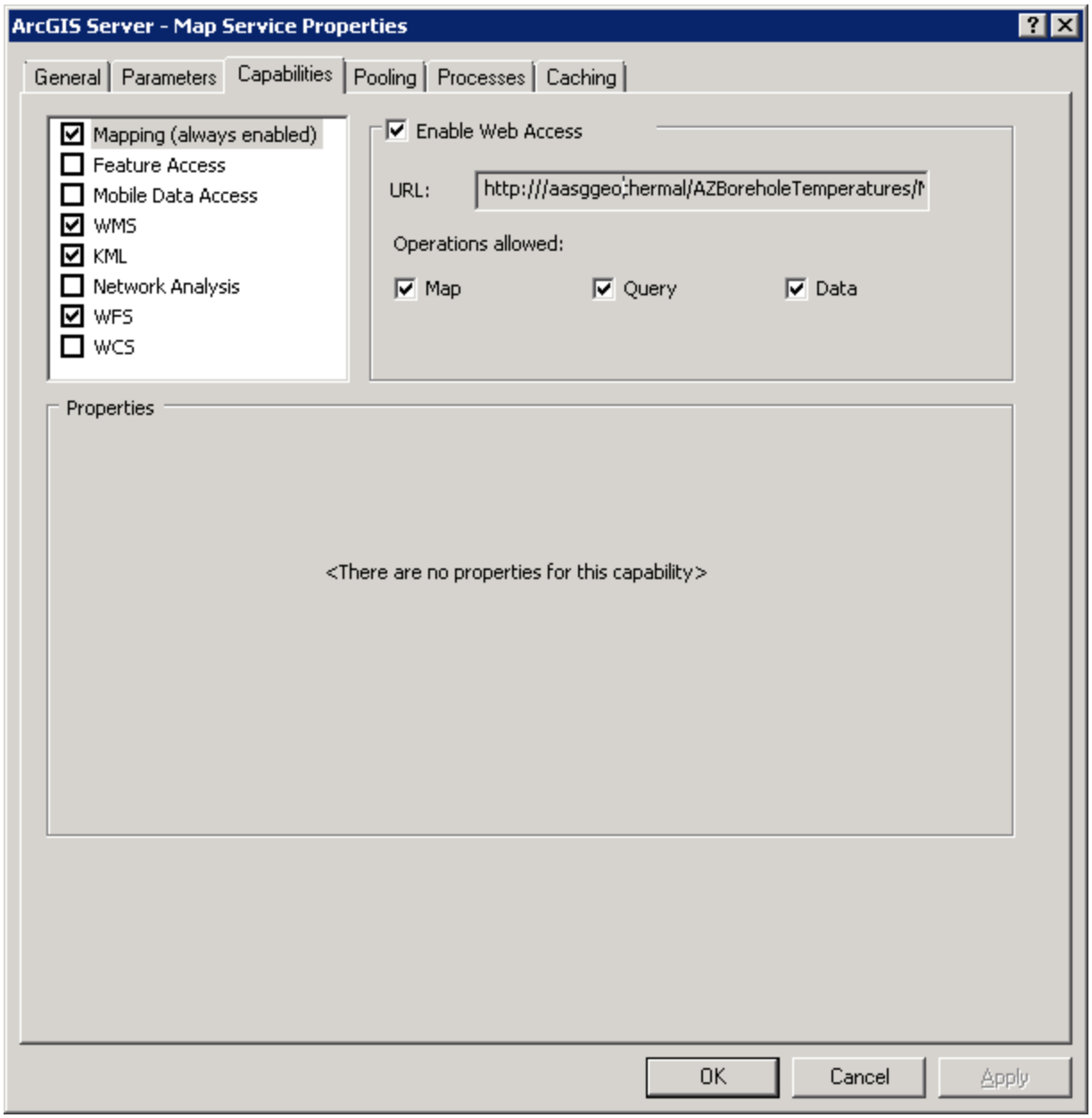
Appendix F: Creating an ArcGIS (v.10) Service



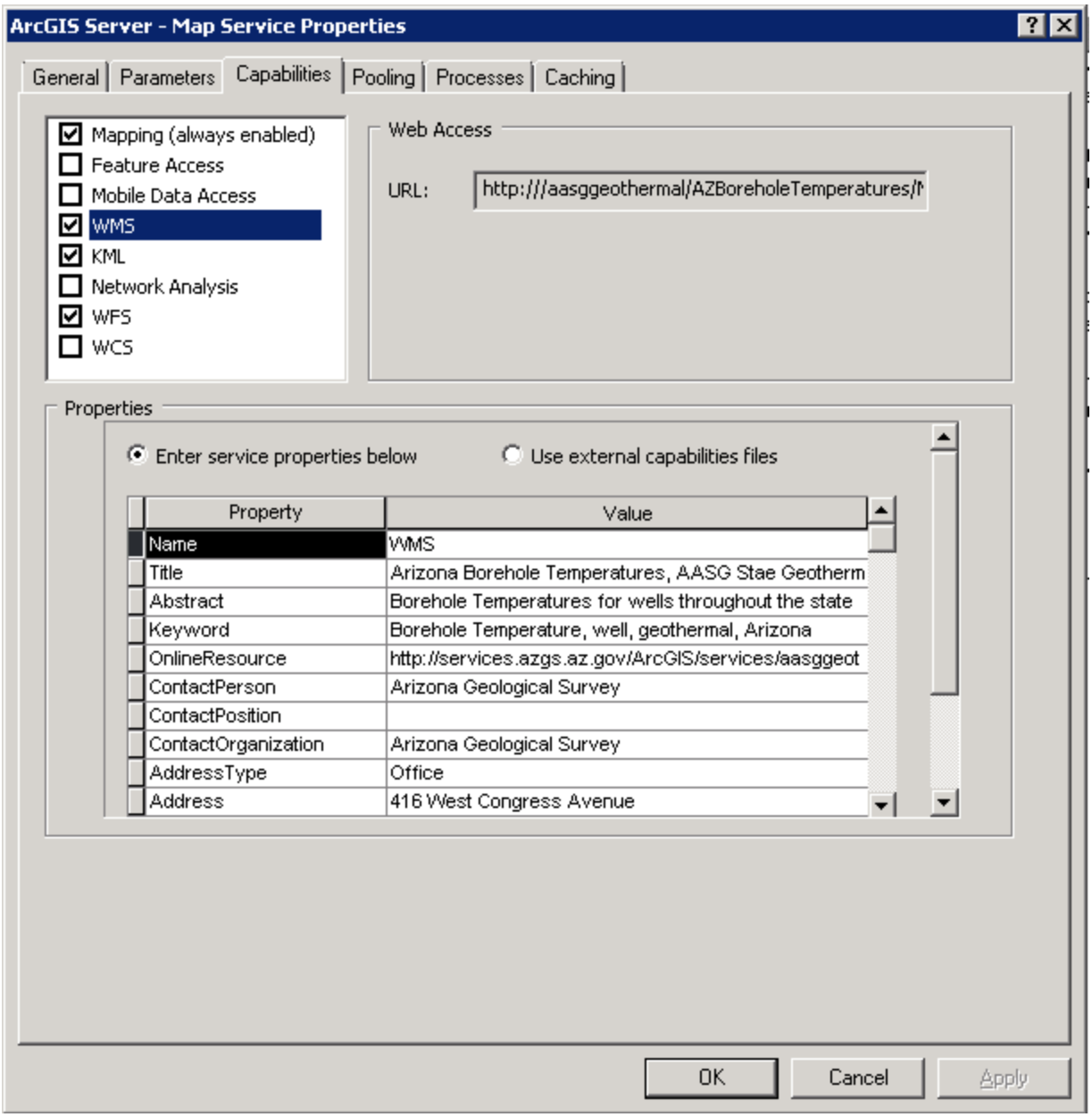
Add ArcGIS Service: Service Name and Description



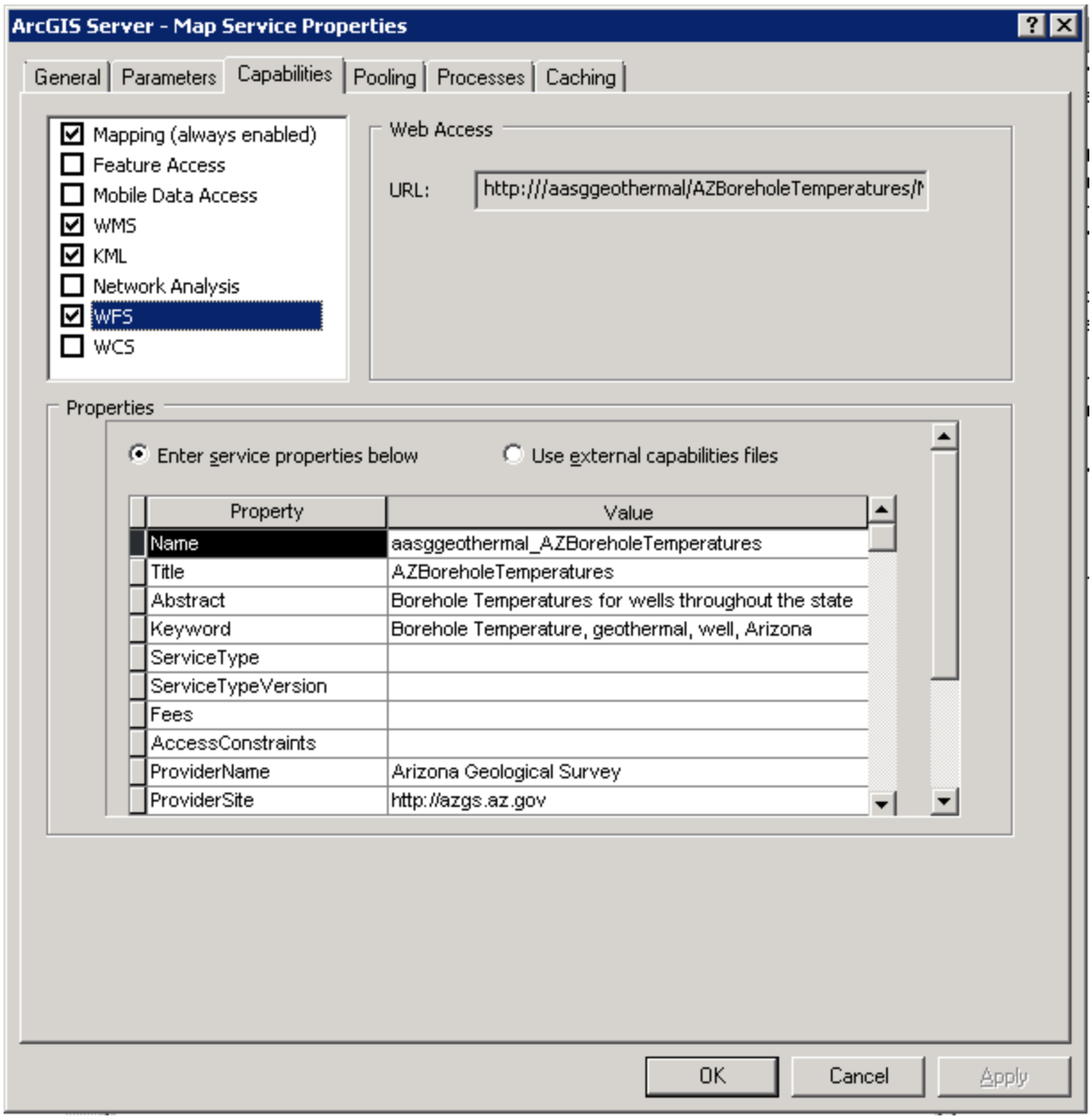
Add ArcGIS Service: Select Map Service Document



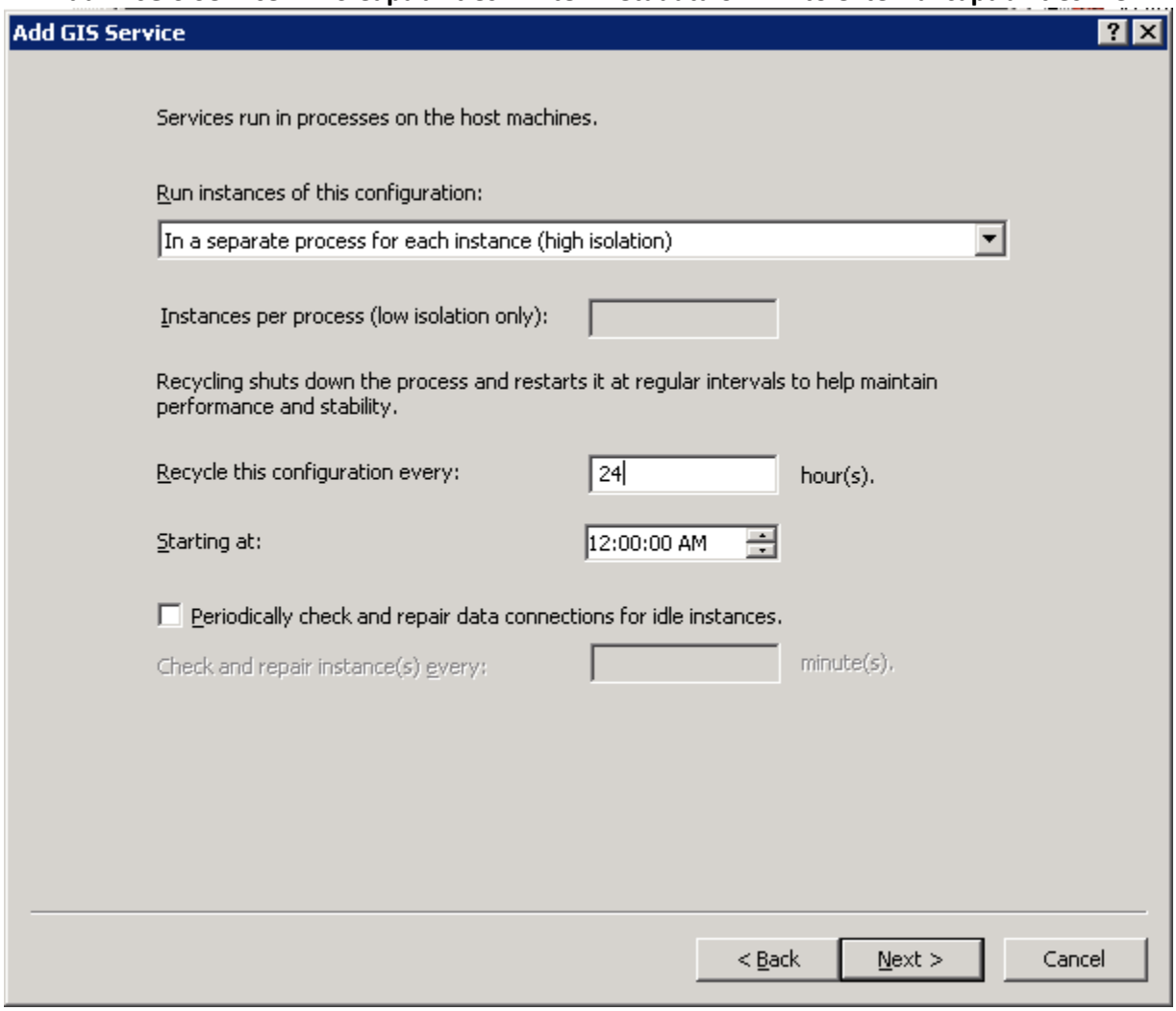
Add ArcGIS Service: Select Capabilities



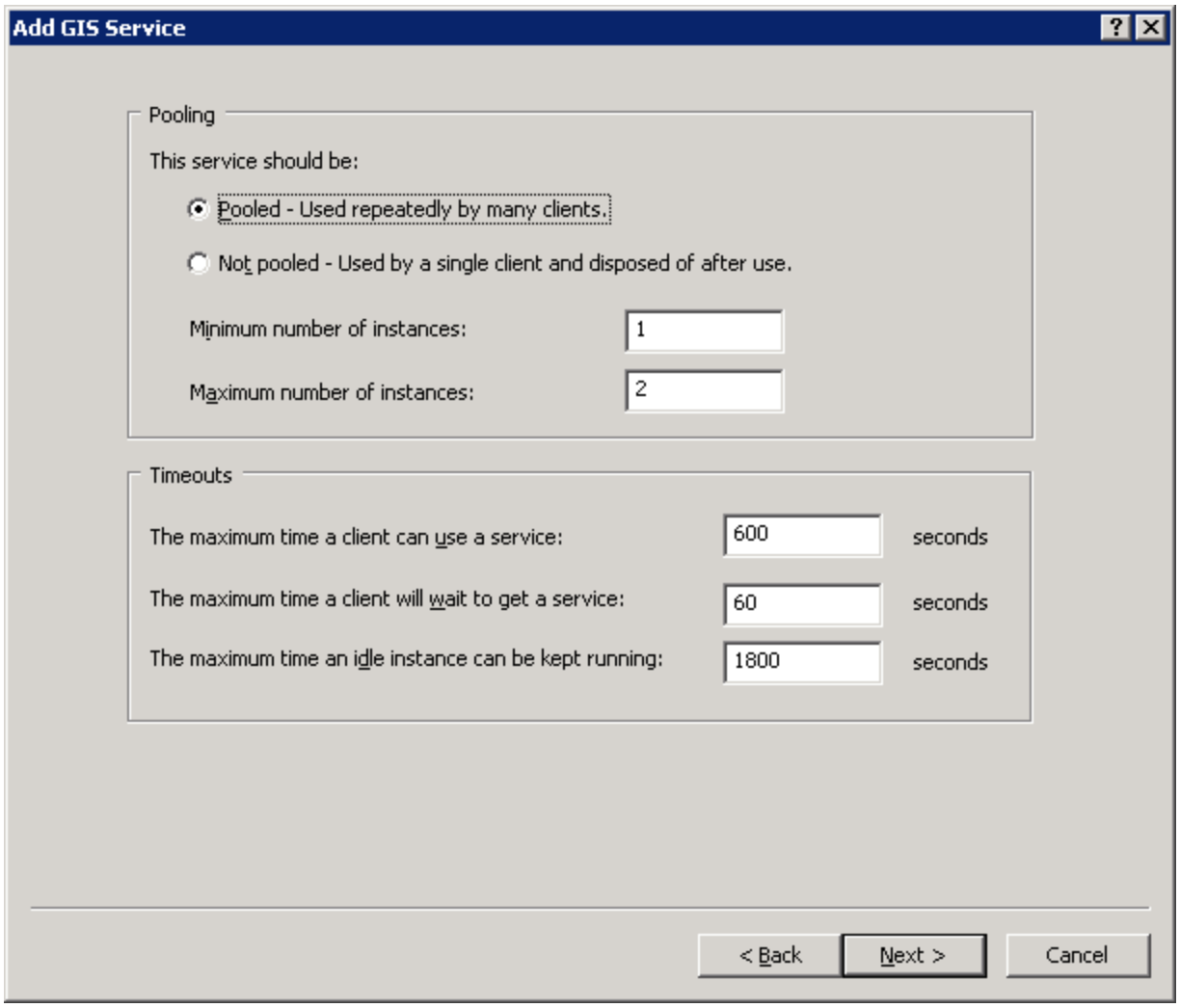
Add ArcGIS Service: WMS Capabilities - Enter metadata or link to external capabilities file


ArcGIS Server - Map Service Properties
General
Parameters
Capabilities
Pooling
Processes
Caching
Mapping (always enabled)
Feature Access
Mobile Data Access
WMS
KML
Network Analysis
WFS
WCS
Web Access
URL:
http:///aasggeothermal/AZBoreholeTemperatures/M
Properties
Enter service properties below
Use external capabilities files
Property
Value
Name
aasggeothermal_AZBoreholeTemperatures
Title
AZBoreholeTemperatures
Abstract
Borehole Temperatures for wells throughout the state
Keyword
Borehole Temperature, geothermal, well, Arizona
ServiceType
ServiceTypeVersion
Fees
AccessConstraints
ProviderName
Arizona Geological Survey
ProviderSite
http://azgs.az.gov
OK
Cancel
Apply

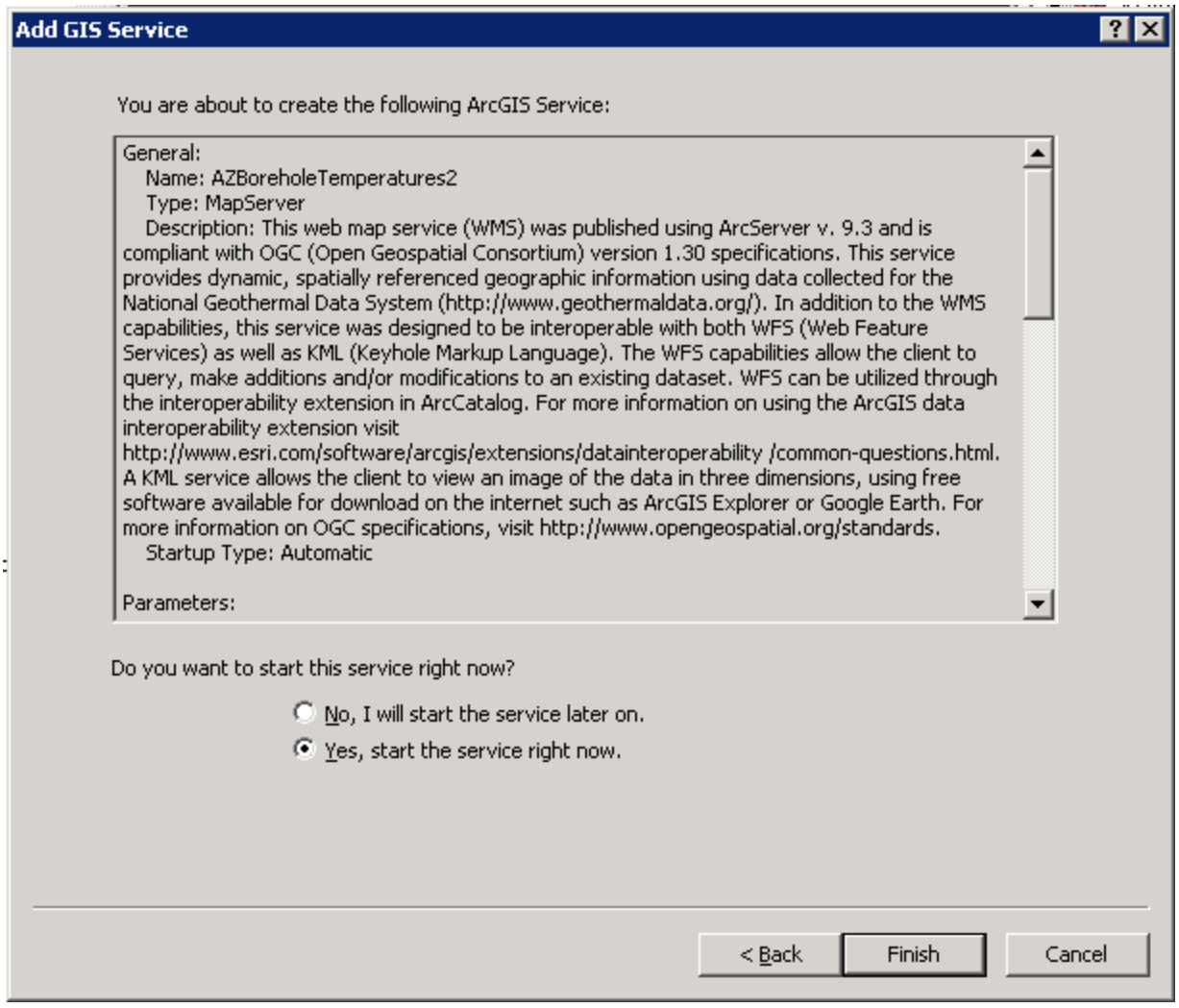
Add ArcGIS Service: WFS Capabilities - Enter metadata or link to external capabilities file



Add ArcGIS Service: Processes



Add ArcGIS Service: Pooling



Add ArcGIS Service: Start Services

Appendix G: Manually Reviewing the Data and Creating the NGDS Service

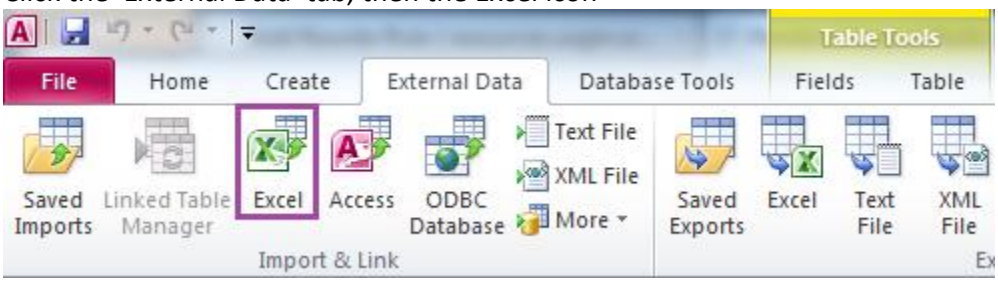
If the Excel to NGDS Service is unable to be used below are the directions to be followed.

Section G1: Import Excel Data Into Access and Review Data Types

Access provides a bridge between Excel and ArcMap, creating a database from the Excel file. Access is also where data types can be specified and Shape and OBJECTID fields can be maintained. Microsoft provides a free runtime version of Microsoft Access which can be installed on any Windows desktop, and run any Microsoft Access database application (http://www.microsoft.com/en-us/download/details.aspx?id=4438). Although use of Access is highly recommended, it is possible to create a Feature Class from an Excel file without use Access or having a shapefile. Please proceed to Section 3a for this procedure. To import the Excel file into Access:

1. Open Access

2. Create a new .mdb database (2002-2003 format)
3. Import data
    h. Click the 'External Data' tab, then the Excel icon



i. Get External Data - Excel Spreadsheet Window opens. Browse to Excel Workbook, select worksheet and click open. Click OK.

j. Import Spreadsheet Wizard Opens. Select the worksheet that corresponds to the tab from the Excel Workbook which contains the data. This worksheet will most likely be suffixed "Template". Click Next.

k. Check the box next to 'First Row Contains Column Headings'. Click Next.

l. Click Next on Field Options window. Field data types will be addressed later.

m. Choose 'No primary key'. Click Next.

n. Import to Table: This is where you can name the imported table. It is recommended that the new Access table is prefixed with the content model type and suffixed with the word Table. For example, if you imported data from the Borehole Temperature Observation content model, name the table "BoreholeTemperatureObservationTable". This convention allows a later differentiation from this table and the Feature Class that will eventually be created from this table.

2. Address Import Errors: If the Excel import has errors in Access, evaluate the error (usually a data type conversion error, where number fields must not contain alpha-numeric characters). Delete the imported table from Access. Re-open the Excel Workbook and fix errors. Re-import the table into Access.
3. Review field names, field order and data types in Access
    h. Right-click table and select Design View. Here you will see the field names, field order and data types for the table. Compare these against the schema (http://schemas.usgin.org/models/).

i. Field names, field order and data types should match the schema exactly. If field names, field order and data types are different in the Access table, make necessary changes.

i. Field names: check spelling and capitalization. Correct any spellings and capitalizations if necessary.

ii. Field order: verify that the fields in the table have the same ordering as the fields in the schema. If not, move fields in Access to match the schema field ordering.

iii. Data types: verify that the data type for a field matches the data type defined in the schema. If it doesn't match, change it.

NOTE: Save the Design View after changing each data type. If a warning appears that records will be deleted, cancel the change and reformat data in Access.

1. If schema data type = string, then field data type in Access table could be one of three data types:

a. Text (255 character limit)
b. Memo
c. Date/Time (format: yyyy-mm-dd\Thh:n)

The data type will depend on the type of information that is going into the field. Text data type should be sufficient for most fields. Memo is recommended for fields that will have a lot of text (i.e. notes, sources). Date/Time data type should be used for date fields.

2. If schema data type = double, then field data type in Access table should be:
a. Number

ii. Reformating data in Access

0. Do this only if you get a warning when saving data type changes in Design View!
. Switch to the Table View in Access.
a. For date fields, it may be necessary to replace the 'T' with a space to keep the correct data format. Do a find and replace for the column, replacing the 'T' with a space. Save and again change the data type & format mask in the Design View.
b. For number fields, ensure that no letter characters exist in the data in those fields. When all letter characters have been removed, so back to the Design View and correct the data type.
c. Return to Design View after you modify the data.
b. Save changes. There may be a warning that some data may be lost. This error pops up when saving the table after changing a data type from 'text' to 'number'. This is due to the restricted size of number fields. Unless the number character string is more than 255 characters, ignore this warning.
3. Replace special characters
h. Certain characters are not accepted by ArcCatalog for SDE Feature Classes. These are usually found in chemical analysis datasets, and must be removed before deploying if using Arc products for service deployment.
i. Open Access Table (Datasheet View)
j. Replace the following characters:

ii. δ (replace with "delta")

iii. μ (replace with "mu")

1. RelatedResources field: This field is meant to correlate other relevant services or other relevant online documents. If other services from the same state are running, find a column of matching data (ID or APINo, for example) to correlate with dataset. Joining the two datasets by this field (under 'Create' then 'Query Design' in Access), populate the RelatedResources field with the main URI from the other service data.

Please continue to Section 2 of this document.

Section G1a: Excel to Geodatabase without Access

This method is not generally recommended due to the inability to choose "Memo" data type. In Access, this data type indicates that the string of text characters may exceed 255. Be aware that when indicating "Text" data type in ArcMap for a given field, the characters in that field will be limited to 255. In this section, you will create a File Geodatabase and a Feature Class, importing the data types from a specified Excel content model using ArcMap. Navigate to the Excel file (table) in ArcCatalog.

Note: the spatial reference for NGDS services is WGS 1984. If the SRS field data indicates another coordinate system (such as NAD 27 or NAD 83), then create a Feature Class using that coordinate system.

1. Right click in the Contents space to create New - File Geodatabase.
2. Right click the new geodatabase and choose Import - Table (Single).

a. Choose the working Excel file created in Sections 1 & 2 for "Input Rows".
b. Rename the new table.
3. Right-click on new table, select Create Feature Class, select From XY Table...
4. Input fields:
a. X Field => LongDegree
b. Y Field => LatDegree
c. Z Field => <None>
d. Coordinate System of Input Coordinates: Click button, navigate to the coordinate system that corresponds to the value entered in the SRS field.
e. Specify output location for the Feature Class to be the same database as the table. Name the Feature Class. If the spatial reference is WGS 84, name the Feature Class the layer name (Appendix B) that will be used in your service. If not, add the spatial reference to the end of the layer name.

> For example, for an Arizona Borehole Temperature Observation service, the Feature Class should be named "AZBoreholeTemperature". If the original data indicated an SRS of NAD 27, then the Feature Class should be named "AZBoreholeTemperatureNAD27".

If the data has a spatial reference other than WGS 84, project and transform the data as described in Section 6 of this document.

5. Create another New Feature Class by right clicking in the Contents space and choosing New - Feature Class.
a. Name the Feature Class suffixed with '_content model' to avoid confusion.
b. Choose the feature type: Point Features.
c. Choose the Geographic Coordinate Systems
i. World
1. WGS 1984
d. XY Tolerance is default.
e. Configuration Keyword is default.
f. In the next window specifying fields, choose Import. Navigate to the Excel file of the relevant content model (these can be downloaded from http://schemas.usgin.org/models/).
g. Modify the data types indicated by the FieldList worksheet in the Excel NGDS content model and the schema (XSD) file (also available from the USGIN model schemas site).
. OBJECTID field will automatically be indicated as 'Object ID' data type
i. SHAPE field will automatically be indicated as 'Geometry' data type
ii. If schema data type = string, then field data type in ArcMap could be one of the following data types:
a. Text (255 character limit)
b. Date/Time (format: yyyy-mm-dd\Thh:n)

ii. If schema data type = double, then field data type in ArcMap should be:

. Double

b. Click Finish.

6. Append the data to the content model to preserve data types.

a. Open ArcToolbox.

i. Choose Data Management Tools.

1. Choose Append.

b. Input Datasets => Feature Class with data created in numbers 1 - 2 of this Section.

c. Target Datasets => empty content model Feature Class created in number 5 of this Section.

d. Choose 'No Test' to map the fields of one Feature Class to the other.

e. Click OK; the Feature Class will now have the correct data types as specified by the schema.

Continue to Section 5 of this document.

Section G2: Register Access Table with ArcGIS Geodatabase

An Access file with the .mdb extension is a personal geodatabase in ArcGIS. To eventually publish data using ArcGIS Server, the Access table must be registered with an ArcGIS geodatabase.

1. Open ArcCatalog.
2. Navigate to the database created in Section 3. Double-click to open the database.

a. Click on Connect to Folder icon if you need to connect to the database folder.

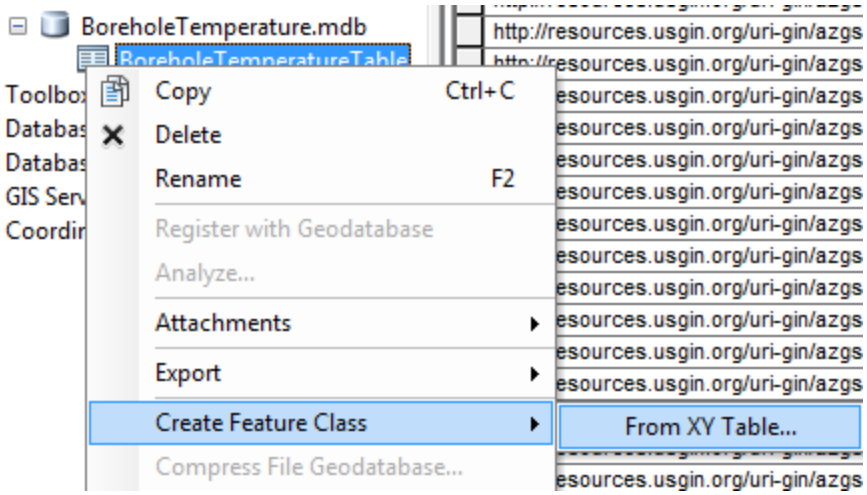
3. Right click the table with the data and select 'Register with Geodatabase.' This creates the OBJECTID field, a require for publishing services with ArcGIS Server.
4. Close ArcCatalog.
5. Re-open the Access database. Open the table in Design View. Move the OBJECTID field so it is the first field. This is a schema specification. Save (right click on the tab) and close the database.

Section G3: Create a Feature Class

In this section, you will create a point Feature Class from the latitude and longitude values in the table you registered with the geodatabase in Section 4. The SRS (Spatial Reference System) field of the table contains the spatial reference system value that should be used when setting the input coordinate system in the process outlined below.

Note: the spatial reference for NGDS services is WGS 1984. If the SRS field data indicates another coordinate system (such as NAD 27 or NAD 83), then create a Feature Class using that coordinate system.

1. Open ArcCatalog.
2. Navigate to same database.
3. Right-click on table, select Create Feature Class, select From XY Table...

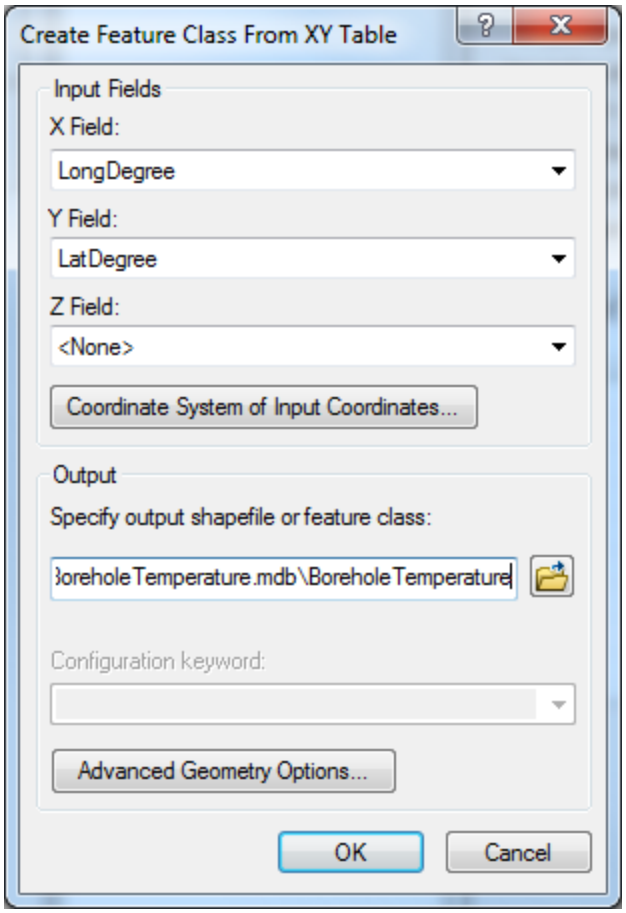


Create Feature Class From XY Table

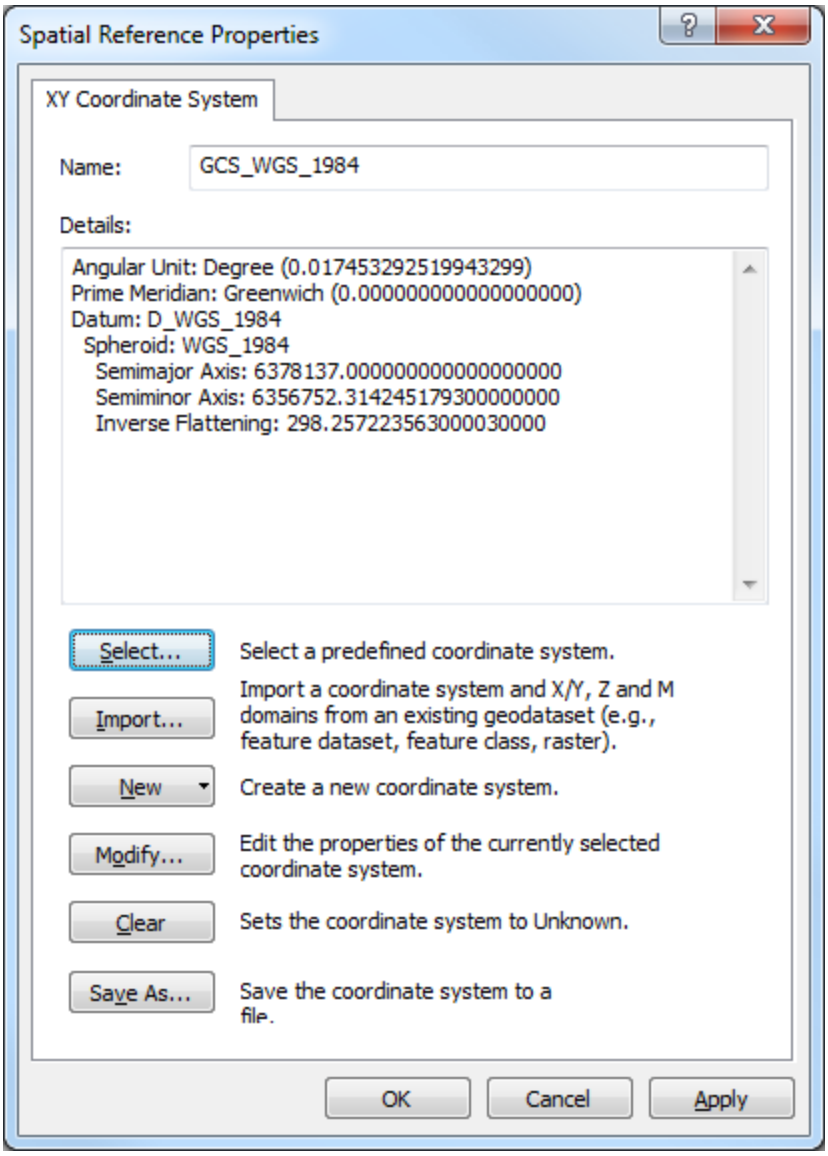
4. Input fields:
    f. X Field => LongDegree or UTM_E
g. Y Field => LatDegree or UTM_N
h. Z Field => <None>
i. Coordinate System of Input Coordinates: Click button.
i. Click Select. . . button in next window and navigate to the coordinate system that corresponds to the SRS. If input coordinates are latitude and longitude, select the appropriate geographic coordinate system. If input coordinates are UTM coordinates, select the projected coordinate system for the appropriate UTM Zone and datum (i.e. NAD 1983 UTM Zone 12N.prj).
j. Specify output location for the Feature Class to be the same database as the table. Name the Feature Class. If the spatial reference is WGS 84, name the Feature Class the layer name (Appendix B) that will be used in your service. If not, add the spatial reference to the end of the layer name.

    For example, for an Arizona Borehole Temperature Observation service, the Feature Class should be named “AZBoreholeTemperature”. If the original data indicated an SRS of NAD 27, then the Feature Class should be named “AZBoreholeTemperatureNAD27”.

If the data has a spatial reference of WGS 84, then proceed to Section 5. Otherwise, project and transform the data as described Appendix E.



Select coordinates and Spatial Reference Properties


Spatial Reference Properties
XY Coordinate System
Name:
GCS_WGS_1984
Details:
Angular Unit: Degree (0.0174532925199433299)
Prime Meridian: Greenwich (0.000000000000000000)
Datum: D_WGS_1984
Spheroid: WGS_1984
Semimajor Axis: 6378137.000000000000000000
Semiminor Axis: 6356752.314245179300000000
Inverse Flattening: 298.257223563000030000
Select...
Select a predefined coordinate system.
Import...
Import a coordinate system and X/Y, Z and M domains from an existing geodataset (e.g., feature dataset, feature class, raster).
New
Create a new coordinate system.
Modify...
Edit the properties of the currently selected coordinate system.
Clear
Sets the coordinate system to Unknown.
Save As...
Save the coordinate system to a file.
OK
Cancel
Apply

SOFTWARE REQUIREMENTS SUMMARY: NATIONAL GEOTHERMAL DATA SYSTEM

based on Software Requirements Specification Version 2.7, dated 11/05/2012

Version History

Version	Author	Date	Reason
	Stephen Richard	2014-01-30	Cull use cases and requirement to create draft summary document. Highlight statements of requirement or use case in red if not achieved, green if achieved to some degree, yellow if not evaluated or there are other issues.
	Christy Caudill	2015-01-22	Updated from newest revisions of general system, workflows, and codebase.

Executive Summary

This document is based on v2.7 of NGDS software requirements as delivered by Siemens, dated 11/05/2012. Most of the explanatory and cover material is removed to reduce the document to a list of use cases and requirements for quick reference and review of progress on meeting requirements as of January 2014

Purpose of NGDS

In the following we list the five basic needs that need to be realized by NGDS:

1. NGDS shall enable data collectors to create and administrate a repository for geothermal data.
2. NGDS shall enable end users/data consumers to search geothermal data across a multitude of repositories
3. NGDS shall enable end users/data consumers to evaluate discovered data
4. NGDS shall enable end users/data consumers to acquire (i.e. download) selected data
5. NGDS shall enable the end users/data consumers to analyze selected data

Table Of Contents

1 Introduction 6
1.1 Vision & Goal 6
1.2 Notation in this document 6
1.3 User Communities and Roles 6
1.3.1 Data Provider Community 7
1.3.2 Data Consumer Community 8
1.3.3 System Administrators 8
1.3.4 Software Developer Community 8
2 Overall system quality attributes 8
2.1 Maintenance 9
2.2 Usability & Accessibility 9
2.3 Performance and Scalability 10
2.4 Security 10
2.5 Supportability 11
2.6 Data Requirements 11
2.7 Design Constraints 11
2.8 Licensing Requirements 12
2.9 Applicable Standards 12
3 GIN-stack Software Package 12
4 End-User/Data Consumer Software 14
4.1 NGDS WebSite 14
4.2 NGDS WebApp 14
5 NGDS Data Access Protocols 14
6 Use Case Models 14
6.1 High-level Use Cases View 15
6.2 Data Provider Use Cases 16
6.2.1 Data Submitter 16
6.2.1.1.1 Upload new files 16
6.2.1.1.2 Update uploaded file 17
6.2.1.2.1 Process data file in NGDS content model template 18
6.2.1.3.1 Create metadata record through a form 19
6.2.1.3.2 Bulk upload metadata from metadata content template table 20
6.2.1.3.3 Browse and manage resource directory and collections 21

6.2.1.3.4 Bulk update metadata records 22
6.2.2 Data Steward 22
6.2.2.1.1 View resource submission and update logs 22
6.2.2.2.1 View catalog audit log?? 24
6.2.2.2.2 View quality assurance reports 24
6.2.2.2.3 Flag resource quality issue 25
6.2.2.2.4 Browse flagged data catalog entries 25
6.2.2.2.5 Perform manual error correction 26
6.2.2.2.6 Clear quality flag 26
6.2.2.2.7 Catalog cleanup 27
6.2.2.2.8 Notify data submitter of data or metadata problem 28
6.2.2.3.4 Make resource public 28
6.2.3 Node Administrator 29
6.2.3.1.1 Administer users 29
6.2.3.1.2 Enrollment 30
6.2.3.1.3 Add users 31
6.2.3.1.4 Delete users 31
6.2.3.1.5 Administer user roles 32
6.2.3.1.6 Backup 32
6.2.3.1.7 Restore 33
6.2.3.2.4 Register node 33
3.2.3.3.1 e-mail (NGDS or GIN-stack) administrator 34
6.2.3.3.4 Upgrade GIN-stack application 34
6.3 Software Developer Use Cases 35
6.3.1 Develop Applications 35
6.3.1.1.1 Develop apps using NGDS standard protocols (needs reformulating) 35
6.3.2 Register New Application with NGDS 36
6.3.2.1.1 Contact NGDS admin to provide link to application 36
6.4 End-User Use Cases 37
6.4.1 Discover and Gather Data 39
6.4.1.1.1 Map-based search 39
6.4.1.1.2 Landmark-based search 40
6.4.1.1.3 Coordinate-based search 40

6.4.1.2.1 Keyword content-based search 41
6.4.1.3.1 Filter results by type 41
6.4.1.3.2 Filter results by metadata attributes 42
6.4.1.3.3 Filter results [geographically] on map 42
6.4.2 Validate and Evaluate Data 43
6.4.2.1.1 Browse/view metadata search results 43
6.4.2.2.1 View metadata record 44
6.4.2.2.2 Provide peer ratings 44
6.4.2.3.1 Triangulate with other sources 45
6.4.2.4.1 View data content 46
6.4.2.4.2 Save selected search criteria 47
6.4.2.4.3 Load previous search criteria 47
6.4.2.4.4 e-mail metadata record URI to third party users 48
6.4.2.4.5 Subscribe to new data 48
6.4.3 Analyze and Visualize Data 49
6.4.3.1.1 Export metadata 50
6.5 System Administrator Use Cases 51
6.5.1.1.1 Register new nodes into NGDS 51
6.5.1.1.2 Delete nodes from NGDS network 52
6.5.1.1.3 Communicate with the Node-in-a-box admin 52
6.5.1.1.4 Manage NGDS user accounts 53
6.6 Use cases common to all users 53
6.6.1.1.1 Login 53
6.6.1.1.2 Logout 54
7 Acronyms, and Abbreviations 54

1 Introduction

This Software Requirements Specification (SRS) collects, organizes and describes requirements for the NGDS software system captured through use-case models, and from natural language requirements statements from a sample of system users. These include functional requirements, non-functional requirements (NFR), design constraints, and other factors needed to provide a comprehensive picture of the software's operation.

1.1 Vision & Goal

The ultimate goal of the National Geothermal Data System (NGDS) is to support the discovery of geothermal sources of energy. The NGDS will provide online access to important geothermal-related data from a network of data providers in order to:

I. Increase the efficiency of exploration, development and usage of geothermal energy by providing a basis for financial risk analysis of potential sites

II. Assist state and federal agencies in making land and resource management assessments

III. Foster the discovery of new geothermal resources by supporting ongoing and future geothermal-related research

IV. Increase public awareness of geothermal energy

1.2 Notation in this document

[Comments from SMR 2014-01-30 are in brackets]

Red highlight -- not done or not working

Green highlight -- working to some degree

Yellow highlight -- other considerations, or SMR can't evaluate

1.3 User Communities and Roles

User groups include the three main target groups defined in the System Vision document, and Administrators who are responsible for maintenance of the searchable catalog and entry-point web-application, as well coordinating management of system-wide standards and protocols. User groups or roles and their relationships are illustrated in Figure 1.

In Figure 1 we also outline the three main target user communities – Data Provider, Software Developer, and End User. These communities are discussed in general terms, with more detailed descriptions of their respective use cases outlined in the next section.

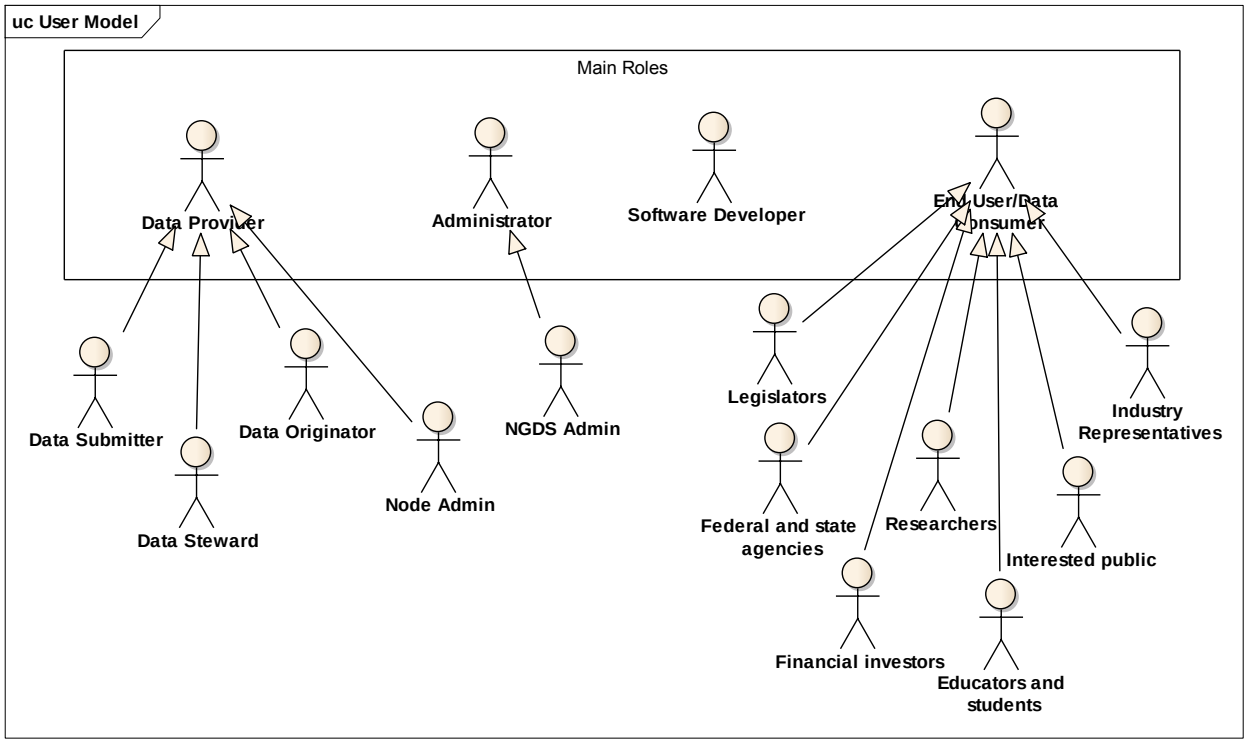


Figure 1 Main user roles and their relationships

The National Geothermal Data System will be a network consisting of four linked communities:

- **Data providers** who will expose information to the system through standardized, internet-accessible interfaces and interchange formats
- **Data consumers** who will utilize the software and information provided by the system in order to understand and develop geothermal resources.
- **Administrators** who are responsible for administrating and monitoring the system. Typical tasks are installation and configuration, user management, node monitoring, or system backup.
- **Software developers** who will build applications that utilize the data in the system, and make it easier for end-users to interact with the system.

1.3.1 Data Provider Community

Data providers represent the collection of users who will work together to publicize information to NGDS through standardized, internet-accessible interfaces using one of the supported interchange formats. This community of users can be further broken down into several distinct roles.

- **Data Submitter**: the user who contributes any resources to the NGDS, preferably using the NGDS protocols and services to publish a piece of data
- **Data Steward**: the user who maintains the quality of a piece of published data
- **Data Originator:** the person that created a piece of data, e.g. a publication. The data originator is not an active actor in any of the use cases involved in the NGDS: as soon as the originator wishes to contribute to NGDS, they take on the role of Data Submitter

- **Node-in-a-Box (or simply Node) Administrator**: the user who is responsible for operating and maintaining an organization's data repository and insures that the data is provided according to NGDS standards and protocols

1.3.1.1 Data Submitter Role

Data submitters will publish geothermal related data to NGDS.

1.3.1.2 Data Steward Role

The data steward user role will verify and maintain the quality of published data. The data steward will have write access to data under his or her responsibility.

1.3.1.3 Node Administrator Role

The node administrator is responsible for the administration of one of the data repositories, or nodes, in the system.

1.3.2 Data Consumer Community

Data consumer will utilize the software and information provided by the system in order to understand and develop geothermal resources.

Specifically, Data Consumers will:

- Search for the entire NGDS system utilizing the federated NGDS catalog service either via the NGDS WebApp application, or potentially any other catalog search application that operates with NGDS search protocols and metadata interchange formats.
- Visualize and explore data in map, text, or other graphic presentations (as prioritized by user research findings).
- Select and acquire data via service interfaces using NGDS interchange formats
- Access files in NGDS data repositories
- Save and re-load search queries
- Set filters to be notified as new data of interest is available via the system

1.3.3 System Administrators

The system administrator is responsible for the administration of system components. Node administrators are responsible for their own node (repository, WebApp and catalog), while the NGDS administrator is responsible for the whole NGDS set of nodes, and the federated catalog service. The system administrator is also responsible for coordinating review and adoption of system protocols and interchange formats and for the registration of new interchange formats and associated schema.

1.3.4 Software Developer Community

Software developers may build applications that access NGDS resources using the protocols and standards outlined as part of the system architecture.

2 Overall system quality attributes

In this section, we discuss overall system quality attributes that further describe the conditions on which the system will function.

2.1 Maintenance

For those components that are designed within the project we require unit tests to be available that test at least the minimum functionality. In addition system integration tests are recommended but due to the high costs of designing a system test toolkit this might not be possible with the given time and resources.

NFR001 All project-developed source code shall have comments at least on a per-class level.

NFR002 The System's architecture shall be documented. [document exists, needs maintenance]

NFR003 The System's configuration parameters shall be documented.

NFR004 The System's source code shall be covered by unit tests to at least 50% of coverage. Regression tests will be run as part of the software process. [?]

2.2 Usability & Accessibility

NFR005a The System shall provide a reasonably simple to use installation tool: The tool shall install all required components (potentially with the exception for Java and/ or Python) and guide the administrator through the initial configuration steps.

NFR005 The system shall be cloud-ready: It must be available as an archived Virtual Machine (or VM) that is ready to use after a few configuration steps. Our main target is EX2(from Amazon) and other providers that are able to boot a VM image)

NFR006 The system shall include detailed instructions that guide the user through the process of installation of one node and joining a grid. [Document exists]

NFR008 The project-developed graphical user interfaces shall use a uniform look-and-feel for web applications, defined by the UX team. Minor customizations will be possible by adjusting Cascading Style Sheets, for example.

NFR009 The project-developed applications shall provide online help explaining how to perform user-related functions. [some help is available, not context sensitive, no tooltips]

NFR010 The project-developed applications shall present the user with clear, understandable and accurate information explaining each task that can be performed using the software.

NFR011 The project-developed applications shall present the user with human understandable error messages explaining the errors that occur during user interactions. [many error messages are not helpful]

NFR012 The key data import operations should be transactional. i.e. The user shall be able to abort operations before completion, without any negative consequences. [?]

NFR013 The project-developed applications shall provide a status indicator showing the progress towards completion of user triggered processing, search queries, exports and downloads.

NFR013b The system shall comply with the section 508 Amendment to the Rehabilitation Act of 1973 section related to Web-based Intranet and Internet Information and Applications. [not tested]

NFR013c The system shall comply with the ISO/TS 16071, "Ergonomics of human-system interaction – Guidance on accessibility for human-computer interfaces. [not tested]

2.3 Performance and Scalability

We assume the system will be designed to support universities and data providers within USA, with a maximum of 1000 concurrent users.

NFR014 Each data provider node must be capable to maintain a list of at least 100 other NGDS nodes for harvest or distributed search.

NFR015 Each data provider node shall indicate it has taken action in response to all user operations within 2 (two) seconds.

NFR016 Each data provider node shall be capable of supporting up to 50 simultaneous authenticated, logged-in users.

NFR017 Each data provider node shall be capable of handling at least 50 (fifty) HTTP requests every 1 (one) minute.

NFR018 Each data provider node shall respond to every request from the NGDS in no more than 10 (ten) seconds.

NFR019 The System shall be able to handle the import of data files up to 2GB in size.

NFR020 The System shall be able to handle the import of up to 1000 data files in any one import operation.

NFR021 The System shall support the storing of up to 100000 data files in the import directory of each data provider.

NFR022 The System shall support the storing of up to 500GBs of data files in the import directory of each data provider.

2.4 Security

NFR024 The System shall embody a security plan and process to ensure that unauthorized users are denied access.

NFR025 Valid login authentication is required for all data submitter, steward, and administrator functions.

NFR026 The System shall only allow users access to write data they have permissions to write

NFR027 The System shall only allow users access to download data files they have permissions to download

NFR028 The System shall only allow the data steward for a resource permission to delete it.

NFR029 The communication between end-users and the services of the system will be encrypted using HTTPS protocol.

NFR030 Data communications between the External Systems and NGDS applications shall be secured by message authentication where applicable/necessary.

NFR031 A data provider node shall only transmit data to clients when the data is published by an authenticated data submitter or steward for access by that client's permission group.

NFR032 The data provider node shall maintain the integrity and availability of all data stored in its local data store.

NFR033 The data provider node shall maintain the integrity of all files stored in the node's local file repository.

NFR034 The data provider node shall maintain a log of activities for auditing purposes. [Data submission, updates to metadata are logged, not clear if anything else is logged]

NFR035 NGDS applications shall be developed considering good security coding practices, thus minimizing vulnerability to attacks. In particular, it should comply with FIPS (Federal Information Processing Standards)

2.5 Supportability

Here follows the requirements that enhance the supportability and maintainability of NGDS project-developed software.

NFR036 NGDS software shall be written using the standard coding style for the used programming languages. For example, Oracle Java Coding Styles, if Java turns out to be the selected language.

NFR038 NGDS software shall be designed utilizing the concept of encapsulation. Components shall be created that encapsulate related functionality within them, and nothing else.

NFR039 All software shall be modular to minimize the time and complexity involved in maintaining and extending the platform and application. [needs refactoring]

NFR040 NGDS software shall not contain any statically detectable dead code. [Adrian reports significant dead code; no static detection tests run]

2.6 Data Requirements

The Data Assessment Team will inventory additional data to be submitted by project subcontractors, and project management will prioritize datasets for delivery based on recommendations from the Domain Steering Committee and User-Centered design team. A prioritized list of important geothermal data types will be ascertained and used as the basis for populating NGDS data assets. See the Data Requirements Specification /P02/, for a description of data categories and attributes that will be stored in the System's database.

2.7 Design Constraints

The follow requirements are derived from design decisions that represent constraints that are mandated and must be adhered to.

NFR041 The NGDS components shall interface to NGDS data provider nodes via NGDS web services. [CKAN API is used between web application and server]

NFR042 The NGDS components shall use the API provided by NGDS web services for data functions, for example WCS, WFS. [Data browse and view from WFS and WMS is supported, uses CKAN API for tabular (csv) datasets]

NFR043 The NGDS participants shall use a data abstraction layer for access to databases used for metadata management and management of data in NGDS content models. [Unclear how this is to be tested]

NFR044 The system components shall use web services for communication with NGDS client software and other NGDS nodes.

NFR045 The NGDS components shall use web services to send data to NGDS client applications, including the NGDS portal application. [CKAN API?]

NFR046 The NGDS components shall use web services for the querying of data from NGDS nodes.

NFR047 The system shall provide requested metadata to the NGDS as XML files [transactions from web portal to server use CKAN JSON??]

NFR048 NGDS metadata catalog services shall provide metadata search results using the USIGN ISO XML encoding profile in response to CSW 2.0.2 protocol. [need to test CKAN csw more]

NFR049 The software shall be designed with Linux as its main target platform. However, as much as possible, it shall utilize portable technologies such as Java, that will facilitate its porting to other operating systems and platforms.

NFR050 Under no condition should the failure of one node be capable of crippling or rendering the entire NGDS system useless. The system must be capable of adding or removing nodes while maintaining normal operations.

2.8 Licensing Requirements

The software system being developed as part of the project shall have an open source license variant. The license will include provisions to allow users to copy, distribute and transmit the software, to adapt the software for other applications, and to make commercial use of the software, under the condition that the following attribution for the source of the software is included in any copy or derived work:

In order to preserve the original NGDS licensing terms, the use of third party libraries and application servers that violate these terms will be vetted.

The data provided by the data providers will not be made available to the public domain until it is released and published through the “publish” feature of the System. Other access control constraints may be applied by individual nodes at their discretion.

2.9 Applicable Standards

See the WSS ***/P05/*** for the details of the web services standards to be applied.

3 GIN-stack Software Package

A redistributable, free/open-source node or aggregator software package will be developed (composed of an NGDS repository, an NGDS catalog and an NGDS Desktop as shown in Figure 2).

The SCR development team will deliver the NGDS node-in-a-box software stack (WepAPP, NGDS Repository and NGDS catalog) to BSU project management team for deployment, testing and acceptance on a server of their choice, in accordance to the project’s schedule.

The NGDS node-in-a-box software **will be delivered** as an installable application to BSU project management for deployment, testing and acceptance on a server of their choice, in accordance to the project’s schedule.

The software package **will include** an installation program in the form of a yum shell script of an apt package for CentOS Linux, due to its versatility and support of Red Hat Enterprise Linux.

This GIN-stack software application will

1. give data providers a simple way to register data sources, load data and expose those data as a node in the NGDS network
2. support batch import and upload of shared datasets in supported formats adhering to standard content models
3. WebApp, user interface will be provided to help users upload data to the system.

In this arrangement, different GIN-stack instances can co-exist in the system. This requires their catalogs to be synchronized via a federation service or some sort of aggregating catalog.

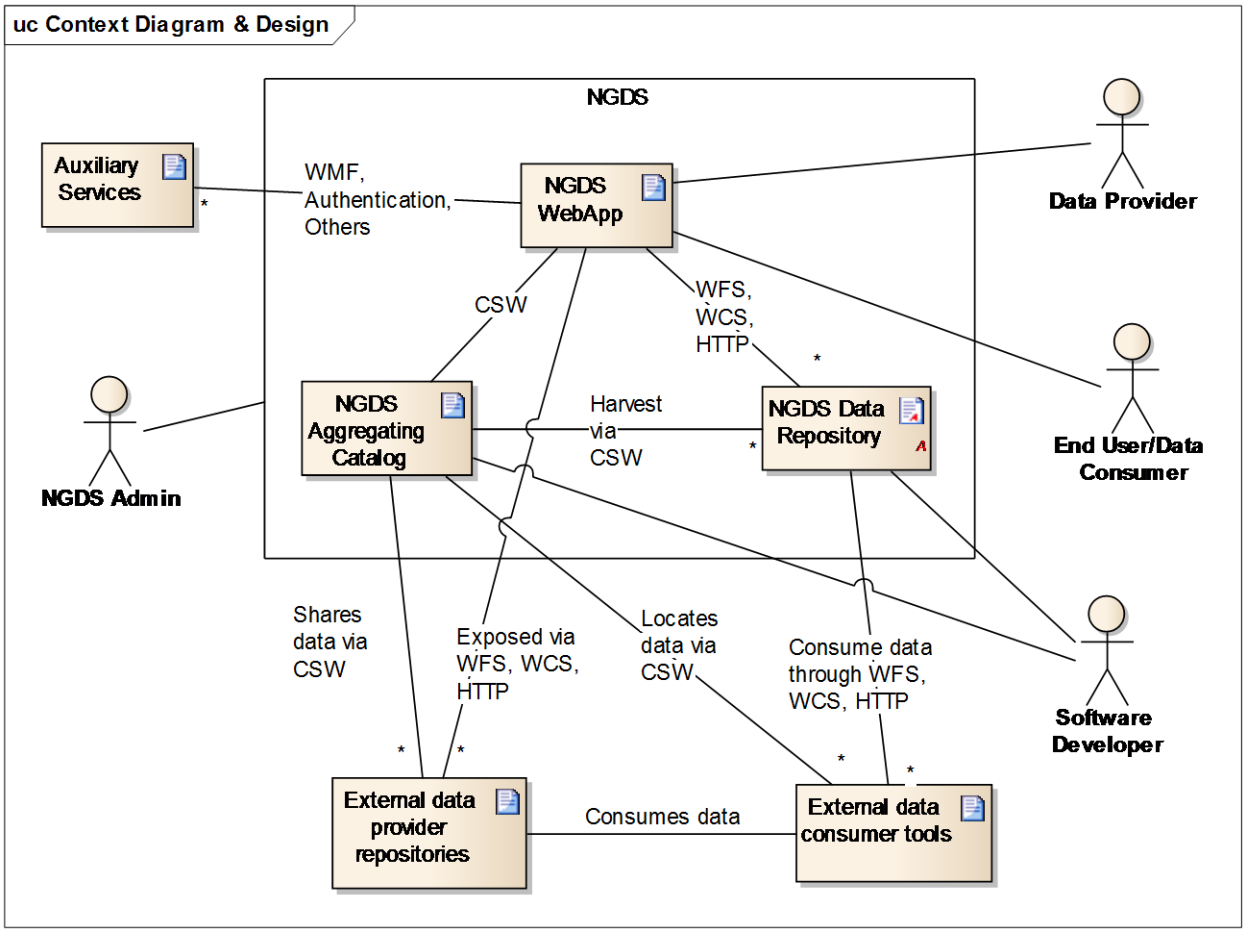


Figure 2 NGDS Data Provider Software Package

NGDS will

1. Facilitate publication, visualization and discovery of geothermal data using services and applications.
2. Provide a catalog, a web top UI and a data repository application.

3. NGDS System Administrator can install and manage data providers, granting them the ability to publish data to the system
4. Enable users to easily update via the same yum shell script with no re-installation.

4 End-User/Data Consumer Software

The project will implement two primary access points: a NGDS website and a NGDS WebApp application. As much as possible, these will be integrated in order to appear to the end-user as a single web-based experience. The primary access point to the system will be the NGDS WebApp application.

4.1 NGDS WebSite

The website will be designed to

1. provide information about all the NGDS participants
2. serve as gateway to the system to discover data and applications that utilize NGDS resources
3. include information on the project's progress, NGDS specifications, access to the map-centric search application, other software applications utilizing NGDS services, NGDS presentations, documentation and tutorials, a link to the catalog of NGDS nodes, and any other results as they become available.

4.2 NGDS WebApp

A user-friendly, web-based application will be created in order to support

1. finding, visualizing, mapping, and acquisition of data by end-users/data consumers
2. provide a user interface that allows end-users/data consumers to discover and access resources made available across all NGDS nodes
3. search for data across the system based on topic, location, time or other criteria
4. provide the user with the information necessary to determine the utility of that dataset for their purposes.
5. visualize Geographic datasets through a map interface that will also allow users to inspect the details of individual data points (e.g. wells, temperature measurements, etc.) from properly formatted datasets.
6. provide a user interface for node-in-a-box data providers to publish data to their NGDS data repositories.

Note that the map UI is not intended for analysis and comparison of different data layers. It works only as a graphical way for searching data in the map and inspecting individual elements metadata.

5 NGDS Data Access Protocols

The OGC Catalog Service for the Web (CSW 2.0.2) will be used to enable catalog search via a Web API. Data services will be implemented using OGC WMS, WFS, NetCDF services, as well as other services adopted by the technical and steering committee as the system evolves. File-based resources will be accessed using standard HTTP GET requests.

6 Use Case Models

In general, the use cases for NGDS system are organized by user category as follows.

- Data Provider Use Cases
- End-user/Data Consumer Use cases
- System administration Use Cases
- Software Developer Use Cases

When certain tasks or workflows require multiple users to collaborate, they will be represented as workflow use cases, such as Quality Assurance Workflow.

6.1 High-level Use Cases View

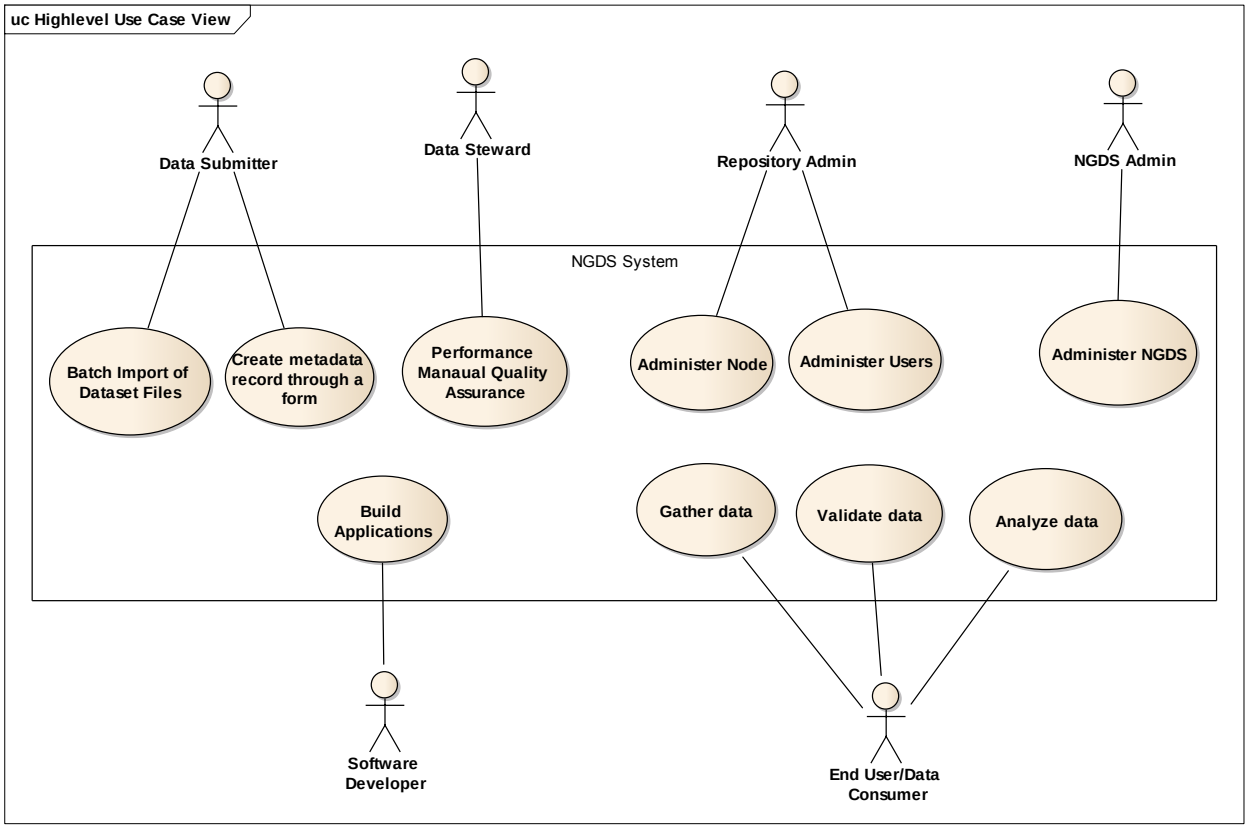


Figure 3: High-level Use Case View

6.2 Data Provider Use Cases

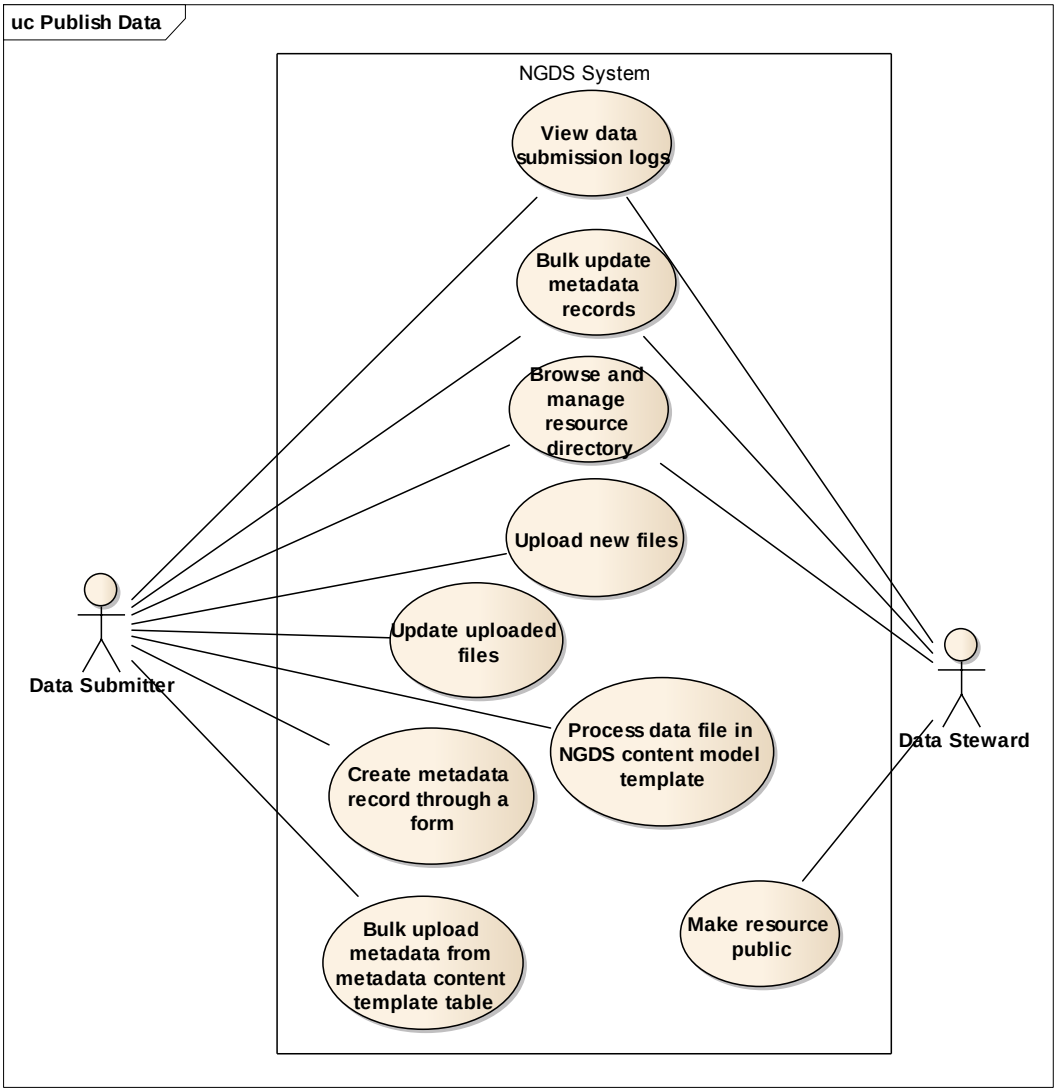


Figure 4 : Data Entry and Submission Use Case View

6.2.1 Data Submitter

6.2.1.1 File Uploads

Use Case ID	UC_001a
Use Case Name	*6.2.1.1.1 Upload new files*
Short Description	The goal of this use case is to allow data submitters to upload one or more files to be stored in the NGDS data repository. After the upload, the submitter will also update the metadata record of that file, thus allowing it to be cataloged. We assume the file is opaque, i.e. stored as is, with no further content parsing.
Actors	Data submitter
Pre-Conditions	Data submitter is properly authenticated;

Success End Conditions	The files are successfully uploaded and stored in the NGDS repository The metadata record for the provided file is successfully created The metadata remains "private", waiting to be made public by a data steward
Data	Metadata attributes for the specific data type as input to the form Files to be uploaded Geographical location of files
Functions	• Upload files • Form-based metadata input for specific data type • Auto-complete of user contact information • metadata validation • URI creation • Metadata duplicate detection • Tagging of metadata with geographical information • Converting non-standard location coordinates into latitude/longitude and shapes. • Log changes to metadata log file

Variants

Step	Actor	Description
7b	NGDS System	In case of duplicates, or incomplete information in the form, the system will provide a failure notification message, indicating the type of error. The user-provided metadata form will be presented to the Data Submitter for correction
7c	Data Submitter	Will correct the form data and resubmit for validation or will quit the import procedure.

Exceptions

Step	Actor	Description
	NGDS System	In case of internal file upload/metadata record creation failure, the system will roll back all existing transactions, returning to its previous state.

Open Issues (Please use this field to indicate questions/comments on the use case)

ID	Issue Description
1	Can't it be assumed that the submitter is working in cahoots with the data steward? Can the approval step be short circuited by assigning both submitter and steward roles to an individual? Yes, it is possible.
2	IN step 9, perhaps the data steward should be notified that there is a new record to be approved? This could be done by the NGDS system once the data submitter confirms the upload is completed?

Use Case ID	**UC_001b**
Use Case Name	***6.2.1.1.2 Update uploaded file***
Short Description	The goal of this use case is to allow data submitters to update a file that has been uploaded. This process includes the update of the metadata record of the file. We assume the file is opaque, i.e. stored as is, with no further content parsing.
Actors	Data submitter
Pre-Conditions	Data submitter is properly authenticated;
Success End Conditions	The file is successfully updated and new content stored in the NGDS repository The metadata record for the provided file is successfully updated The entry for the new file is also propagated & updated in the NGDS catalog.
Data	Possible new metadata record changes Files to be uploaded

Functions	• Upload files • update metadata • replace existing file. • Log changes to metadata log file
Variants	
Step	**Actor**	**Description**
4b	NGDS System	In case of duplicates, or incomplete information in the form, the system will provide a failure notification message, indicating the type of error. The user-provided metadata form will be presented to the Data Submitter for correction
4c	Data Submitter	Will correct the form data and resubmit for validation or will quit the import procedure.
Exceptions		
Step	**Actor**	**Description**
	NGDS System	In case of internal file upload/metadata record update failure, the system will roll back all existing transactions, returning to a valid state.
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	What happens to the old file? Is it deleted in the database or DOEs it remain there? Is a history of all metadata changes kept? DN: This is a good question and one that probably has no right answer for everyone. Does the new item get a new UUID or is it a revision (version X++) of the existing item? Earlier in the document there is a statement that says "all" node interactions are logged. This might require that the old version stay there. Probably best to ask domain experts.	

6.2.1.2 Data Processing

Use Case ID	**UC_006**
Use Case Name	***6.2.1.2.1 Process data file in NGDS content model template***
Short Description	This allows data submitter to process (upload/parse) a file to NGDS data provider nodes. The difference between the upload/update file use cases is that the file here is formatted according to an existing template, and therefore, can be parsed and checked for correctness. Processing will include validation of data schema, loading data into a data store on the provider node.
Actors	Data Submitter
Pre-Conditions	Requires authentication, access permission to edit metadata records. File must be formatted according to one of the NGDS content model templates (See /P02/ data specification from Arizona State University, for supported file formats and content models
Success End Conditions	File is submitted to a repository, being accessible through a valid URI
Data	Files properly formatted according to NGDS supported formats
Functions	• import data files • Validate data file content & formats • Log changes to metadata log file

Variants		
Step	**Actor**	**Description**
3b	Data Submitter	Is notified of suboptimal data content Chooses to submit the data anyways
3c	NGDS system	Accepts the file, flags the problems with content or metadata Hand file over to the data steward
Exceptions		
Step	**Actor**	**Description**
	NGDS System	In case of internal import failure, the system will roll back the existing transaction, returning to a valid previous state.
Open Issues (Please use this field to indicate questions/comments on the use case)		

6.2.1.3 Metadata Generation

Use Case ID	**UC_003**
Use Case Name	***6.2.1.3.1 Create metadata record through a form***
Short Description	The goal of this use case is to allow data submitters to create a metadata record describing a resource (tier1, tier2 data) by input of information manually through a form interface for inclusion in the NGDS catalog.
Actors	Data submitter
Pre-Conditions	Data submitter is properly authenticated; Data is available through one of the NGDS data repositories, and is identified through a valid URI. Metadata includes the geological location of the geological feature associated to the data.
Success End Conditions	The meta-data for the provided geological feature is successfully imported into the NGDS catalog The data remains “private”, waiting to be made public by a data steward
Data	Metadata attributes for the specific data type as input to the form
Functions	• Form-based metadata input for specific data type • metadata validation • URI validation • Metadata duplicate detection • Tagging of metadata with geographical information • Converting non-standard location coordinates into latitude/longitude and shapes. • Log changes to metadata log file
Variants	
Step	**Actor**	**Description**
7b	NGDS System	In case of duplicates, invalid URIs or incomplete information in the form, the system will provide a failure notification message, indicating the type of error. The original metadata form will be presented to the Data Submitter for correction
7c	Data Submitter	Will correct the form data and resubmit for validation or will quit the import procedure.

Exceptions		
Step	**Actor**	**Description**
	NGDS System	In case of internal import failure, the system will roll back all existing transactions, returning to a valid state.
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
2	In case of duplicates, which one is the most important the new instance of the data or the existing version of it? How to resolve these conflicts? DN: The domain experts said if records are overlapping, they would want both. If they are truly identical, it doesn't really matter does it (from a pure logic perspective anyways).	
3	What if the user DOEs not provide a geo-location to the data? DN: Given the UCD feedback, a geospatial reference is mandatory for every record for map based searching.	

Use Case ID	**UC_004a**
Use Case Name	***6.2.1.3.2 Bulk upload metadata from metadata content template table***
Short Description	The goal of this use case is to allow data submitters to import a CSV file containing metadata into the NGDS catalog. New records will be marked 'submitted'; Data Steward will then have option to review metadata through forms interface; records will become visible to public when marked 'published'. Metadata will be validated for content completeness, URLs checked for http 200 responses, and new metadata record will be run through a duplicate-detection process to identify existing metadata that may already describe the resource.
Actors	Data Submitter
Pre-Conditions	Data submitter is properly authenticated; Individual data items, referenced to in the CSV file are available through one of the NGDS data repositories, and is identified through a URI. Metadata includes the geological location of the feature associated to the data. CSV file containing metadata in table form according to the NGDS Compilation template Metadata Excel workbook
Success End Conditions	The meta-data for the provided geological feature is successfully imported into the NGDS catalog. The imported data remains private, waiting for the Data Steward to make it public.
Data	CSV file containing resources metadata, formatted according to the NGDS compilation template metadata excel workbook.
Functions	• process NGDS metadata compilation template csv file to produce individual metadata records • metadata validation • URI validation • Metadata duplicate detection • Converting non-standard location coordinates into latitude/longitude and shapes. • Log changes to metadata change log file

Variants		
Step	**Actor**	**Description**
11b	NGDS System	Rejects data based on invalid records
Exceptions		
Step	**Actor**	**Description**
	NGDS System	In case of internal import failure, the system will roll back the transaction, returning to a previous valid state.
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	SMR 2014-01-30 the bulk uploader does not use the NGDS metadata compilation template; the functions in the original 'requirements' document were copied from other metadata use cases, not specific to processing bulk metadata listing.	

Use Case ID	**UC_008**
Use Case Name	***6.2.1.3.3 Browse and manage resource directory and collections***
Short Description	Allows the User to view & manage all of the resources (or datasets) under their stewardship (data steward) or that they have submitted (Data submitter), based on metadata describing the resources. The resource listing will be presented to the User in a tree view directory structure. Users can define collections (folders, subdirectories) to organize the listing according their needs. A resource may be assigned to multiple collections. Access control may be assigned at the collection level. The display should indicate clearly any resources that have quality issue flags attached
Actors	Data submitter, Data Steward
Pre-Conditions	Requires authentication, access permission to view & edit individual user metadata records & datasets.
Success End Conditions	Users can adequately manage (create, rename, delete) resources Users can assign resources to collections (updating metadata) Users can delete existing resources (and their associated metadata) Metadata is kept in synchrony with changes in the collection
Data	Resources: metadata records or files
Functions	• Visualize flagged resources with quality issues • Create, delete, reorganize resource collections • delete resources • Visualize resources & collections
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	It seems to me that, in this view, the only operation that users can do with resources is delete. (Importing and modification of resources must be done in other views). Is it correct?
2	Missing requirements: tree-view portrayal of resource hierarchy. More clarity required as to who can create collections and sub-collections at what levels in the hierarchy. [Data steward assigns permissions to create public collections; perhaps allow authenticate users to define personal collections that they persist in a user workspace to use between sessions]. How do permissions work if metadata can belong to more than one collection? [Metadata record has only one owner; do we need a use case for reassigning ownership of a metadata record?] If a record is part of one collection and I have edit privileges to that collection, but it is also included in another collection to which I do not have privilege, what are my options? [Looks like assigning access control at collection level won't work; work access control

	through users and group membership, with permissions at group level to share edit capabilities]
	SMR 2014-01-30. Only part of this implemented is to see list of resources for a given account and to delete metadata (not tested--does uploaded file or deployed service get deleted?) and edit metadata. No collection/group functionality implemented. Add 'and collections' to title; that should be a separate use case but is bundled in here.

Use Case ID	**UC_005**	
Use Case Name	***6.2.1.3.4 Bulk update metadata records***	
Short Description	Metadata records may be selected based on a content-based filter query based on fields as contact information or linkage URIs. Once selected, those records can be replaced using regular expressions. User selects metadata content item to update, value to replace, and new value to use.	
Actors	Data submitter, Data steward, System administrator	
Pre-Conditions	Requires authentication, access permission to edit metadata records.	
Success End Conditions	All relevant metadata record attributes that match the search criteria are replaced with the new value defined by the user	
Data	Specific fields of metadata records: contact information and URIs	
Functions	• Search & replace of metadata record attributes based on simple regular expressions • Log changes to metadata log file	
Variants		
Exceptions		
Step	**Actor**	**Description**
	NGDS System	In case of internal import failure, the system will roll back all existing transactions, returning to a valid state.
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	How interactive should be the process of search & replace?	

6.2.2 Data Steward

6.2.2.1 Activity Logs

Use Case ID	**UC_007 & UC_046**
Use Case Name	***6.2.2.1.1 View resource submission and update logs***
Short Description	Allows authorized users to view the logs for metadata record creation, file uploads to file repository, and processing of NGDS files that conform to valid content model templates. These logs are created during any data submission or update activity.
Actors	Data submitter, Data Steward, System administrator
Pre-Conditions	Requires authentication, access permission view metadata records import logs. These logs are created during any data submission or update activity, so at least one data submission must have occurred.
Success End Conditions	Successful and correct visualization of data submission logs. No missing logs

Data	Data submission logs containing details such as time of activity, actions taken, data submitter, Data Steward, size of data, submitter comments, etc
Functions	• Submission log capture • Submission log visualization
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	Missing requirements detailing the logs and their creation. What exactly is logged? Also, how is the log portrayed to the user, or is that too much in the realm of design?
	SMR 2014-01-30 currently see list of create and update events via the dashboard when logged in; this is working in at least a rudimentary way.

6.2.2.2 Validation and Quality Assurance Logs

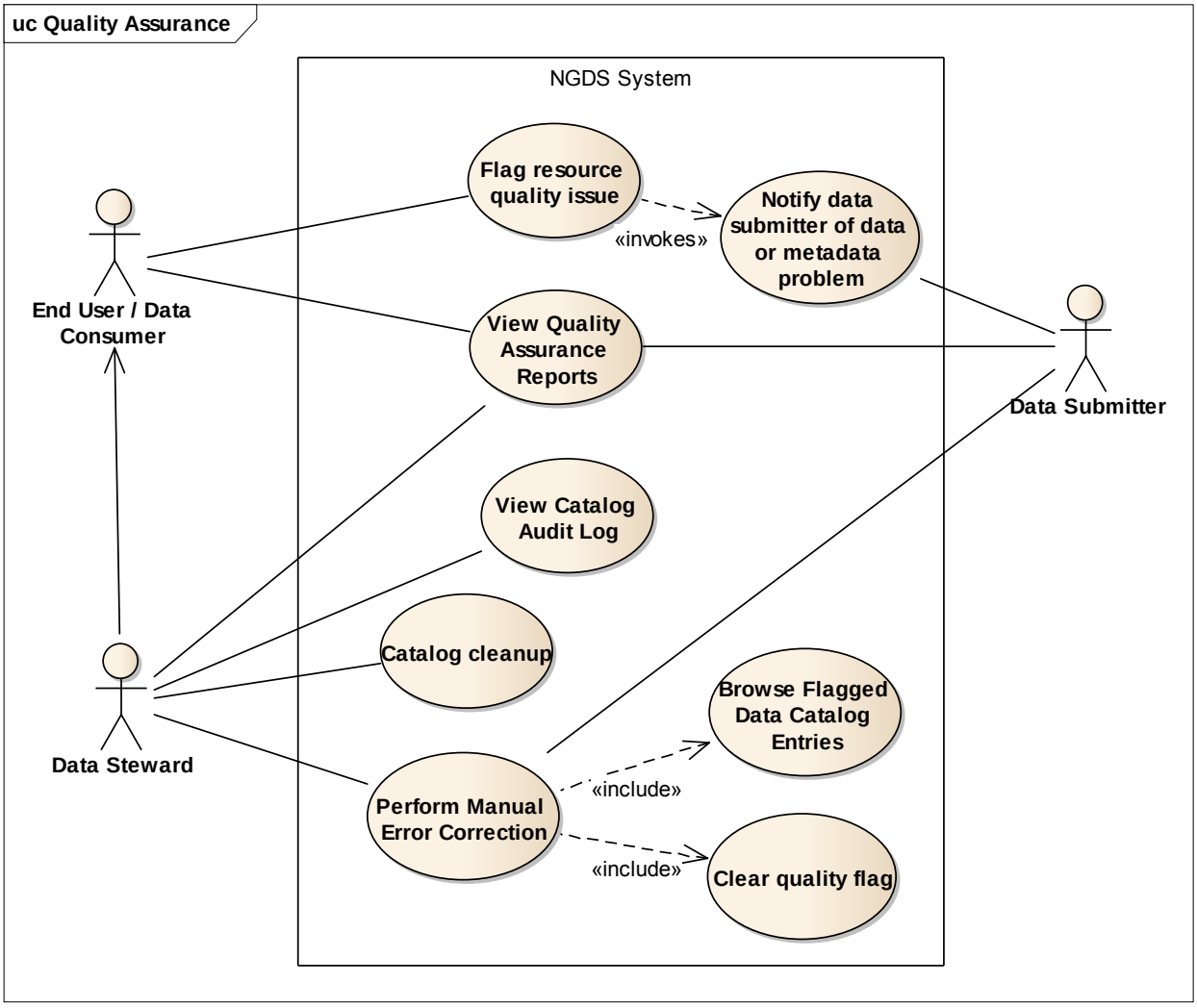


Figure 5 Quality Assurance Use Case View

Use Case ID	UC_046
Use Case Name	***6.2.2.2.1 View catalog audit log??***
Short Description	At every change the system creates a log for a given catalog item. This use case allows authorized users to view the log of all changes made to a data or metadata record, changes made both manually and automatically.
Actors	Administrator, Data Steward
Pre-Conditions	An audit log of catalog changes has been created and actively updated by the system
Success End Conditions	Users are able to visualize the change logs for a given record
Data	Catalog audit log
Functions	• View catalog audit log
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	Issue Description
1	Missing requirements detailing the validation and generation of the validation log. Here are two, but where are ones about URL checking? Syntax? Is the validation log attached to the activity log for a particular resource?
2	DN: I am also confused if this is happening at the node or NGDS system level? If another node change must be recorded in every other node, this could get quite messy fast (math suggests that it would be equal to ((N * (N^2 – 1) * R) * MR) where N is the number of nodes, R is the number of records and MR is the number of metadata records and that is assuming a 1:1 relationship between instances or R and MR.
	SMR 2014-01-30 I can't tell what this is supposed to be about, and how its different from **UC_007 & UC_046 (3.2.2.1.1)**

Use Case ID	UC_042
Use Case Name	***6.2.2.2.2 View quality assurance reports***
Short Description	Allows the Data submitter or Data Steward to view quality assurance reports for resources they submit or maintain. Quality assurance can be automatic or manual. As an automated process, it is performed by NGDs during import, flagging possible typos and simple input errors in the metadata being managed by the system As a manual process, it involves users that see possible issues and flag them in the data records.
Actors	Data Submitter, Data Steward, End User/Data Consumer
Pre-Conditions	The metadata record has been imported into NGDS catalog and a set of quality assurance functions were automatically ran in those records. As a result, the metadata records were flagged for possible errors. Users have write access to the data i.e. they are either submitters or stewards of that data
Success End Conditions	Users are able to identify the flagged errors in the metadata they provide or maintain
Data	Metadata records, Quality Report
Functions	• Visualize quality assurance report

Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	Should consumers be able to see QA reports? If they are not addressed, it seems like it would serve as a good "warning" to analysts. The Quality Report is for steward and submitter consumption, with lots of details that may not be relevant to End Users, e.g. wrong measures, duplicated fields. We assume the steward will fix them before making the data public. However, some stewards may want to make that data public anyway, in that case, consumers could benefit from this information.

Use Case ID	**UC_043**	
Use Case Name	***6.2.2.2.3 Flag resource quality issue***	
Short Description	Allows authorized & authenticated End-User/Data Consumer to create a flag, indicating that some issue exists with the quality of a data or metadata record. The process creates an annotation record documenting the dataset, URI for the resource in question, identity of the user raising the flag, timestamp. The annotation record should also include notes on resolution process, who, when, what. Flagged resources should be clearly marked	
Actors	Data Steward. End User/Data Consumer	
Pre-Conditions	The metadata record has been imported into NGDS catalog. Metadata QA records can be modified by any user in the system	
Success End Conditions	Users are able visualize metadata records, and to provide quality assurance feedback to these records as they see fit.	
Data	Metadata records (read only) Metadata QA (quality attribute) records (read/write) User information	
Functions	• Manually flag resource quality issues	
Variants		
Step	**Actor**	**Description**
1b	End User/Data Consumer	End user can flag data records directly on the metadata visualization page, without going to the quality assurance page.
Exceptions		
Open Issues (Please use this field to indicate questions/comments on the use case)		

Use Case ID	**UC_046c**
Use Case Name	***6.2.2.2.4 Browse flagged data catalog entries***
Short Description	Allows Users to navigate through the catalog of resources that have had their quality flagged for correction. A special view of the resource browse list, filtered for records that have quality flags raised from those that do not have them
Actors	Data Steward, Data Submitter
Pre-Conditions	a list of metadata resources is available, allowing filtering by quality attribute flag
Success End Conditions	Users are able to identify flagged entries in the metadata catalog
Data	Metadata record
Functions	• Browse data catalog entries

	• Filter catalog entries by flagged attribute
Main Sequence	
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

Use Case ID	**UC_044**
Use Case Name	***6.2.2.2.5 Perform manual error correction***
Short Description	Allows a Data submitter or Data Steward to manually address issues present in a data or metadata record that has been flagged for having quality issues. Note that this process may require resubmission of data files. Quality check and error editing will be built into the data submission process, but will require the steward to access the data in the appropriate environment for quality issues recognized after a submission is complete. The process will require the reviewer to make a note that is recorded as annotation on the record, along with the identity of the reviewer and a time stamp.
Actors	Data Submitter, Data Steward, Administrator
Pre-Conditions	The metadata record has been imported into NGDS catalog.
Success End Conditions	Users are able visualize metadata records, and to provide quality assurance feedback to these records as they see fit.
Data	Metadata records (read/write) Metadata quality attribute records (read/write) User information
Functions	• Edit metadata record • Edit data • Log metadata record change
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

Use Case ID	**UC_045**
Use Case Name	***6.2.2.2.6 Clear quality flag***
Short Description	Allows authorized User to clear a quality flag on a data or metadata record. This will occur when the User believes that the flagged issue is actually valid in the context of that record. The process will require the reviewer to make a note that is recorded with the flag, along with the identity of the reviewer and a time stamp.
Actors	Administrator, Data Steward
Pre-Conditions	Metadata records having quality flag attributes are available in the system The users have corrected the quality attribute issues, especially those that are automatically detected.
Success End Conditions	Users are able to remove data quality attribute flags
Data	Metadata quality attribute records (read/write)
Functions	• Remove quality issue flag • Log metadata record change

Variants		
Step	**Actor**	**Description**
3	NGDS System	If some quality attribute is found, the metadata record flag is turned back on and the user is notified of the issue.
Exceptions		
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	DN: Who has ultimate authority to determine the true data in the event of a dispute?	

Use Case ID	**UC_004b Catalog cleanup**
Use Case Name	***6.2.2.2.7 Catalog cleanup***
Short Description	The NGDS catalog will periodically and automatically verify the liveliness of the data sources referenced in its metadata catalog. That information will be used to mark the "unreachable", or "non-public" data items in the metadata catalog. The data steward can use that information to delete metadata records that are invalid, or to fix those records with valid URIs.
Actors	Data Steward
Pre-Conditions	Requires authentication, access permission to edit metadata records. The catalog has metadata records for which URIs that will be checked for liveliness There is also a predetermined threshold time period above which the metadata record will be marked as having "broken links" There is also a period of time parameter used by the NGDS catalog to periodically sweep the catalog, looking for those links
Success End Conditions	The NGDS catalog is cleared up from metadata records that have invalid URIs or those records have their URIs fixed.
Data	All metadata records of the NGDS catalog
Functions	• Automatic verification of broken links • Automatic verification of "private links" from third party repositories • Automatic tagging of broken metadata records • Deletion of catalog metadata records based on broken link attributes
Variants	
Exceptions	

Step	**Actor**	**Description**
	NGDS System	In case of internal import failure, the system will roll back all existing transactions, returning to a valid state.
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	Is there any undo capability? All deletions are final? DN: I would defer this question to the domain experts. IN some cases, I would imagine that knowing there once was a record might be useful if it can be tracked down via other means (or if they want to talk to the data submitter directly). Good question	
2	Before any data can be deleted or removed (potentially via the duplicate detection process) the node must ensure that there are no metadata records existing for it.	

Use Case ID	UC_004c	
Use Case Name	***6.2.2.2.8 Notify data submitter of data or metadata problem***	
Short Description	The goal of this use case is to allow the system to notify data submitters whenever other users other than the submitter herself, detect problems in the data or metadata.	
Actors	Data Submitter, Data Steward, End User	
Pre-Conditions	The metadata record has been imported into NGDS catalog.	
Success End Conditions	The data submitter receives e-mail notifications with quality issues detected by other users.	
Data	Metadata records (read/write) Metadata QA (quality attribute) records (read only) Data submitter information: e-mail	
Functions	• Send e-mail to data submitter with metadata and data issues	
Variants		
Exceptions		
Step	**Actor**	**Description**
	NGDS System	Bounding of e-mail may occur if the data submitter e-mail is invalid
Open Issues (Please use this field to indicate questions/comments on the use case)		

6.2.2.3 Resource Publication

Use Case ID	UC_009
Use Case Name	***6.2.2.3.4 Make resource public***
Short Description	This allows Data Steward to indicate that a resource is available for public discovery and access once they are satisfied with the data quality.
Actors	Data Steward
Pre-Conditions	Metadata record is properly created in the catalog, waiting to be made public by a data steward Data is properly uploaded to a data repository, waiting to me made public.
Success End Conditions	The metadata record is publicized AND the data is made available through a public URI.
Data	Metadata records in the catalog or files in the repository
Functions	• Make data public though a URI • Make metadata record public

Variants
Exceptions
Open Issues (Please use this field to indicate questions/comments on the use case)

6.2.3 Node Administrator

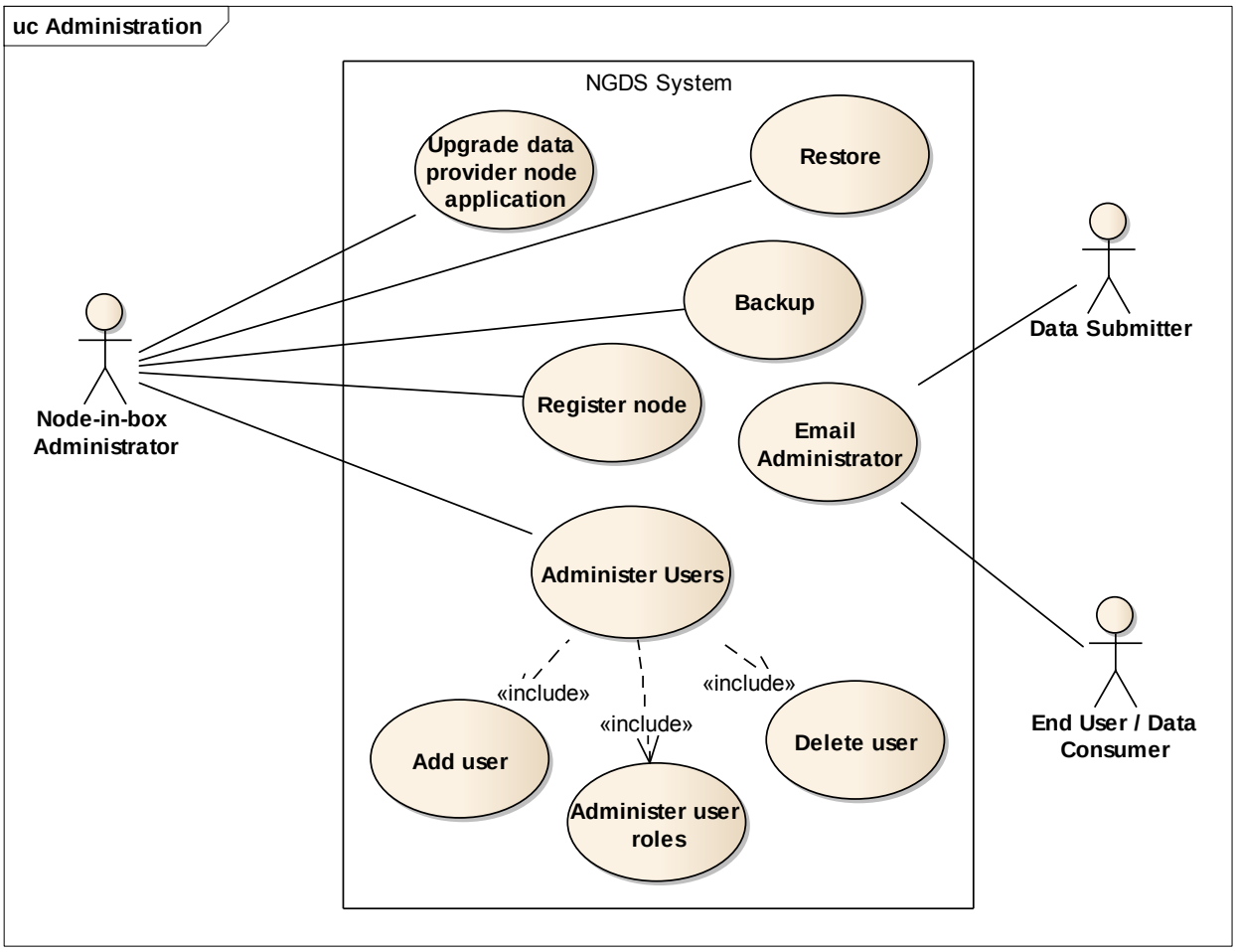


Figure 6 Administration Use Case View

6.2.3.1 User Account Management

Use Case ID	UC_029a
Use Case Name	*6.2.3.1.1 Administer users*
Short Description	Allows the system administrator to manage data provider node users. This will allow the system administrator to add and remove users on the administered node, and assign user roles and group membership.
Actors	GIN-stack administrator
Pre-Conditions	GIN-stack is properly installed and configured
Success End Conditions	The administrator is able to perform the main administration operations
Data	User records

Functions	• Add user • Delete user • Modify user permissions and roles
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	Issue Description
1	DN: Within this requirement, it will be necessary to perform a check to ensure that a user being deleted does not have custody of any records otherwise we may end up with orphaned records (unless they default to the node administrator). Just a though that came to mind.
2	DN: A user should only be deleted once all their records (data and metadata) is transferred to another user. Otherwise there will be orphaned data and metadata.
	SMR 2014-01-30 can't add or remove users from user Management page, only change roles; no way to do stuff like reset passwords.

Use Case ID	**UC_029b**
Use Case Name	***6.2.3.1.2 Enrollment***
Short Description	Allow users to self enroll. i.e. to create their account in the NGDS system for the purpose of supporting NGDS data import/export and exploration activities.
Actors	End User/Data Consumer, Data Submitter
Pre-Conditions	GIN-stack is properly installed and configured
Success End Conditions	The users have their accounts created, and their user data and credentials accessible throughout the system
Data	User records
Functions	• Enroll user • Un-enroll user
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

Step	Actor	Description
1b	User	Navigates to the enrollment screen of the system Requests deletion of her profile -> un-enrolment
2b	NGDS System	Responds by: Deleting the user record and its credentials Includes <<delete user>> use case

ID	Issue Description
1	Do we really want to allow users to self enroll as data providers? DN: My answer would be yes since the system has been designed with the data steward as a failsafe against publishing bad data. The less human involvement the better.
2	DN: Same caveat here WRT orphaned data and metadata. Do not allow a user to remove themselves if they have data submitted until the ownership of the data is re-assigned or the data is removed.
	SMR 2014-01-30 user who creates self is not assigned a role; admin should get notification of new user.

Use Case ID		**UC_030**
Use Case Name		***6.2.3.1.3 Add users***
Short Description		Creates a new user account in the system, allowing the storage of important user information such as e-mail, name, login, password, address, and enabling features such as subscription and search saving, and auto complete.
Actors		GIN-stack administrator, Users
Pre-Conditions		GIN-stack is properly installed and configured
Success End Conditions		A new user account is created
Data		User records
Functions		• Add user • Add users via invitation
Variants		
Step	**Actor**	**Description**
1b	Administrator	Navigates to user management screen Starts new user creation Inputs user information Send invitation to user
2b	NGDS System	Responds by checking for repeated user credentials and if positive, send invitation to user
3	User	Responds by accepting or rejecting invitation
4	NGDS System	Creates user account or does nothing if the invitation was rejected
Exceptions		
Open Issues (Please use this field to indicate questions/comments on the use case)		

Use Case ID		**UC_031**
Use Case Name		***6.2.3.1.4 Delete users***
Short Description		Removes a user, its credentials and associated information from the system
Actors		GIN-stack administrator
Pre-Conditions		GIN-stack is properly installed and configured
Success End Conditions		A new user account is deleted, together with its data including subscriptions, saved searches, history, etc.
Data		User records
Functions		• remove user
Variants		
Exceptions		
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	What happens if a user that is a data steward is deleted? Are her metadata records unpublished, assigned to a system pseudo user?	
2	DN: QUESTION: Should the system remove the associated information or should certain info persist (logs, metadata etc)?	

Use Case ID	**UC_032**
Use Case Name	***6.2.3.1.5 Administer user roles***
Short Description	The goal of this use case is to allow the administrator to assign different roles to users. These roles control the users' abilities to publish data in the system or to administer system functions.
Actors	GIN-stack administrator
Pre-Conditions	GIN-stack is properly installed and configured
Success End Conditions	User role assignment is updated according to administrator needs
Data	User records
Functions	• assign role to user • remove use role
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	**Issue Description**
1	Are users federated across nodes? RSSF: Stewards and submitters are local users, end user/consumers are global users.
2	DN: DERIVED REQUIREMENT: Every Node must have at least one administrator, one steward correct?

Use Case ID	**UC_032b**
Use Case Name	***6.2.3.1.6 Backup***
Short Description	The goal of this use case is to allow node administrators to backup the data and meta-data being stored in the system.
Actors	GIN-stack administrator
Pre-Conditions	GIN-stack is properly installed and configured
Success End Conditions	The data, metadata and indexes of a node-in-a-box is successfully backed up
Data	User records Metadata records Data indexes
Functions	• manual backup NGDS node • automatic backup of NGDS node
Variants	

Step	**Actor**	**Description**
1	Administrator	Navigates to user management screen Selects backup option Set ups automatic backup option by providing a destination and a day of week/time
2	NGDS System	Responds by backing up node data to the assigned repository at the given day of week/time

Exceptions
Open Issues (Please use this field to indicate questions/comments on the use case)

Use Case ID	UC_032b
Use Case Name	*6.2.3.1.7 Restore*
Short Description	The goal of this use case is to allow node administrators to restore the backed up data and meta-data being stored in the system.
Actors	GIN-stack administrator
Pre-Conditions	GIN-stack is properly installed and configured
Success End Conditions	The data, metadata and indexes of a node-in-a-box is successfully backed up
Data	Backed up blob with important node data
Functions	• restore NGDS node
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

6.2.3.2 System Management (Meet NGDS System Requirements)

Use Case ID	UC_034
Use Case Name	*6.2.3.2.4 Register node*
Short Description	The goal of this use case is to allow new nodes to be added to the NGDS grid. These nodes can be of different types included “GIN-stack” installations, or third party data provider repositories.
Actors	GIN-stack Administrator, NGDS Administrator
Pre-Conditions	GIN-stack properly installed as a NGDS node management hub Client node-in the box properly installed but not yet registered
Success End Conditions	A new node is registered in the NGDS network, and the data it provides becomes available to the other nodes, and searchable through the by the NGDS catalog
Data	e-mails, node credentials
Functions	• add new node • index new node
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	Issue Description
1	Who manages the grid? Is the grid configured centrally? NGDS Administrator in a special user in a node elected to be the main node.
	SMR 2014-01-30 Any aggregator may register any publisher, so the network is defined by which aggregators register which publishers. What is missing is any kind of 'self' declaration manifest by a node that would allow it to be discoverable. See 'How to be an NGDS node' design document. We haven't implemented a process as part of

	deploying a node that would request the manager of the new node to contact one or more aggregators to request harvesting, or any kind of 'registry' of known nodes that can be harvested from or are actively aggregating.

6.2.3.3 Routine Maintenance

The node administrator must also perform routine maintenance tasks pertaining to their node. These include upgrading software and responding to user questions and requests.

Use Case ID	**UC_033**
Use Case Name	***1.1.1.1.1 e-mail (NGDS or GIN-stack) administrator***
Short Description	The goal of this use case is to allow users to send e-mail to administrators of the system to handle matters such as granting of especial access rights, or to register new nodes in the NGDS network, or other issues. There are two types of administrators: node-in-a-box administrators, and NGDS administrators.
Actors	All users, Administrators
Pre-Conditions	GIN-stack is properly installed and configured Administrator has registered her e-mail information
Success End Conditions	Users can communicate their needs/issues with the node-in-a-box administrator
Data	e-mails
Functions	• send email to administrator
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
	SMR 2014-01-30 I don't see any 'contact administrator' path from the current publisher node or aggregator node--just stuff for generic contact NGDS. Contact link in LL of page doesn't seem to work, and looks like it will send to ngdsweb@geothermaldata.org.

Use Case ID	**UC_035**
Use Case Name	***6.2.3.3.4 Upgrade GIN-stack application***
Short Description	The goal of this use case is to upgrade the software that implements the GIN-stack application. This can potentially include the data repository, the catalog service, and the WebApp application
Actors	Node-in-a-box administrator
Pre-Conditions	GIN-stack is properly installed
Success End Conditions	The software components(s) of the node are updated without data loss and will minimum impact on other nodes of the NGDS network
Data	Possibly all the data stored in the data repository and index
Functions	• update software components • shutdown node • restart node

Variants
Exceptions
Open Issues (Please use this field to indicate questions/comments on the use case)

6.3 Software Developer Use Cases

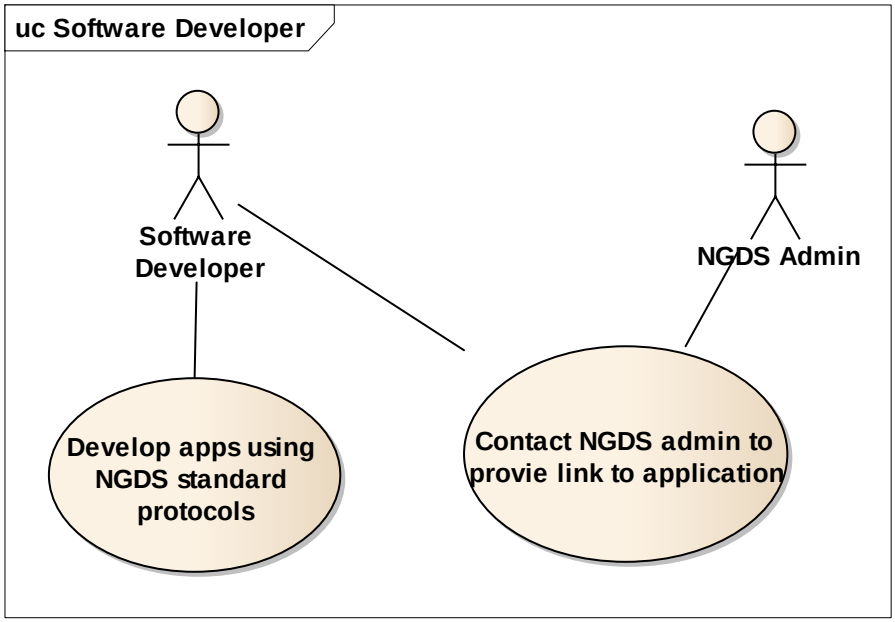


Figure 7 Software developer use cases

6.3.1 Develop Applications

Use Case ID	UC_040
Use Case Name	*6.3.1.1.1 Develop apps using NGDS standard protocols (needs reformulating)*
Short Description	Through the use of web protocols, in particular: CSW, WFS and HTTP, software developers can build applications that utilize the data and meta-data stored in NGDS. They can also use NGDS to locate referenced data (data that is not stored in NGDS components but is referenced by the NGDS catalog). They can, for example, develop applications that display that information on maps, perform data analysis and discover geological information of geothermal sites within US
Actors	Software developer
Pre-Conditions	NGDS data repositories and catalogs are available and accessible through standard internet protocols. Software developers may need an account to interact with the system if their application involves the updating of information in repositories and catalog, or if they utilize services such as posting of reviews.
Success End Conditions	Software developers can successfully build applications that utilize the NGDS system resources
Data	Data models, metadata schemas, key system URIs, e.g. catalog service main URI, protocol and data model documentation

Functions	• Standard protocols • Ability to interact with NGDS via APIs • Logging of API usage statistics and accesses
Main Sequence	
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	DN: It might be a good idea to issue API keys to control the number of queries and use of data or at least have a metric to understand who is using it and for what. DN: Discussion about developer API keys. Not critical but something to keep in mind. Given there are not concrete plans on how to pay for maintenance after the system is up and running, providing unlimited API access might not be something that can be afforded.
	SMR 2014-01-30 pretty marginal here-- there is a CSW but hasn't been tested much, also the CKAN API is available, but there isn't much to help the user know this… this use case is not really actionable. Need more specifics on the API functionality that needs to be exposed.

6.3.2 Register New Application with NGDS

Furthermore, the software developer will wish to register their new application with the NGDS in order to promote its use throughout the system.

Use Case ID	**UC_041**
Use Case Name	***6.3.2.1.1 Contact NGDS admin to provide link to application***
Short Description	The goal of this use case is to capture the need for a way to software developers request the inclusion of application links in the NGDS web site. They will do so via a public e-mail address or via a form that allow them to submit this information.
Actors	Software developer, NGDS administrator
Pre-Conditions	NGDS website is available and its URI is know by the software developer
Success End Conditions	NGDS software developers can provide a link to their applications and this information is incorporated in the list of applications hosted in the NGDS website
Data	URIs from software developers e-mail or URI to website location where developers can post those requests
Functions	• Publish new application links via the NGDS system

Variants
Exceptions
Open Issues (Please use this field to indicate questions/comments on the use case)

6.4 End-User Use Cases

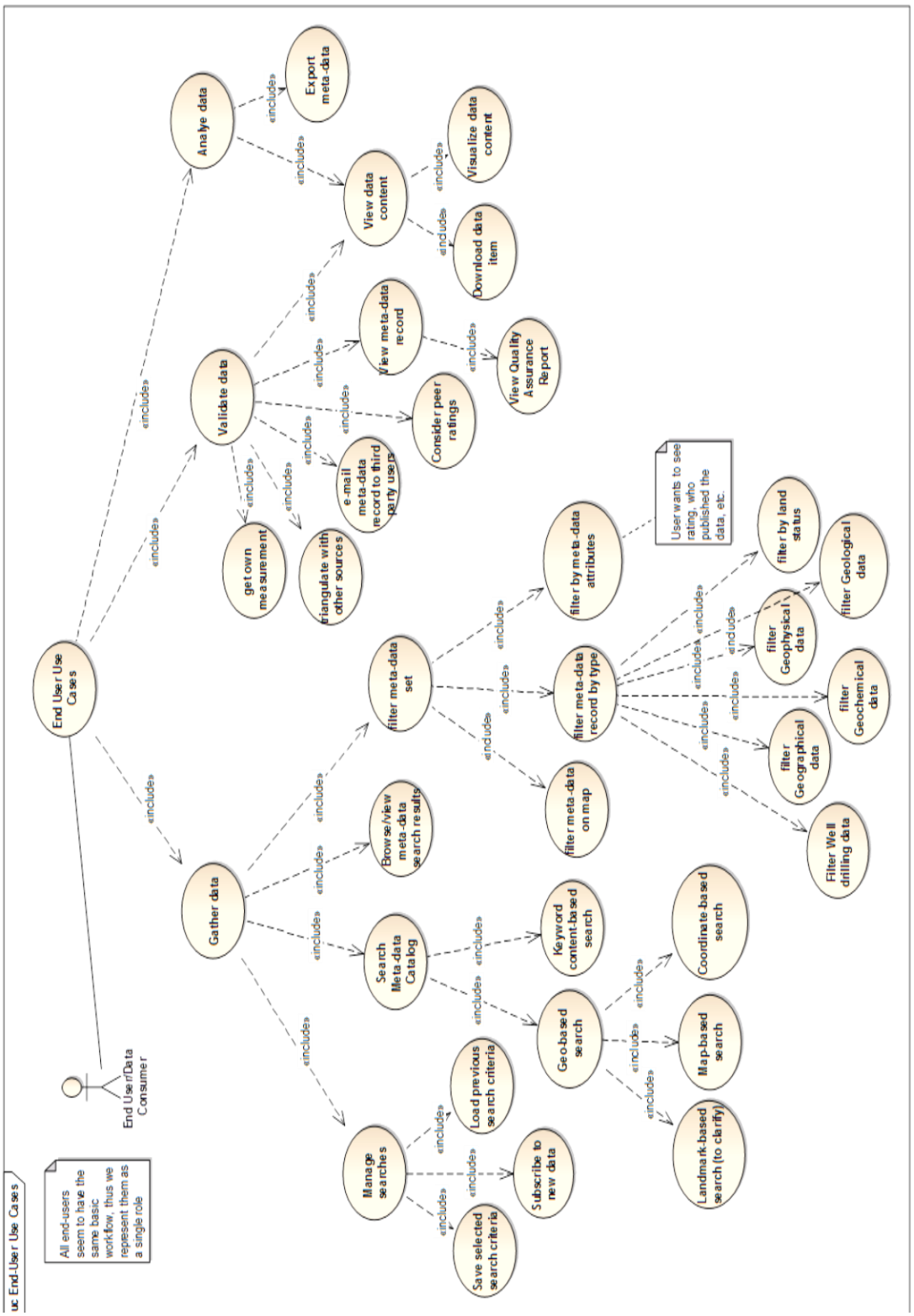


Figure 8 End-user use cases overview

6.4.1 Discover and Gather Data

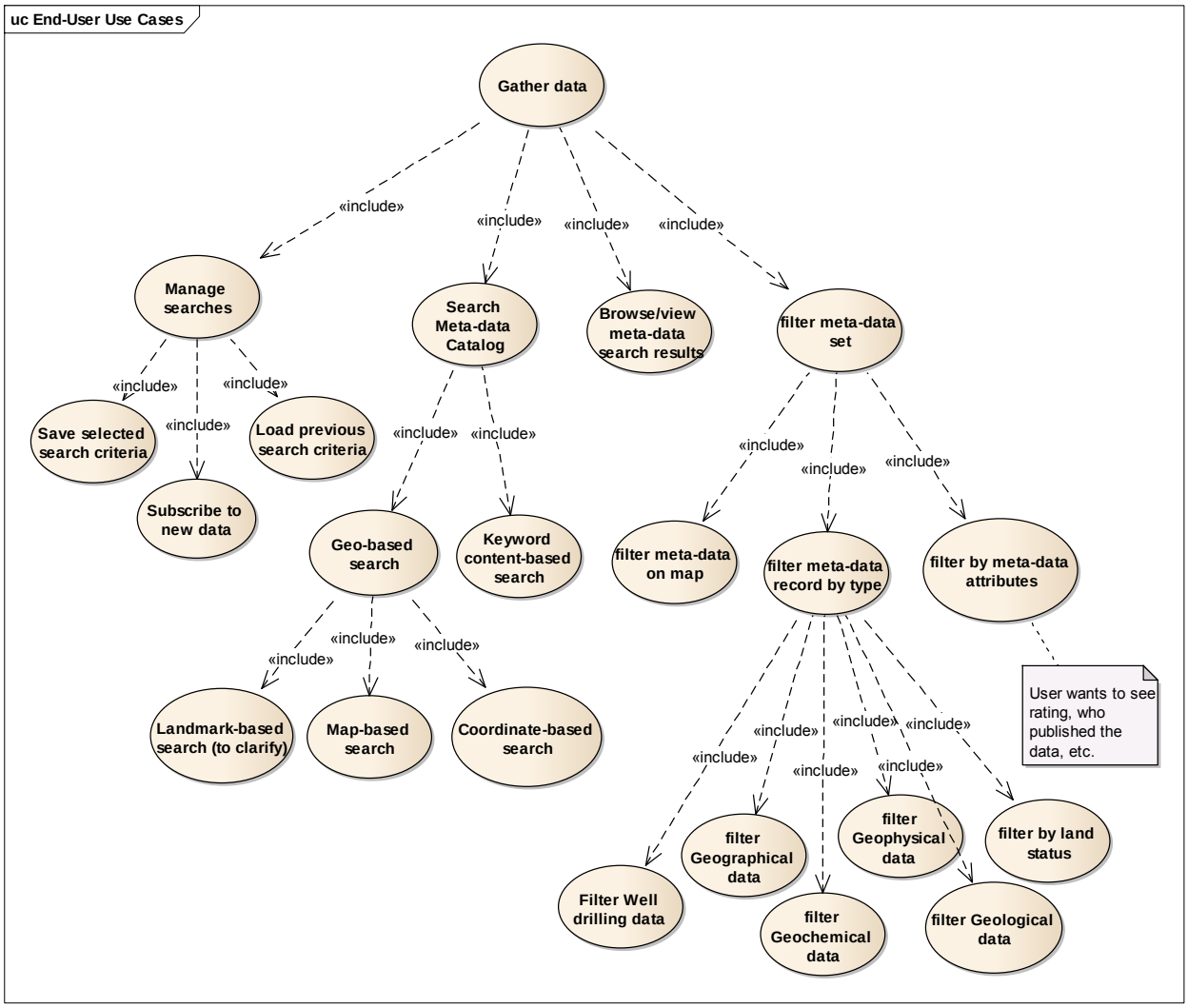


Figure 9 Data gathering supporting use cases

6.4.1.1 Map-Based Search

Use Case ID	UC_014
Use Case Name	***6.4.1.1.1 Map-based search***
Short Description	The goal of this use case is to support users discovery of metadata by using maps. Map-based search consist in zooming, panning and selecting a region in a map. This search method can be used together with other search methods, to narrow down the data in a search.
Actors	End User/Data Consumer
Pre-Conditions	System is correctly installed and operational
Success End Conditions	The user can visualize a result of a search as layers in a map.
Data	Metadata features shown as layers (WFS protocol), metadata summary of selected data items on map (when users click on a data point in the map)

Functions		• Map navigation capability: panning, zooming, selecting sub-regions • Retrieve metadata record for selected elements in a map
Variants		
Step	**Actor**	**Description**
1a		Include other types of search as landmark-based, coordinate-based, content-based, etc.
Exceptions		
Step	**Actor**	**Description**
	NGDS System	In case of failure to display layers due to network errors, for example, the system should notify user of the situation
Open Issues (Please use this field to indicate questions/comments on the use case)		

Use Case ID	**UC_013**
Use Case Name	***6.4.1.1.2 Landmark-based search***
Short Description	The goal of this use case is to allow users to utilize landmarks (state, city, county, district, known geothermal sites) to narrow down the search in the catalog. This search method can be used together with other search methods, to narrow down the data in a search.
Actors	End User/ Data Consumer
Pre-Conditions	A database of landmarks and their geo-locations must be available for search
Success End Conditions	The user finds information based on landmarks
Data	Landmark name
Functions	• Find landmark • Retrieve landmark geo-location
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	As noted by one of the monitors this needs to be clarified with Antro-tech. what is a landmark? What are examples of landmarks? How difficult would it be to support landmark-based search in our approach? Can we reuse functionality from search engines as Google map services to do this? DN: Agree. I think pushing that off to the map vendor is a good approach as funding is limited within NGDS. Still, a request from Anthro-tech for clarification would be good.
	SMR 2014-01-30 doesn't seem to work

Use Case ID	**UC_015**
Use Case Name	***6.4.1.1.3 Coordinate-based search***
Short Description	The goal of this use case is to allow users to utilize geographical coordinates to narrow down the search in the catalog.
Actors	End User/Data Consumer
Pre-Conditions	
Success End Conditions	The user finds information based on geographical coordinates
Data	Geographical coordinates

Functions	• Validate coordinates • Retrieve data from the system based on proximity or containment within geographical coordinates
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	As noted by one of the monitors this needs to be clarified with Antro-tech. do the users need to type coordinates? Is there a UI-based way to support this search without requiring users to type those coordinates, for example, by drawing a box in a map? I think the onus should be on the end user to figure out where the landmark is and locate it on the map. It is very unclear what the landmark means.

6.4.1.2 Keyword-Based Search

Use Case ID	**UC_016**
Use Case Name	***6.4.1.2.1 Keyword content-based search***
Short Description	The goal of this use case is to allow users to search data by its metadata content registered in the catalog. If data comes in tier 3 format, it includes its indexed content; if it comes in Tier1 and Tier2 formats, the search is based on whatever could be extracted/converted/indexed into data or meta-data.
Actors	End User/Data Consumer
Pre-Conditions	There is metadata in the catalog
Success End Conditions	The user finds information based on metadata content
Data	All metadata stored in the NGDS catalog
Functions	• Content-based search of metadata records • Content-based search of data content
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	DN: There may be a need for a basic thesaurus but hard to make with a limited budget.
	SMR 2014-01-30: uploaded content is not indexed. Search is entirely based on metadata content, which includes tagging by data providers. Resources are not consistently keyword from a controlled vocabulary; such a vocabulary was developed fall 2013, and tags in State geothermal data were mapped to the keywords and metaedata updated to begin move to formal keyword, but it is incomplete and needs additional curation work.

6.4.1.3 Refining Results and Faceted Search

Use Case ID	**UC_019**
Use Case Name	***6.4.1.3.1 Filter results by type***
Short Description	The user can also narrow down its search results by specifying certain data types of interest, thus ruling out all other data that do not belong to these types from the returned list of search results.
Actors	End User/Data Consumer

Pre-Conditions	A search was performed and a subset of the metadata from the catalog was retrieved by the NGDS catalog. The search result is displayed as a list of metadata records
Success End Conditions	The user can narrow down the search results
Data	A subset of metadata obtained by a search A list of metadata types present in the subset of metadata under consideration
Functions	• Filter metadata set by type
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
	SMR 2014-01-30: works best on user defined tag processing built into CKAN; doesn't work for category/facet based filters, generally doesn't return results.

Use Case ID	**UC_020**
Use Case Name	***6.4.1.3.2 Filter results by metadata attributes***
Short Description	Different content models prescribe different attributes to different types of data, these attributes can be used to further refine the search result, for example, excluding data points for which their metadata record do not have certain attribute content values.
Actors	End User/Data Consumer
Pre-Conditions	A search was performed and a subset of the metadata from the catalog was retrieved by the NGDS catalog. The search result is displayed as points in a map and as a list of metadata records
Success End Conditions	The user can narrow down the search results
Data	A subset of metadata obtained by a search A list of metadata types present in the subset of metadata under consideration
Functions	• Filter metadata set by attribute content
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	Missing Filters: by location terms, geothermal thematic terms, source organization, publication date, popularity and user ratings. Requirements need to clearly indicate that these vocabularies will exist on which such faceted filtering can be performed. ND: This should be a requirement for the metadata team
	SMR 2014-01-30 this only works if the user knows the index field names and the Lucene syntax to use.. There is no 'refine search' function that allows fielded or free text filter restrictions.

Use Case ID	**UC_018**
Use Case Name	***6.4.1.3.3 Filter results [geographically] on map***
Short Description	Once a search is made and search results is displayed as points on a map, the user can narrow down its search by selecting a sub-area in the map, thus filtering out all data points that are outside that geographical region.
Actors	End User/Data Consumer

Pre-Conditions	A search was performed and a subset of the metadata from the catalog was retrieved by the NGDS catalog. The search result is displayed as points in a map
Success End Conditions	The user can narrow down the search results
Data	A subset of metadata obtained by a search
Functions	• Select region in a map • Filter search results based on a bounding box on a map
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

6.4.2 Validate and Evaluate Data

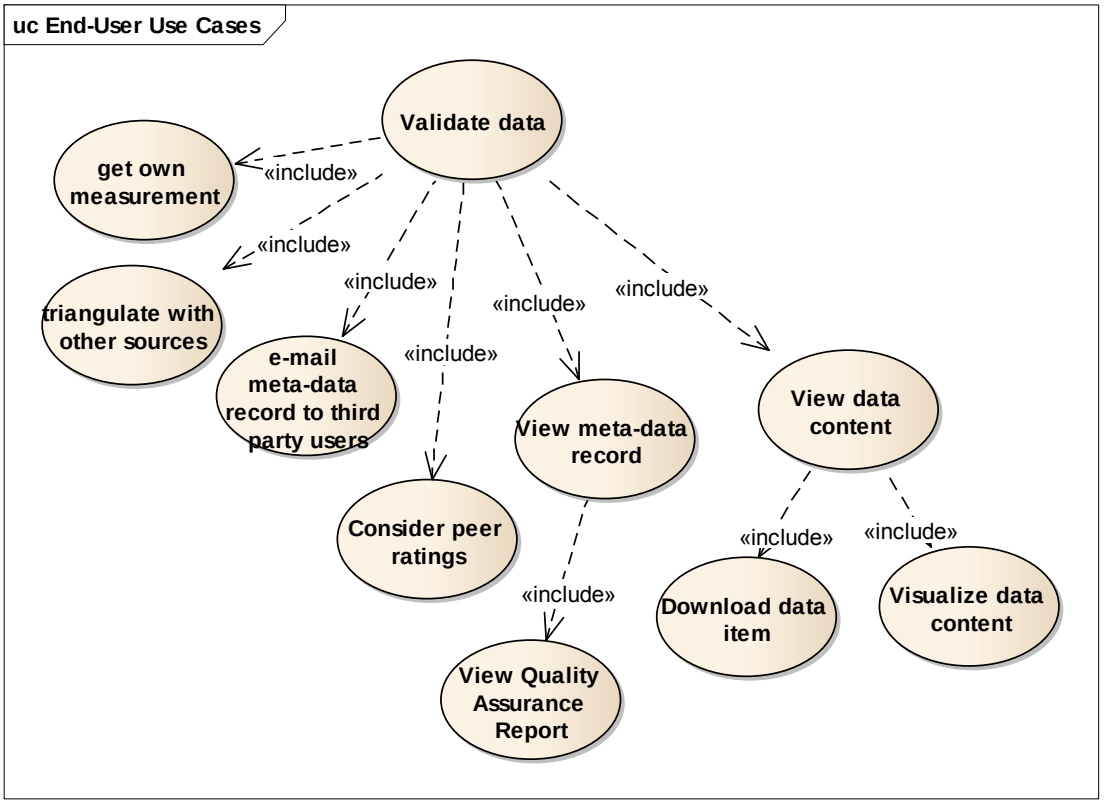


Figure 10 Data validation supporting use cases

Use Case ID	**UC_017**
Use Case Name	***6.4.2.1.1 Browse/view metadata search results***
Short Description	The goal of this use case is to allow users to visualize the results of a search and inspect its content. This visualization is supported by metadata lists and maps.
Actors	End User/Data Consumer

Pre-Conditions	A search was performed and a subset of the metadata from the catalog was retrieved by the NGDS catalog.
Success End Conditions	The user can find what he/she is looking for
Data	A subset of metadata obtained by a search
Functions	• Browse search results • Inspect elements in a map ??
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
	SMR 2014-01-30 see center points of bounding boxes, clidking on result shows BB for that result. Can display WMS for data offered through services, and CKAN provides some simple graphing function for CSV files. This use case is too poorly described to be actionable.

6.4.2.2 Metadata Evaluation

Use Case ID	**UC_026**	
Use Case Name	***6.4.2.2.1 View metadata record***	
Short Description	After locating a piece of metadata in the catalog, the user inspects the metadata record in more detail, for example, to decide if it refers to the data she is looking for.	
Actors	End User/Data Consumer	
Pre-Conditions	A search was performed and a metadata record was selected for further inspection	
Success End Conditions	The user is able to access and view the contents of the metadata	
Data	Documents and structured records stored in GTDA repository or third party repositories. The data is located through a URI	
Functions	• Retrieve metadata record • Visualize metadata record	
Variants		
Exceptions		
Step	**Actor**	**Description**
	User	In case the URI pointed by the metadata record becomes unavailable during the execution of this procedure, the system must provide an error message. The metadata record may be marked as invalid.
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	Missing requirements for other “views” of metadata, for example as human-readable HTML, or as an ATOM entry. These are just low-hanging fruit.	
	SMR 2014-01-30 display of metadata is very poor presentation; the HTML needs A LOT of work	

Use Case ID	**UC_024**
Use Case Name	***6.4.2.2.2 Provide peer ratings***
Short Description	By inspecting data reviews posted by other users, a user can gauge the accuracy and validity of data. Peer ratings can include textual description, star ratings or both. This UC captures the fact that the system must provide support for peer ratings. As part of this use case, users can also post peer ratings. There is no restriction of who can

		post those ratings. The user log-in information is used to identify the peer rating poster
Actors		End User/Data Consumer
Pre-Conditions		The existence of ratings posted by other peers for the case of review None if the user will be the first to post a rate
Success End Conditions		The user is able to view peer ratings text and star ratings in the metadata posted in the NGDS catalog.
Data		Peer ratings and their textual description that are attached to metadata in features of the map and the search results list
Functions		• Provide peer ratings star rating and text for a given metadata record • Post a peer rating • Store peer reviews
Variants		
Step	**Actor**	**Description**
2b	User	Include use case <<view metadata content>>, <<view document content>> Post a peer rating
Exceptions		
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	This feature leads me to suggest that optional user profile fields indicating who a reviewer is (e.g Steve Richard of AZGS rates this 3 stars)	

6.4.2.3 Data Comparison

Use Case ID		**UC_022**
Use Case Name		***6.4.2.3.1 Triangulate with other sources***
Short Description		Users may compare the metadata returned by the system with external data layers, provided by external WFS sources such as demographics, topological, weather, and so on, thus helping them to make inference on the quality of the information. These use cases can also be performed by third party applications that consume the data published in NGDS
Actors		End User/Data Consumer
Pre-Conditions		A search was performed and a subset of the metadata from the catalog was retrieved by the NGDS catalog.
Success End Conditions		The user compares the returned metadata with third party data sources and comes to a conclusion
Data		External data sources as WFS providers for map layers Metadata from the system
Functions		• Download data • Export data via standard protocols to third party applications
5	User	Compares the data manually with their own, or with the help of a CAD tool.

Variants		
Step	**Actor**	**Description**
3b	User	Uses CAD tool, via WFS to read NGDS repository data
4b	NGDS System	Exports data via WFS for third party CAD tool.
Exceptions		
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	It is not clear if the scope of the system includes the mashing up of information with external data sources. This seems to be an advanced feature that may be better achieved by utilizing third party CAD tools. DN: Agree!	
	RC: The general idea is that you compare the data in a particular dataset to some other known data. This should be fleshed out as requirements for accessibility and download of data.	
	MM: There is an opportunity here to enable crowd sourced information about resources via this use case.	

6.4.2.4 Storing and Sharing Search Results

Use Case ID	**UC_025**	
Use Case Name	***6.4.2.4.1 View data content***	
Short Description	After locating a piece of data, the user inspects the data content by URI element that the metadata refers to.	
Actors	End User/Data Consumer	
Pre-Conditions	A search was performed and a metadata record was selected for further inspection	
Success End Conditions	The user is able to access the data pointed by the metadata record	
Data	Documents and structured records stored in NGDS repository or third party repositories. The data is located through a URI	
Functions	• Retrieve URI document • Open and display document content to user	
Variants		
Exceptions		
Step	**Actor**	**Description**
	User	In case the URI pointed by the metadata record becomes unavailable during the execution of this procedure, the system must provide an error message. The metadata record may be marked as invalid.
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	WMS can be shown on map. WFS, CSV data can be displayed in tabular view. This provides rudimentary data browse/evaluate capability.DN: Long URLs can be aliased via a URL shortener. This may be better as it does not require the system to preserve the state of a specific search yet allows the search to be shared. A discussion to have. Rather than save the search criteria, maybe it should just provide a URL encoded string that can be used to represent the same state. This is less expensive from a systems perspective. Example - https://www.google.ca/search?q=NGDS&oq=ngds &sugexp= chrome, mod=0 &sourceid=chrome&ie=UTF-8	

Use Case ID	**UC_010**	
Use Case Name	***6.4.2.4.2 Save selected search criteria***	
Short Description	The goal of this use case is to allow users to save searches, to be reused in a later time, and for setting up subscriptions to content changes. In this use case, after the user performs a search, she saves that search parameters for further use. This search then can be used to subscribe to new data, and to continue a previous discovery activity. Searches are saved on the end-user accounts, for their private use. In the future they may be shared among other users. When saving a search, users can opt to make search public so others can reuse it.	
Actors	End User/Data Consumer	
Pre-Conditions	In order to allow saving and retrieval, the user must be identifiable; hence, there is a need for users (in particular the end user) to be logged in using their unique account.	
Success End Conditions	The search criteria is properly validated and saved into the system under a given name.	
Data	Search criteria	
Functions	• Save search criteria • Validate search criteria • Record search parameters by monitoring user input • Make search public to other users	
Main Sequence		
Variants		
Step	**Actor**	**Description**
1b	Users	Input search criteria in a separate content-based search form, instead of using the included use cases
2b	NGDS System	Uses form-based search as criteria
Exceptions		
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	Should the system support form-based search only or should it record data as shown above? Which option is better? DN: I would adhere to the architectural principles known as REST. A URI represents the state of a specific resource etc.	
2	Should the saved searches from one user be visible to other users as Antro-tech indicated that searching is a collaborative process? Maybe searches are not saved under a use id but treated as a resource. They could then be saved under the resource name (URI) and the (anonymous) user could create a bookmark to access the search later. This bookmark/URL can also be shared with other users. DN: Agree.	

Use Case ID	**UC_012**
Use Case Name	***6.4.2.4.3 Load previous search criteria***
Short Description	The goal of this use case is to support users in loading previously saved search criteria. They do so by browsing through their list of saved searches.
Actors	End User/Data Consumer
Pre-Conditions	In order to allow saving and retrieval, the user must be identifiable; hence, there is a need for users (in particular the end user) to be logged in using their unique account.

Success End Conditions	The loaded search criteria is loaded and executed, displaying results in the system
Data	Saved search criteria
Functions	• Load saved search criteria
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	Issue Description
1	DN: Again, about the need to authentication in order to save a search, this is possibly not true. A saved search can be represented by a URI. Who made it is not relevant.

Use Case ID	**UC_023**
Use Case Name	***6.4.2.4.4 e-mail metadata record URI to third party users***
Short Description	After a search, users can also choose to e-mail the metadata set URI to other users in order to collect opinions on the quality of the metadata and possibly the data also.
Actors	End User/Data Consumer
Pre-Conditions	A search was performed and a subset of the metadata from the catalog was retrieved by the NGDS catalog.
Success End Conditions	An e-mail is sent out to a recipient with a URI to the report on the current metadata set obtained through the system
Data	URI to Metadata from the system Search information
Functions	• E-mail metadata report • Build metadata report
Variants	

Step	Actor	Description
1b	NGDS System	Runs periodic subscription query Goes to step 4

Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	Issue Description
1	This ought to include not only email, but RSS, Facebook, Twitter, Google+ and maybe Reddit and/or Yammer?
2	DN: DERIVED REQUIREMENT: If you do this, you must also add in mechanisms to prevent this system from being used to spam people or abuse it in other ways.
	SMR 2014-01-30 allows posting to Google+, Facebook, Twitter, but not e-mail to specific address. This is out of box CKAN function I think.

Use Case ID	**UC_011**
Use Case Name	***6.4.2.4.5 Subscribe to new data***
Short Description	The goal of this use case is to allow users to utilize saved search criteria as subscriptions to new content published in the catalog that matches specific criteria. Users will be notified via e-mail, when new data that has been input in the system, matching

	that subscription criteria was published
Actors	End User/Data Consumer
Pre-Conditions	In order to allow subscriptions and notifications to occur, the user must be identifiable; hence, there is a need for users (in particular the end user) to be logged in using their unique account. The search criteria used in the subscription is properly validated and saved into the system under a given name
Success End Conditions	A subscription is successfully performed
Data	saved search/subscription criteria, user e-mail obtained from user profile subscription name
Functions	• Load search criteria • Subscribe to search criteria • System notification service that periodically notifies users of matched searches
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	Issue Description
1	DN: Is it possible this can be done without using saved search criteria?

6.4.3 Analyze and Visualize Data

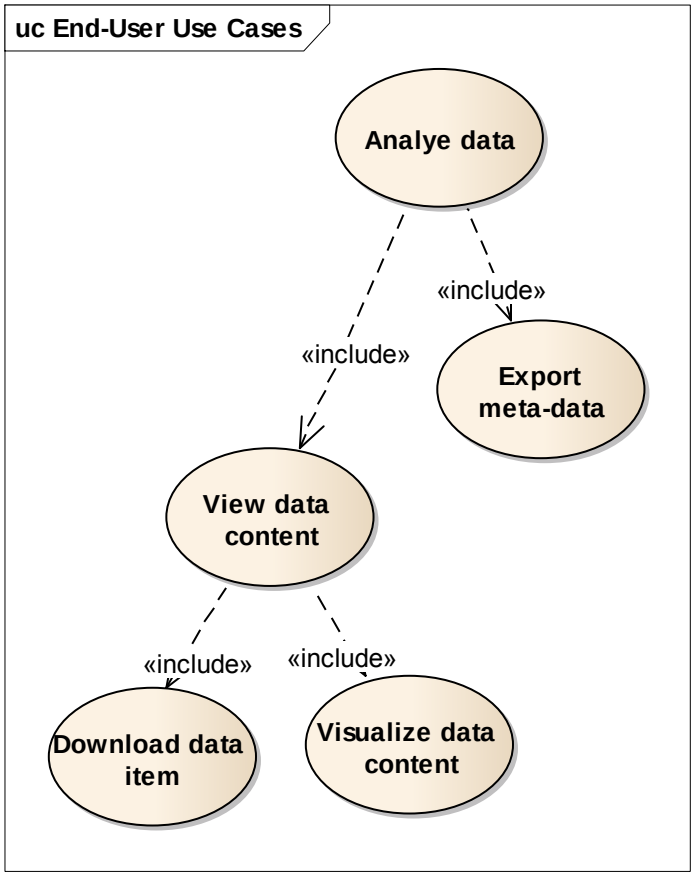


Figure 11 Data analysis

Use Case ID	**UC_027**
Use Case Name	***6.4.3.1.1 Export metadata***
Short Description	Users can export metadata records for different purposes, for example, to integrate them into their reports and spreadsheets, to further analyze these records with a CAD system, etc.
Actors	End User/Data Consumer
Pre-Conditions	Metadata was gathered and filtered
Success End Conditions	The user is able to save the metadata records
Data	Metadata records managed by GTDA
Functions	• Retrieve metadata records • Visualize metadata records
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	**Issue Description**
1	If there is more than one metadata record type, how the user would like to have the report? One CSV file per metadata type?
2	It is beyond the scope of this project the development of tools or providing support for detailed analysis of data. Hence, the need for exporting the data for further analysis.
3	DN: Is CSV the only option here? Would it be possible that some users want JSON or XML? Would suggest yes. JSON is probably the most useful but it depends on the data models (content models). I would hate to try and represent complex binary data as CSV.
	SMR 2014-01-30 Typical Siemens confusion of data and metadata that plagued project development. Metadata can be harvested and tabular data can be accessed using CKAN API, but that's not tested or documented; I think the idea here is the kind of functionality implemented by the NGDS data explorer.

6.5 System Administrator Use Cases

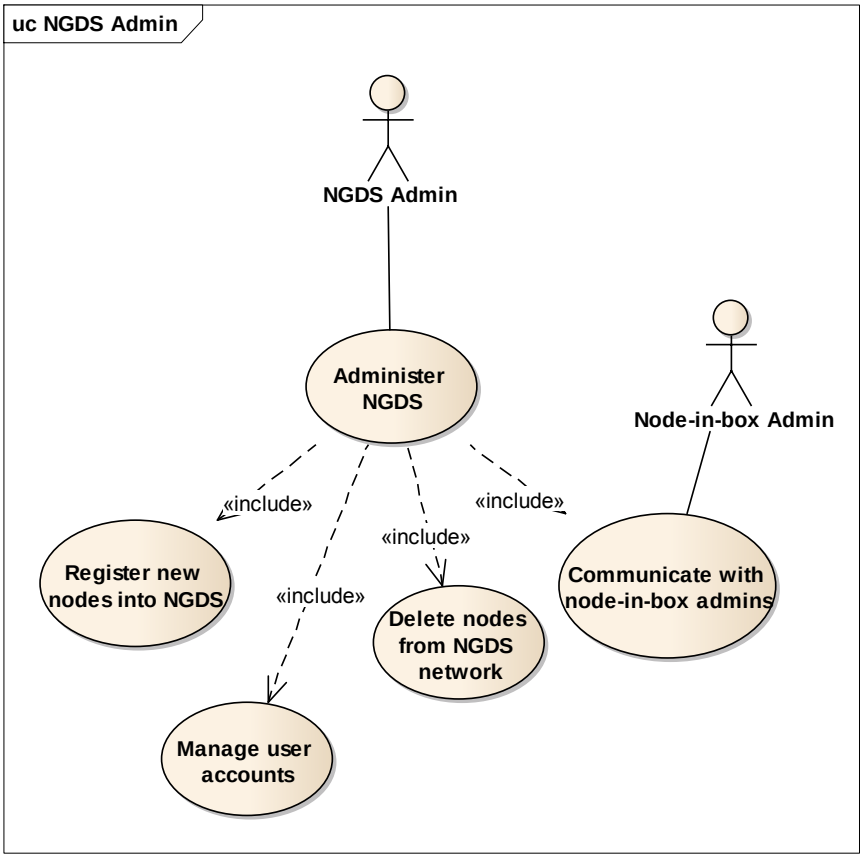


Figure 12 NGDS Administration use cases

Use Case ID	UC_036
Use Case Name	***6.5.1.1.1 Register new nodes into NGDS***
Short Description	The goal of this use case is to allow NGDS administrators to respond to new node requests. The administrator should evaluate the validity of the request, accepting, or rejecting it.
Actors	NGDS administrator
Pre-Conditions	GIN-stack properly installed as a NGDS node management hub Client node-in the box properly installed but not yet registered The new node must have at least one unique data or metadata record not currently in the system.
Success End Conditions	A new node is registered in the network, and the data it provides becomes searchable in the by the NGDS catalog
Data	e-mails, NGDS nodes registry
Functions	• add new node • index new node

Variants		
Step	**Actor**	**Description**
2b	NGDS Administrator	Rejects request based on external criteria
Exceptions		
Step	**Actor**	**Description**
1	NGDS System	Raise exception in case of incompatible/invalid protocols. Unless they support the common standards and protocols that they cannot join.
Open Issues (Please use this field to indicate questions/comments on the use case)		
	SMR 2014-01-30 Since we have no registry of nodes, there is nothing to register nodes into…Setting up a harvest from another node is about as close as it gets.	

Use Case ID	**UC_038**
Use Case Name	***6.5.1.1.2 Delete nodes from NGDS network***
Short Description	The goal of this use case is to allow NGDS administrators to respond to remove previously registered nodes from the system.
Actors	NGDS administrator
Pre-Conditions	GIN-stack properly installed as a NGDS node management hub The node to be removed is currently registered
Success End Conditions	A currently registered node is removed from the network, and the metadata records referencing this node are removed from the NGDS catalog
Data	NGDS nodes registry, NGDS catalog data
Functions	• Remove existing node • Remove index for node
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

Use Case ID	**UC_039**
Use Case Name	***6.5.1.1.3 Communicate with the Node-in-a-box admin***
Short Description	The goal of this use case is to allow node-in-a-box admins and NGDS admins to communicate in the handling of administration issues such as request node removal, check node registration information, etc.
Actors	NGDS administrator, node-in-a-box administrator
Pre-Conditions	GIN-stack is properly installed and configured Administrator has registered her e-mail information
Success End Conditions	Administrators can communicate with each other
Data	e-mails
Functions	• send email to administrator

Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	DN: I would rank the implementation of this use case as a low priority. This can be done by simply placing a mailto:xxx@xxx.com link on a page and letting the users own email client handle it. This seems a bit much to build into the system.

Use Case ID	**UC_037**
Use Case Name	***6.5.1.1.4 Manage NGDS user accounts***
Short Description	Allows the system administrator to manage NGDS users. This will allow the system administrator to add and remove users on the administered node, and assign user roles and group membership.
Actors	NGDS administrator
Pre-Conditions	NGDS network is properly installed and configured
Success End Conditions	The administrator is able to perform the main administration operations
Data	User records
Functions	• Add user • Delete user • Modify user permissions and roles
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	
ID	**Issue Description**
1	DN: Not sure if this is a use case I agree with. I think it would be better to delegate this to node administrators. If there is a user that the NGDS super admin wants removed, he or she can communicate that to the node admin perhaps? The worry is that the super NGDS admin would have to understand the metadata and data that the user is associated with before deleting to ensure no data or metadata is left orphaned. The decision is not mine but I encouraged discussion on this point.

6.6 Use cases common to all users

Use Case ID	**UC_001**
Use Case Name	***6.6.1.1.1 Login***
Priority [1-3] higher is more important	3
Reason for priority	User authentication is a basic activity that is a pre-condition for many other use cases. Failing to implement it is a show-stopper.
Short Description	The goal of this use case is to uniquely identify and authenticate a user, allowing the system to enforce access policies, and to use the user information to automatically fill in forms data, save searches and subscriptions, identify comments, etc.
Actors	Data Submitter, End User, Data Steward, System Administrator
Pre-Conditions	The user is logged out of the system

Success End Conditions		The user is logged in and authenticated with the system
Data		User login and password, or credentials collected in a third party authentication service
Functions		• Authenticate user using system credentials • Authenticate user using third party services, for example: Facebook, Gmail, and others
Variants		
Step	**Actor**	**Description**
		1. User forgets password but remembers username 2. User forgets username but remembers password 3. User forgets both username and password 4. User forgets username/password and the email they used to register.
Exceptions		
Open Issues (Please use this field to indicate questions/comments on the use case)		
ID	**Issue Description**	
1	DN: Password recovery tools could be added. Depending on the level of sophistication, enforcing a inimum set of standards for password might also be prudent.	

Use Case ID	**UC_002**
Use Case Name	***6.6.1.1.2 Logout***
Priority [1-3] higher is more important	3
Reason for priority	User authentication is a basic activity that is a pre-condition for many other use cases. Failing to implement it is a show-stopper.
Short Description	This allows a logged-in user to gracefully end and close their session
Actors	Data Submitter, End User, Data Steward, System Administrator
Pre-Conditions	The user is logged in the system
Success End Conditions	The user is logged out with no negative side effects to the system
Data	Current user section and credentials
Functions	• Logout user
Variants	
Exceptions	
Open Issues (Please use this field to indicate questions/comments on the use case)	

ID	**Issue Description**
1	DN: This implies session management. If that is the case, it is not noted as a requirement. Otherwise the user would not need to end a session. They could just close the browser window and walk away.

7 Acronyms, and Abbreviations

The following table lists the abbreviations used in this document, in order to promote their unique and unambiguous usage throughout the document and the Project.

Abbreviations	Definition
DOE	Department of Energy
NGDS	National Geothermal Data System
SDD	Software Design Description
DIS	Data Import Schema
OGC	Open Geospatial Consortium
WSS	Web Services Specification
ORM	Object-Relational Mapping
WMS	Web Map Service
WFS	Web Feature Service
CSW	Catalogue Service for the Web
WCS	Web Coverage Service
NetCDF	Network Common Data Form
API	Application Programming Interface
CSV	Comma-Separated file format
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
OAI-PMH	Open Archives Initiative – Protocol for Metadata Harvesting
UTM	Universal Transverse Mercator coordinate system

Table 1: Abbreviations

Placeholder

Placeholder

Placeholder

Geoserver Installation Instructions

Introduction

GeoServer is a free, open source software server written in Java that allows users to share and edit geospatial data that is published in Open Geospatial Consortium (OGC)-compliant services, including Web Feature Services (WFS) and Web Map Services (WMS) (http://geoserver.org/display/GEOS/Welcome). OGC services published as part of the National Geothermal Data System (NGDS, http://geothermaldata.org/) can use a variety of software servers and platforms, and GeoServer is an ideal candidate.

This document is meant to assist in simple GeoServer set-up using OpenGeo Suite (http://opengeo.org/products/suite/) as well as deployment of NGDS data services using GeoServer, with some additional configuration considerations for use of application schemas in GeoServer. It is assumed that schema-valid data tables already exist in a database management system. For more information on USGIN data-interchange schemas, visit http://schemas.usgin.org/models/.

Installation and Set-Up

Install GeoServer and Postgres/PostGIS together by getting OpenGeo Suite. Download the OpenGeo Suite, community or paid editions, from OpenGeo (http://opengeo.org/). The OpenGeo Suite not only includes Postgres and PostGIS but also includes GeoServer, OpenLayers and other packages, configured to work together. Also install the pgAdmin III administration tool. This is an incredibly useful tool as it provides the interface for creating new databases, tables, views, and users, execute SQL statements, modify user permissions, and much more. Although these tasks can still be accomplished via the psql command shell, pgAdmin III makes life much easier.

GeoServer Installation

GeoServer is built on the Java framework and therefore supported on many different platforms, such as Linux, Mac OS X and Windows. Java applications such as GeoServer need to be run through a Java servlet engine. The servlet engine is an application that provides the host environment within which Java applications run; the servlet engine provides the wrapper that allows the Java program to run on different operating systems.

The two primary ways to install GeoServer on Windows are: 1) by launching the Windows exe executable file, or 2) by deploying the web archive (war) file. The Windows exe installer will install the Jetty servlet engine and create a Windows Service for automatic startup. The war file must be deployed in a previously existing servlet engine; this typically means copying the file into a specific file directory specified by the servlet engine configuration.

The latest version of GeoServer can be found at http://geoserver.org/display/GEOS/Download. It is recommended to install the latest stable version. No GeoServer extensions/plugins are required. Installing GeoServer using the Windows exe installer is a straight-forward process: download the latest stable version of GeoServer and launch the executable file, which starts a GeoServer set-up wizard. Once GeoServer is installed and the service running, point your web

browser to http://localhost:8080/geoserver/ to confirm the installation. Log in using the username and password set during installation. The default username and password is admin/geoserver. After installation, be sure to resolve any notices about security related issues that are displayed on your local GeoServer homepage. Then, click Contact Information to specify the service provider information that will serve as your metadata contact information in service requests.

Deploying Services

The following section outlines how to use OpenGeo Suite on Windows platform to deploy WMS/WFS services, using a connection to PostGIS database tables as an example.

1. Using the OpenGeo Suite Dashboard, click Start and choose GeoServer to open the GeoServer interface.
2. Under Workspaces, click Add new workspace. The workspace specifies a namespace, so each data type will need its own separate workspace.
    a. Choose a name to identify your workspace (i.e., ThermalSprings)
    b. Fill in the namespace, given in the first line of an NGDS schema, or XSD, (http://schemas.usgin.org/models/)
    c. Click Submit
3. Under Stores, click Add new Store. The Store will house the connection parameters to the data table in PostGIS database.
    a. If using PostGIS, choose this from the Vector Data Sources list. Choose the workspace you've just created from the drop-down list. The Data Source Name* will be the name of the data table as it appears exactly in your PostGIS database. (Using pgAdmin III administration tool is especially helpful here. When you understand the file structure and exactly where the data tables are in your PostGIS database, you will be better able to input the correct parameters to connect to that database in the Store page on Gerserver. It will also allow you to see the exact name and capitalization of a given data table.)
    b. Be sure to click the box next to "Enabled"
    c. Once connection parameters are filled in, click Save.
4. Under Layers, click Add a new resource.
    a. From the drop-down list, choose the workspace:Store you've just created.
    b. If connected to PostGIS database, this should give all the available tables in the specified folder. Navigate through the list to the data table (i.e., azthermalsprings) and click Publish to the right of that data table name.
    c. This takes you to Edit Layer. Keep the lower-case name.
    d. In the Data tab, under Bounding Boxes, choose 'Compute from data' under 'Native Bounding Box; choose 'Compute from native bounds' under 'Lat/Lon Bounding Box'. This will pull locational bounding boxes from the data table.
    e. Click Save at bottom.
5. Under Layers, click Add a new resource.
    a. Choose the workspace:Store you created previously.

b. When the available data tables pop up, there is an bolded option at the top labeled ‘Configure new SQL view...’ Click this.

i. We’ve found that when connecting to a PostGIS database, GeoServer will not recognize the correct SRID (Spatial Reference System Identifier). This needs to be indicated manually by creating a new SQL view of the layer. The SQL view will serve as the final layer for your published service.

c. In the Create new SQL view page, fill in the ‘View Name’ that will serve as the final layer name, as indicated by the appropriate schema.

d. In ‘SQL statement’ use a SQL statement that will call all the fields from the desired data table as such:

select * from azthermalsprings

Click Refresh.

e. When the fields appear, check objectid as the Identifier. For the shape field, type is indicated as Geometry. The SRID box must be filled in with the correct EPSG code (i.e., 4326). Click Save.

f. This takes you to Edit Layer.

g. In the Data tab, under Bounding Boxes, choose ‘Compute from data’ under ‘Native Bounding Box; choose ‘Compute from native bounds’ under ‘Lat/Lon Bounding Box’. This will pull locational bounding boxes from the data table.

h. In the publishing tab, under WMS Settings, a Style (or symbology) may be chosen.

i. To add a Style: under Styles from GeoServer home, click Add a new style.

ii. Choose a name and the workspace for this style (i.e., ThermalSprings)

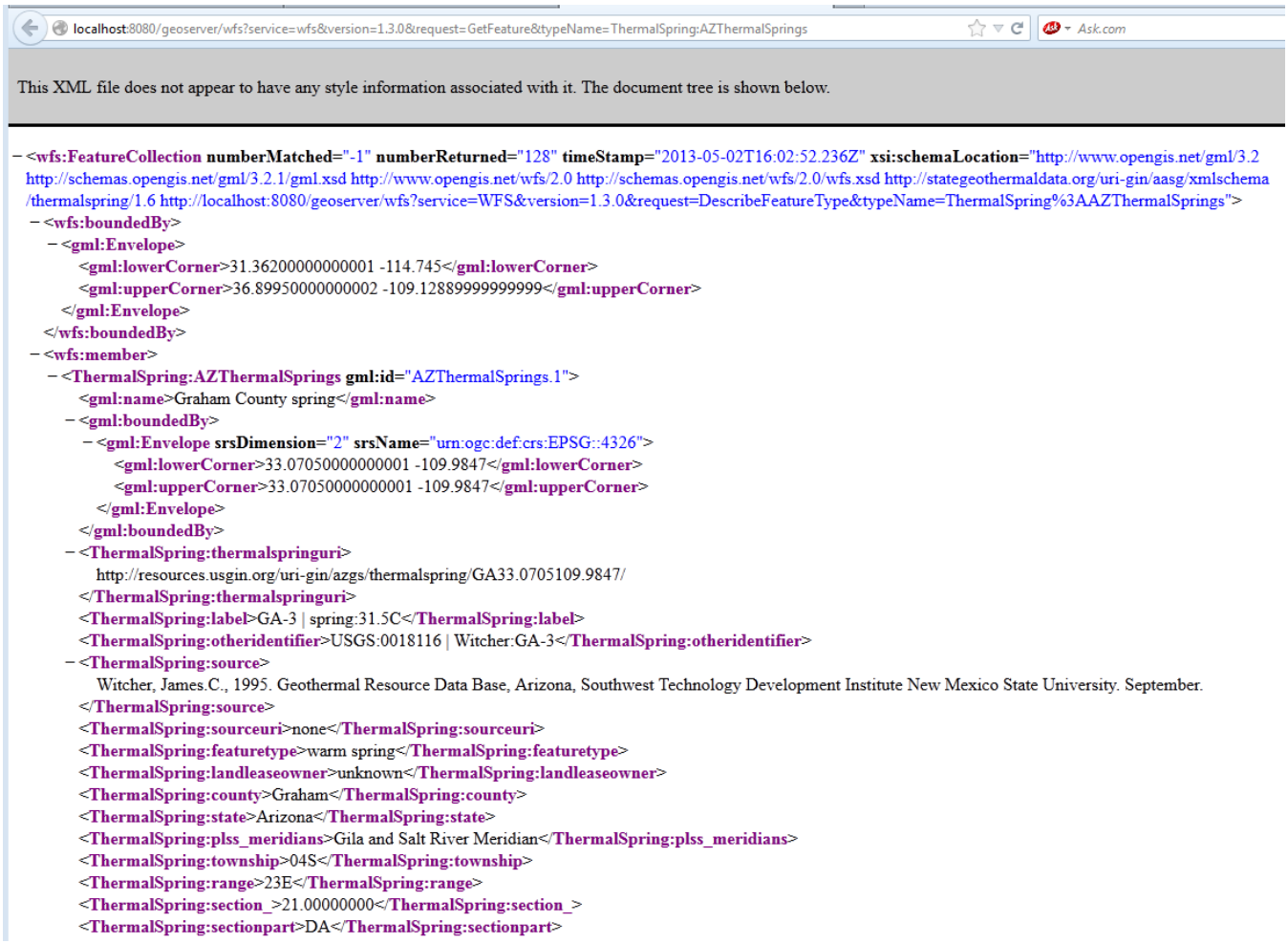
iii. To use the SLDs for the NGDS project, scroll down to SLD file and click Browse. Browse to your downloaded SLD copies, avaialbe at http://repository.stategeothermaldata.org/repository/resource/50ec3aefb656b70647f32e38bc1b7479/

iv. Click Upload. Click Submit.

i. Click Save.

6. Perform a GetFeature request to confirm data return. Construct a GetFeature URL, where the typeName=workspace:layer. For example:

http://localhost:8080/geoserver/wfs?service=wfs&version=1.3.0&request=GetFeature&typeName=ThermalSpring:ThermalSpring



Application Schemas

One hurdle to using GeoServer to deploy NGDS Geothermal Data services is the resolution of schemas. Creating additional configuration files for GeoServer allows the use of "app-schemas" which will automatically download the indicated web-accessible schema (.xsd) and be referenced in the GetFeature request. This solves a problem with service deployment using ArcGIS, where custom schema locations is necessary. The following documentation is based on Tomcat running GeoServer, set up on an Oracle VM VirtualBox. The data for the services are pulled from a PostGIS database. Likely, the configuration files will be the same if running GeoServer on Windows platform

Install the App-Schema .jar files

See this page http://sourceforge.net/projects/geoserver/files/GeoServer%20Extensions/ for app-schema versions and downloads.

- Download the app-schema plugin zip file for the same version of your GeoServer instance.
- Unzip the app-schema plugin zip file to obtain the jar files inside. Do not unzip the jar files.
- Place the jar files in the WEB-INF/lib directory of your GeoServer installation.
- Restart GeoServer to load the extension.

Configuration files

App-Schema GeoServer documentation is available on thier site here: http://docs.geoserver.org/stable/en/user/data/app-schema/index.html.

Two configuration files are necessary for app-schemas, one of which will have to be manually created. The mapping-file.xml file will be created, and the datastore.xml will be edited. GeoServer automatically creates the datastore.xml file, it just needs to be edited to specify that the layer will use app-schema configuration.

Mapping file:

This file will contain the database connection information, the web address to the schema (.xsd), and the TypeMapping which correlates the fields in your database table to those in your schema. Please see these two blogs for FeatureTypeMapping examples: http://lab.usgin.org/node/116 and http://lab.usgin.org/groups/building-geosciml-wfs-server/ncgmp-geosciml-mappedfeature. See GeoServer documentation here: http://docs.geoserver.org/stable/en/user/data/app-schema/mapping-file.html

Example mapping file: https://github.com/ccaudill/GeoServerWorkspace/archive/4.1.2013.zip under folder hf - azheatflow - HeatFlow.xml

NOTE: The name of your layer **must** match exactly the name of the targetElement in your mapping file, which is the element indicated in the schema. Thus, the layer name in GeoServer used for app-schema configuration must match the name specified in your schema document.

datastore.xml file:

The datastore.xml file points to the file location of the mapping file and also indicates that an app-schema is to be used. See GeoServer documentation here: http://docs.geoserver.org/stable/en/user/data/app-schema/data-stores.html#app-schema-data-stores

Example mapping file: https://github.com/ccaudill/GeoServerWorkspace/archive/4.1.2013.zip under folder hf - azheatflow - datastore.xml

- The datastore.xml file **must** keep the letter and number string that geoserver gives as its internal id in the datastore id and workspace id fields. (Try to rename this as something human-readable, and GeoServer will not be able to open when restarting.)
- As you remove the type=PostGIS and enabled=true from the datastore.xml, the layer no longer shows as available, and now points to the app-schema. When the type field under dataStore is left as "PostGIS", the store will try to look for kyanite, no matter what else is in the file. Remove type field completely and indicate under connectionParameters dbtype=app-schema and store will look at the app-schema.

File structure

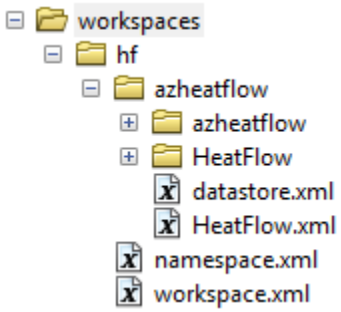
Simply insert these configuration files into the correct location in your GeoServer file structure and GeoServer should automatically download the schemas, pulling those schemas from a cache each time a GetFeature request is performed on the given service.

See this blog http://lab.usgin.org/node/117 for an example GeoServer directory structure. The following example of the file structure begins with the workspaces folder.

In **OpenGeo Suite GeoServer on Windows platform**, the folder path to the workspaces folder was: C:\Program Files (x86)\OpenGeo\OpenGeo Suite\webapps\geoserver\data\workspaces.

In the **Oracle VM VirtualBox on Linux platform**, the folder path to the workspaces folder was: /var/lib/tomcat6/webapps/geoserver/data/workspaces. In the following example,

1. hf is the name of the Workspace (created in the GeoServer interface)
2. azheatflow is the name of the Store (created in the GeoServer interface)
3. HeatFlow is the name of the Layer (SQL view was created in the GeoServer interface)
4. HeatFlow.xml is the name of the mapping file (stored in the same location as the datastore.xml file)



It is not necessary to edit the namespace.xml or workspace.xml files, nor the files that are in the HeatFlow folder (which are the featuretype.xml and layer.xml files).

One other important file location is the cache of the schemas, which GeoServer creates. In the Oracle VM VirtualBox on Linux platform, the folder path to the schema cache location is /var/lib/tomcat6/webapps/geoserver/data/app-schema-cache. In OpenGeo Suite GeoServer on Windows platform, the folder path to the schema cache location is C:\Program Files (x86)\OpenGeo\OpenGeo Suite\webapps\geoserver\data\app-schema-cache. In these locations, the filing of the downloaded schemas is a little convoluted. For example, the online link to the heat flow schema is http://schemas.usgin.org/files/heat-flow/1.23/HeatFlow.xsd. In GeoServer, this schema is filed in as series of folders from the app-schema-cache folder as such: /org/usgin/schemas/files/heat-flow/1.23/HeatFlow.xsd

Making the final location of the schema in Linux:

/var/lib/tomcat6/webapps/geoserver/data/app-schema-cache/org/usgin/schemas/files/heat-flow/1.23/HeatFlow.xsd

GetFeature request example using app-schema configuration

```
−<wfs:FeatureCollection numberOfFeatures="2" timeStamp="2013-03-28T16:31:18.703Z"
 xsi:schemaLocation="http://stategeothermaldata.org/uri-gin/aasg/xmlschema/boreholetemperature/1.5
 http://schemas.usgin.org/files/borehole-temperature-observation/1.5/BoreholeTemperature.xsd http://www.opengis.net/wfs
 http://schemas.opengis.net/wfs/1.1.0/wfs.xsd">
 −<gml:featureMember>
  −<aasg:BoreholeTemperature gml:id="azboreholetemperatures.1">
   −<aasg:ObservationURI>
      http://resources.usgin.org/uri-gin/azgs/bhtemp/1932/
    </aasg:ObservationURI>
    <aasg:WellName>AZ943</aasg:WellName>
   −<aasg:HeaderURI>
      http://resources.usgin.org/uri-gin/azgs/well/AZ:MA33.6383112.4092/
    </aasg:HeaderURI>
    <aasg:Label>AZ943</aasg:Label>
    <aasg:SpudDate>1900-01-01 00:00:00.0</aasg:SpudDate>
    <aasg:WellType>nil:missing</aasg:WellType>
    <aasg:Status>nil:missing</aasg:Status>
    <aasg:CommodityOfInterest>nil:missing</aasg:CommodityOfInterest>
    <aasg:Function>nil:missing</aasg:Function>
    <aasg:Production>nil:missing</aasg:Production>
    <aasg:County>Maricopa</aasg:County>
    <aasg:State>Arizona</aasg:State>
    <aasg:PLSS_Meridians>Gila and Salt River Meridian</aasg:PLSS_Meridians>
    <aasg:TWP>T03N</aasg:TWP>
    <aasg:RGE>02W</aasg:RGE>
    <aasg:Section_>1.00000000</aasg:Section_>
    <aasg:UTMDatumZone>NAD83 zone 12</aasg:UTMDatumZone>
    <aasg:LatDegree>33.63819675</aasg:LatDegree>
    <aasg:LongDegree>-112.40992191</aasg:LongDegree>
    <aasg:SRS>WGS 84</aasg:SRS>
   −<aasg:LocationUncertaintyStatement>
      Witcher TRS location matches topo; BLM TRS location information changed to match (ERP).
    </aasg:LocationUncertaintyStatement>
    <aasg:DrillerTotalDepth>600.0</aasg:DrillerTotalDepth>
    <aasg:LengthUnits>ft</aasg:LengthUnits>
    <aasg:ElevationDF>0.0</aasg:ElevationDF>
    <aasg:ElevationGL>-99999.0</aasg:ElevationGL>
```

App-schema errors and resolution

"java.lang.IllegalArgumentException: Filter Function problem for function strConcat argument #1 - expected type String" when preforming a GetFeature request."

This error resulted from the first element in AttributeMapping in the mapping file, where an OCQL expression was not accepted. We attempted using a OCQL statement here because the GetFeature returned an arbitrary string as the gml:id instead of the OBJECTID. This was because the OBJECTID field in our PostGIS data table was not specified as the primary key. Once this field was specified as such, GeoServer could read this as the gml:id by the appropriate expression for the element in the mapping file:

```
<AttributeMapping>
    <targetAttribute>aasg:HeatFlow</targetAttribute>
    <idExpression><OCQL>getId()</OCQL></idExpression>
</AttributeMapping>
```

According to GeoServer developers, "For performance reasons, we strongly recommend avoiding using CQL to generate gml:id in the mapping file as this prevents the use of database indices and the generation of efficient SQL queries."

It must be noted that one important caveat to creating tables in any database structure using ArcMap is that a primary key is **not** indicated. We found this when the return of a GetFeature request in GeoServer using app-schemas showed not only an arbitrary string for the gml:id, but the data was not correctly mapped to the fields (or columns from the PostGIS data table). It appeared that the first column, the OBJECTID, was acutally being mapped as data and therefore put into the first column from the mapping file, which was ObservationURI. This caused more choas; as the schema was being read by GeoServer with a one-field offset, it misinterpreted some fields as not required and completely left them out of the GetFeature request. To fix this, we used pgAdmin III to specify the primary key as the objectid in the table. Then, GeoServer was able to acknowledge this and leave that field's integer out of the data when the GetFeature was returned, thus giving a correct return of the data from that PostGIS table. For those using pgAdmin III, specify a primary key by right-clicking the PostGIS database table and selecting "Properties". Under the "Constraints" tab, click the Add button beside "Primary Key". Give a name, such as "bht_p_key" then click the "Columns" tab to choose the appropriate field (here, the objectid). Click OK.

"java.lang.VerifyError: Cannot inherit from final class"

Be certain that the version of the app-schema download is the same version is the same version of GeoServer you're using. Those specific .jar files in the app-schema plug-in are essential. Using all the same versions cleared this error. See this page http://sourceforge.net/projects/geoserver/files/GeoServer%20Extensions/ for app-schema versions and downloads.

"Can't parse complex datastore config."

This error was resolved with an awful lot of tweaking the mapping file. See the above for a polished example.

"org.xml.sax.SAXParseException: Content is not allowed in prolog."

This error was caused by characters that were unintentionally added to the first line of the mapping file. This had two causes - one was the copy and paste method used to create the file, and the second was the use of XML Copy Editor. This free XML editor evidently inserts it's own code into the file. Every time I opened and saved the file with this software, I would get this error. Removing characters in the mapping file that were before the first line (i.e., <?xml version="1.0" encoding="UTF-8"?>) and using another free editor solved these problems.

Helpful housekeeping tips

catalina.sh

Is GeoServer thinking forever to take you to the Layers page or on a request, or not able to return a request at all? Try editing the **catalina.sh** file in the GeoServer data directory to indicate that more of Tomcat's memory needs to be allocated to running GeoServer. To this file, add:

JAVA_OPTS="-Djava.awt.headless=true -Dfile.encoding=UTF-8 -server -Xms512m -Xmx1024m -XX:NewSize=256m -XX:MaxNewSize=256m -XX:PermSize=256m -XX:MaxPermSize=256m -XX:+DisableExplicitGC"

So that a portion of the file will now look like this:

```
# JSSE_HOME (Optional) May point at your Java Secure Sockets Extension
# (JSSE) installation, whose JAR files will be added to the
# system class path used to start Tomcat.
#
# CATALINA_PID (Optional) Path of the file which should contains the pid
# of catalina startup java process, when start (fork) is used
#
# $Id: catalina.sh 609438 2008-01-06 22:14:28Z markt $
# -----------------------------------------------------------------------------

JAVA_OPTS="-Djava.awt.headless=true -Dfile.encoding=UTF-8 -server -Xms512m -
Xmx1024m -XX:NewSize=256m -XX:MaxNewSize=256m -XX:PermSize=256m -
XX:MaxPermSize=256m -XX:+DisableExplicitGC"


# OS specific support. $var _must_ be set to either true or false.
cygwin=false
os400=false
darwin=false
```

Setting up Hyper-V 2012 Replication on Workgroup Servers

A Guide

5/10/2013

Metadata	
Author(s)	The Arizona Geological Survey
Current Version	1.0
Initial Publication Date	5/10/2013

Introduction

This document provides instructions regarding the use of Hyper-V 2012 to virtualize and replicate servers on x86 and x64 workgroup servers. This process makes it easy to back up entire servers and their contents.

Hyper-V is a *Windows Server Role* that can be installed on Windows Server 2012 platforms.

Prerequisites

1. A minimum of two **Windows Server 2012** Hyper-V host servers
    a. One host server will act as the *primary* server, of which copies will be made
    b. Another host server will act as a *replica* server used to receive replica copies of virtual machines from the primary host server

Installation and Setup

Step 1). Set up a Windows Server 2012 instance in a workgroup; do not join the server to a domain

Step 2). To install the **Windows 2012 Hyper-V role** on a Windows Server 2012 host server, open the **Windows Server 2012 Server Manager**

Step 3). Select **Local Server** (Figure 1); click **Manage**, then select **Add Roles and Features**; this will open the **Before you Begin** page of the **Add Roles and Features** wizard (Figure 2). Click **Next** to continue.

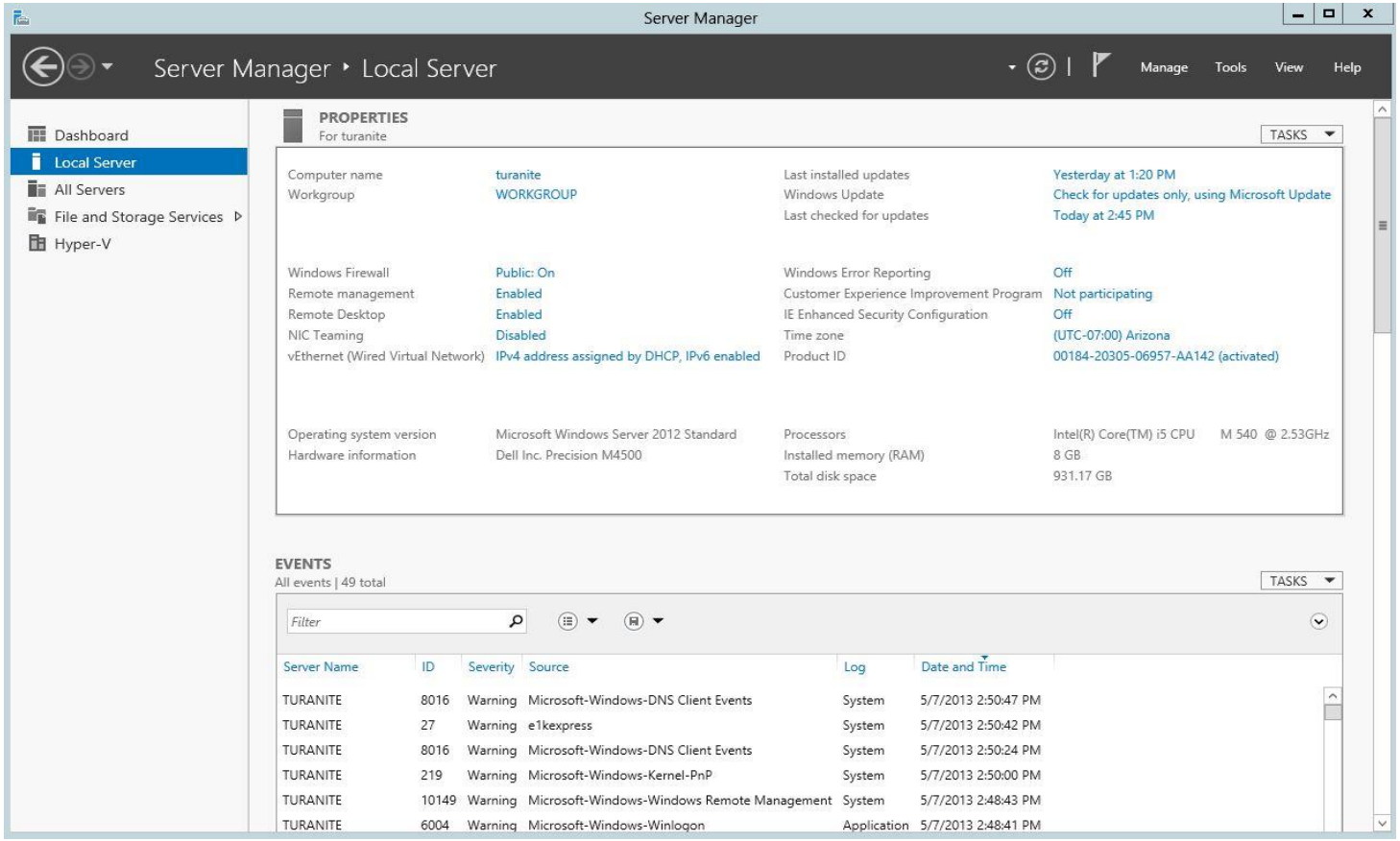


Figure 1: The Local Server window

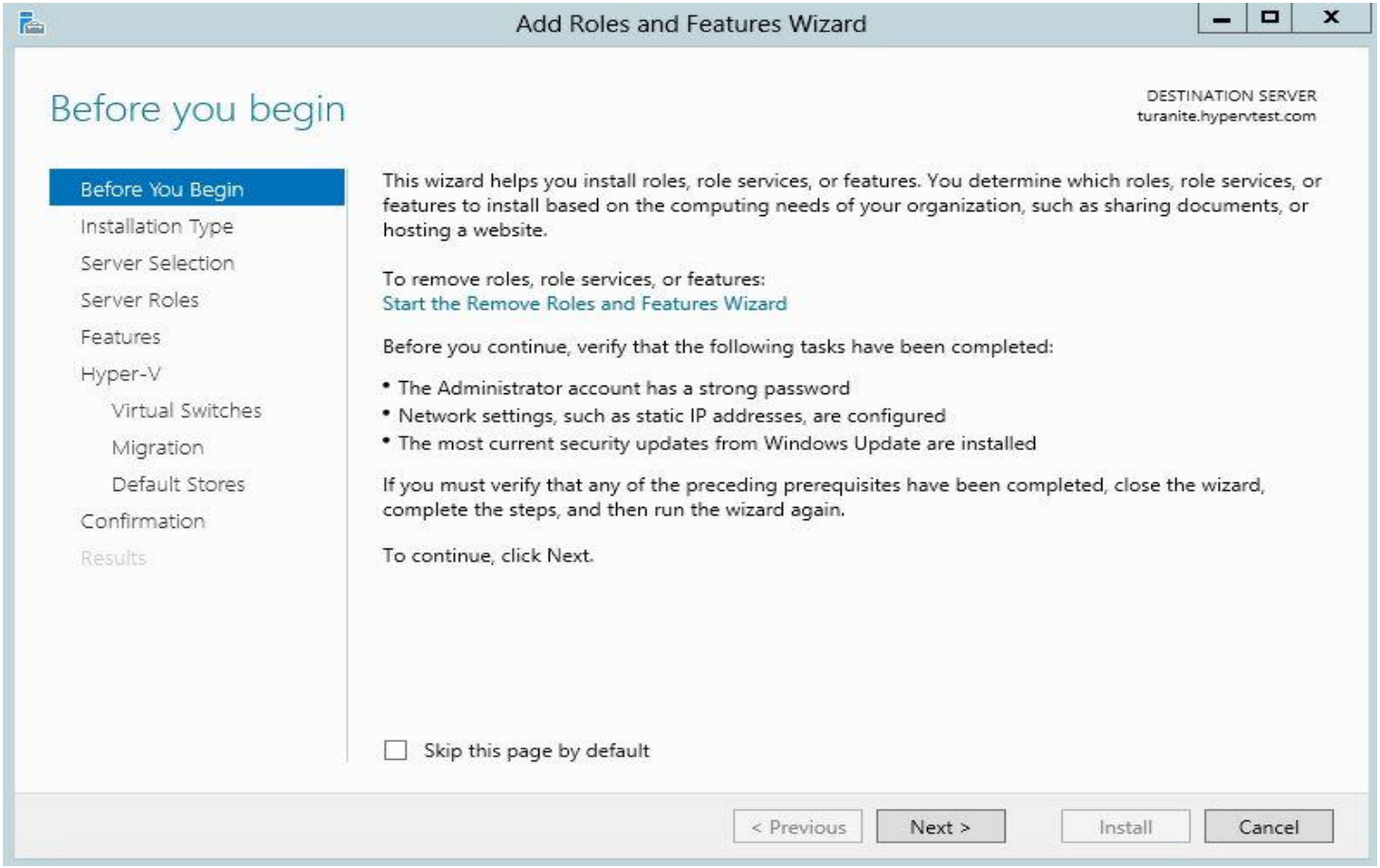


Figure 2: Before you Begin

Step 4). On the **Installation Type** page of the **Add Roles and Features** Wizard (Figure 3), click the Role-based or feature-based installation radio button; click **Next** to continue.

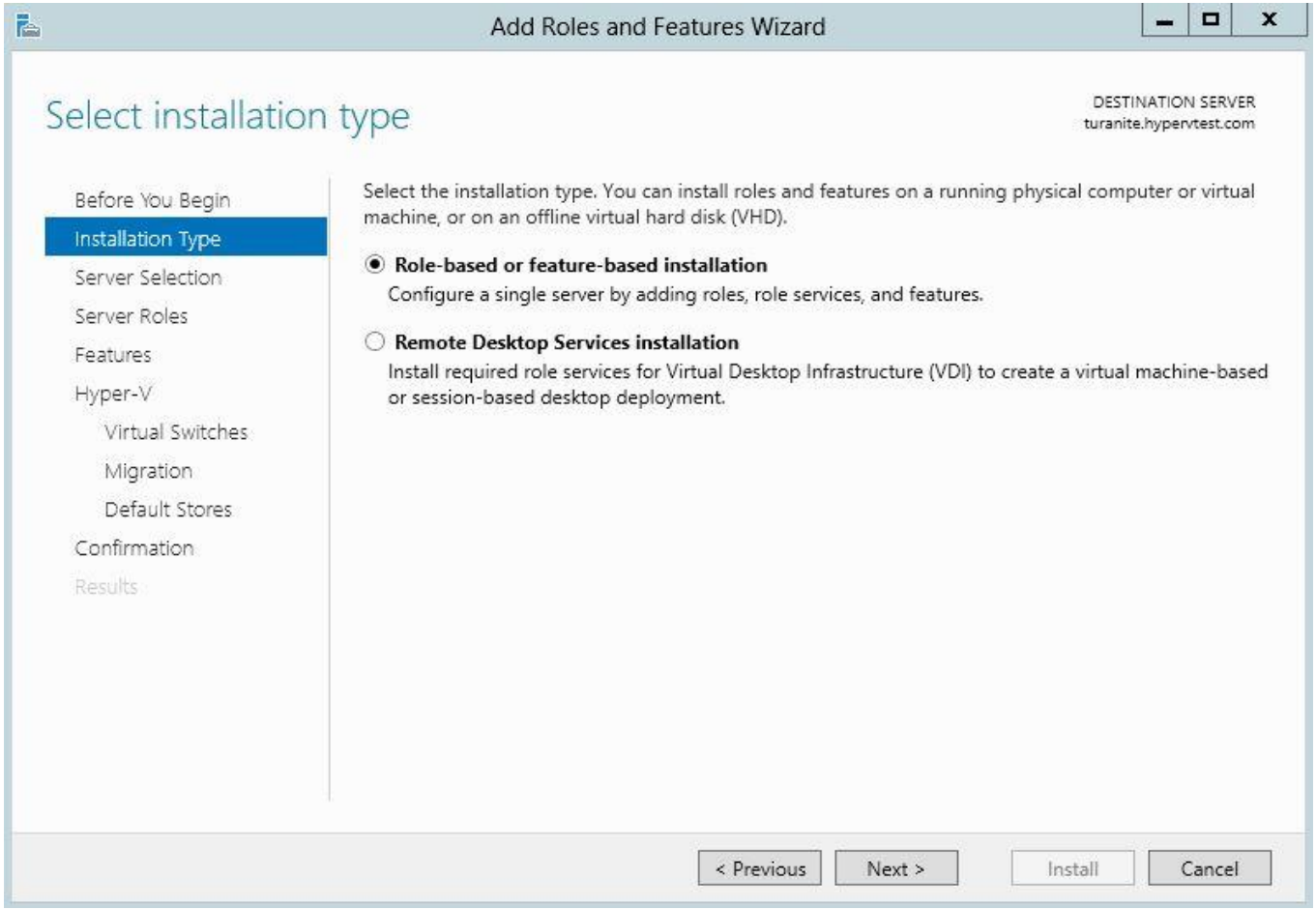


Figure 3: The Installation Type page of the Add Roles and Features Wizard

Step 5). On the **Server Selection** page of the **Add Roles and Features** Wizard (Figure 4), select the destination server to which you wish to install Hyper-V and click **Next** to continue.

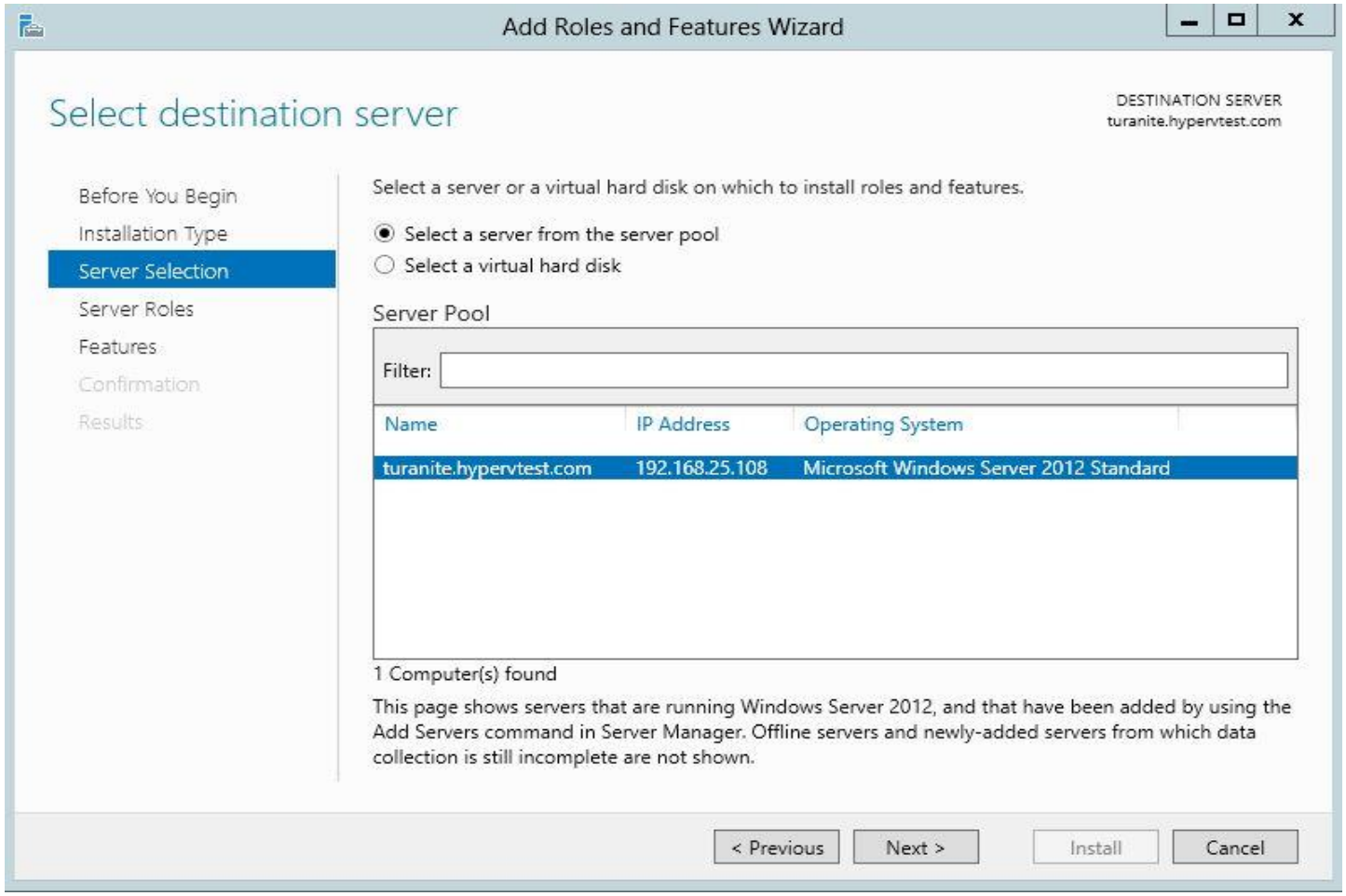


Figure 4: Server Selection

Step 6). On the **Server Roles** page of the **Add Roles and Features** Wizard (Figure 5), select the Hyper-V server role and click **Next** to continue.

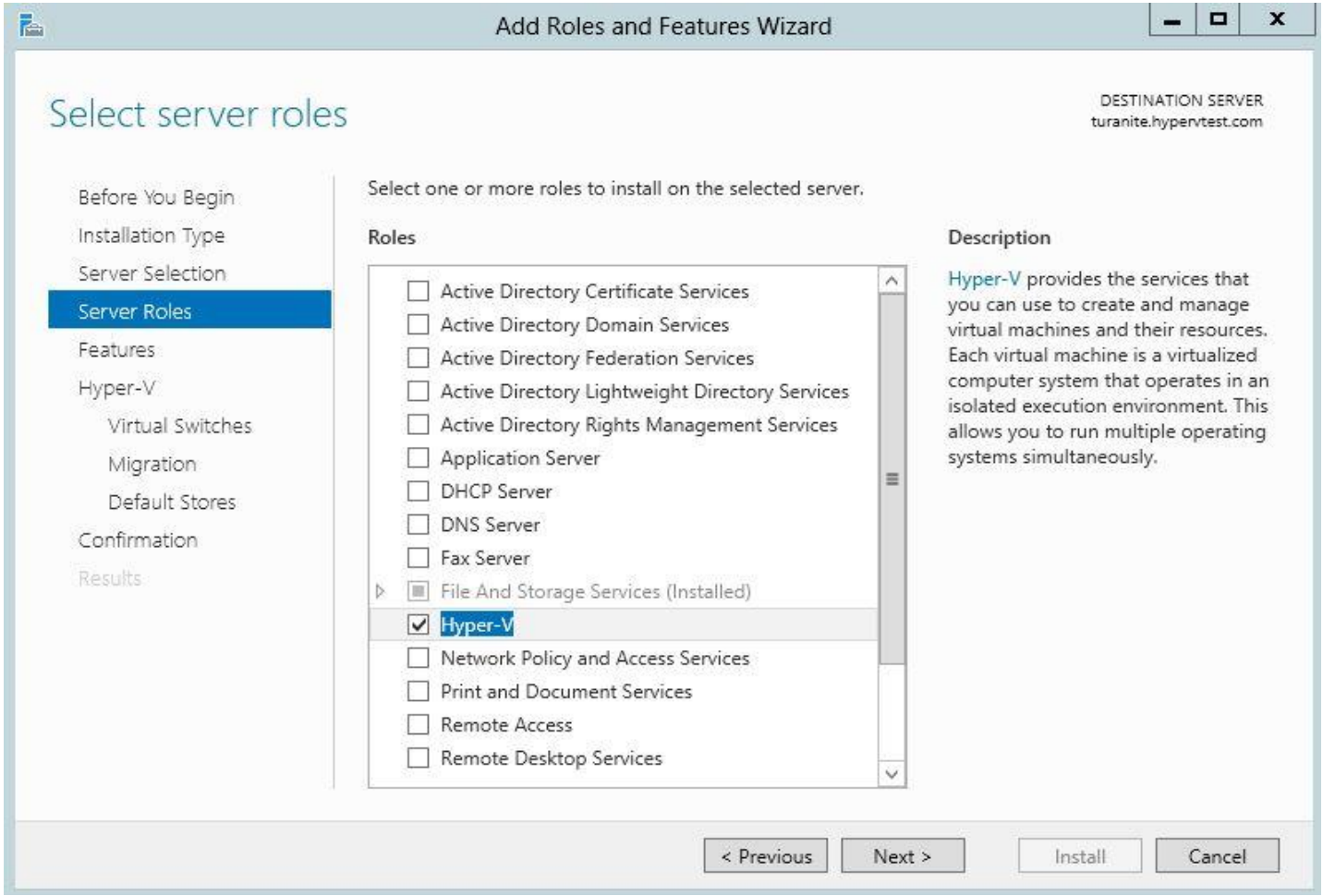


Figure 5: Server Roles

Step 7). On the **Features** page of the **Add Roles and Features** Wizard (Figure 6), select any desired features to be installed; note that for the purposes of this tutorial, the default features are sufficient. When finished, click **Next** to continue.

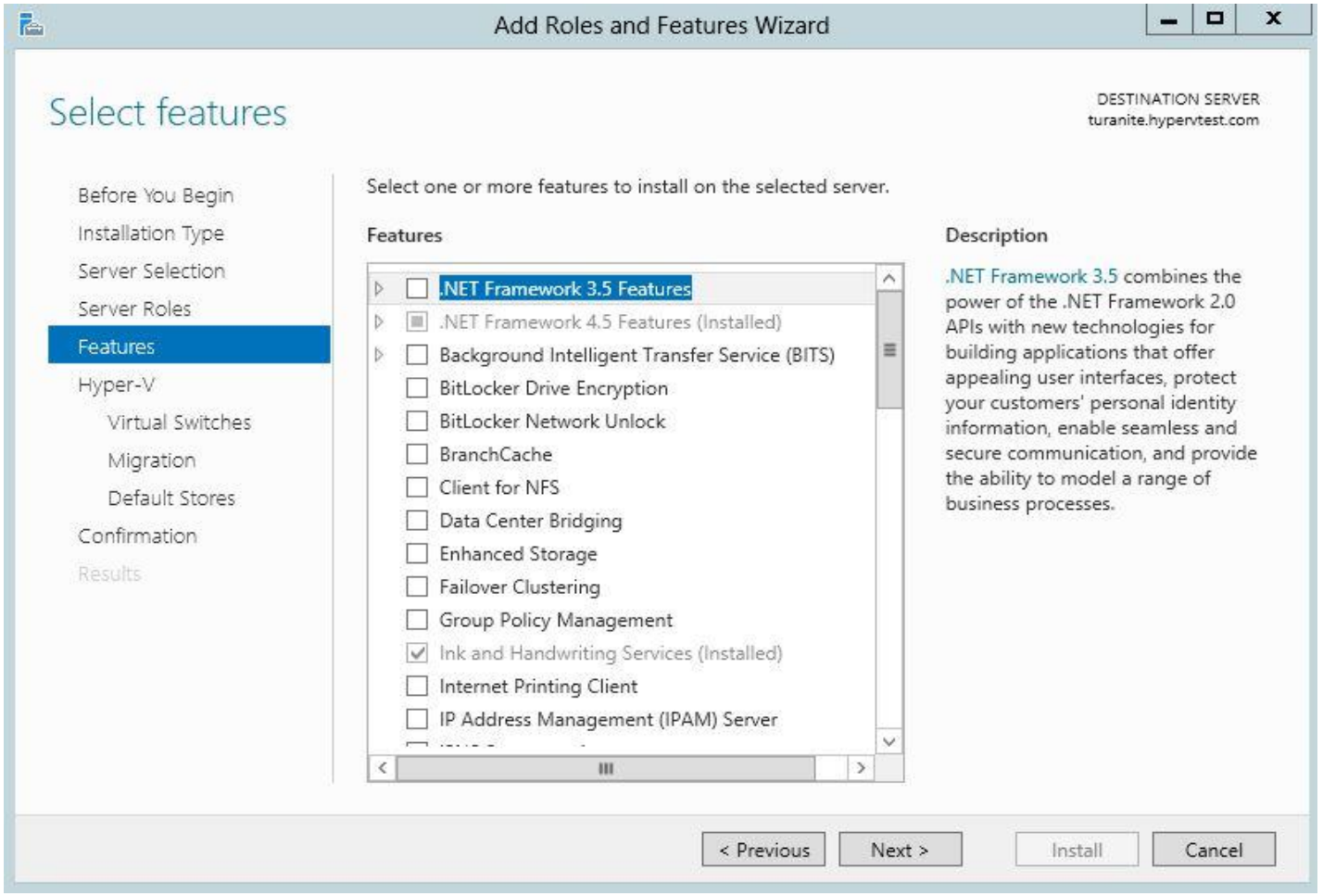


Figure 6: Features

Step 8). On the **Hyper-V** page of the **Add Roles and Features** Wizard (Figure 7), peruse information about Hyper-V; when you are finished, click **Next** to continue.

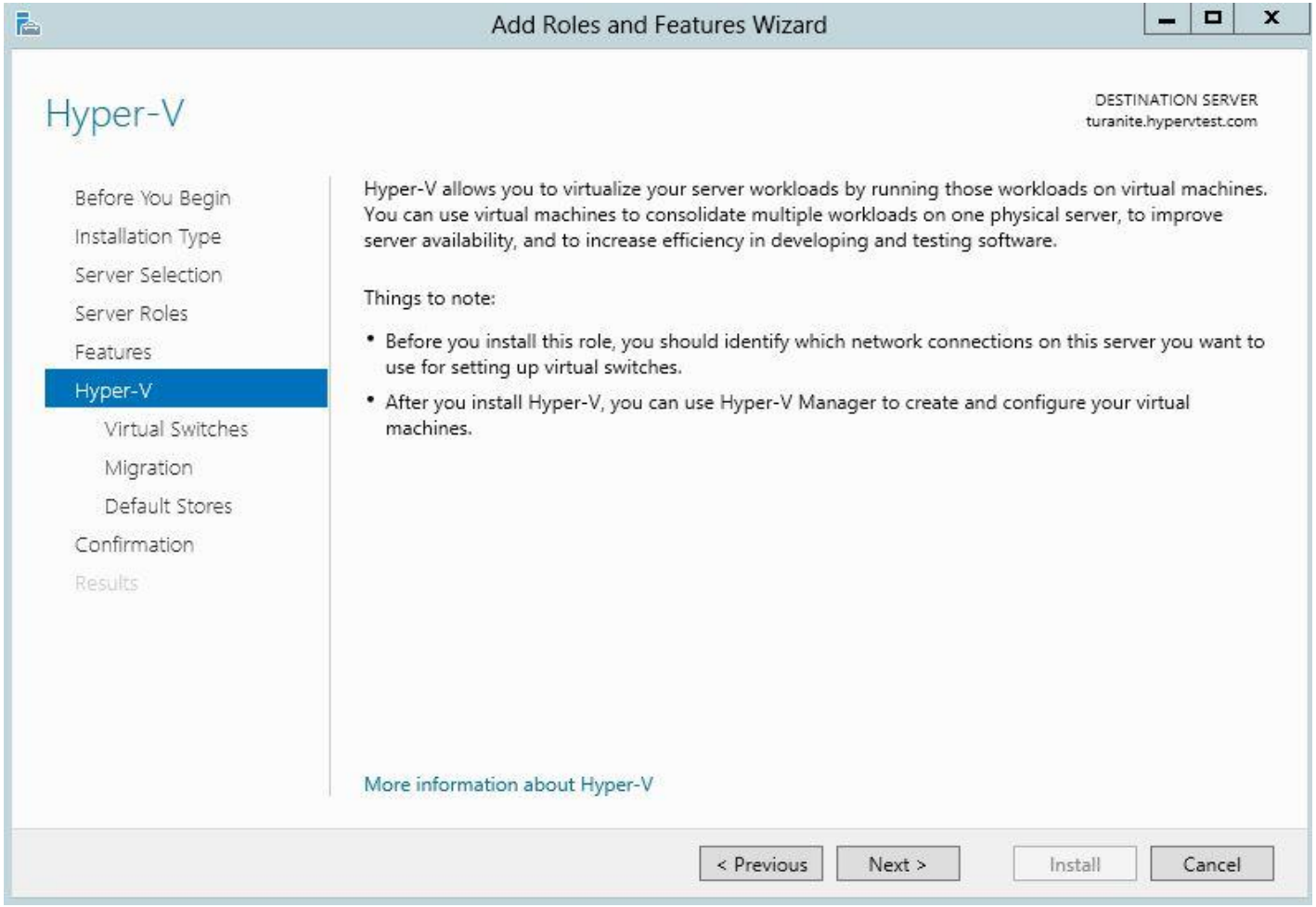


Figure 7: Information about Hyper-V

Step 9). In the **Virtual Switches** page of the **Add Roles and Features** wizard (Figure 8), create a virtual switch for your Hyper-V installation by clicking the checkbox that corresponds with the network adaptor you wish to emulate. When you have made your selection, click **Next** to continue.

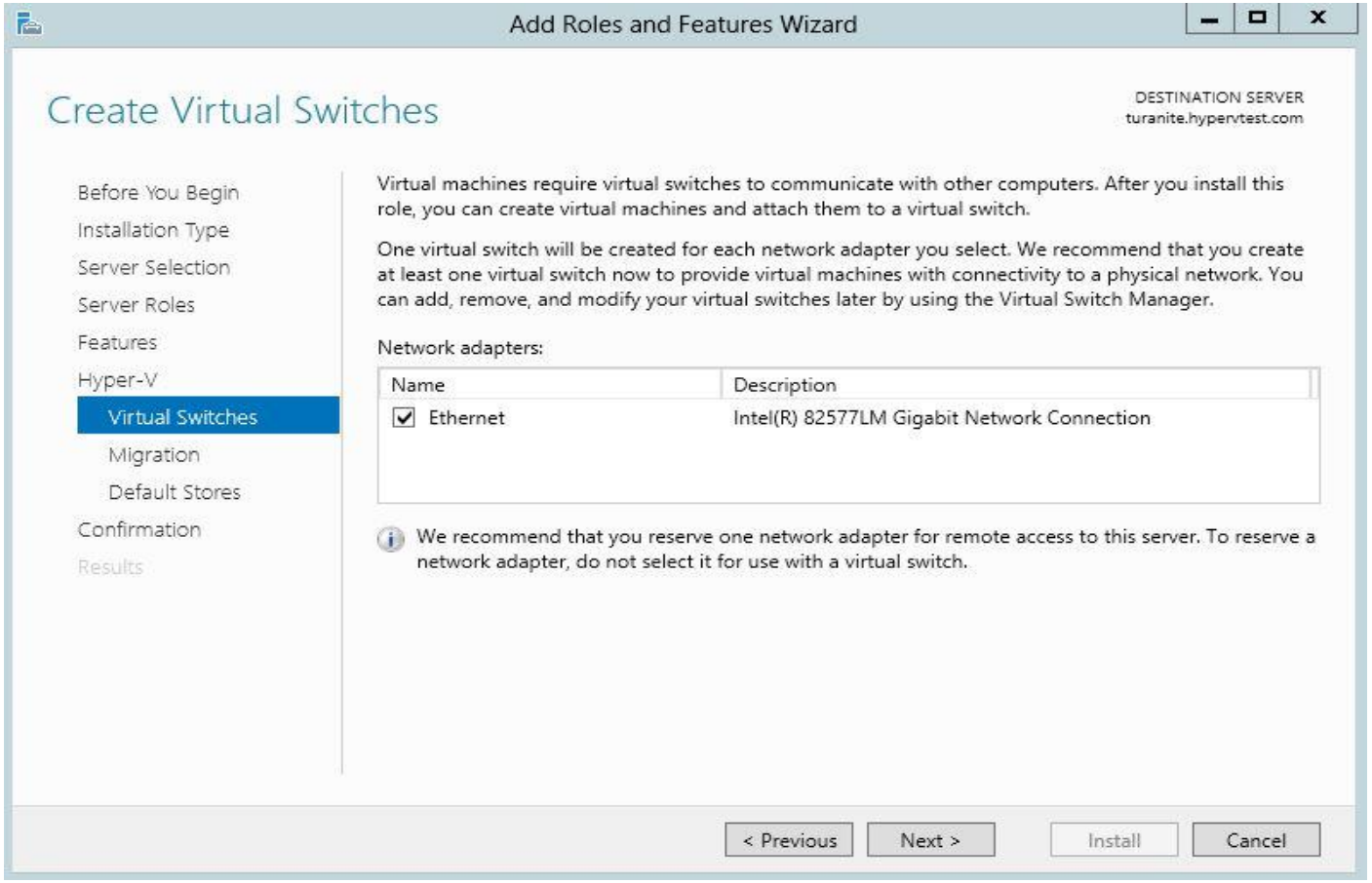


Figure 8: Virtual Switches

Step 10). On the **Migration** page of the **Add Roles and Features** Wizard (Figure 9), *do not* check the box enabling your Hyper-V installation to be able to send and receive live migrations of virtualized machines. Click **Next** to continue.

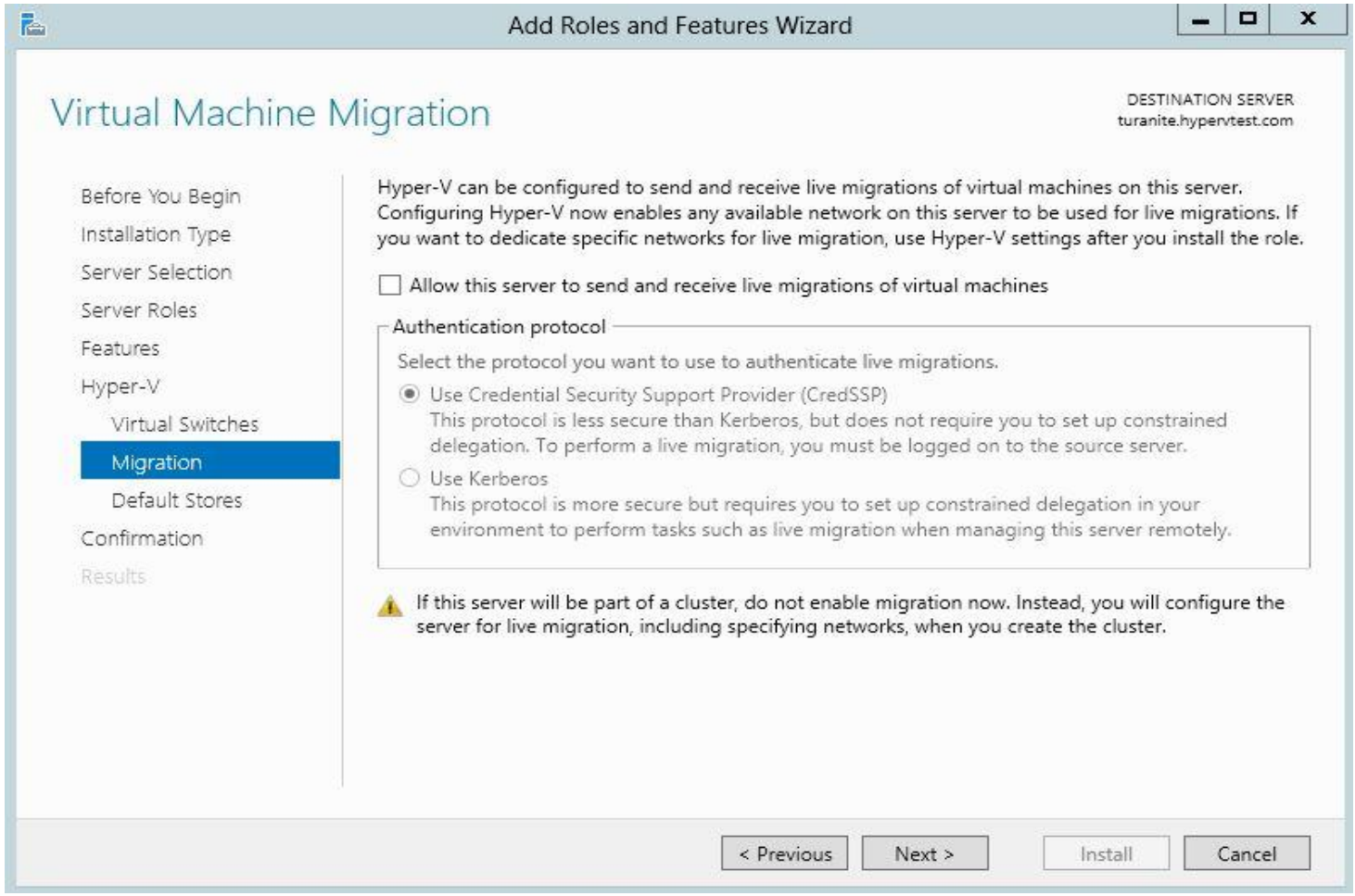


Figure 9: Virtual Machine Migration

Step 11). On the **Default Stores** page of the **Add Roles and Features** Wizard (Figure 10), click **Browse** to navigate to the location on which you wish to store virtualized hard disks and configuration files. Click **Next** to continue.

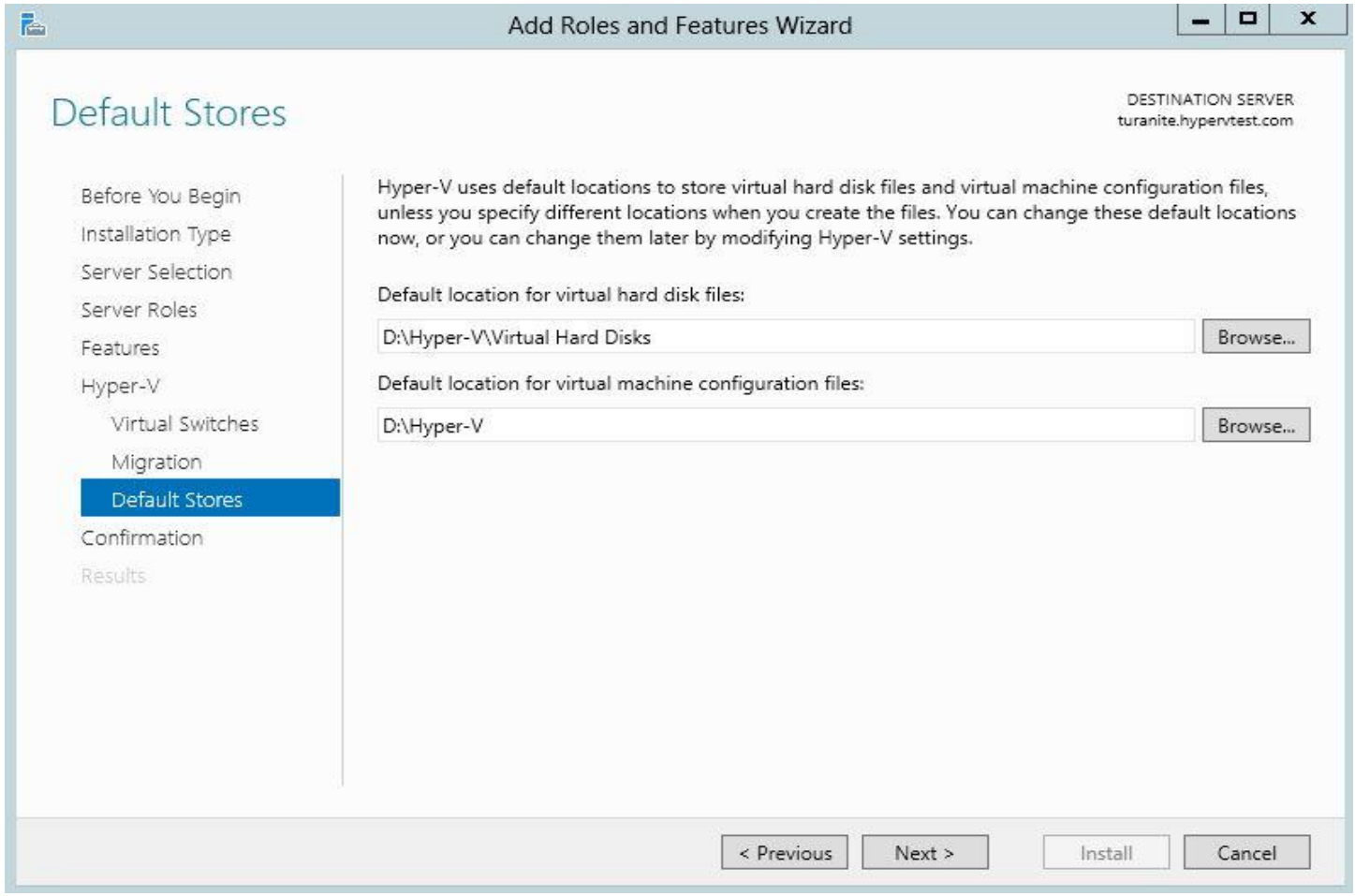


Figure 10: Default Stores

Step 12). On the **Confirmation** page of the **Add Roles and Features** wizard (Figure 11), review the roles, role services, and features you wish to install; when you are finished, click **Install** to begin the installation process.

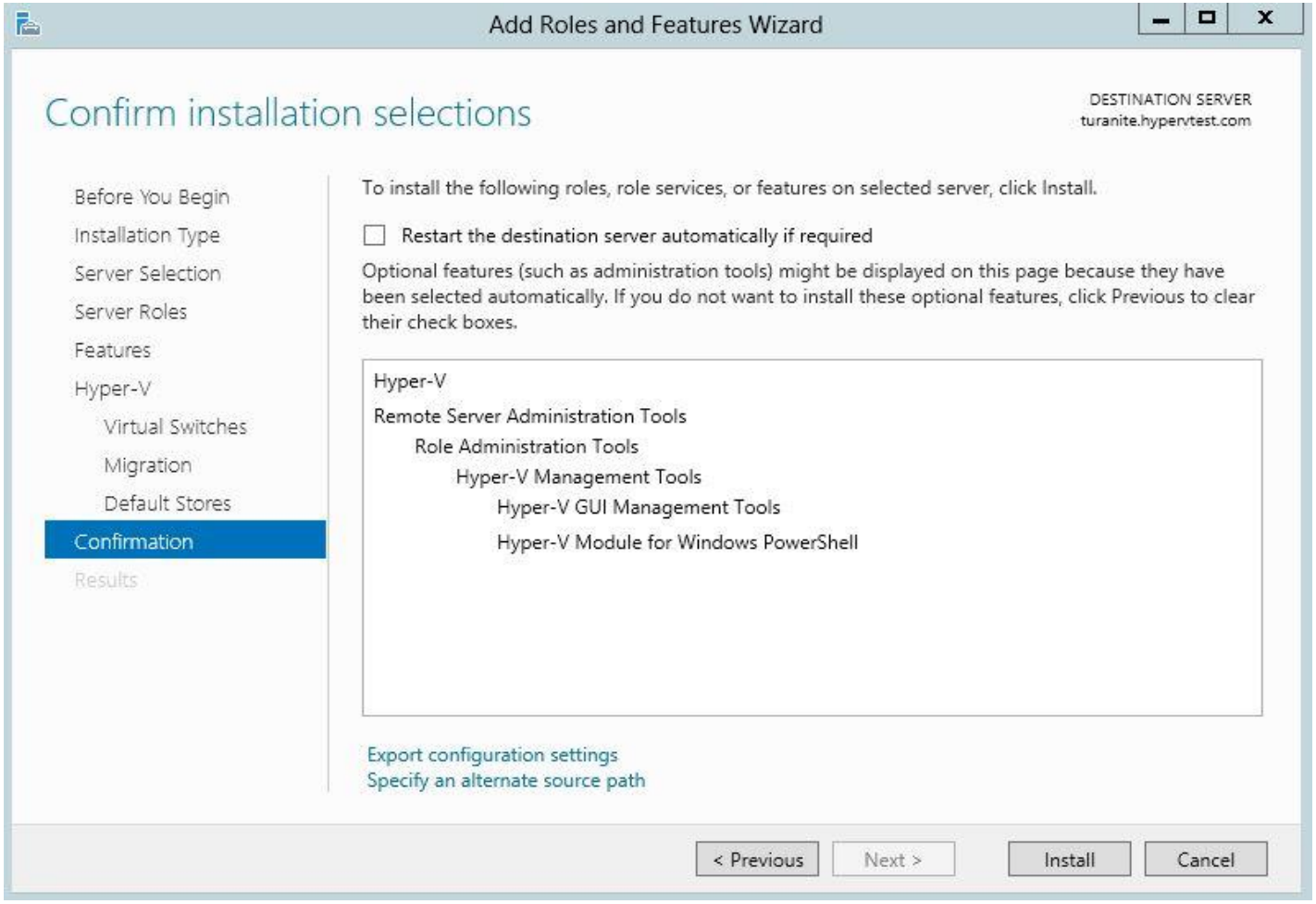


Figure 11: Confirmation

Step 13). After the installation is complete, restart the server on which the **Hyper-V** role has been installed.

Step 14). After the server has restarted, access the **Control Panel**. Open **Administrative Tools** and right-click **Hyper-V Manager**; in the context menu that appears, select **Send to Desktop** to create a desktop shortcut that will simplify access to the **Hyper-V Manager**.

Step 15). Provide each Windows Server 2012 instance with a Fully Qualified Domain Name (FQDN)

This step furnishes each Hyper-V host server with a domain name *without* joining a domain or registering a domain with a domain name server on the Internet.

1. Open the **hosts** file[1] on each Hyper-V host server:
    a. Locate the **hosts** file
    b. Right-click the **hosts** file; in the context menu that appears, click **Open**
    c. In the **Open with** window that appears, select **Notepad** (or an equivalent text editor) and click **OK**
2. In each **hosts** file, map each IP address on the network to the appropriate domain name by appending entries[2] to the **hosts** file. For example:

    192.168.25.108 turanite.hypervtest.com
    192.168.25.100 scottyite.hypervtest.com

3. Append the primary and replica server IP addresses; map them to FQDNs of your choice
4. Save your edits
5. Remove any file extensions appended by the act of saving the **hosts** file[3]

[1] The **hosts** file is typically found at the following location:
C:\Windows\System32\drivers\etc

[2] Entries should not be commented out (there should be no # signs in front of them) and should use the following syntax:
NNN.NNN.NNN.NNN hostname.domainname.com

[3] Note: The act of editing the **hosts** file using a text editor can be problematic because saving your edits will prompt the text editor to assign a *file extension* to the **hosts** file; but the **hosts** file *should not have a file extension;* consequently, if you plan to edit the **hosts** file using a text editor, you will need to *remove* any file extensions assigned to the **hosts** file after editing it. To do so:

1. In the Windows Explorer menu bar, click **Tools > Folder options…**
2. In the **Folder Options** window, click the **View** tab
3. In the **View** tab, *uncheck* the **Hide extensions for known file types** checkbox and click **OK**
4. Rename the **hosts** file; delete the file extension (such as ".txt")
5. In the **Rename** dialogue box that appears, click **Yes**

Step 16). Open the **Windows Server 2012 Server Manager** and select **Local Server** (Figure 12); click on the name of the server in the **Computer name** field to access the **System Properties** window (Figure 13).

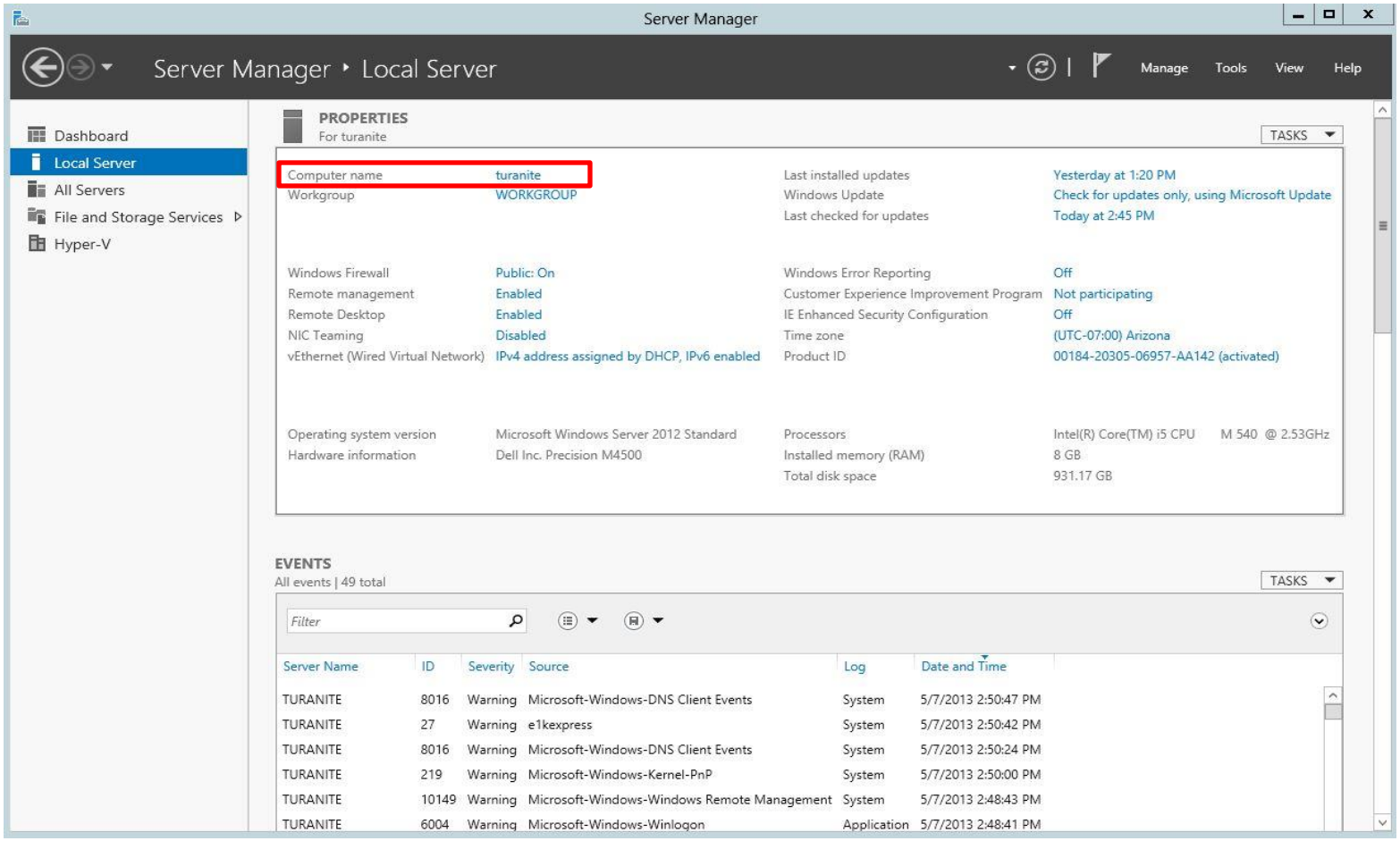


Figure 12: The Windows Server 2012 Server Manager

Step 17). In the **System Properties** Window (Figure 13), click **Change…** to bring up the **Computer Name/Domain Changes** window (Figure 14). Here, click **More…** to bring up the **DNS Suffix and NetBIOS Computer Name** window (Figure 15).

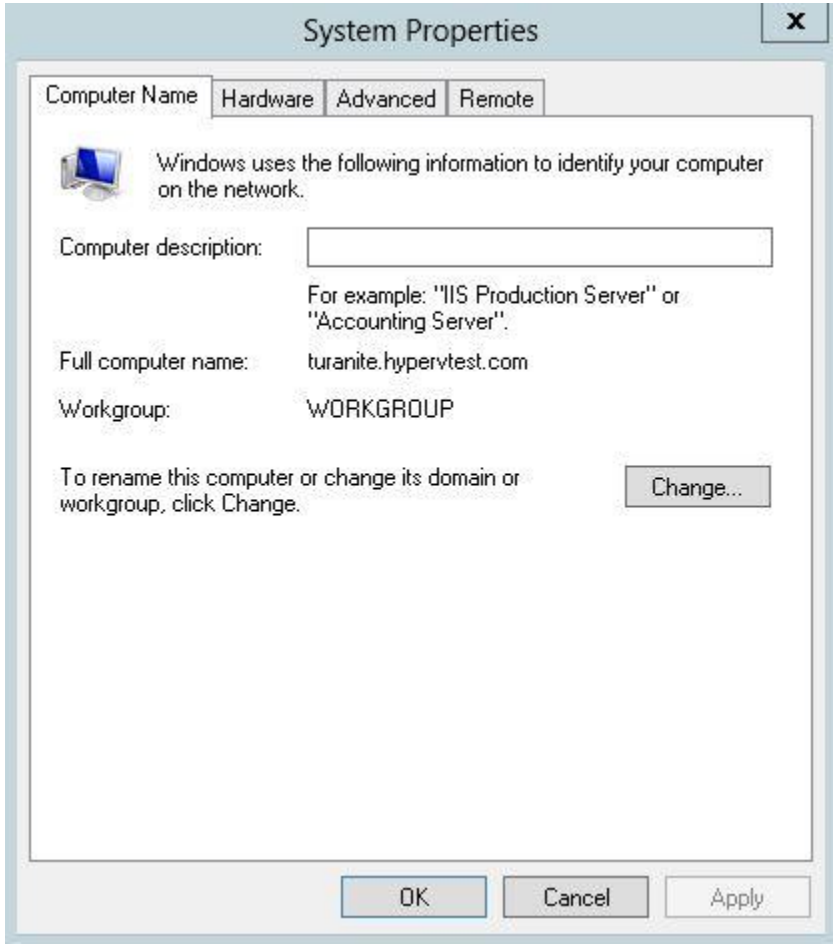


Figure 13: The System Properties window

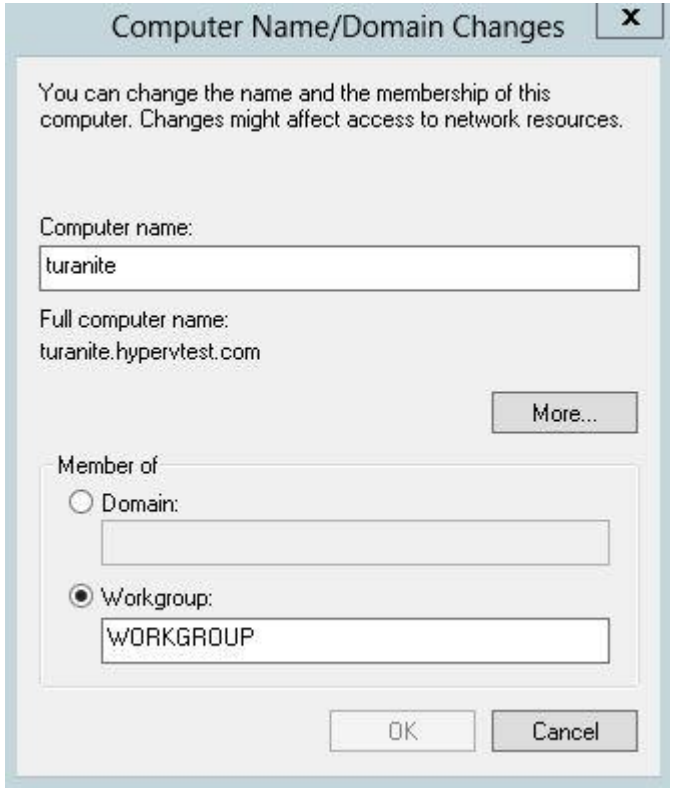


Figure 14: The Computer Name/ Domain Changes window

Step 18). In the **DNS Suffix and NetBIOS Computer Name** window (Figure 15), enter the domain name you established in Step 15 into the **Primary DNS suffix of this computer** field. Click **OK** on all three preceding windows (Figures 15-13).

Step 19). Restart the server. After the server is restarted, the server name will include the FQDN.

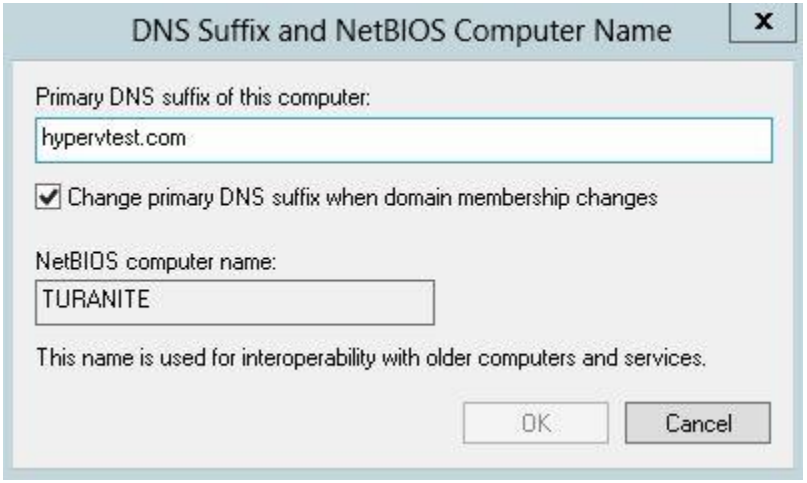


Figure 14: The DNS Suffix and NetBIOS Computer Name window

Step 20). Install the **Windows SDK for Windows 7** on each server

1. Download the **Windows SDK for Windows 7 ISO** (http://www.microsoft.com/en-us/download/details.aspx?id=8442)
2. Install **Windows SDK for Windows 7**

Step 21). Create a directory in which you will store certificate files; this directory will be used later in this tutorial

Step 22). Install SSL certificates for both the primary server and the replica server

1. Open a command prompt
2. On the taskbar, right-click the icon representing the command prompt you just opened; in the context menu that appears, right-click **Command Prompt**; click **Run as administrator**
3. In the elevated command prompt you just established, run a series of MS-DOS commands; make the substitutions listed below:
    a. Substitutions:
        i. Replace **primaryhostname** with the **FQDN** of the primary server created in Step 15
        ii. Replace **replicahostname** with the **FQDN** of the replica server created in Step 15

b. On the replica server, navigate to the directory you created in Step 21 and run the following commands:

```
makecert -pe -n "CN=ReplicaTestRootCA" -ss root -sr LocalMachine -sky signature -r
"ReplicaTestRootCA.cer"

makecert -pe -n "CN=replicahostname" -ss my -sr LocalMachine -sky exchange -eku
1.3.6.1.5.5.7.3.1,1.3.6.1.5.5.7.3.2 -in "ReplicaTestRootCA" -is root -ir LocalMachine -sp "Microsoft
RSA SChannel Cryptographic Provider" -sy 12 PrimaryTestCert.cer
```

c. On the primary server, navigate to the directory you created in Step 21 and run the following commands:

```
makecert -pe -n "CN=PrimaryTestRootCA" -ss root -sr LocalMachine -sky signature -r
"PrimaryTestRootCA.cer"

makecert -pe -n "CN=primaryhostname" -ss my -sr LocalMachine -sky exchange -eku
1.3.6.1.5.5.7.3.1,1.3.6.1.5.5.7.3.2 -in "PrimaryTestRootCA" -is root -ir LocalMachine -sp "Microsoft
RSA SChannel Cryptographic Provider" -sy 12 PrimaryTestCert.cer
```

4. On both the primary and replica servers, use the **Run** command to run **mmc** (the *Microsoft Management Console*); once in the console, click **File > Add/Remove Snap-in…**
   a. In the **Add or Remove Snap-ins** window, select **Certificates** from the **Available Snap-ins** list
   b. Click **Add >**; the **Certificates snap-in** window will appear
   c. In the **Certificates snap-in** window, click the **Computer account** radio button; click **Next** to continue
   d. In the **Select Computer** window, make sure the **Local computer** radio button is clicked; then click **Finish**
   e. In the **Add or Remove Snap-ins** window, click **OK**
5. In the Microsoft Management Console on the primary server, expose the contents of *Certificates (Local Computer),* which can be found under the **Console Root** directory
   a. Expose the contents of the **Personal** directory; click the **Certificates** directory
   b. Right-click *ReplicaTestRootCA.cer*; in the context menu that appears, mouse over **All Tasks >**; in the submenu that appears, click **Export…**
   c. In the **Certificate Export Wizard** that appears, click the **Yes, Export the private key** radio button and click **Next**
   d. On the next page of the **Certificate Export Wizard**, click **Finish**
6. Copy the **PrimaryTestRootCA.cer** file and the **RecoveryServer.pfx** file you just created from the primary server to the directory you created in Step 21 on the replica server, and run the following command in an elevated command prompt on the replica server:

```
certutil -addstore -f Root "PrimaryTestRootCA.cer"
```

7. In the Microsoft Management Console on the replica server, expose the contents of *Certificates (Local Computer),* which can be found under the **Console Root** directory
    a. Expose the contents of the **Personal** directory
    b. Right-click on the **Certificates** directory; in the context menu that appears, mouse over **All Tasks >**; in the submenu that appears, click **Import**…
    c. Locate the **RecoveryServer.pfx** file
8. Copy the **ReplicaTestRootCA.cer** file from the replica server to the directory you created on the primary server in Step 21, in an elevated command prompt, navigate to the directory you created on the primary server in Step 21 and run the following command:

```
certutil -addstore -f Root "ReplicaTestRootCA.cer"
```

9. On both the primary and the replica servers, run an elevated command prompt, navigate to the directory you created in Step 21, and run the following commands:

```
reg add "HKLM\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Virtualization\FailoverReplication" /v
DisableCertRevocationCheck /d 1 /t REG_DWORD /f

reg add "HKLM\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Virtualization\Replication" /v
DisableCertRevocationCheck /d 1 /t REG_DWORD /f
```

Step 22). Enable replication on the Primary and Replica servers

1. Open the Hyper-**V Manager** (Figure 16)

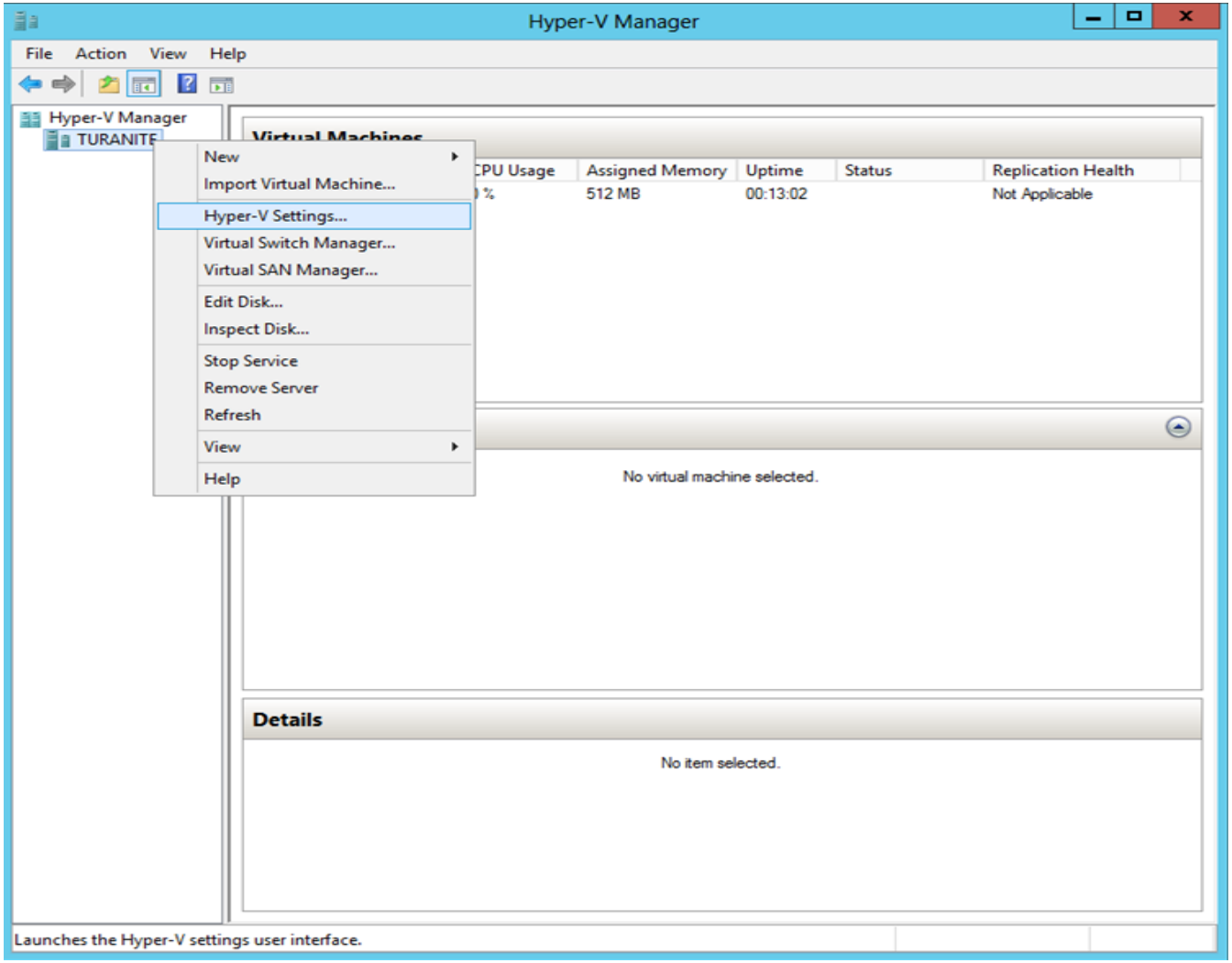


Figure 16: The Hyper-V Manager

2. In the **Hyper-V Manager** (Figure 16), right-click the primary server; in the context menu that appears, click **Hyper-V Settings…**
3. In the **Hyper-V Settings** window (Figure 17), click **Replication Configuration**
    a. Click the **Enable this computer as a Replica server** checkbox
    b. Click the **Use certificate-based Authentication (HTTPS)** checkbox

c. Click the **Select Certificate…** button; specify the SSL certificate you installed in Step X.

d. Click **Apply**; a dialogue box will appear prompting you to open **Port 443** in your local fire wall; *do so*

4. Repeat Steps 21-2 through 21-3 for each Hyper-V host server that you want to enable replication on

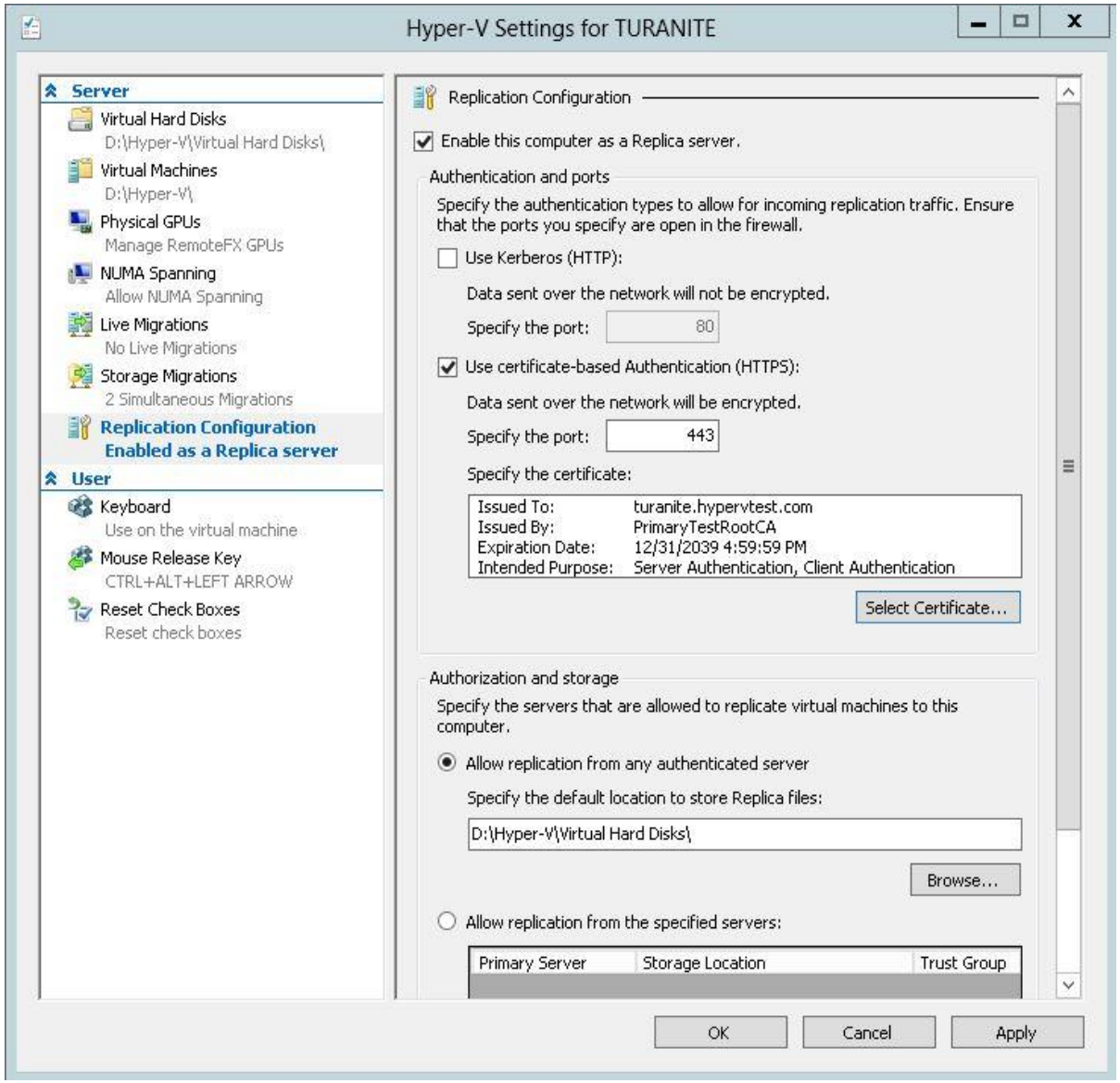


Figure 17: The Hyper-V Settings window

Step 23). Enable replication

1. In **Hyper-V Manager**, right-click the virtual machine that you wish to replicate; in the context menu that appears, click **Enable Replication**. This will bring up the **Enable Replication** Wizard.
2. Read the **Before You Begin** page of the **Enable Replication** wizard (Figure 18); when you are ready, click **Next** to continue

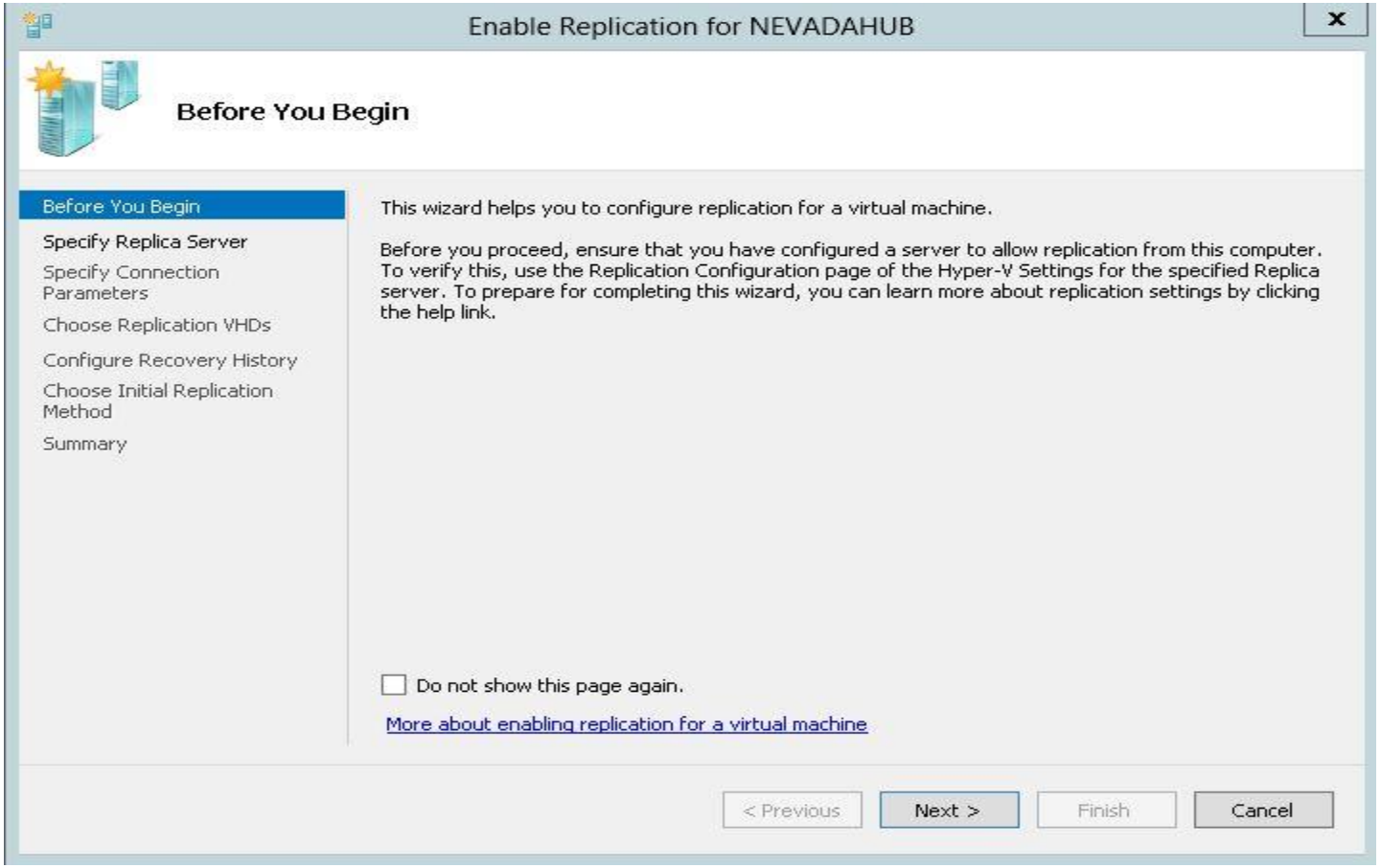


Figure 18: Before You Begin

3. On the **Specify Replica Server** page of the **Enable Replication** Wizard (Figure 19), enter the host name and FQDN of the replica server, as specified in Step 15 and Step 22; click **Next** to continue

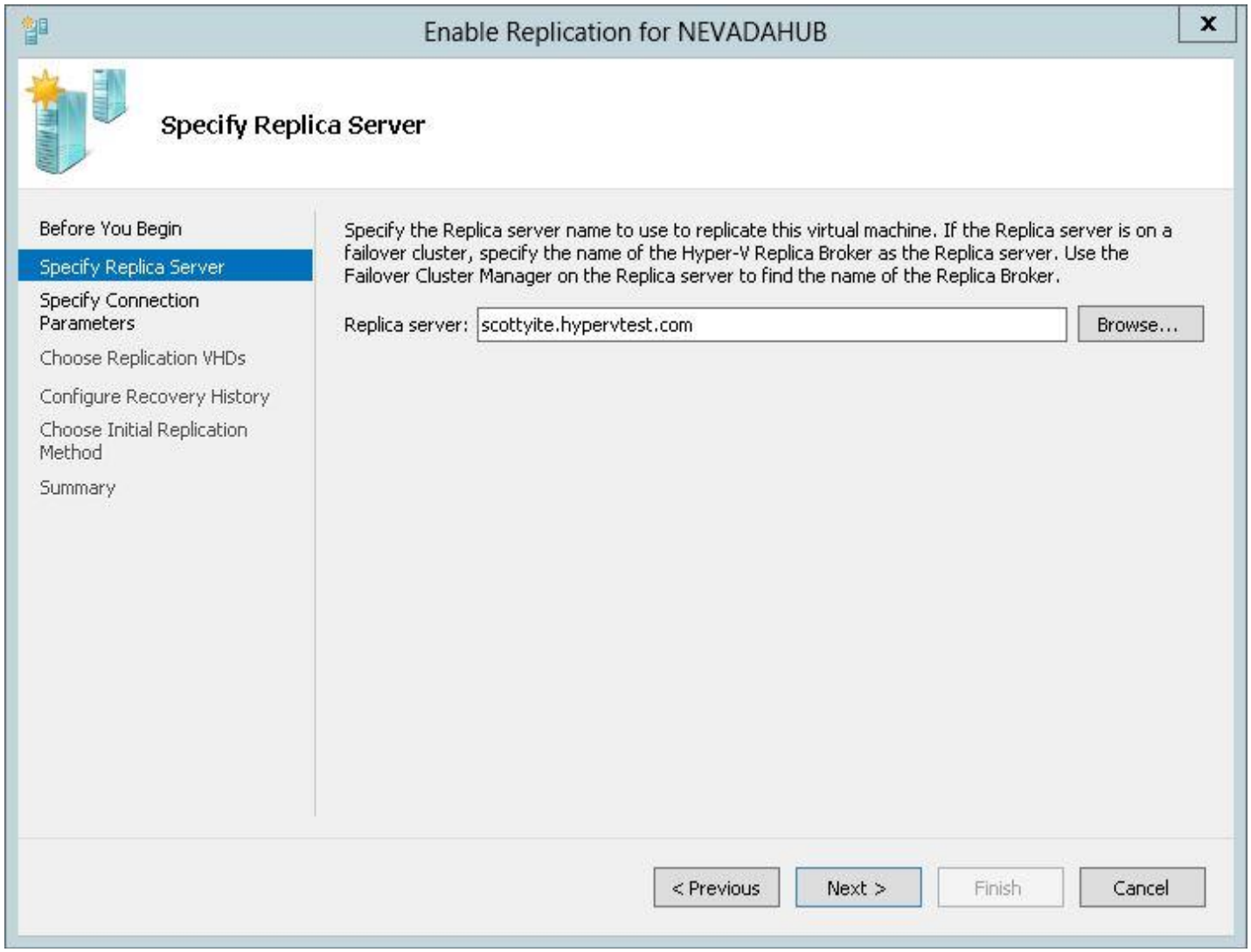


Figure 19: Specify Replica Server

4. On the **Specify Connection Parameters** page of the **Enable Replication** Wizard (Figure 20), click **Select Certificate…** to locate the SSL certificate you created in Step 22; click **Next** to continue

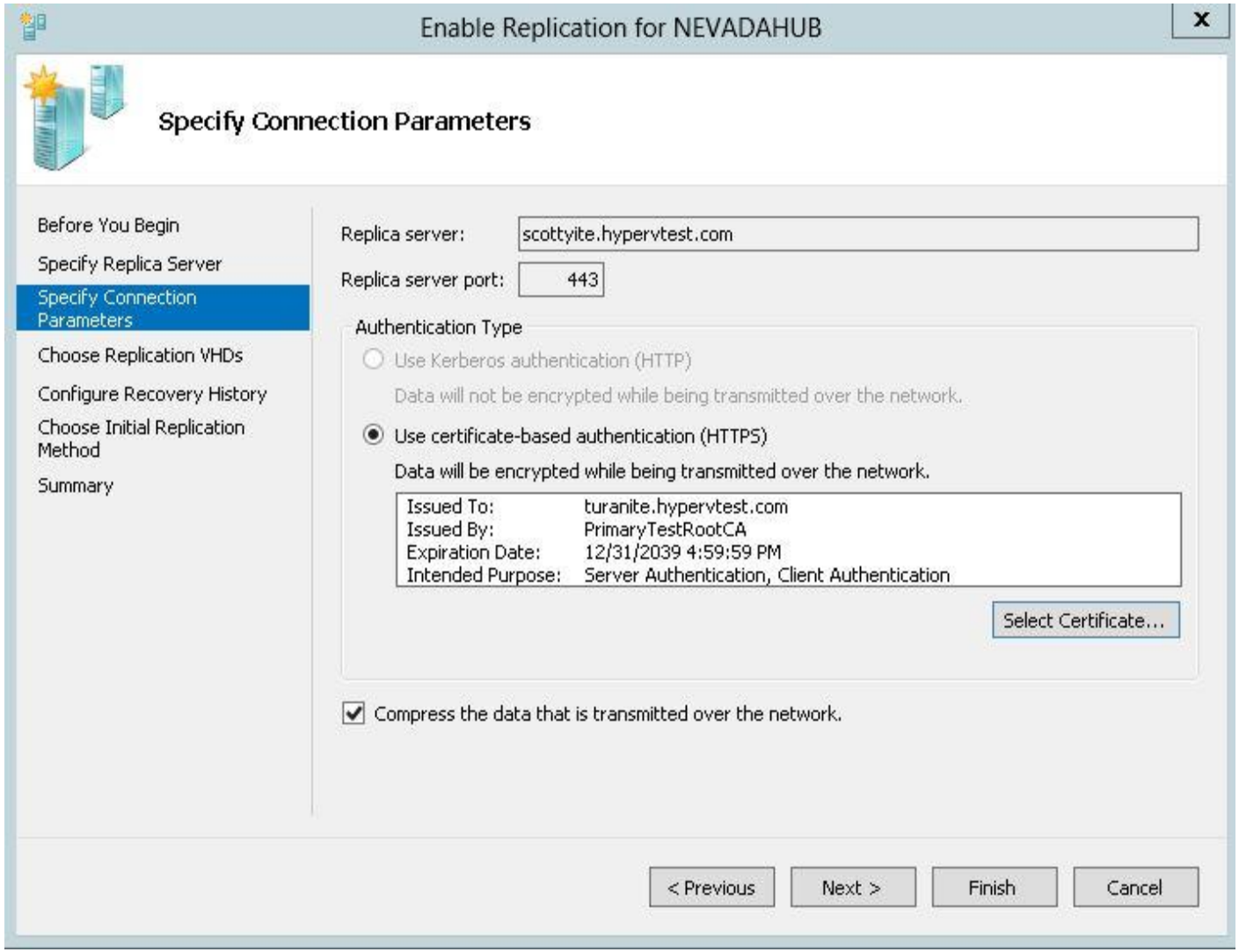


Figure 20: Specify Connection Parameters

5. On the **Choose Replication VHDs** page of the **Enable Replication** Wizard (Figure 21), click the checkboxes indicating the virtual hard disks (VHDs) you wish to replicate; typically, all VHDs are selected. Click **Next** to continue.

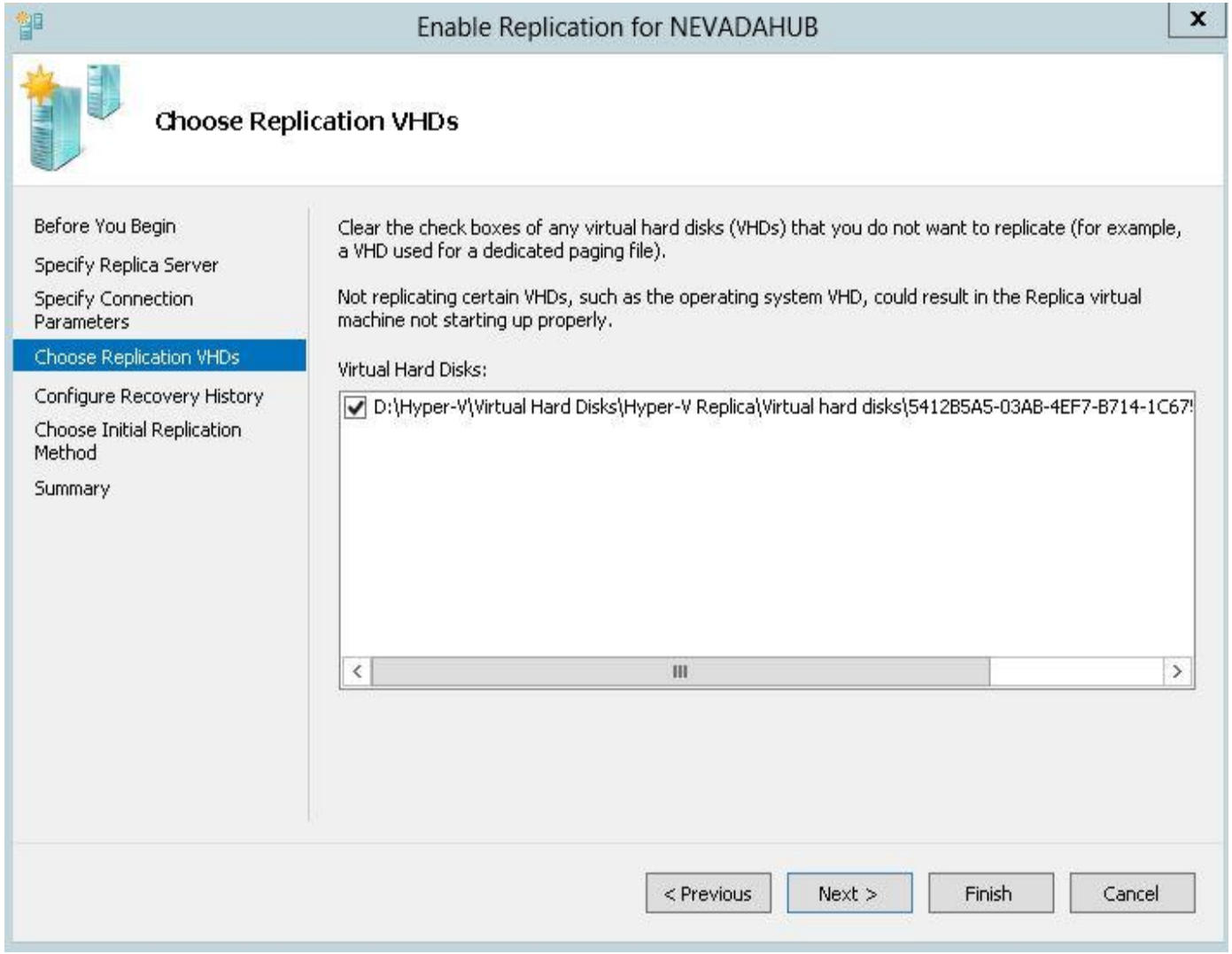


Figure 21: Choose Replication VHDs

6. On the **Configure Recovery History** page of the **Enable Replication** Wizard (Figure 22), specify the recovery points you wish to replicate. In this example, only the latest recovery point will be replicated, but it is possible to specify the replication of additional recovery points.

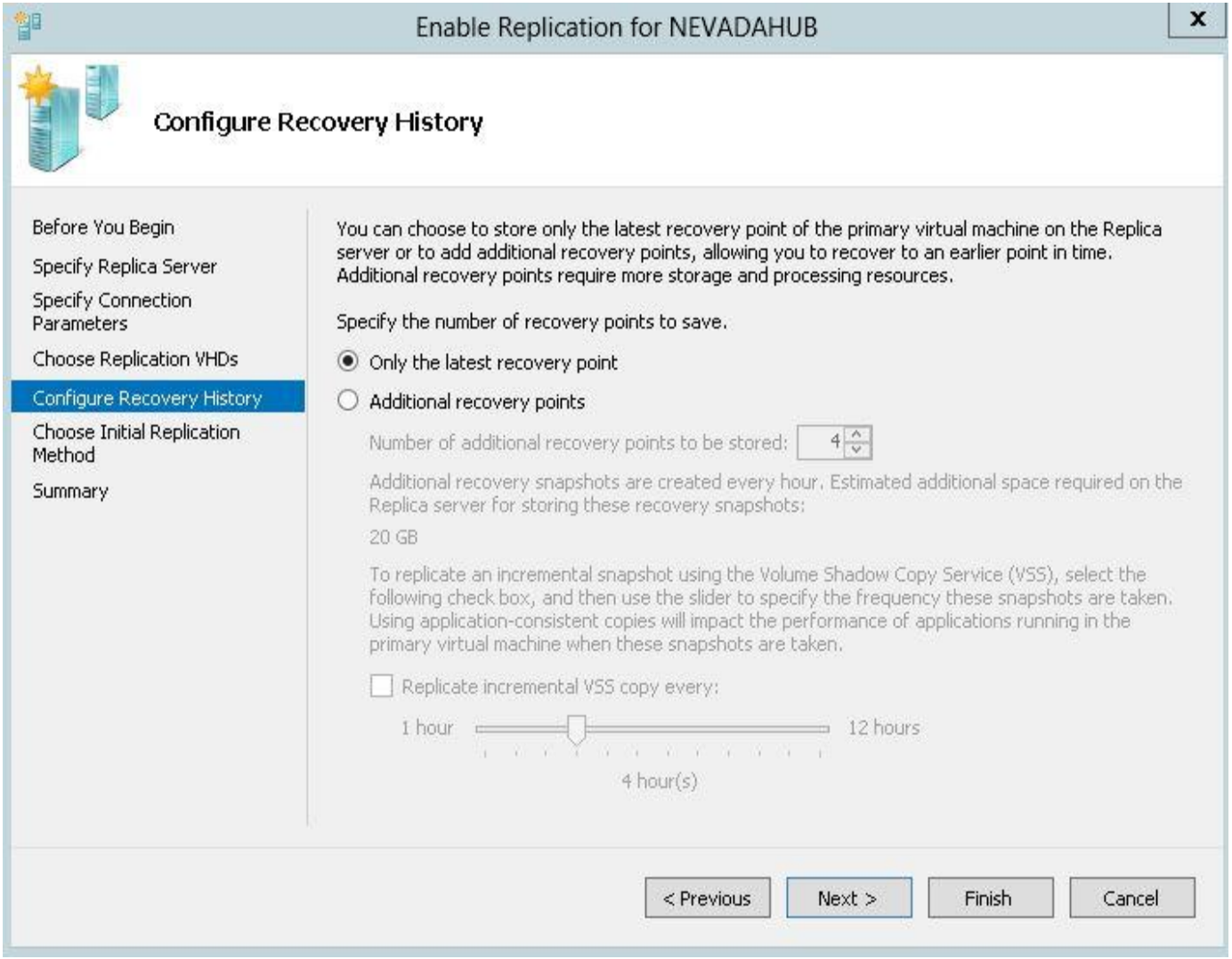


Figure 22: Configure Recovery History

Click **Next** to continue.

7. On the **Choose Initial Replication Method** page of the **Enable Replication** Wizard (Figure 23), click the **Send initial copy over the network** radio button. If the virtual machine being replicated is very large, you might want to consider using external media to seed the initial copy.

   In addition, specify the date and time at which the initial replication will occur. If you elect to start replication immediately, replication will begin as soon as this wizard is finished. Click **Next** to continue.

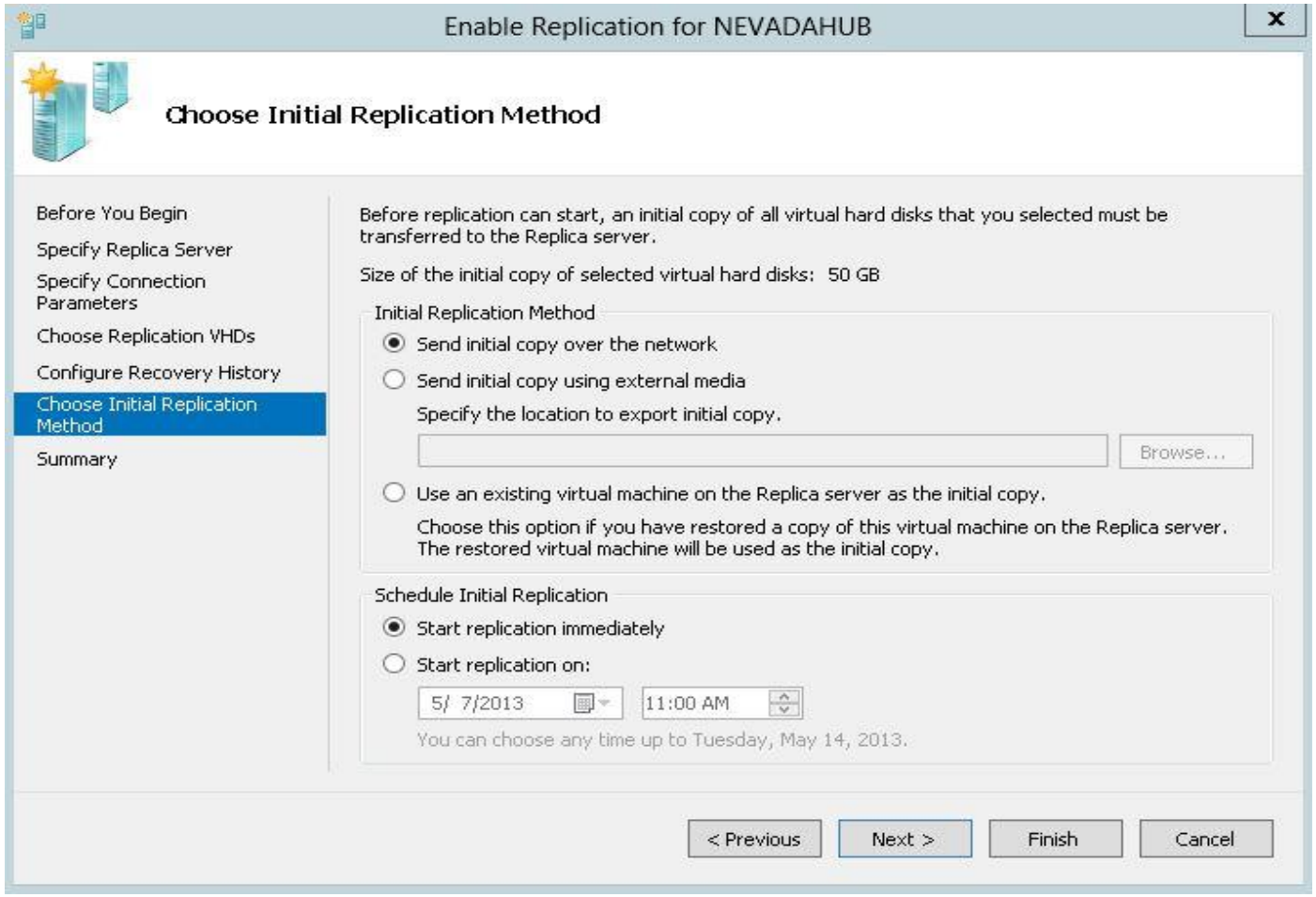


Figure 23: Choose Initial Replication Method

8. On the **Summary** page of the **Enable Replication** Wizard, review your settings and click **Finish**.

Appendix F: Subrecipient Training, Outreach, and Support

1. Workflow Tools at State Geothermal Data site: http://www.stategeothermaldata.org/ (uploaded to OSTI site)
2. Tools and Tutorials
    a. Repository Tutorial……2
    b. Workflow Tools Description……16
    c. Deduplication Workflow……23
3. Subrecipient regional workshop materials
    a. List of training webinars and PowerPoint presentations……26

The AASG Geothermal Data Repository

A Primer

6/18/2012

Table of Contents

Section 1: Introduction 2
Section 2: Context 2
Section 2.1: AASG Geothermal Data Organization 2
Section 2.2: AASG Geothermal Data Architecture 3
Section 2.3: Metadata 3
Section 2.4: Web Services 4
Section 2.4.1: Web Service Flavors 4
Section 3: The AASG Geothermal Data Repository 5
Section 3.1: AASG Geothermal Data Repository Metadata Workflow Summary 5
Section 4: Initial Steps 6
Section 4.1: Data Resources 6
Section 4.2: Collections 6
Section 5: Accessing the AASG Geothermal Data Repository 7
Section 6: AASG Geothermal Data Repository Collections 8
Section 7: Logging In 10
Section 7.1: Creating an Account 10
Section 8: Creating a Metadata Record 11
Section 8.1: Schemas 12
Section 8.1.1: The AASG Geothermal Data Repository Schema 12
Section 8.2: Field List and Schema Requirements 13

Section 1: Introduction

The **AASG Geothermal Data Repository** (http://repository.stategeothermaldata.org/) is a web-accessible digital data repository for metadata records and static or legacy geothermal-relevant documents pertaining to the **AASG Geothermal Data** project (http://www.stategeothermaldata.org/).

Metadata records entered into the **AASG Geothermal Data Repository** are periodically harvested by the **AASG Geothermal Data Catalog** (http://search.stategeothermaldata.org/). Metadata records harvested by the catalog can be used to locate and access the distributed resources described by the metadata records in the catalog. Consequently, the **AASG Geothermal Data Catalog** acts as a metadata nexus from which the distributed resources described by metadata records in the catalog can be located and accessed.

This tutorial describes the process of using the **AASG Geothermal Data Repository** to manage metadata records describing existing data resources.

The first few pages of this document are devoted to the background and context of the **AASG Geothermal Data Repository**. If you have no need for this information, feel free to skip to Section 7.

Section 2: Context

The **AASG Geothermal Data** project exists within the context of the **National Geothermal Data System (NGDS)**, a Department of Energy (DOE)-funded initiative to digitize large amounts of data describing geothermal resources in the United States and host these data in web-accessible and interoperable ways.

The **National Geothermal Data System (NGDS)** is nominally composed of two Department of Energy grants, NGDS I and NGDS II:

- **NGDS I Objective**: design, test, and implement the interface for a nationwide geoscience data network to be populated by geothermal data
- **NGDS II Objective**: Populate the NGDS architecture with geothermal data from geological surveys nationwide

NGDS II constitutes the AASG Geothermal Data project.

Section 2.1: AASG Geothermal Data Organization

With regard to Department of Energy (DoE) **NGDS II** grant, contract No. DE-EE0002850, the Arizona Geological Survey (http://azgs.az.gov/) is the prime contractor, responsible for the following:

- Administration and management of sub recipients (state agencies housing relevant geologic data)
- Quality assurance and quality control of submissions by **NGDS II** sub recipients

Section 2.2: AASG Geothermal Data Architecture

Web-accessible data resources hosted by the **AASG Geothermal Data** project will typically fall into two categories:

- Static, analog, or legacy documents (flat files) that cannot be web-accessible in an interoperable manner; these documents are uploaded directly to digital repositories
- Online, interoperable web services (Section 2.4)

In both cases, *metadata records* (Section 2.4) describing web-accessible **AASG Geothermal Data** data resources will be the principle means by which these web-accessible data resources are located and accessed. These metadata records can be searched via the **AASG Geothermal Data Catalog** (http://search.stategeothermaldata.org/).

The relationship between **AASG GEOTHERMAL DATA** project data resources, metadata records, and the catalog is described in **Figure 1**:

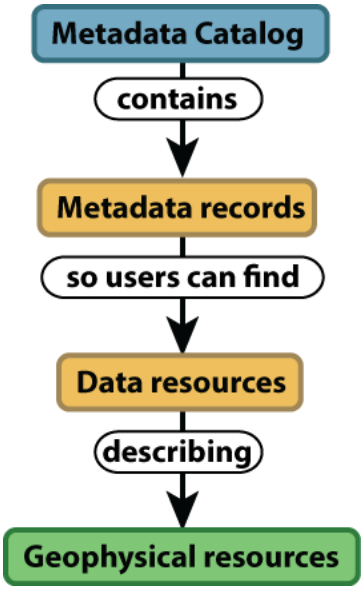


Figure 1: The relationship between the AASG Geothermal Data Catalog and geothermal resources

Section 2.3: Metadata

As indicated above, the **AASG Geothermal Data Catalog** uses metadata records to keep track of registered data resources. These records act like cards in a library catalog: each metadata record in the **AASG Geothermal Data Catalog** describes a discrete data resource. Metadata records in the **AASG Geothermal Data Catalog** provide users with the information necessary to access the data resource described by the metadata record.

For a detailed introduction to metadata, see the USGIN Metadata Tutorial (http://usgin.org/content/usgin-metadata-tutorial/).

Section 2.4: Web Services

Whenever possible, the **AASG Geothermal Data** project uses *web services* to host data resources in an interoperable manner.

Under the client-server relationship, a web service is a server-side protocol that responds to specific requests from client software applications. In response to these requests, web services access appropriately configured web-accessible resources on the server. As long as a request for web-accessible resources uses appropriate syntax, web services can respond. Web service syntax defines required input parameters, operation output, and the results of any operations performed.

Web services act as an intermediary between the client and the server, allowing the two systems to develop independently. This means that regardless of client or server make and model, regardless of changes to content on the server, a web service will be able to respond to requests as long as those requests are made using correct syntax.

Etymological Note: "hosting data as a service" is considered synonymous with "hosting data in accordance with a web service protocol." Likewise, data hosted in accordance with a web service protocol is often referred to as a web service.

Section 2.4.1: Web Service Flavors

The Open Geospatial Consortium has provided specifications for several different flavors of web service that are relevant to geographic information systems, **USGIN**, and the **AASG GEOTHERMAL DATA** project. These include **web map services (WMS)** and **web feature services (WFS).**

Web Map Services (WMS) provide a georeferenced map image within a geographic bounding box. Existing georeferenced map images are acquired by **getMap** requests sent by client software to the web map service; **getMap** requests specify the geographic area of the bounding box and the map that is desired. A variety of other parameters allow control over image size, map projection, and other details. The server responds by returning an image file (typically .tif, .jpg or .png) for the requested area.

For example, a **getMap** request could be used to request a georeferenced satellite photo of Arizona; if the satellite photo exists on the server as a web-accessible image file, the web service will respond to the **getMap** request by providing the client with the desired image.

Web map service image files are often based on a source shapefile or feature class. Map services support **FeatureInfo** requests at point locations; **FeatureInfo** request returns are not standardized and depend on server configuration. Map services are most useful for client applications designed to visually explore geographic relationships.

Web Feature Services (WFS) provide collections of georeferenced features described by attributes. Attributes are essentially a collection of database records that describe the location and characteristics of associated features.

For example a feature service containing data about river systems might provide features with linear geometry representing river segments; each feature might be described by attributes such as average flow rate, width, and depth for the segment.

Feature services are useful for geospatial analysis performed using data describing individual geographically located entities. For example, a client application used to calculate the most efficient route between points in a city utilizing real-time traffic information would require a representation of the streets as features described by attributes.

Owing to the large amount of data associated with WFS, these web services tend to require fast Internet connections.

Section 3: The AASG Geothermal Data Repository

The **AASG Geothermal Data Repository** (http://repository.stategeothermaldata.org/) performs the following functions with regard to the **AASG Geothermal Data** project:

- Stores relevant non-interoperable static, legacy, or analog data resources
- Provides users with an interface that facilitates adding, editing, and managing metadata records that describe **AASG Geothermal Data** project data resources
- Stores metadata records until they are harvested by the **AASG Geothermal Data Catalog** (http://search.stategeothermaldata.org/), from which metadata records can be searched

Consequently, users of the **AASG Geothermal Data Repository** will usually be either uploading relevant data resources or managing metadata records describing existing data resources.

The following sections provide operating instructions for the **AASG Geothermal Data Repository**.

Section 3.1: AASG Geothermal Data Repository Metadata Workflow Summary

The following steps constitute the overall process of managing metadata records within the **AASG Geothermal Data Repository**.

1. Ascertain the existence of a data resource you wish to describe with a metadata record
2. **Browse Collections** to locate and ascertain appropriate collection(s) for your metadata record(s); if appropriate collection(s) do not exist, you will need to create them by clicking **Create a New Resource > Manage Collections**
3. Access the **AASG Geothermal Data Repository** at the following web location:
    a. http://repository.stategeothermaldata.org/
4. Log in to an account with the appropriate permissions to edit the desired collection(s) by clicking **Login**; if you do not have an account, you must create one by clicking **Register** and request the permissions necessary to edit the desired metadata records
5. Create a new metadata record by clicking **Create a New Resource**
    a. Populate the required fields in accordance with the appropriate schema (Section 8.1)
    b. Specify the desired collections for your metadata record

    c. Manage any desired attachments
    d. Link your metadata record to the resource it describes (either a web service or a document within the repository)
6. **Validate** your metadata record against the appropriate schema (Section 8.1)
7. **Save** your metadata record; if you want your metadata record to be visible to the public, click the **Published / Visible to the Public** checkbox

The following tutorial will provide all the necessary information to complete these steps.

Section 4: Initial Steps

Before you begin adding metadata records to the **AASG Geothermal Data Repository**, you will need two things:

- A data resource to describe with your metadata record
- One or more collections to which to add your metadata record when you create it

Section 4.1: Data Resources

AASG GEOTHERMAL DATA metadata records are created to describe, locate, and facilitate access to existing data resources.

A classic example is a library catalog: each card in the catalog is a metadata record describing a data resource in the library (usually, a book). One can use cards in the catalog to search for, and access, the desired book. By comparison, each metadata record added to the **AASG Geothermal Data Repository** is used to search for, and access, the desired data resource.

As indicated in Section 2.2, relevant web-accessible data resources in the **AASG Geothermal Data Repository** fall into two categories:

- Static, legacy, or analog documents uploaded to an online repository
- Web services (described in Section 2.4)

Consequently: in order for you to create a metadata record describing a web-accessible data resource, you must have a web-accessible data resource that will be described by your metadata record; this resource can be either a web service or a document available from a digital repository.

Section 4.2: Collections

Metadata records within the **AASG Geothermal Data Repository** are organized by *collection*.

Collections are user-defined groups of metadata records: the **Source Organization** collection groups metadata records by the organization responsible for providing the data resources these metadata records describe. For example, metadata records describing data resources provided by the Arizona Geological Survey are grouped distinctly from metadata records describing data resources provided by the California Geological Survey.

Another example: the **Resource Accessibility** collection groups metadata records by the methods by which described data resources can be accessed, e.g., metadata records describing downloadable data resources are grouped distinctly from metadata records describing ESRI web services.

There are three important things to remember about collections:

1. Collections can be *nested*: it is possible for collections to exist within collections within collections (and so on, *ad infinitum*), much like folders can be created within folders on a computer
2. A single metadata record can be listed under *multiple* relevant collections. For example, a metadata record describing a **web map service (WMS)** provided by the Arizona Geological Survey could be listed under at least two collections simultaneously: the **WMS Services** collection (a sub-collection of the **Resource Accessibility** collection) and the **Arizona Geological Survey** collection (a sub-collection of the **Source Organization** collection).
3. Removing a metadata record from a collection by clicking **Delete** does not delete the metadata record; it simply removes the record from the collection from which it was deleted, much like removing something from a list. If a metadata record thus deleted does not belong to any collections, it remains in the repository.

When creating a metadata record within the **AASG Geothermal Data Repository**, it is important to be aware of whether or not your metadata record will fall within existing **AASG Geothermal Data Repository** collections.

For a complete list of collections, click **Browse Collections** on the **AASG Geothermal Data Repository** website (link below; see **Figure 2** below for a screenshot).

Section 5: Accessing the AASG Geothermal Data Repository

To access the **AASG Geothermal Data Repository**, open a web browser and visit the following web location:

http://repository.stategeothermaldata.org/

The website will appear as follows:

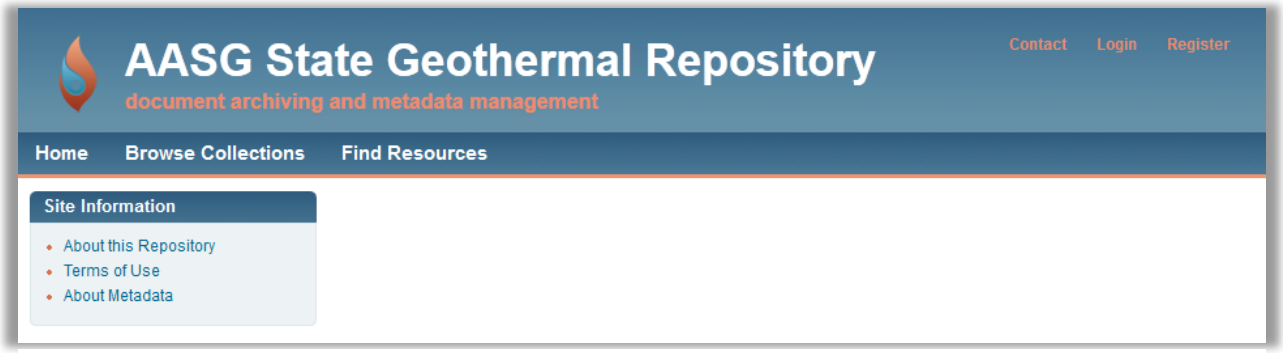


Figure 2: The AASG Geothermal Data Repository home page

From the **AASG Geothermal Data Repository** home page (**Figure 2**), you can:

- Log in to an existing AASG GEOTHERMAL DATA Repository account via the **Login** button
- Create an AASG GEOTHERMAL DATA Repository account via the **Register** button; registration is free
- Browse existing collections; for more information about collections, see Section 4.2
- Search for metadata records within the repository
- Send a message to the site administrator via the **Contact** button

In order to add or edit metadata records within the repository, you will need an **AASG Geothermal Data Repository** account with the appropriate permissions. Registration is free.

Permissions limit the agency and authority of a given account by delimiting the operations an account can perform, most prominently the extent to which a given account can edit or delete existing resources in the repository. When your account is created and confirmed by the web administrator, your account will be furnished with default permissions; to receive more expansive permissions, you will need to contact the site administrator.

Section 6: AASG Geothermal Data Repository Collections

To view existing **AASG Geothermal Data Repository** collections, click **Browse Collections** on the **AASG Geothermal Data Repository** home page (**Figure 2**).

As indicated in Section 4.2, collections are user-generated groups of metadata records within the **AASG Geothermal Data Repository**. A single metadata record may fall within multiple collections. Collections are also nested: collections may exist within collections. Consequently, **AASG Geothermal Data Repository** collections are organized hierarchically.

The top three **AASG Geothermal Data Repository** collections, and their sub-collections, are as follows; sub-collections are indented:

- **AASG Geothermal Data**: Metadata records describing datasets collected as part of the AASG Geothermal Data System
    - **By Data Type**: Records organized by the type of dataset described
        - **Active Faults**: Records describing data resources that describe active faults
        - **Aqueous Chemistry**: Records describing data resources that containing the results of water chemistry analyses
        - **Borehole Temperature Observations**: Records describing data resources containing the results of borehole temperature analyses
        - **Direct Use Features**: Records describing data resources describing direct-use geophysical resources and sites
        - **Drill Stem Tests**: Records describing data resources containing the results of drill stem tests
        - **Fault Features**: Records describing data resources that describe fault features

        - **Geologic Contact Features**: Records describing data resources that describe geologic contacts
        - **Geologic Unit Features**: Records describing data resources that describe known geologic units
        - **Geothermal Data Metadata**: Records describing other metadata records, themselves describing geothermal data resources
        - **Heat Flow**: Records describing data resources that describe geothermal heat flow
        - **Lithology Logs**: Records describing data resources that contain lithology data
        - **Miscellaneous Geologic Maps**: Records describing geologic maps that do not fall into other categories
        - **Seismic Event Hypocenters**: Records describing data resources that describe documented seismic hypocenters
        - **Thermal Conductivity**: Records describing data resources that describe thermal conductivity
        - **Thermal Springs**: Records describing data resources that describe known thermal springs
        - **Volcanic Vents**: Records describing data resources that describe known volcanic vents
        - **Well Headers**: Records describing data resources that contain well header information
        - **Well Log Observation Data**: Records describing well logs
    - **By State**: Records describing datasets that describe geophysical resources within a given state; sub-collections that correspond with each state in the United States

- **Resource Accessibility**: Metadata records describing data resources categorized in terms of the mechanisms by which they can be accessed, either online or offline
    - **Downloadable Files**: Records describing data resources that may be downloaded as discrete files
    - **ESRI Services**: Records describing data resources that are available as ESRI web services (see Section 2.4 for more information about web services)
    - **Offline Access**: Records describing data resources that may be accessed offline
    - **OPENDaP Services**: Records describing data resources that may be accessed via OPENDaP requests
    - **WCS Services**: Records describing data resources that are available as **web coverage services** (see the corresponding USGIN Glossary entry for more information about web coverage services)
    - **Web Applications**: Records describing data resources hosted as web applications
    - **WFS Services**: Records describing data resources available as **web feature services** (see Section 2.4.1 for more information about web feature services)

- **WMS Services**: Records describing data resources available as **web map services** (see Section 2.4.1 for more information about web map services)

- **Source Organization**: Metadata records describing data resources categorized in terms of the organization responsible for providing said data resources; sub-collections correspond with data resources provided by specific organizations

It is important to verify the presence of appropriate collections for any metadata record you wish to add to the repository: during the process of creating a new metadata record, you will be prompted to specify any collections to which you wish to add your metadata record. If an appropriate collection does not exist, you will need to contact the site administrator.

Section 7: Logging In

To log in to the **AASG Geothermal Data Repository** with an existing account, access the **AASG Geothermal Data Repository** (http://repository.stategeothermaldata.org/) with a web browser and click **Login** (Figure 2) and enter your username and password. If you do not have a username and password, you will need to create an account (Section 7.1).

After you log in, new options will become available on the home page (**Figure 3**):

Figure 3: The AASG Geothermal Data Repository home page after logging in

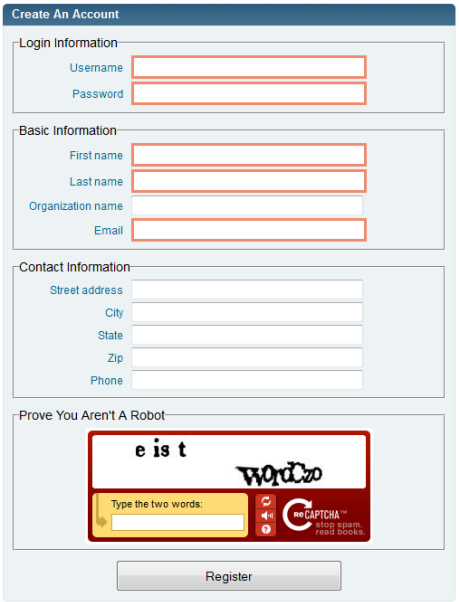


Figure 4: The AASG Geothermal Data Repository account registration form

As indicated in **Figure 4**, you may now create new metadata records by clicking **Create a New Resource**; you may also import metadata records provided at external locations by clicking **Harvest Remote Metadata**.

Section 7.1: Creating an Account

To create an **AASG Geothermal Data Account**, click **Register** on the home page (**Figure 2**) and fill out the AASG GEOTHERMAL DATA Registration Form (**Figure 4**); fields surrounded by an orange box are required. Registration is free.

At the bottom of the registration form is a Captcha™ challenge-response test designed to prevent automated software from generating spurious **AASG Geothermal Data Repository** accounts. Human operators must prove they are human by interpreting the deliberately obscured graphical representation and providing a text-based rendition thereof.

When you have populated the registration form and provided the appropriate response to the Captcha test, click **Register** to continue.

Section 8: Creating a Metadata Record

After logging in to an existing account (Section 7), click **Create a New Resource (Figure 3)** and fill out the **New Resource** form (Figure 5) to create a new metadata record.

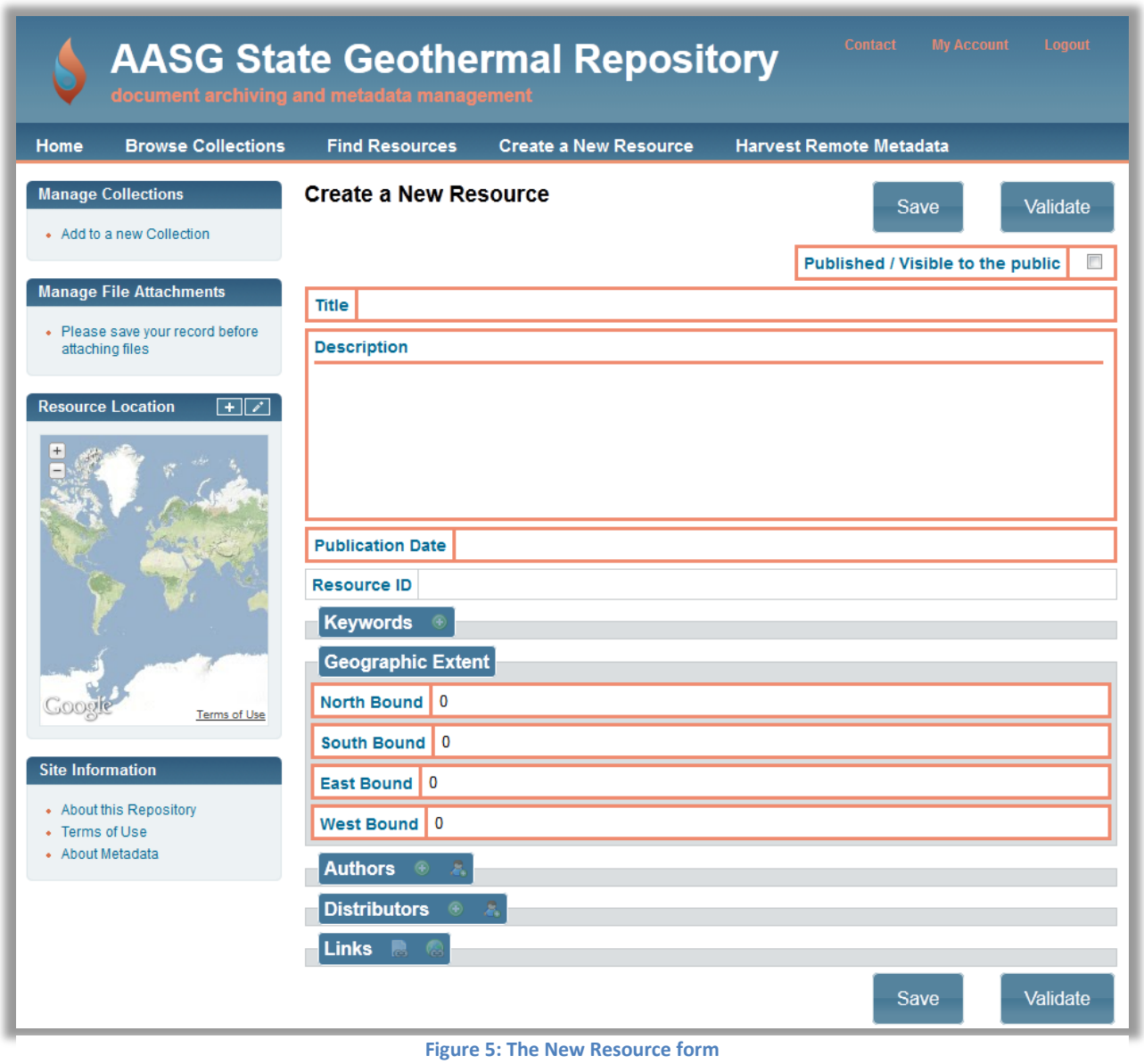


Figure 5: The New Resource form

Fields surrounded by an orange box are required.

When you have populated all required fields (see Section 8.2 for a complete list of fields), click **Validate** to determine whether or not you have populated the fields in the **New Resource** form in accordance with the appropriate schema (see Section 8.1). If your metadata record is not valid, **Validation Errors** requiring correction will appear above the metadata form. Make all necessary corrections until your metadata record is populated correctly.

After you have successfully validated your metadata record, click the **Published / Visible to the public** checkbox. If you have not completely populated the form, you can save without publishing; your metadata record will remain in the repository, but it will be only visible with the appropriate permissions. Click **Add to a new collection** to choose the **AASG Geothermal Data Repository** collection(s) to which you wish to add your metadata record (Section 6); **Save** your metadata record.

Section 8.1: Schemas

From a practical perspective, schemas structure documents.

For example: a database schema is a description of the tables and fields in a database, describing the specific properties of each field and table in the database: the nature of the data that should be entered into each table or field, and the manner in which this data should be formatted.

Database schemas thereby dictate where and how data should be entered into a database.

A database record containing data that has not been entered in accordance with the appropriate schema, such as data that has not been entered into the appropriate column or formatted in the appropriate way, is invalid.

For example, if a database table were structured so as to contain a **DATE** field, in which dates should be entered according to the format **mm/dd/yyyy**, a database record would not be valid if the **DATE** field contained *anything but* a date entered according to the **mm/dd/yyyy** format. Likewise, in order for a given database record to be valid, the **DATE** field must contain *only* dates entered according to the specified format (**mm/dd/yyyy**).

Section 8.1.1: The AASG Geothermal Data Repository Schema

As **AASG Geothermal Data Repository** metadata records are effectively database records, **AASG Geothermal Data Repository** metadata records are structured by a schema. The schema used by the **AASG Geothermal Data Repository** to structure metadata records is a custom JSON-schema developed specifically for the **AASG Geothermal Data Repository**.

Briefly, the **AASG Geothermal Data Repository** schema indicates the following validation rules:

- Any orange input boxes on the form are required and must be populated
- Data entered into the **Description** field must be at least 50 characters long
- The **Author** field must contain at least one author
- The **Distributor** field must contain at least one distributor

- Any contact must list either a **Name** or **Organization Name**, or both

In order for a given **AASG Geothermal Data Repository** record to be valid, the above conditions must be met.

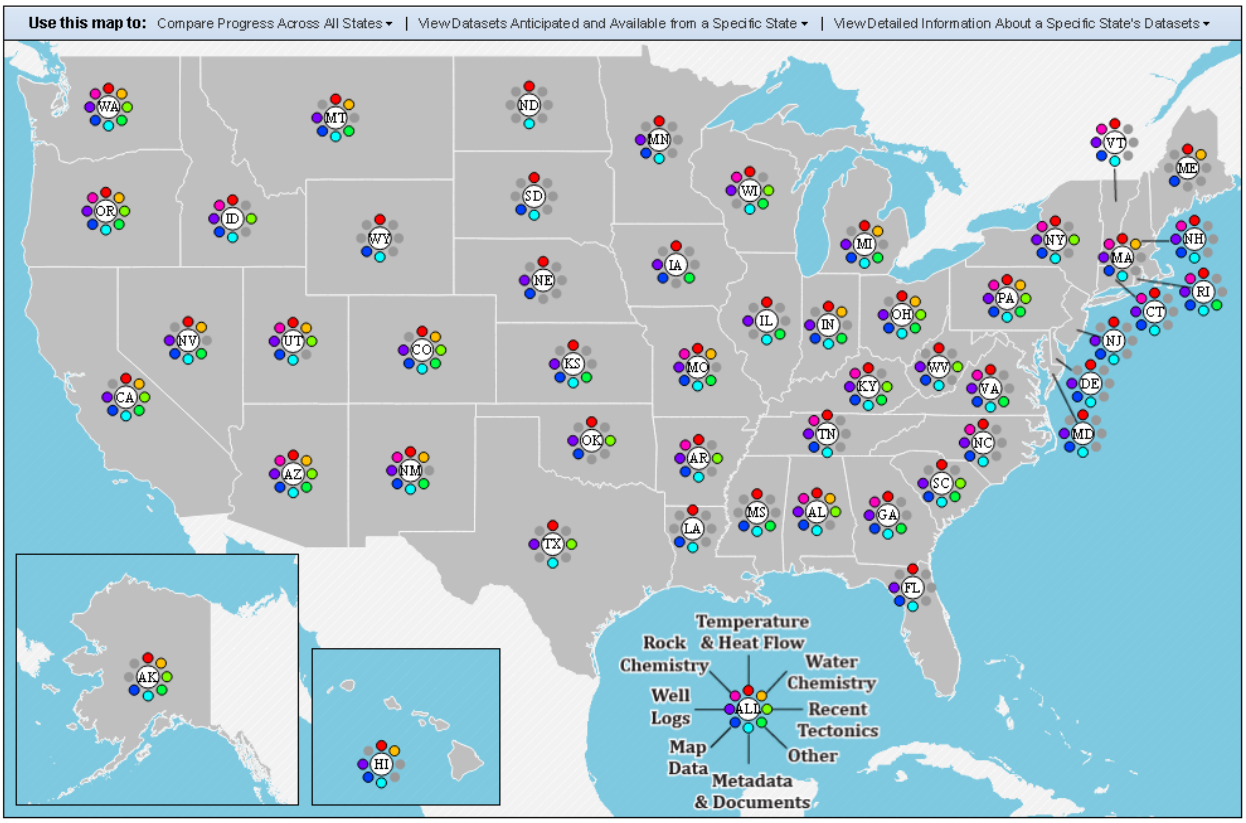
Complete documentation for the **AASG Geothermal Data Repository** custom JSON-schema may be found at the following web location: https://github.com/usgin/metadata-server/blob/master/src/schemas.coffee

Section 8.2: Field List and Schema Requirements

1. **Title**: The name of the data resource described by your new metadata record
2. **Description**: An abstract describing the data resource described by your metadata record
    a. The description must be at least 50 characters long to be valid
3. **Publication Date**: The date on which the data resource described by your metadata record was published
4. **Resource ID**: Any identifier used to systematically label the data resource described by your metadata record
5. **Keywords**: Any keywords used to describe the data resource described by your metadata record
6. **Geographic Extent**: A rectangular geographic area completely surrounding any physical resources described by the data resource described by your metadata record. For example: if the data resource described by your metadata record constituted a map of boreholes in Arizona, then the geographic extent of the data resource would be a box reflecting the entire geographic distribution of these boreholes. A geographic extent may be entered manually by providing lines of latitude and longitude (in decimal degrees) that correspond with the geographic extent of any geophysical resources described by the data resource described by your metadata record; or you may provide a bounding box my clicking the **Pencil** button above the map on the left side of the **New Resource** form (**Figure 5**) and boxing in the extent of features on the interactive map.
7. **Authors**: Anyone personally responsible for creating the data resource described by your metadata record
    a. There must be at least one author listed
8. **Distributors**: Any party responsible for publishing, housing, or distributing the data resource described by your metadata record
    a. There must be at least one distributor listed
9. **Links**: Any hyperlinks that can be used to access the data resource described by your metadata record. Links can either be *direct links* to static data resources stored in a digital repository; or they can be links to live web services.

Deliverables Management Site

This map provides the ability to view data collection status across all 50 states.



Data being provided by each state has been categorized into eight different themes, as indicated by the legend near the bottom of the map. Clicking on one of the colored dots on the legend will cause many of the states to become colored. The range of colors that appear indicate how far along each state is in the completion of their data deliverables in the theme that was selected.

Over each state, a colored dot indicates that the state intends to provide data falling under that particular theme, while a grey dot means no data in that theme is currently anticipated. By clicking one of the colored dots on a state, you can view the specific deliverables expected from that state in the selected theme, and the progress of those deliverables.

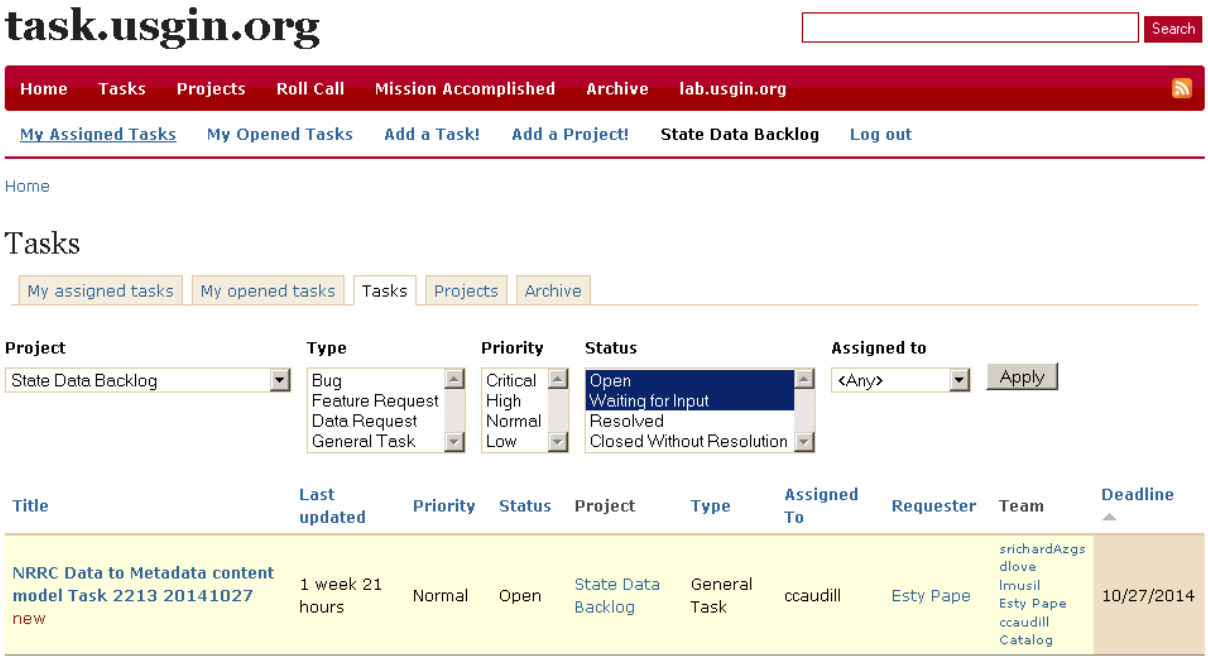
Location: http://stategeothermaldata.org/content/interactive_data_contribution_map

Data and Metadata Processing Management and Workflow Tasks Site

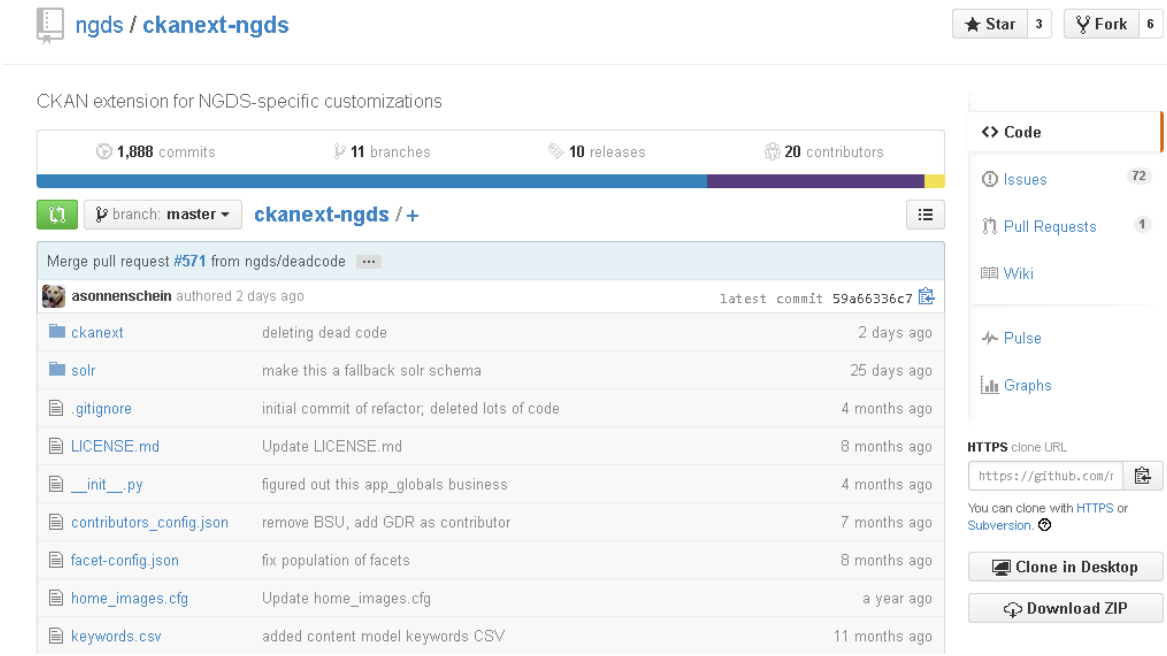
Task management for NGDS took place at two different locations:

- http://task.usgin.org/
- http://github.com/ngds

At the USGIN Task Management website, emergent tasks are announced on a peer-to-peer basis and resolved as opportunity permits by the users to whom announcements are sent.



On GitHub, tasks are initiated on the basis of issue reports, pull requests, and the presence of branches or recent commits.



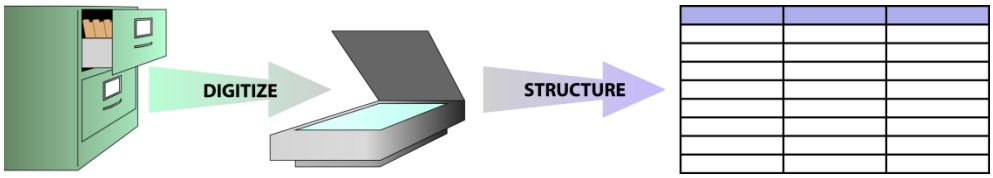
Much of the development of GIN-Stack (also referred to as the NGDS CKAN Extension and the Node-in-a-Box) and the NGDS Catalog took place over GitHub.

Documentation and forum discussions also exist to facilitate workflow and best practices.

Documentation

USGIN Data Provider Workflow

USGIN data providers are technical users who have data they wish to share on a data-sharing network created according to USGIN specifications. This page provides a brief overview of the workflow undertaken by prospective USGIN data providers. For more information about performing these tasks, see the USGIN Best Practices page.



Step 1: Data Assessment

If you are interested in becoming a USGIN data provider, consider the following questions:

1. Do you have geoscience data to share?
2. Who is the intended audience or consumer of your data?
3. Do you have any computing resources at your disposal to make your data web-accessible?
4. Can you implement your data as a live, interoperable Open Geospatial Consortium (OGC) web service?

a. Which tier of interoperability does your data occupy?
b. What steps must be taken before your data conform to the third tier of interoperability?

If you have data that meets the above conditions, move on to Step 2.

Step 2: Data Integration

Data integration is the process of preparing your data for deployment as part of a USGIN data-sharing network. To do so, your data must conform to the third tier of interoperability. Typically, data integration will involve some combination of the following:

- **Digitization**
- **Schema mapping**: mapping your data from its native schema to an interchange format that is designed for maximum interoperability
- **Vocabulary mapping**: furnishing your dataset with definitions for all specialized vocabulary terms used in the dataset; ideally, this is accomplished by using identifiers in your dataset that link to a web-accessible vocabulary service

Step 3: Data Deployment

Data deployment is the process of publishing your data as part of a USGIN data-sharing network such as the National Geothermal Data System (NGDS).

To deploy your data as a geospatial web service:

1. Secure access to the following:
    a. A server; examples of web server software include:
        - Actively maintained free-and-open-source web server software:
            - Apache HTTP server (http://httpd.apache.org/)
            - Nginx (http://www.nginx.org)
        - Proprietary web server software:
            - Windows Server 2012 (http://www.microsoft.com/en-us/server-cloud/windows-server/default.aspx)
    b. Database software compatible with geospatial web service; examples include:
        i. Actively maintained free-and-open-source database software:
            - MySQL (http://www.mysql.com/)
            - PostgreSQL (http://www.postgresql.org/)
                - The **PostGIS** extension for PostgreSQL (http://postgis.refractions.net/) adds support for geographic objects to PostgreSQL
        ii. Proprietary database software:
            - Oracle Database (http://www.oracle.com/technetwork/database/enterprise-edition/downloads/index.html)
    c. Web service server software; examples include:

i. Actively maintained free-and-open-source web service server software:
    - Geoserver (http://geoserver.org/display/GEOS/Welcome)

ii. Proprietary web service server software:
    - ESRI ArcGIS Server (http://www.esri.com/software/arcgis/arcgisserver)

d. GIS software (optional)
    - Actively maintained free-and-open-source GIS software
        - GrassGIS (http://grass.osgeo.org/)
        - GvSIG (http://www.gvsig.org/web/)
        - Quantum GIS (http://www.qgis.org/)
        - uDIG (http://udig.refractions.net/)
    - Proprietary GIS software:
        - ArcGIS (http://www.esri.com/software/arcgis/arcgis-for-desktop)

2. Configure your server
3. Deploy your data on the server
4. Create a metadata record describing your data; enter this metadata record into the catalog of the data-sharing network to which your data is being submitted. If you are creating, rather than joining, a data-sharing network, you should create and configure a catalog service to manage access to your metadata records.

See the USGIN Best Practices page to read about optimal procedures for each step in this workflow.

USGIN Best Practices

This page contains a list of best practices for each step of the USGIN Data Provider Workflow. Though not mandatory, these practices are highly recommended.

Best Practices for Data Assessment

Best practices for data assessment include the following:

- Computing resources used to share data should be compatible with Open Geospatial Consortium (OGC) web services
- Data should achieve Tier-3 interoperability
- Raster data should be deployable as web map services
- Digital databases should be deployable as web feature services
- Data to be shared should meet the needs of consumers in terms of the breadth, depth, and amount of information provided

Best Practices for Data Integration

Data integration can be divided into two steps: **schema mapping** and **vocabulary mapping**.

During **schema mapping**, datasets should be mapped to an interchange format structured by the GeoSciML-Portrayal schema. It is further recommended that individual feature classes in be mapped to GeoSciML-Portrayal Simple Feature Schemas. For more information, see the GeoSciML-Portrayal Cookbook.

The process of **vocabulary mapping** is slightly more complicated. Though the ultimate goal is the provision of web-accessible definitions for each vocabulary term used in a given dataset to facilitate clear communication and understanding, there are several ways to do this, listed in order from most optimal to least optimal:

1. Usage of GeoSciML CGI vocabulary terms, identified by HTTP-URIs that dereference to vocabulary definitions; optimally, these vocabulary terms are accessible as a vocabulary service
    - A web-accessible directory containing the *.rtf and *.ttl files used to create a CGI vocabulary service can be found here
    - A sub-directory containing *.HTML representations of CGI Simple Feature vocabulary term definitions can be found here; these *.HTML files can be used to gain a sense of the way in which the CGI Simple Feature vocabulary terms work. For example, the CGI Simple Lithology Categories *.HTML document contains definitions for the CGI simple lithology vocabulary definitions
2. If GeoSciML CGI vocabulary terms are not used, then any vocabulary terms used in the dataset should be furnished with hyperlinks to a glossary website containing definitions of any and all vocabulary terms used
3. If it is not possible for the user to provide a glossary website, users should to furnish vocabulary terms in their dataset with hyperlinks to one or more generic websites containing definitions for vocabulary terms in the dataset; a website such as Wikipedia is acceptable here
4. If a suitable generic website containing vocabulary terms cannot be found, users are encouraged to add one or more fields to the dataset; when applicable, these fields will contain vocabulary definitions for terms present in each record in the dataset

Best Practices for Data Deployment

Data deployment can likewise be broken down into two parts:

- Hosting web-accessible data via web service server software
- Registering web-accessible data via a metadata record in the USGIN Catalog

A recommended software package is the OpenGeo Suite (http://opengeo.org/), which includes Postgres and PostGIS (database management systems) and OpenLayers (a JavaScript library that provides JavaScript applications with the ability to perform map-related functions, including data consumption from web map services and web feature services).

Deduplication Workflow

Because NGDS is a system of largely independent, federated data providers, it is likely that a given resource (document, dataset, application) will be registered in more than one catalog, and that particular data items might be published in different datasets. The first case is referred to as metadata duplication--multiple metadata records that describe the same resource. The second case is data duplication, in which the same feature or observation is described by records in different datasets.

Ideally metadata records that document the same resource would be merged such that there is only a single metadata record (identified by a unique gmd:fileIdentifier) in the system that documents the resource, and that record would include all the available distributions for the resource. Given the distributed nature of the system and the cost of merging such records and propagating the updates, this is unlikely to happen. Practical options for aggregating catalogs to deal with metadata duplication in harvested metadata include identifying duplicate records and merging them in their metadata collection, or keeping both harvested records and adding an explicit link between them to indicate that they document the same resource.

Data duplication is a trickier issue. Features like wells may be described in multiple databases, with information collected at different times by different observers. Data compiled into a dataset that is to be deployed as a feature service may be sourced from one or more such databases. Different data compilers may assemble data for a feature service from different source databases. The problem is exacerbated by the lack of intrinsic identifiers for things like wells, rock samples, or observations--different observers may use equally valid schemes to identify things for their intended application. Feature location and observation time stamps are about the best thing we have to go on, but those may be reported with different precision or be inaccurate, making establishment of equivalence a probabilistic exercise. Our proposed solution in this situation is to provide a link between records that describe features that might or are known to be the same feature.

This document describes a workflow to identify and correlate duplicate metadata in harvested metadata collections and duplicate data in related feature datasets. We refer to this process as deduplication. Multiple iterations of deduplication will likely be necessary, in which each iteration uses different approaches to detecting duplication.

Identifying Duplicate Data

Iteration 1

With tier 3 data services or documents metadata, a CSV dump of a feature type of interest is an easy way to begin deduplication. As these datasets are standardized, expected field headings exist in all federated catalogs which can be queried for correlation. Correlation of databases can begin in Microsoft Access or SQL, joined based on a foreign key. By successive tests, search for duplication by:

1. IDs (API numbers, labels, other IDs)
2. Titles (if assessing documents metadata)

3. Location information (Latitude, Longitude)
4. Names (well names, operator names, spring names)
5. Dates

This will generate a list of *likely* and *certain* duplicates. Some may remain uncertain due to slight differences in records due to data collection techniques, for example. Location information is inevitably expressed at varying levels of certainty and with a varying number of significant digits. To cull the list of certain duplicates down further, additional iterations of deduplication efforts are at this point necessary.

Iteration 2

Place shapefiles of the list of possible duplicates into ArcMap or another spatial analysis tool. If using ArcMap, perform a Spatial Join between the two datasets. This correlates data from two source databases based on relative spatial location. A series of joins can be performed:

1. The default 'match_option' of a join is INTERSECT, which only matches features based on intersection. These duplicates will likely have been caught in the operations from Iteration 1.3.
2. WITHIN_A_DISTANCE 'match_option' allows features to be joined if they are within a specified distance. This is likely the most useful option for NGDS deduplication based on the location uncertainty that is sometimes associated with field data. This is also the most time-consuming option, which requires the worker to manually inspect the data in matching records to assess if locations of nearby point data are likely to be the same. For example, this may be evidenced by a well drilling pad visible in ArcMap base layers or GoogleEarth imagery near a well header data point.

A spatial join can also be performed in an open-source GIS application such as QGIS. With both datasets loaded into QGIS, from the Vector menu, choose *Data Management Tools* and *Join attributes by location*.

Relating Duplicates

Once a list of *certain* duplicates is found within the system, a correlation table must be set up to capture those relationships. See http://modeling.sourceforge.net/UserGuide/design-rels-many-to-many-principle.html for a visual representation of a correlation table. Since NGDS is structured so that each resource or data point has a URI (unique resource identifier) which represents an individual record, these will be entered into the correlation table and used as foreign keys to any other associated URI. This table will allow complex queries, such as "show everything that is related to x". Finally, the found and indexed related URIs may be indicated as an additional related resource for end users to access easily, being exposed to the foreign key's additional documents and other available data.

NGDS tier 3 data services have a RelatedResouces field which houses additional documents (usually live links) as well as foreign keys to related URIs within the system. For these data types, the related URI can be entered into the actual data in the RelatedResources field. This increases the amount and quality of data for any particular resource or data point available to the end user, accessible by browser.

Webinars documenting how to use, submit and host NGDS data

- AZGS Webinar: “Metadata tools and workflows for AASG Geothermal Data,” February 15, 2011 [32 attendees]
- AZGS Webinar: “Metadata tools and workflows for AASG Geothermal Data,” February 16, 2011 [35 attendees]
- AZGS Webinar: “Web Services for Bottom-Hole Data,” March 16, 2011 [36 attendees]
- AZGS Webinar: “AZGS HUBS – Data and Serving Data,” April 12, 2011 [6 attendees]
- AZGS Webinar: “AZGS HUBS – Data and Serving Data,” April 26, 2011 [4 attendees]
- AZGS Webinar: “Reviewing Data for Online Delivery,” May 17, 2011
- AZGS Webinar: “Reviewing Data for Online Delivery,” May 17, 2011 *https://www.youtube.com/watch?v=IUVA7TKCUJk&index=3&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx*
- AZGS Webinar: Richard, Stephen M. “USGIN and WRP” (linking WRP’s catalog to the GIN network), September 12, 2011
- Webinar: Lehnert, Kerstin. “Introduction to IGSN and SESAR,” March 8, 2012 [31 attendees] *https://www.youtube.com/playlist?list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx*
- Webinar: Kay, Martin, Gosnold, W., & Lovenitti, J., “Geothermal Data: How’s it Used & Why it’s Important”. April 26, 2012. [37 Registered Participants] *https://www.youtube.com/watch?v=CRPfSyjX6cs&index=2&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx*
- AZGS Webinar: “Updated Data Delivery Tools for AASG NGDS,” October 4, 2012.
- USDOE Virtual Workshop: Allison, M. Lee, “The National Geothermal Data System and the Energy-Water Nexus,” Energy/Water Datasets and Models Workshop, U.S. DOE, November 28, 2012.
- Allison, M. Lee, moderator, “Web Accessibility Strategies,” Association of American State Geologists Annual Meeting, Breakout Session, Deadwood, SD, June 10, 2013. LiveStream & YouTube. *http://www.youtube.com/watch?v=_rggaq-hCs8&feature=share&list=UU8wP7yOz8_ebyNkSoUGmdSg*
- Webinar/live presentation: Caudill-Daugherty, Christy, “Excel to NGDS Conversion,” US DOE Geothermal Technologies Office Lunch and Learn Seminar, Washington DC, August 2, 2013.
- Webinar, DOE GTO: AZGS presenters, “NGDS Demo and How to Adopt NGDS Node-in-a-Box (NIAB) for Geothermal and Other Applications,” January 27, 2014.
- Webinar, DOE GTO: AZGS presenters, "How to Utilize the National Geothermal Data System,” January 28, 2014.
- AZGS Webinar: “Final Science Advisory Board Review,” May 29, 2014, Tucson, AZ.

Appendix G: Outreach Materials

1. List and Sample of Publications, Presentations, Media Coverage, and Social Media Sites……2
2. NGDS Atlas……184
3. Access Tools and Tutorials
    a. Scenario card tutorial videos……245
    b. NGDS Data Explorer: http://data.geothermaldatasystem.org/- Submitted to OSTI

State Geothermal Survey Contributions to the National Geothermal Data System (NGDS) Award DE-EE0002850 Education, Outreach, & Training

Contents

Publications & Proceedings 2
Talks/Presentations 3
Online Presentations 12
Exhibits 13
Workshops 15
Testimony/Written Responses 15
Annual Reports 15
Published Technical Reports 16
Awards and Honors 16
News Media Interviews 16
Non-technical Articles 17
Websites, Social Media 18
Subrecipients Presentations, News Interviews, and More 18

Publications & Proceedings

2010

Allison, M. Lee, Richard, S. M., Gundersen, L., & Jackson, I. (2010). *U.S. Geoscience Information Network (GIN) and Convergence towards Global Data Integration in the Geosciences*, Microsoft Environmental Research Workshop 2010, Redmond, Washington. July 14, 2010, (Proceedings, p4, 48p). (Abstract Attached)

Walker, J. D., Gundersen, L.C., & Allison, M.L. (convenors). 2010. *Workshop on Working towards a National Geoinformatics Community (NGC), USGS Denver Federal Center, Denver, Colorado, September 23-24, 2010*, 39p.

2011

Allison, M. Lee, Gundersen, L. & Richard, S.M. (2011). *Geoinformatics in the Public Service: Building a Cyberinfrastructure Across the Geological Surveys*, in Geoinformatics, R. Keller & C. Baru, eds, Cambridge University Press, (pp. 342-349). (Copyrighted material- not attached)

Richard, Stephen M., Clark, R., & Grunberg, W. (2011). *Application of the U.S. Geoscience Information Network to deploying a National Geothermal Data System*, in Geoinformatics, R. Keller & C. Baru, eds, Cambridge University Press, (pp.350-370). (Copyrighted material – not attached)

Anderson, Arlene, Cuyler, D., Snyder, W., Allison, M.L., Blackwell, D. & Williams, C. 2011. *National Geothermal Data System*, Geological Society of America Abstracts with Programs, v. 43, n. 5, abstract 9-10 (Sunday, October 9, 2011, 10:50 AM) (Abstract Attached)

Allison, M. Lee and Gallagher, K.T. 2011. *U.S. Geoscience Information Network: A Critical Path for Data Integration in the U.S. Earth Sciences* [Poster]. Geological Society of America Abstracts with Programs, v. 43, n. 5, abstract 170-10, Tuesday, October 11, 2011, 9:00 AM – 6:00 PM)(Abstract Attached)

Jackson, Ian, Broome, J. & Allison, M.L. 2011. *Delivering Geoscience Knowledge in Federal Systems* Geological Society of America Abstracts with Programs, v. 43, n. 5, abstract 223-10 (Wednesday, October 12, 2011, 11:00 AM) (Abstract Attached)

Love, Diane S., Coleman, C., Pape, E., Clark, R., Richard, S., &Allison, M.L. 2011, State Geothermal Survey Contributions to the National Geothermal Data System: Geological Society of America Abstracts with Programs, v. 43, n. 5, abstract 9-10 (Sunday, October 9, 2011, 11:05 AM)

Patten, Kimberly, Allison, M.L., & Richard, S.M. 2011 U.S. Geoscience Information Network: Distributed Deployment Across 50 States: [Poster] Geological Society of America Abstracts with Programs , Vol. 43, No. 5, abstract 170-11, Tuesday, October 11, 2011, 9:00 AM – 6:00 PM)

2012

Allison, M. Lee, Richard, S.M., Clark, R., Patten, K., Love, D., Celia C., Genhan C., Matti, J., Day, J., Pape, E., & Musil, L. 2012. *Interactive assessment of geothermal energy potential in the U.S.* Proceedings, Thirty-Seventh Workshop on Geothermal Reservoir Engineering. Stanford University, Stanford California, January 30-February 1, 2012. –Paper uploaded to e-link 241.3

Allison, M. Lee, Richard, S.M., Clark, R., Patten, K., Love, D., Coleman, C., Chen, G., Matti, J., Day, J., Pape, E., Musil L., & Caudill C. 2012. *A Geothermal Data System for Exploration and Development*, Geothermal Resources Council *Transactions*, Geothermal Resources Council Annual Meeting, Reno, Nevada, September 30-October 3, 2012. Paper uploaded to e-link 241.3

2014

Anderson, Arlene, Allison, M.L., Richard, S.M., Daugherty C.C., & Patten, K. 2014. *National Geothermal Data System: Case Studies on Exploration and Development of Potential Geothermal Sites Through Distributed Data Sharing*, Geothermal Resources Council Transactions, Vol. 38, P835-838, 2014. Paper uploaded to e-link 241.3

Talks/Presentations

2010

Allison, M. Lee, "AASG Update." AIPG Mid-year Board Meeting. Tucson, AZ. February 12, 2010.

Allison, M. Lee, "AZGS Update." AIPG Arizona Chapter Annual Meeting. Tucson, AZ. February 13, 2010.

Allison, M. Lee, "Towards a Global Data Network for the Geosciences." Society of Mining, Metallurgy, and Exploration Annual Meeting Environmental Division annual luncheon. Phoenix, AZ. March 2, 2010.

Allison, M. Lee, "Building a Global Data Network for the Geosciences." Project Management Institute, Tucson Chapter, Tucson, AZ, March 9, 2010.

Allison, M. Lee, "Building a Data Network for the Geosciences." Arizona Hydrologic Society. Tucson, AZ. April 13, 2010.

Allison, M. Lee, "Geothermal Energy Potential in Arizona." Black Canyon City Rockhounding Group Meeting. Black Canyon City, AZ. April 15, 2010.

Richard, Stephen, "Web services to assemble pieces of a geoscience information network," U.S. Geological Survey seminar. Menlo Park, CA. April 15, 2010.

Allison, M. Lee, "Everything Digital, Online, and Integrated." Arizona Geological Society Meeting. Tucson, AZ. May 4, 2010.

Allison, M. Lee, "Data Integration in the U.S."*OneGeology Initiative Roundtable*. European Geophysical Union Annual Conference, Young Earth Scientists (YES) Network, Vienna Austria. May 7, 2010 by webcast.

Clark, Ryan, Richard, S., & Grunberg, W., "The National Geothermal Data System: The Geoscience Information Network in Action." Digital Mapping Techniques Workshop. Sacramento, CA. May 17, 2010. http://lab.usgin.org/sites/default/files/group/file/u6/NGDSPoster.pdf

Allison, M. Lee, "State Geological Survey Contributions to the National Geothermal Data System." US Dept. of Energy Geothermal Technologies Program Peer Review. Arlington, VA. May 18, 2010.

Allison, M. Lee, "A Digital Revolution in the Geosciences," Association of Engineering & Environmental Geologists (AEG). Tucson, AZ. July 8, 2010.

Allison, M. Lee, Richard, S.M., Clark, R., & Grunberg, W., "The National Geothermal Data System: An Implementation of the Geoscience Information Network." Webservices Blast, 2010 USGS Community for Data Integration Workshop. Denver, CO. August 10-13, 2010.

Allison, M. Lee, "AZGS Contributions to the NGDS," NGDS Project Annual Meeting, Salt Lake City, UT. August 16, 2010.

Allison, Lee, Spencer, J., Young, J., & Richard, S.M., "Digital Data Integration of Uranium Resources of Northern Arizona." Arizona Hydrological Society Symposium, Tucson, AZ. September 1-4, 2010.

Allison, M. Lee, "Perspectives from the U.S. on Data Interoperability." OneGeology-Europe final workshop. Paris, France. October 27, 2010.

Allison, M. Lee, and Richard, S.M., "State Geological Survey Deployment of the National Geothermal Data System." Geological Society of America Annual Meeting, Denver, CO. November 2, 2010

Allison, M. Lee, Richard, S.M., Gundersen, L., & Jackson, I., "U.S. Geoscience Information Network (GIN) and Convergence towards Global Data Integration in the Geosciences." Geological Society of America Annual Meeting, Denver, CO. October 30 - November 3, 2010,

Allison, M. Lee, "A Digital Revolution in Resource Exploration," Society of Mining, Metallurgy, and Exploration Tucson Chapter dinner meeting, Tucson, AZ. November 10, 2010.

Allison, M. Lee, "Global Data Integration in the Geosciences." Council of State Regulatory Officials, Interstate Oil & Gas Compact Commission Annual Meeting. Tucson, AZ. November 17, 2010.

Allison, M. Lee, Richard, S.M., Clark, R.J., & Grunberg, W., "Application of the U.S. Geoscience Information Network to deploying a National Geothermal Data System." [poster] American Geophysical Union Fall Meeting, San Francisco, CA. December 17, 2010. (Poster Attached)

Gundersen, L., Whitmeyer, S.J., Walker, D., Allison, M.L., Babaie, H., Cervato, C., Fils, D., Richard, S.M., & Arrowsmith, R., "New Initiatives in the Development of a National Geoinformatics Community." American Geophysical Union Fall Meeting, San Francisco, CA. December 15, 2010. (Abstract Attached)

2011

Allison, M. Lee, "Meeting the challenges of the 21st century with data integration." Geo-Data Informatics: Exploring the Life Cycle, Citation and Integration of Geo-Data, National Science Foundation Workshop ("GeoData2011"). Broomfield, CO. March 3, 2011.

Allison, M. Lee, "Online Data for Mining and Mineral Resources – Arizona Leads the Way," ARPA Critical Issues Conference, AZ Rock Products Association. Phoenix Airport Marriott, Phoenix, AZ. April 1, 2011.

Allison, M. Lee, "Everything Digital, Online, and Interoperable." Utah Geographic Information Council, 20th annual meeting. Riverwoods Conference Center, Logan, UT. April 6, 2011. http://gis.utah.gov/about/annual-conference

Keynote Address.

Allison, M. Lee, Richard, S.M., Anderson, A., & Cuyler, D., "National Geothermal Data System and Global Geosciences Data Integration." ESRI Petroleum User Group (PUG) Conference, Data Management & GIS Technology - Data & Standards session. George R. Brown Convention Center, Houston, TX. April 19, 2011. (Presentation Attached)

Allison, M. Lee, "Everything Digital, Online, and Interoperable," Americas Petroleum Survey Group, 25th Annual Conference. Houston, TX. May 6, 2011.

Allison, M. Lee, "Geothermal Data Project Annual Meeting Review." Association of American State Geologists Annual Meeting, Breakout Session. Dubuque, IA. June 15, 2011.

Allison, M. Lee, "Web Accessibility." Association of American State Geologists Annual Meeting, Breakout Session. Dubuque, IA. June 13, 2011.

Allison, M. Lee, "US GIN and Related Activities." OneGeology Operational Management Group Annual Meeting. Edinburgh UK. June 30, 2011.

Allison, M. Lee, and Richard, S.M., "Geoscience Information Network (GIN) Town Hall." Earth Science Information Partners Summer Meeting. Santa Fe, NM. July 12, 2011. http://wiki.esipfed.org/index.php/July_12,_2011

Richard, Stephen M., Allison, M.L. & Clark, R.J., "Distributed web services for geospatial geologic information." ESRI User Conference. San Diego, CA, July 13, 2011.

Richard, Stephen M., and Hutchinson, V., "Sustaining Community Efforts in the Geosciences: The Future of the US Geoscience Information Network." USGS Community on Data Integration workshop. Denver CO. August 17-18, 2011.

Patten, Kimberly, and Allison, M.L., "State Geological Survey Contributions to the National Geothermal Data System." State of Arizona, Governor's Office of Energy Policy, Geothermal Heat Pumps in Arizona Policy Workshop. September 22, 2011.

Allison, M. Lee, "Renewable Energy, Geologic Hazards, and Risk." Arizona Land Subsidence Group and Association of Environmental & Engineering Geologists (AEG) Conference on "Opportunities for Alternative Energy Development in Arizona and the Southwest – Geologic/Hydrologic Considerations." Tempe, AZ. October 27, 2011.

Wunsch, David, and Allison, M.L., "Web-based Information Services Available from State Geological Surveys to Assist Your Well Contracting Business." National Ground Water Association Annual Meeting and Expo. Las Vegas, NV. November 30, 2011.

Andersen, Arlene F., Cuyler, D., Snyder, W.S., Allison, M.L., Blackwell, D.D., & Williams, C.F., "National Geothermal Data System." American Geophysical Union Annual Conference. San Francisco, CA. December 7, 2011.

Allison, M.L.,. Richard, S.M., Clark, R., Coleman, C., Love, D., Pape, E., & Musil, L., "Online, interactive assessment of geothermal energy potential in the U.S." American Geophysical Union Annual Conference. San Francisco, CA. December 7, 2011. (Abstract Attached)

Gallagher, Kevin T. and Allison, M.L., "A Critical Path for Data Integration in the U.S. Earth Sciences." [Poster] American Geophysical Union Annual Conference, San Francisco, CA. December 9, 2011. (Poster Attached)

Jackson, Ian, Broome, H.J., & Allison, M.L., "Delivering Geoscience Knowledge in Federal Systems: What Can the Old and New Worlds Learn from Each Other?" Poster session presented at the American Geophysical Union Annual Conference. San Francisco, CA. December 9, 2011.

2012

Allison, M. Lee, Richard, S.M., Clark, R., Patten, K., Love, D., Coleman, C., Chen, G., Matti, J., Day, J., Pape, E. & Musil, L., "Online, interactive assessment of geothermal energy potential in the U.S." Thirty-Seventh Workshop on Geothermal Reservoir Engineering. Sanford University, Stanford, CA, January 30-February 1, 2012. (Abstract Attached)

Anderson, Arlene, Cuyler, D., Snyder, W., Allison, M.L., Blackwell, D., & Williams, C., "National Geothermal Data System." [Poster]. American Association of Petroleum Geologists Annual Conference. Long Beach, CA, April 24, 2012. (Poster Attached)

Allison, M. Lee, Richard, S.M., Clark, R., Patten, K., Love, D.S., Coleman, C., Chen, G., Matti, J., Day, J., Pape, E., and Musil, L., "Online, Interactive Assessment of Geothermal Energy Potential in the U.S." [Poster], American Association of Petroleum Geologists Annual Conference. Long Beach, CA. April 24, 2012.

Allison, M. Lee, "Geothermal Exploration: Everything Digital, Online, and Interoperable." Energy Minerals Division Annual Meeting, American Association of Petroleum Geologists Annual Conference. Long Beach, CA. April 25, 2012.

Allison, M. Lee, Atkinson, R., Arctur, D., Cox, S., Jackson, I., Nativi, S., Wyborn, L., "International Convergence on Geoscience Cyberinfrastructure." European Geosciences Union, Vienna, Austria, April 25, 2012, Stefano Nativi presenting.

Allison, M. Lee, Gallagher, K., Richard, S.M., Hutchison, V., "Strategic Roadmap for the U.S. Geoscience Information Network." [Poster]. European Geosciences Union, Vienna, Austria, April 25, 2012. (Abstract Attached)

Baru, C., Allison, M.L., Fox, P., Keane, C., Keller, R., & Richard, S.M., "GET21: Geoinformatics Training and Education for the 21st Century Geoscience Workforce." European Geophysical Assembly. Vienna, Austria. April, 2012. (Abstract Attached)

Allison, M. Lee, "State Geological Survey Contributions to the National Geothermal Data System." U.S. Dept. of Energy Geothermal Technologies Program Annual Peer Review. Westminster, CO. May 9, 2012.

Day, Janel, "Moving toward a new geologic map database standard, NCGMP: the good, the bad, and the ugly." Digital Mapping Techniques. Champaign, IL. May 21, 2012.

Clark, Ryan, "Online Geologic Maps: A Simple Application for Publishing NCGMP09 Databases." Digital Mapping Techniques. Campaign, IL. May 21, 2012.

Allison, M. Lee, Richard, S.M., Clark, R., Patten, K., Coleman, C., Caudill, C., Chen, G., "Implementing USGIN, a Distributed Data Network for Geoscience Information." Earth Science Information Partners Federation Summer Meeting. Madison, WI. July 17, 2012.

Allison, M. Lee, "EarthCube: How do we Organize, Manage, and Govern?" seminar, Earth Science Information Partners Federation Summer Meeting. Madison, W. July 19, 2012.

Allison, M. Lee, "Building Cyberinfrastructure in the Geosciences - the USGIN Experience." Geoscience Australia. Canberra, Australia. July 30, 2012.

Allison, M. Lee, Gallagher, K., Richard, S.M., & Hutchison, V., "Strategic Roadmap for the U.S. Geoscience Information Network." 34th International Geological Congress. Brisbane, Australia. August 6, 2012.

Allison, M. Lee and Richard, S.M., "U.S. Geothermal Data System." 34th International Geological Congress. [Poster]. Brisbane, Australia. August 6, 2012. (Poster Attached)

Allison, M. Lee, Atkinson, R., Arctur, D., Cox, S., Jackson, I., Nativi, S., & Wyborn,L., "International Convergence on Geoscience Cyberinfrastructure." 34th International Geological Congress. Brisbane, Australia. August 9, 2012. (Abstract Attached)

Richard, Stephen M. and Allison, M.L., "Access to State Geological Survey data for geothermal energy development in the United States." 34th International Geological Congress. [Poster]. Brisbane, Australia. August 7, 2012. (Poster Attached)

Allison, M. Lee, "Arizona Contributions to the National Geothermal Data System." Arizona Governor's Office of Energy Policy's Geothermal Workshop. Phoenix, AZ. September 6, 2012.

Allison, M. Lee, "Big Data," American Geoscience Institute Leadership Forum. [Presentation and panel discussion]. Washington DC. September 10, 2012.

Allison, M Lee, Richard, S.M., Clark, R., Patten, K., Love, D., Coleman, C., Chen, G., Matti, J., Day, J., Pape, E., Musil, L., & Caudill, C., "A Geothermal Data System for Exploration and Development." [Paper and oral presentation]. October 1, 2012.

Allison, M. Lee, and Pearthree, G., "Gathering Community Input on Governance for EarthCube." Poster session presented at the EarthCube End-user Domain Workshop for EarthScope. Arizona State University, Tempe, AZ. October 29, 2012.

Allison, M. Lee, Richard, S.M., Clark, R., Patten, K., Love, D., Caudill, C., Coleman, C., Matti, J., Day,J., Chen, G., & Musil, L., "National Geothermal Data System: State Geological Survey Contributions to Date." Geological Society of America Annual Meeting. Charlotte, NC. November 6, 2012.

Patten, Kim, Allison, M.L., Richard, S.M., Clark, R., Love, D., Coleman, C., Caudill, C., Matti, J., Musil, L., Day, J., & Chen, G., "National Geothermal Data System: State Geological Survey Contributions to Date." American Geophysical Union Fall Meeting. San Francisco, CA. December 3, 2012. (Abstract Attached)

Van Daalen, Tirza, and Allison, M.L. (presenter), "OneGeology – A Geoscience Exemplar for Worldwide Cyberinfrastructure Capacity-building and Scientific Innovation." [Poster]. American Geophysical Union Fall Meeting. San Francisco, CA. December 4, 2012. (Abstract Attached)

Allison, M. Lee, Co-convenor, "Towards a Global Cyberinfrastructure for the Geosciences." American Geophysical Union Fall Meeting, Union Session. San Francisco, CA. December 5, 2012.

Richard, Stephen M., "GeoSciML and EarthResourceML Update, 2012," Interoperability Working Group Commission for the Management and Application of Geoscience Information, IUGS, Poster session presented at the American Geophysical Union Fall Meeting. San Francisco, CA. December 5, 2012. (Abstract Attached)

Moosdorf, Nils, and Richard, S.M., "A Lithology Based Map Unit Schema For Onegeology Regional Geologic Map Integration." Poster session presented at the American Geophysical Union Fall Meeting, San Francisco, CA. December 7, 2012. (Abstract Attached)

2013

Blackman, Harold, Boyd, S., Patten, K., and Zheng, S., "NGDS User Centered Design – Meeting the Needs of the Geothermal Community." 38th Stanford Geothermal Workshop on Reservoir Engineering. Stanford, CA., February 11, 2013.

Anderson, Arlene, Blackwell, D., Chickering C., Boyd, T., Horne, R., MacKenzie, M., Moore, J., Nickull, D., Richard, S.M., & Shevenell, L., "National Geothermal Data System (NGDS) Geothermal Data Domain: Assessment of Geothermal Community Data Needs." 38th Stanford Geothermal Workshop on Reservoir Engineering, Stanford, CA. February 11, 2013.

Clark, Ryan, Kuhmuench, C., & Richard, S.M., "NGDS Node Deployment Adoption of CKAN for Domestic and International Data Deployment." 38th Stanford Geothermal Workshop on Reservoir Engineering, Stanford, CA. February 11, 2013.

Allison, Lee, "An NGDS model for Canada." CanGEA Annual Meeting. Calgary, Canada, March 21, 2013. (PowerPoint Presentation Attached)

Richard, Stephen, "Future of Canadian & Global Databases." panel discussion, CanGEA Annual Meeting, Calgary, Canada, March 21, 2013.

Richard, Stephen, "Mapping & Database Workshop." CanGEA Annual Meeting, Calgary, Canada, March 21, 2013.

Richard, Stephen (presenting), Anderson, A., Blackwell, D., Chickering, C., Boyd, T., Horne, R., MacKenzie., M., Moore, J., Nickull, D., & Shevenell, L., "Geothermal Data in the US National Geothermal Data System." CanGEA Annual Meeting, Calgary, Canada, March 21, 2013. (PowerPoint Presentation Attached)

Allison, M. Lee, "The National Geothermal Data System: A National Sustainabile, Distributed, interoperable Data Network." Research Data Access & Preservation Summit 2013, Baltimore, MD, April 4, 2013.

Richard, Stephen M., Clark, R., Chen, G., & Anderson, A., "U.S. National Geothermal Data System: Web Feature Services and System Operations." European Geophysical Union General Assembly, Vienna, Austria, April 8, 2013. (presented by Kim Patten)

Allison, M. Lee, "NGDS Overview." Cornell Energy Institute & Atkinson Center for a Sustainable Future. Cornell University, Ithaca, NY. May 1, 2013.

Allison, M. Lee, "The Digicene: the Age of Digital Geosciences." Cornell University Geological

Sciences Lecture Series. Ithaca, NY. May 1, 2013.

Clark, Ryan, Chen, G., & Kuhmuench, C., "An Open Source Stack for the National Geothermal Data System." Free and Open Source Software for Geospatial North America (FOSS-4G). Minneapolis, MN. May 22, 2013.

Richard, Stephen M., "Terminology Development to Support Geoscience Information Interoperability: the CGI Geoscience Terminology Working Group." IUGS Commission for Management of Geoscience Information (CGI) GeoSciML and GTWG workgroup meetings plenary session. St. Petersburg, Russia. June 1, 2013.

Clark, Ryan, "Faster, More Responsive Web GIS: Three technologies to watch." Digital Mapping Techniques Meeting. Denver, CO. June 3, 2013.

Day, Janel, "Integrating GeoSciML with OneGeology." Digital Mapping Techniques Meeting. Denver, CO. June 4, 2013.

Allison, M. Lee, "National Geothermal Data System – An AASG Success Story." Association of American State Geologists Annual Meeting, Plenary Session. Deadwood, SD. June 10, 2013.

Allison, M. Lee, moderator, "Web Accessibility Strategies," Association of American State Geologists Annual Meeting, Breakout Session. Deadwood, SD. June 10, 2013.

Allison, M. Lee, "NGDS Project Annual Meeting: Sustainability and Additional Funding Options." Association of American State Geologists Annual Meeting, Breakout Session. Deadwood, SD. June 11, 2013.

Patten, Kim, "NGDS User Interface." Association of American State Geologists Annual Meeting, Breakout Session. Deadwood, SD. June 11, 2013.

Allison, M. Lee, "Arizona Geothermal Energy Potential." Arizona Master Energy Task Force, Natural Resources and Land Use Work Group. Phoenix, AZ. July 10, 2013.

Richard, Stephen M., "Machine actionable links: a hypermedia format for data." Breakout session, Earth Science Information Partners Federation Summer Meeting. Chapel Hill, NC. July 12, 2013.

Richard, Stephen M., "National Geothermal Data System." Breakout session, Earth Science Information Partners Federation Summer Meeting. Chapel Hill, NC. July 12, 2013.

Allison, M. Lee, Anderson, A., Chickering, C., Patten, K., & Richard, S.M., "Sustaining the National Geothermal Data System: Considerations for a System-Wide Approach and Node Maintenance." Geothermal Resources Council Annual Meeting. Las Vegas, NV. October 2, 2013.

Allison, M.L., Richard, S.M., Blackman, H., Anderson, A., "National Geothermal Data System: an Exemplar of Open Access to Data." Poster session presented at the American Geophysical Union Fall Meeting. San Francisco, CA. December 12, 2013.

Patten, Kimberly, Allison, M.L., Richard, S.M., Blackman, H., Anderson, A., "National Geothermal Data System: An Exemplar of Open Access to Data." Geological Society of America Annual Meeting. Denver, CO. October 29, 2013. https://gsa.confex.com/gsa/2013AM/webprogram/Paper233220.html

Allison, M.L. and Lehnert, K.A., Conveners, "Digital Geosciences: A Framework for Data-Intensive, Multi-Disciplinary Research and Education." GSA Geoinformatics Division Session T90. Geological Society of America Annual Meeting. Denver, CO. October 29, 2013.

Gurney, R. and Allison, M.L. (presenter), "Towards a global e-infrastructure and data management plan." Geological Society of America Annual Meeting. Denver, CO. October 29, 2013. https://gsa.confex.com/gsa/2013AM/webprogram/Paper233089.html

Patten, Kim, Richard, S.M., & Allison, M.L., "National Geothermal Data System: An Open Data Access Exemplar." Poster session presented at the NSF EarthCube Rock Deformation Workshop. Washington, DC. November 12, 2013.

Allison, M. Lee, Richard, S.M., Blackman, H., Anderson, A., "National Geothermal Data System: An Open Data Access Exemplar." Poster session presented at the American Geophysical Union Fall Meeting. San Francisco, CA. December 12, 2013.

Richard, Stephen M., "Geoscience terminology for data interchange: the CGI Geoscience Terminology Work Group." Geoscience Terminology Work Group, American Geophysical Union Fall Meeting. San Francisco, CA. December 13, 2013 [presented by Christy Caudill-Daugherty].

Allison, M.L., Richard, S.M., Blackman, H., & Anderson, A., "National Geothermal Data System: an exemplar of open access to data." Poster session presented at the American Geophysical Union Fall Meeting. San Francisco, CA. December 12, 2013.

2014

Bristol, Sky, Allison M.L., & Richard, S.M., "Cross-Agency Panel Discussion - US Open Data Initiatives." Plenary Session, Earth Science Information Partners Winter Meeting. Washington, DC. January 9, 2014.

Richard, Steve, "Accessibility of Earth systems science data." breakout session, Earth Science Information Partners Winter Meeting. Washington, DC. January 10, 2014.

Cuyler, David, Anderson, A., & Patten, K., "How to Use the National Geothermal Data System (NGDS) and NGDS Node Deployment." 39th Stanford Geothermal Workshop. Stanford, CA. February 24, 2014. (Referenced at https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2014/2014program.html)

Allison, M.L., Richard, S.M., Blackman, H., Anderson A., & Patten, K., "National Geothermal Data System (USA): an Exemplar of Open Access to Data." European Geophysical Union General Assembly. Vienna, Austria. April 29, 2014. (Abstract Attached)

Allison, M. Lee, "Natural Resources Review Council," Arizona Mining Alliance, Tucson, AZ, May 9, 2014.

Allison, M. Lee and Patten, Kim, "Geologic Data Informatics." Breakout Session, Association of American State Geologists Annual Meeting. Lexington, KY. June 10, 2014.

Richard, Stephen M., "Information Exchanges and Interoperability Architecture." Earth Science Information Partners Summer Meeting. Frisco, CO. July 9, 2014. http://commons936.rssing.com/chan-17576118/all_p2.html

Richard, Stephen M., "Cyberinfrastructure for Science: data management and documentation." University of Arizona Department of Mines Colloquium. Tucson, AZ. September 24, 2014.

Anderson, Arlene, Allison, M.L., Richard, S.M., Caudill Daugherty, C., & Patten, K., "National Geothermal Data System: Case Studies on Exploration of Potential Geothermal Sites Through Distributed Data Sharing." [Oral presentation and poster session]. Geothermal Energy Association Annual Conference & Expo. Portland, OR. October 1, 2014. (Poster Attached)

Caudill, Christy, Richard, S.M., Musil, L., Sonnenschein, A., & Good, J., "National Geothermal Data System: Open Access to Geoscience Data and Digital Technologies." Poster presentation presented at the Geological Society of America Annual Meeting. Vancouver, Canada. October 20, 2014.

Allison, M. Lee, Richard, S.M., & Patten, K., "Providing Free and Open Source Access to Geoscience Data Worldwide." Oral presentation presented at the Geological Society of America Annual Meeting. Vancouver, Canada. October 21, 2014.

Caudill, C., Richard, S.M., Musil, L., Sonnenschein, A., Good, J., "National Geothermal Data System: Open Access to Geoscience Data, Maps, and Documents." Poster presentation at the American Geophysical Union Fall Meeting, San Francisco, CA. December 16, 2014.

Allison, M.L., Richard, S.M., & Patten, K., "Geoscience Information Network (USGIN) Solutions for Interoperable Open Data Access." Poster presentation presented at the American Geophysical Union Fall Meeting. San Francisco, CA. December 17, 2014.

Briefings

2010

Allison, M. Lee, Project Partners Meeting, "State Geological Survey Contributions to the National Geothermal Data System," Arlington, VA, May 17, 2010.

Allison, M. Lee, Geothermal Breakout Session, moderator, AASG Annual Meeting, New Brunswick, NJ, June 30, 2010.

Allison, M. Lee, "State Geological Survey Contributions to the National Geothermal Data System," AASG Annual Meeting, New Brunswick, NJ, June 30, 2010.

Allison, M. Lee, "Overview of the NGDS," Arizona Geographic Information Council, Data Committee, [briefing and demo], Phoenix, AZ, July 9, 2010. [by web conference]

Allison, M. Lee, Richard, S.M., "Geoscience Information Network Contributions to the National Geothermal Data System," NGDS Annual Meeting, Boise, ID, August 16, 2010.

Lovasz, Gabe, Hamilton, M., & Allison, M.L., "GIS Committee Report," Western Regional Partnership, Principals Meeting, Albuquerque, NM, August 17, 2010.

Allison, M. Lee, "GIN as a data integration framework for the Southern Arizona Template," Western Regional Partnership, Southern Arizona Template Working Group, Phoenix, AZ, November 8, 2010.

2011

Richard, Stephen M., U. S. Geological Survey Community for Data Integration, ScienceBase planning meeting, "US Geoscience Information Network", web meeting, Jan. 24, 2011.

Allison, M. Lee, American Institute of Professional Geologists, Mid-year Board Meeting, "AASG Strategic Directions," Tucson, AZ, February 11, 2011.

Allison, M. Lee, American Institute of Professional Geologists, Arizona Section, "Arizona Geological Survey," Tucson, AZ, February 12, 2011.

Richard, Stephen M., System for Earth Science Sample Registration Workgroup Meeting, "Metadata for Geoscience Resources", San Diego, CA, Feb. 22, 2011.

Allison, M. Lee, U.S. Geoscience Information Network (GIN) and Convergence Towards Global Data Integration in the Geosciences, iPlant GIS Advisory Committee, Univ. of Arizona, Tucson AZ, February 25, 2011.

Clark, Ryan, Western Regional Partnership GIS Committee Meeting, "Brief overview of US Geoscience Information Network", Reno, NV, Mar. 8, 2011.

Richard, Stephen M. "USGIN and NGDS status report". IUGS CGI Interoperability Workgroup Meeting, Edinburgh, Scotland, July 4, 2011.

Allison, M. Lee, Western Regional Partnership 4th Principals Meeting, Salt Lake City, UT: update on NGDS, GIN and progress on working agreement for data integration with WRP, September 16, 2011.

2012

Allison, M. Lee, "Geoscience foundation for geothermal energy utilization in the US," Environmental Planning Advisory Committee, Pima Association of Governments, Tucson, AZ, January 6, 2012.

Allison, M. Lee, "AASG Update," AIPG National Board Mid-Year Meeting, Tucson, AZ, February 10, 2012.

Allison, M. Lee, "AZGS Update," AIPG Arizona Section Annual Business Meeting, Tucson, AZ, February 11, 2012.

Allison, M. Lee, "Geoscience Information Network: A Critical Path for Data Integration," Western Regional Partnership Energy Committee, Reno NV, January 10, 2012.

Allison, M. Lee, "Sustainability of USGIN and NGDS," AASG Liaison Meeting, Washington, DC, September 10, 2012.

Allison, M. Lee, "Energy/Water Datasets and Models Workshop", U.S. Department of Energy, National Renewable Energy Lab, GoToMeeting, November 28, 2012.

2013

Allison, M. Lee, "AASG State Geological Survey Contributions to the NGDS," U.S. Department of Energy, Geothermal Technologies Office Annual Peer Review, Denver, CO, April 22, 2013.

Patten, Kim, "Arizona's Geothermal Energy Potential and the NGDS," Governor's Office of Energy Policy's Geothermal Energy Workshop IV, Phoenix, AZ, May 2, 2013.

Allison, M. Lee, "NGDS Sustainability Briefing," Geothermal Resources Council Board of Directors, Las Vegas, NV, September 28, 2013.

Allison, M. Lee, Chickering, C., Patten, K., Anderson, A., & Richard, S.M., "Sustaining the National Geothermal Data System: Considerations for a System-Wide Approach and Node Maintenance," Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 2, 2013.

Boyd, Suzanne, Zheng, S., Patten, K., & Blackman, H., "NGDS User Centered Design: Meeting the Needs of the Community," Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 2, 2013.

Anderson, Arlene, Blackwell, D., Chickering, C., Boyd, T., Horne, R., MacKenzie, M., Moore, J., Nickull, D., Richard, S.M., & Shevenell, L., "National Geothermal Data System (NGDS) Geothermal Data: Community Requirements and Information Engineering," Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 2, 2013.

Allison, M. Lee and Pearthree, G., "E-Infrastructure Activities in the USA," Belmont Forum E-Infrastructure Steering Committee, Windsor, UK, October 15, 2013. (no travel charged to this project, an event of opportunity)

Allison, M. Lee, "E-Infrastructure Activities in the USA," OneGeology Consortium Meeting, Paris, France, October 21, 2013. (no travel charged to this project, an event of opportunity)

Allison, M. Lee, "National Geothermal Data System: An Exemplar of Open Access to Data," Association of American State Geologists, Mid-Year Business Meeting, Denver, CO, October 29, 2013.

Allison, M. L., "Global e-infrastructure and data management plan," Physical Samples in the Digital Age Town Hall, American Geophysical Union Fall Meeting, San Francisco, CA, December 9, 2013.

Online Presentations and Webinars

(Select online locations available online at Youtube, Geothermal Data channel, NGDS Webinars playlist: **https://www.youtube.com/playlist?list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx)**

AZGS Webcast: "Thermal Profiling of wells for thermal conductivities and heat fluxes in active sedimentary aquifers," Dr. Franklin Horowitz, Research Professor, at University of Western Australia, School of Earth & Environment, and the Western Australia Geothermal Center of Excellence, January 28, 2011.

AZGS Webcast: "Geothermal air conditioning opportunities in hot sedimentary Aquifers," Dr. Franklin Horowitz, Research Professor, at University of Western Australia, School of Earth & Environment, and the Western Australia Geothermal Center of Excellence, January 28, 2011.

AZGS Webinar: "Metadata tools and workflows for AASG Geothermal Data", February 15, 2011. [32 attendees]

AZGS Webinar: "Metadata tools and workflows for AASG Geothermal Data", February 16, 2011. [35 attendees] Combined: *https://www.youtube.com/watch?v=a9xkGtCzFcE&index=6&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx*

AZGS Webinar: "Web Services for BoreHole Data," March 16, 2011. [36 attendees]

https://www.youtube.com/watch?v=9tJCpYcUx78&index=5&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx

AZGS Webinar: "AZGS HUBS – Data and Serving Data," April 12, 2011. [6 attendees]

AZGS Webinar: "AZGS HUBS – Data and Serving Data," April 26, 2011. [4 attendees]

AZGS Webinar: "Reviewing Data for Online Delivery," May 17, 2011. *https://www.youtube.com/watch?v=IUVA7TKCUJk&index=3&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx*

Richard, Stephen M. "USGIN and WRP" (linking WRP's catalog to the GIN network), AZGS Webinar, September 12, 2011.

Allison, M. Lee, "How GIS works as a User App for Digital Data Networks," GIS in Mining and Exploration Online Summit 2012, Session 8, Webinar series hosted by Mining IQ, January 29, 2012.

Webinar. Lehnert, Kerstin. "Introduction to IGSN and SESAR," March 8, 2012. [31 attendees] *https://www.youtube.com/playlist?list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx*

Webinar. Kay, Martin, Gosnold, W., & Lovenitti, J., "Geothermal Data: How's it Used & Why it's Important," April 26, 2012. [37 Registered Participants]

https://www.youtube.com/watch?v=CRPfSyjX6cs&index=2&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx

AZGS Webinar: "The New AASG Data Repository and Metadata Entry Tool," October 4, 2012.

https://www.youtube.com/watch?v=EYcG9pe1oXI&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx&index=8

Allison, M. Lee, "The National Geothermal Data System and the Energy-Water Nexus," Energy/Water Datasets and Models Workshop, U.S. DOE, Virtual, November 28, 2012.

Allison, M. Lee, Caudill, C., Patten, K., "The National Geothermal Data System," Arizona Mining Review, Webcast, Tucson, AZ, May 31, 2013. *http://www.youtube.com/watch?v=TpAPvJ7E-Rs*

Allison, M. Lee, moderator, "Web Accessibility Strategies," Association of American State Geologists Annual Meeting, Breakout Session, Deadwood, SD, June 10, 2013. LiveStream & YouTube. *https://www.youtube.com/watch?v=_rggaq-hCs8&list=PLxRc6-3iV2qdublF52bE9VO2v6TwMxThx&index=7*

Caudill-Daugherty, Christy, "Excel to NGDS Conversion," US DOE Geothermal Technologies Office Lunch and Learn Seminar, Washington DC, Webinar presentation, August 2, 2013.

AZGS, presenters, DOE GTO Webinar, "NGDS Demo and How to Adopt NGDS Node-in-a-Box (NIAB) for Geothermal and Other Applications," January 27, 2014.

AZGS, presenters, DOE GTO Webinar, "How to Utilize the National Geothermal Data System," January 28, 2014.

AZGS Webinar: "Final Science Advisory Board Review," May 29, 2014, Tucson, AZ* SAB lightening talks available at *https://www.youtube.com/playlist?list=PLxRc6-3iV2qf5lUhvTujIdVjkmY_k7LEr*

Exhibits

Geothermal Resources Council annual meeting and expo in Sacramento, CA. A new high-definition (HD) multi-monitor display system was unveiled at the GRC annual meeting and expo to run demonstrations of live Web services and data integration from the project. Oct.24-Oct.27, 2010.

Geological Society of America Annual Meeting in Denver, CO. The new HD multi-monitor display system was subsequently set up in the AASG booth in the exhibit hall for the Geological Society of America Annual Meeting with live demos run continuously. Oct.30-Nov.3, 2010.

Geothermal Energy Association, Geothermal Energy Technology and International Development Forum, May 4, 2011, Wash. DC [AZGS State Geothermal Data exhibit booth]

Geothermal Energy Association National Geothermal Summit, Reno, NV August 16-17, 2011

Geothermal Resources Council 35th Annual Meeting and GEA Trade Show, San Diego, CA, October 23-26, 2011 [AZGS State Geothermal Data exhibit booth]

Geological Society of America Annual Meeting, Minneapolis, MN, October 9-12, 2011 [AZGS State Geothermal Data exhibit booth]

American Geophysical Union Fall Meeting, San Francisco, CA, December 5-12, 2011 [AZGS State Geothermal Data exhibit booth]

Geothermal Resources Council Annual Meeting, Reno, NV, September 30- October 3, 2012. [Exhibited the new data search tool, new NGDS website, and new data collection from UT]

Geological Society of America Annual Meeting, Charlotte, NC, November 4-7, 2012 [AZGS State Geothermal Data exhibit booth and GSA Geoinformatics Division booth (the second booth is leveraging external funds)]

American Geophysical Union Fall Meeting, San Francisco, CA, December 3-7, 2012 [AZGS State Geothermal Data exhibit booth]

"Geothermal Energy in Arizona: the role of the National Geothermal Data System," Science in the City, Tech in Tucson Showcase, Tucson Electric Power Headquarters, Tucson, AZ, February 16, 2013.

U.S. Department of Energy, Geothermal Technologies Office Annual Peer Review, Denver, CO, April 22-24, 2013. [NGDS Booth]

American Association of Petroleum Geologists (AAPG) Annual Meeting, Pittsburgh, PA, May 19-22, 2013. [NGDS Booth]

National Geothermal Data System and User Interface Testing, Geothermal Energy Association Expo, September 29-October 2, 2013, Las Vegas, NV. [NGDS Booth]

Geological Society of America Annual Meeting 2013, Denver, CO, October 27-30, 2013 [in conjunction with the AASG booth]

American Geophysical Union Fall Meeting 2013, San Francisco, CA, December 9-13, 2013 [NGDS Booth]

National Geothermal Data Systems exhibit booth, Stanford Geothermal Workshop, Stanford, CA, February 24-26, 2014.

National Geothermal Data Systems exhibit booth, AAPG Annual Convention, Houston, TX, April 6-8, 2014.

NGDS Exhibit Boot, White House Energy Datapalooza, Washington, DC, May 28, 2014.

National Geothermal Data Systems/USGIN exhibit booth, National Geothermal Summit, Reno, NV, August 5-6, 2014.

National Geothermal Data Systems/USGIN exhibit booth, Arizona Geographic Information Council Annual Convention, Prescott, AZ, September 3-4, 2014.

National Geothermal Data Systems/ USGIN exhibit booth, Geothermal Energy Association, Geothermal Energy Expo, Portland, OR, September 28-October 1, 2014.

National Geothermal Data Systems/USGIN exhibit booth, Geological Society of American Annual Meeting, Vancouver, Canada, October 19-22, 2014.

National Geothermal Data Systems/USGIN exhibit booth, American Geophysical Union Fall Meeting, San Francisco, CA, December 15-19, 2014.

Workshops

Governor's Renewable Energy Forum, April 15, 2010 Scottsdale, AZ [M. Lee Allison – organizing committee] – funded by State of Arizona

Allison, M. Lee, Member, Organizing Committee, Workshops on Interoperability/Technology and Education/Outreach aspects of Geoinformatics working towards a National Geoinformatics Community Organization, September 23-24, 2010, USGS-Denver Federal Center, Denver, CO

AASG State Geological Survey's Contributions to the National Geothermal Data System Hub Training Workshop. Tucson, AZ August 2-3, 2011

Governor's Office of Energy Policy, Geothermal Heat Pumps in Arizona Policy Workshop, Phoenix, AZ, September 22, 2011

AASG State Geological Survey's Contributions to the National Geothermal Data System Hub Training Workshop. Tucson, AZ September 5-6, 2012

Governor's Office of Energy Policy, "Geothermal Workshop," Phoenix, AZ, September 6, 2012.

Governor's Office of Energy Policy, "Geothermal Workshop," Phoenix, AZ, May 2, 2013.

AZGS, "AASG NGDS Western Regional Hub," Reno, NV, May 14, 2013.

AZGS, "AASG NGDS Eastern Regional Hub," Champaign-Urbana, IL, July 25-26, 2013.

NGDS Southeast Regional Workshop, hosted by the Virginia Department of Mines, Minerals, and Energy, Charlottesville, VA, August 7-8, 2013.

"Drilling Down: How Legacy and New Research Data Can Advance Geothermal Development" National Geothermal Data System for End-Users, Geothermal Resources Council Annual Meeting, Las Vegas, NV, October 2, 2013.

"Drilling Down: The National Geothermal Data System" workshop, in conjunction with the Geothermal Resources Council/Geothermal Energy Association Annual Meeting & Expo, October 2, 2014.

Testimony/Written Responses

Allison, M. Lee, Response to "Spatial Data Enabling USGS Strategic Science in the 21st Century," National Research Council, February 16, 2010, 5p

Annual Reports

Allison, M. Lee, Richard, Stephen M., 2011 "State Geological Survey Contributions to the National Geothermal Data System: Annual Report 2010," OFR-11-01, http://repository.azgs.az.gov/uri_gin/azgs/dlio/948

Allison, M. Lee, Richard, Stephen M. 2012 "State Geological Survey Contributions to the National Geothermal Data System [2011 Annual Report]," OFR-12-06, http://repository.azgs.az.gov/uri_gin/azgs/dlio/1406

Allison, M. Lee, Richard, Stephen M. 2013 "State Geological Survey Contributions to the National Geothermal Data System [2012 Annual Report]," OFR 13-11, http://repository.azgs.az.gov/uri_gin/azgs/dlio/1528

Published Technical Reports

Richard, Stephen M., 2010, "Technical Contributions of the Geoscience Information Network for Design and Deployment of the National Geothermal Data System," Arizona Geological Survey OFR-10-04, http://repository.azgs.az.gov/uri_gin/azgs/dlio/800

Awards and Honors

Geothermal Resources Council, Best Presentation Award, GRC Annual Meeting 2012: "A Geothermal Data System for Exploration and Development"

News Media Interviews

Matthew DeMerritt, *ESRI Arc User News,* April, 2010.

Kerry Klein. (2010, May 28). "Geothermal Technologies Program Celebrates Industry Adoption of the National Geothermal Data System," [EERE Geothermal Technologies Blog]. Retrieved from http://www.eereblogs.energy.gov/geothermaltechnologies/post/Geothermal-Technologies-Program-Celebrates-Industry-Adoption-of-the-National-Geothermal-Data-System.aspx.

Steve Ayers, *Verde Valley News*, Prescott, AZ, May 10, 2010

Pam White, KUAZ-KUAT FM, Tucson, AZ, May 10, 2010

Jonathan DuHamel. (2010, May 11). [*Wry Heat* blog column, *Tucson Citizen*.] Retrieved from http://tucsoncitizen.com/wryheat/tag/geothermal-energy/

Anne Marie de Grosbois. (2010). Emerging cyberinfrastructure for the Earth sciences forged by the US GIN project: http://usgin.org and http://lab.usgin.org. *Environmental Earth Sciences*, 61: 1763-1765. http://www.springerlink.com/content/p14421645w561555/fulltext.pdf,

Daryl Bjoraas. (2011, January 27). *Geothermal Energy in Arizona*. Phoenix, AZ: Arizona Capitol Television.

M. Lee Allison, live hour-long interview on Geothermal Energy with host Brian Hageman and former Congressman Barry Goldwater Jr, Deluge Energy Report Radio Show, KFNX -1100AM, Tucson AZ (and webcast), June 21, 2011

Allison, M. Lee, August 18, 2011, with Dylan Tussel, Columbus [Ohio] Dispatch: re geothermal energy, State Geothermal Data, and NGDS http://www.dispatch.com/content/stories/science/2011/08/28/dirty-past-clean-future.html

Durham, Louise S. (2012). "U.S. Geothermal Database Being Created" with Chacko John (Louisiana Geological Survey), American Association of Petroleum Geologists *Explorer,* January 2012. Retrieved from http://www.aapg.org/explorer/2012/01jan/hot_water0112.cfm

Geothermal Energy Association's Geothermal Energy Weekly, "National News: National Geothermal Data System Aggregates Nationwide Info on Wells," Week of December 6, 2012.

Associated Press. (2013). "Oregon's Geothermal Resources on Display in New Online Interactive Map," February 12, 2013. Retrieved from http://www.oregonlive.com/environment/index.ssf/2013/02/oregons_geothermal_resources_o.html.

University of Idaho News, "Idaho Geological Survey and U-Idaho Explore for Geothermal Energy," January 10, 2013, http://www.uidaho.edu/newsevents/item?name=idaho-geological-survey-and-u-idaho-explore-for-geothermal-energy and EERE GTO http://www1.eere.energy.gov/geothermal/news_detail.html?news_id=18923.

KTVZ News Sources, "Spotlight on Oregon's Geothermal Potential," KTVZ, February 24, 2013, http://www.ktvz.com/news/Spotlight-on-Oregon-s-geothermal-potential/-/413192/19066336/-/1lw7lnz/-/index.html.

Ferguson, William. (2013) "New Geothermal Data System Could Open Up Clean-Energy Reserves," *Scientific American,* February 25, 2013, Retrieved from http://www.scientificamerican.com/article.cfm?id=new-geothermal-data-system.

Geothermal Energy Association, "Geothermal Energy Weekly," March 1, 2013. Retrieved from http://geo-energy.org/Newsletter/2013/Geothermal%20Energy%20Weekly%20March%201%202013.pdf?utm_source=March+1+GEW&utm_campaign=March+1+GEW&utm_medium=email.

Arizona Public Media, *Arizona Illustrated*, "Tracking Geothermal Energy Sources," Tucson, AZ, May 7, 2013. Retrieved from https://ondemand.azpm.org/p/segments/2013/5/7/24206-tracking-geothermal-energy-sources/

North American Clean Energy, "The National Geothermal Data System," July/August 2013. Retrieved from http://www.nacleanenergy.com/articles/16454/the-national-geothermal-data-system

Pratt, Sarah E. (2013). "Digitizing Earth: Developing a Cyberinfrastructure for the Geosciences," *EARTH Magazine*, a publication of the American Geosciences Institute, September, 2013. Retrieved from http://www.earthmagazine.org/tags/september-2013.

Non-technical Articles

Patten, Kim, "New Geothermal Data Collection Funded in Fifteen States," *Arizona Geology* Spring 2011, http://www.azgs.az.gov/arizona_geology/spring11/article_geotherm.html

Patten, Kim, "Positive Feedback on the State Geological Survey Contributions to the National Geothermal Data System," *Arizona Geology* Fall 2011, http://azgeology.azgs.az.gov/article/geothermal/2011/09/positive-feedback-state-geological-survey-contributions-national.

Patten, Kim, "State Geological Survey Contributions to the NGDS Update," *Arizona Geology* Fall 2012, http://azgeology.azgs.az.gov/article/geothermal/2012/08/state-geological-survey-contributions-national-geothermal-data-system.

Patten, Kim, "Success Stories from the National Geothermal Data System," *Arizona Geology Magazine*, December 2012, http://azgeology.azgs.az.gov/article/geothermal/2012/12/success-stories-national-geothermal-data-system.

Anderson, Arlene, "Pinpointing America's Geothermal Resources with Open Source Data," Energy.gov, January 3, 2013, http://energy.gov/articles/pinpointing-america-s-geothermal-resources-open-source-data.

Websites, Social Media

State Geothermal Data, www.stategeothermaldata.org
National Geothermal Data System, www.geothermaldata.org
NGDS GitHub, https://github.com/ngds
NGDS Facebook, https://www.facebook.com/NationalGeothermalDataSystem
NGDS YouTube, http://www.youtube.com/geothermaldata
US Geoscience Information Network (USGIN), http://usgin.org
USGIN developer's blog, http://lab.usgin.org
USGIN GitHub, https://github.com/usgin
USGIN YouTube, http://www.youtube.com/channel/UCw3caw5RRLEErC1S6qFG-DA
State Geothermal Data playlist on the AASG YouTube Channel - http://www.youtube.com/user/StateGeologists

Subrecipients Presentations, News Interviews, and More

(Project Participant listed in parenthesis)

Rancan, Helen, Bi-monthly reports to the Department Commissioner on Project Status (New Jersey Geological Survey)

Thomas, Donald M. "Development of a Digital Geothermal Database: Application to Hawai'i," presented to the Director of the Department of Land and Natural Resources and the Commission on Water Resources Management, November 20, 2010 (University of Hawaii)

Thomas, Donald M. "A Digital Geothermal Database and its Relevance to Groundwater Sources," presented to the Vice Director of Civil Defense, December 3, 2010 (University of Hawaii)

Allis, Rick, "Geothermal Data Collection in Utah," presented to the Bureau of Land Management, January 6, 2011 (Utah Geological Survey)

Smith, Cole and Uddenberg, Matt, "Geothermal Investigation of Southern Crockett County, Val Verde Basin, West Texas," GSA South-Central Section in New Orleans, LA, March 28, 2011 (Texas Bureau of Economic Geology)

Cutright, Bruce, "The Transformation of Tight Shale Gas Reservoirs to Geothermal Energy Production" AAPG National Convention, Houston, TX, April, 2011 (Texas Bureau of Economic Geology)

McConnell, Vicki, and Niewendorp, Clark A., "Oregon ARRA Update," presented at the Twelfth meeting of the Oregon Geothermal Working Group, Oregon Geothermal Update 2011, April 11, 2011 (Oregon Department of Geology and Mineral Industries)

Naylor, Shawn, Hydrogeologic Considerations for Designing Ground-Source Heat Pump Installations in Indiana, Purdue University, April 8, 2011. (Indiana Geological Survey)

Feeney, Chelsea M. "Virginia's Contributions to the National Geothermal Data System," Presented to the Third Annual Virginia Geologic Research Symposium, Virginia Division of Geology and Mineral Resources, April 14, 2011 (Virginia Division of Geology and Mineral Resources)

Mabee, Stephen, "Geothermal Energy in Massachusetts," presented at the American Ground Water Trust Geothermal Workshop, Boxborough, MA, April 19, 2011 (Massachusetts Geological Survey)

Goodhue, N, Koteas, G.C., Rhodes, J.M., and Mabee, S.B., 2011 "Deep geothermal potential of New England granitoids: The Fall River pluton, southeastern, MA." Geological Society of America, Joint Meeting of the Northeastern and North-Central Sections, Abstracts with Programs, 2011. v.43, no.1, p.63. (Massachusetts Geological Survey)

Czajkowski, Jessica, Bowman, Jeff, and Norman, David K., "Collection of New Geothermal and Hydrothermal Data for Washington State," presented at the Eighth Annual Hydrogeology Symposium, Tacoma, WA, April 27, 2011 (Washington Geological Survey)

Cutright, Bruce, "The Transformation of Tight Shale Gas Reservoirs to Geothermal Energy Production" Southern Methodist University Geothermal Conference, June 13, 2011 (Texas Bureau of Economic Geology)

Dylan Tussel, *Columbus Dispatch*, interviewed Tim Leftwich (Ohio Geological Survey), August 29, 2011 http://www.dispatch.com/content/stories/science/2011/08/28/dirty-past-clean-future.html (Ohio Geological Survey)

Ellie Bogue, *Fort Wayne Sentinal*, interviewed Shawn Naylor on the Indiana Shallow Geothermal Monitoring Network, September 21, 2011 (Indiana Geological Survey)

Gosnold, Will D., "Minnesota Geothermal Potential," Geological Society of America Annual Meeting, Minneapolis, MN, October 9, 2011 (University of North Dakota, Minnesota Geological Survey)

Angle, Mike, "Geothermal data compilation in Ohio and potential for abandoned underground coal mines to serve as geothermal reservoirs" at the Ball State University Ground Source Heat Pump Information Sharing Retreat, November 9, 2011 (Ohio Geological Survey)

Naylor, S.C., Gustin, A.R., Ellet, K.M., "Monitoring Near-Surface Thermal Properties in Conjunction with Energy and Moisture Budgets to Facilitate the Optimization of Ground-Source Heat Pumps in the Midwest," American Geophysical Union, Fall Meeting 2011, December 2011.

College of Arts and Sciences, "Michigan Geological Repository for Research and Education has Key Role in Geothermal Energy Search," February 15, 2012 http://www.wmich.edu/cas/alumni/wordpress/2012/02/michigan-geological-repository-for-research-and-education-has-key-role-in-geothermal-energy-search/ (Michigan Geological Survey)

Gustin, Andrew, "The Indiana Shallow Geothermal Monitoring Network: A test bed for optimizing ground-source heat pumps in the glaciated Midwest," North-Central Geological Society of America, Dayton, OH, April 23, 2012 (Indiana Geological Survey)

Wolfe, Mark, Leftwich, Timothy, and Lopez, Dina, "Geothermal potential of abandoned underground industrial mineral mines in Ohio," 48th Forum on the Geology of Industrial Minerals, Scottsdale, AZ May 2, 2012 (Ohio Geological Survey)

Naylor, Shawn, Ellett, Kevin, Gustin, Andrew, "Monitoring near-surface thermal properties in conjunction with energy and moisture budgets to facilitate the optimization of ground-source heat pumps," Indiana Geological Survey's Spring Seminar, Bloomington, IN, May 3, 2012 (Indiana Geological Survey)

Gosnold, Will. Presented a talk at the 2012 Renewable Energy Action Summit, June 25, 2012 in Bismarck, ND (University of North Dakota)

Arkansas Geological Survey, Geothermal Booth at ECOFEST in Conway, Arkansas. September 8, 2012. 190 Visitors; approximately 70% of visitors requested information about ground source heat pump systems for their homes.

Chales Zuppan, Tiffany Proffitt, Kevin Ellett, and Shawn Naylor and collaborators from the Illinois and Kentucky geological surveys presented methods and results from a multi-state effort to compile and refine BHT data for the Illinois Basin. The work was presented at the AAPG Eastern Section's annual meeting as a poster presentation titled: "Development and application of a new geothermal database for the Illinois Basin". September, 2012 (Indiana Geological Survey)

Gosnold, Will. Invited talk at the American Institute of Professional Geologists Annual Meeting in Rapid City, SD. September 23-25, 2012 (University of North Dakota)

Tutolo, Benjamin M., and Saar, Martin O., "Geothermal Energy in Minnesota: Revised Prospects Based on Heat Flow Measurements." Midwest Ground Water Conference, October 1-3, 2012 (Minnesota Geological Survey)

Oberbeck, Steven, "Big Source of Power Looming in Desert?" *Salt Lake City Tribune*. September 28, 2012 (article on the Utah Geological Survey Supplemental Funding Project). A video based on this article and report was created by UGS and shown at the Geothermal Resources Council; it will also be shown at the Geological Society of America meeting. (Utah Geological Survey)

Allis, Rick, et. Al. "Stratigraphic Reservoirs in the Great Basin – the Bridge to Development of Enhanced Geothermal Systems in the U.S." Geothermal Resources Council Annual Meeting in Reno, NV. September 29-October 3, 2012 (Utah Geological Survey)

Hardwick, C. and Chapman, D. "Geothermal Resources in the Black Rock Desert, Utah: MT and Gravity surveys," Geothermal Resources Council Annual Meeting in Reno, NV. September 29-October 3, 2012 (Utah Geological Survey)

Crowell, Anna, Ochsner, Aaron, Gosnold, Will, Correcting Bottom-Hole Temperatures in the Denver Basin, Colorado and Nebraska, Geothermal Resources Council, Transactions 2012, 11pp. Geothermal Resources Council Annual Meeting in Reno, NV. September 29-October 3, 2012 (University of North Dakota)

Majorowicz, Jacek; Gosnold, Will; Gray, Allan; Safanda, Jan; Klenner, Rob; and Unsworth, Martyn. Implications for Post-Glacial Warming for Northern Alberta Heat Flow - Correcting for the Underestimate of the Geothermal Potential. Geothermal Resources Council Trans. Geothermal Resources Council Annual Meeting in Reno, NV. September 29-October 3, 2012 (University of North Dakota)

Gosnold, Will, McDonald, Mark, Klenner, Robert, Merriam, Daniel. Thermostratigraphy of the Williston Basin, Geothermal Resources Council Transactions, 2012, 21 pp. 2012, 21 pp. Geothermal Resources Council Annual Meeting in Reno, NV. September 29-October 3, 2012 (University of North Dakota)

Kolomitz, Christopher, "New Geothermal Discovery made in Utah," Energize Weekly Newsletter, October 10, 2012.

Clough, Jim, "Opportunities for Energy and Geothermal Resources on the Seward Peninsula." October, 2012 (Alaska Geological Survey)

Clough, Jim, "Presentation to the Nome City Council on Energy Opportunities." November, 2012 (Alaska Geological Survey)

Rhodes, J. Michael, Koteas, Christopher, Mabee, Stephen, Thomas, Margaret, Gagnon, Teresa, "Heat Production as a Tool in Geothermal Exploration," December 3, 2012, American Geophysical Union Fall Meeting (Massachusetts Geological Survey, Connecticut Geological Survey)

Crowell, Anna M, and Gosnold, William D., "Available Thermal Energy in the Denver Basin Dakota Group: Colorado and Nebraska," December 3, 2012, American Geophysical Union Fall Meeting (University of North Dakota)

James Crowell and William D. Gosnold, "Using a Divided Bar Apparatus to Measure Thermal Conductivity of Samples of Odd Sizes and Shapes," December 3, 2012, American Geophysical Union Fall Meeting (University of North Dakota)

Gosnold, William D. and Crowell, Anna M., "Synthesis of Bottom Hole Temperatures and Heat Flow Data," December 3, 2012, American Geophysical Union Fall Meeting (University of North Dakota)

Naylor, Shawn, Gustin, Andrew R., Ellett, Kevin M., "Quantifying the Influence of Near-Surface Water-Energy Budgets on Soil Thermal Properties Using a Network of Coupled Meteorological and Vadose Zone Instrument Arrays in Indiana, USA," December 5, 2012, American Geophysical Union Fall Meeting (Indiana Geological Survey)

Thomas, Donald M., Lautze, Nicole C., and Abdullah, Mahany, "Progress and challenges associated with digitizing and serving Hawaii's geothermal data," December 6, 2012, American Geophysical Union Fall Meeting (University of Hawaii Manoa)

Trabish, Herman, "Geothermal's Big 12 Events of 2012,"Seeking Alpha, December 18, 2012, http://seekingalpha.com/article/1071191-geothermal-s-big-12-events-of-2012. (Utah Geological Survey work featured as number 7)

World Wide Geothermal Resource, "Has Utah Discovered a New Source of Geothermal Energy?" December 2012, Volume 3, Number 12, pg. 10.

World Wide Drilling Resource, "Utah Explores New Type of Geothermal Resource," February 2013, Volume 10, Number 2, pg. 36.

Iowa Geological Survey, "Brown bag on geology and geohydrology of ground source heat pump development in Iowa" to Iowa Geological and Water Survey staff, December, 2012.

Iowa Geological Survey, "Geology of Heat Pumps" Presentation to Cedar Valley Rock and Mineral Society, December, 2012.

Bowman, Jeff, "Washington State Contribution to the NGDS," USGS Oregon Water Science Center, Portland, OR, January 8, 2013. (Washington Geological Survey)

Barnard, Jeff, "Interactive map shows Oregon's geothermal resources," Associated Press, February 13, 2013. http://www.katu.com/news/local/Interactive-map-shows-Oregons-geothermal-resources-191044801.html (Oregon Geological Survey)

Gwynn, Mark, "New Geothermal Resource Delineated Beneath Black Rock Desert, Utah," 38th Stanford Geothermal Workshop on Reservoir Engineering, Stanford, California, February 11-13, 2013. (Utah Geological Survey)

Allis, Rick, "Characterizing the power potential of hot stratigraphic reservoirs in the Great Basin, U.S." 38th Stanford Geothermal Workshop on Reservoir Engineering, Stanford, California, February 11-13, 2013. (Utah Geological Survey)

Missouri Geological Survey, Poster Presentations on Depth-to-Bedrock and BHT work, Presented at the Missouri GIS Conference, Chesterfield, MO, February 19-21, 2013. (Missouri Geological Survey)

Bowman, Jeff, "Geothermal Exploration at Your Fingertips: Washington State Geological Survey's Contribution to the NGDS," USGS Washington Water Science Center, Tacoma, WA, February 20, 2013. (Washington Geological Survey)

Gagnon, Thomas, Rhodes, Mabee, 2013, NE GSA Meeting, Abstracts with Programs, v.45, no.1, p.20. "Connecticut Geothermal Map Series: Tools for Exploration and Development," in the session "Getting the Work Done: State Geological Surveys; Partnering and Progress" sponsored by Bob Marvinney, Larry Becker, Jon Kim and Rick Chormann. March 18, 2013. (Connecticut Geological Survey, Massachusetts Geological Survey)

Degnan, James R., Barker, Gregory, Olson, Neil F. and Wilder, Leland, 2013 [Abstract]. **Temperature Logging of Groundwater in Wells to Measure Bedrock Geothermal Gradients in New Hampshire, 2012** in 2013 Abstracts with Programs, vol. 45, no. 1, 48th Annual Meeting of the

Northeastern Section Geological Society of America, p. 50. March 18, 2013. (New Hampshire Geological Survey)

Bowman, Jeff, "Geothermal Exploration at Your Fingertips: Washington State Geological Survey's Contribution to the NGDS," US Army Corps of Engineers, Seattle, WA, March 8, 2013. (Washington Geological Survey)

Ellett, Kevin, Naylor, Shawn, Gustin, Andrew, Waldoch, Connor, "Results from the new Indiana Geothermal Monitoring Network: Implications for how and where soil modifications can improve the performance and affordability of geothermal heat pumps." Purdue University, April 19, 2013. (Indiana Geological Survey)

Washington Geological Survey, "An Update to the Collection of New Geothermal Data for Washington State," 2013 Washington Hydrological Symposium, Tacoma, WA, April 23, 2013, Poster.

Naylor, Shawn C., "Results from the new Indiana Geothermal Monitoring Network: Providing data to help optimize geothermal heat pumps in Indiana," Indiana Geological Survey, Bloomington, IN, May 30, 2013.

Bob Blackett, Utah Geological Survey, Panel Presentation, "State of Geothermal in the States," Geothermal Energy Association's National Summit, Reno, NV, June 27, 2013.

Easley, Elisabeth, Morgan, Paul, *Fluid, Gas, and Isotopic Variation of Thermal Springs in the Southern Rocky Mountains: Colorado,* Geothermal Resources Council Annual Meeting, Las Vegas, NV, September 30, 2013 (Colorado Geological Survey)

Welhan, John, Garwood, Dean, Feeney, Dennis, *The Blackfoot Volcanic Field, Southeast Idaho: A Hidden High-T Geothermal Resource Revealed Through Data Mining of the National Geothermal Data Repository,* Geothermal Resources Council Annual Meeting, Las Vegas, NV, September 30, 2013. (Idaho Geological Survey)

Morgan, Paul, *Advantages of Choosing a Sedimentary Basin as the Site for an EGS Field Laboratory,* Geothermal Resources Council Annual Meeting, Las Vegas, NV, September 30, 2013

Shevenell, Lisa, Zehner, Richard, *Geochemical Data Collection and Evaluation as Part of the National Geothermal Data System,* Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 1, 2013. (Nevada Bureau of Mines and Geology)

Shevenell, Lisa & Zehner, Richard, *Shallow Temperature Surveys at Four Sites in Nevada as Part of the National Geothermal Data System,* Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 1, 2013. (Nevada Bureau of Mines and Geology)

GIS-Based Geothermal Resource Assessment of the Denver Basin: Colorado and Nebraska, Crowell, Anna & Gosnold, Will, Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 1, 2013. (University of North Dakota)

Blackett, Robert & Gwynn, Mark & Hardwick, Christian & Thomas,

Kevin & Weaver, Lance & Allis, Rick, *Utah Contributions to the National Geothermal Data System,* Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 1, 2013 [poster] (Utah Geological Survey)

Detecting Spatial Trends in Thermal Conductivity in the Williston Basin, Crowell, James "Josh" & Gosnold, Will, Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 1, 2013. (University of North Dakota)

GIS-Based Geothermal Resource Assessment of the Denver Basin: Colorado

and Nebraska, Crowell, Anna & Gosnold, Will, Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 1, 2013. (University of North Dakota)

Crowell, James, Gosnold, Will, *Detecting Spatial Trends in Thermal Conductivity in the Williston Basin,* Geothermal Resources Council Annual Meeting, Las Vegas, NV, Oct. 2, 2013. (University of North Dakota)

Clough, Jim, "Alaska Geothermal Sites Map and Database: Bringing together legacy and new geothermal data for research, exploration and development," American Geophysical Union Fall Meeting 2013, December 9, 2013 [poster] (Alaska Geological Survey)

Publications and Proceedings

Microsoft®

Environmental Research Workshop 2010

Microsoft Environmental Research Workshop 2010

July 14
Baker, Microsoft Conference Center, Building 33
16070 NE 36th Way
Redmond, WA 98052

A soft copy of this book can be found at
http://research.microsoft.com/events/environmentalresearch2010/ERW2010.pdf.

To request a hard copy, please email e3env@microsoft.com.

Cover and text designed by Yan Xu

©2010 Microsoft Corporation. All rights reserved.

Microsoft Research

Table of Contents

Introduction – *1*

ERW2010 Agenda – *2*

Position Papers

Multi-touch Interfaces for Content-Based Video Searches
Mark Abbott (Oregon State University) – ***3***

U.S. Geoscience Information Network (GIN) and Convergence towards Global Data Integration in the Geosciences
Lee Allison (Arizona Geological Survey) et al – ***4***

Designing Effective Interfaces for Environmental Data and Residential Climate Control
Cecilia Aragon (Lawrence Berkeley National Laboratory & University of Washington) – ***5***

A Virtual Telescope for Global Monitoring of Earth Surface Changes
Jean-Philippe Avouac and **Sébastien Leprince** (California Institute of Technology) – ***6***

Forecasting Crop Disease Risk Using Mesoscale Weather Forecasts
Kathleen Baker (Western Michigan University) – ***7***

Observing the Dynamics of Marine Organisms
James Bellingham (Monterey Bay Aquarium Research Institute) – ***8***

Live Andes: Endangered wildlife ID and geo-mapping digital images for ecological studies
Cristián Bonacic, et al (Pontificia Universidad Católica de Chile) - ***9***

Home Heating Using GPS-Based Arrival Prediction
A.J. Brush, James Scott, John Krumm, Brian Meyers, Ashish Kapoor (Microsoft Research) – ***10***

Wireless Sensor Networks for Environmental Monitoring: A Case Study in a Brazilian Rainforest
Doug. Carlson, J. Gupchup**, Andreas Terzis** (Johns Hopkins University), and
Rob Fatland (Microsoft Research) – ***11***

OData: The Open Data Protocol
Pablo Castro (Microsoft Corporation) – ***12***

A Global Tree Database
You-Wei Cheah (Indiana University, Bloomington) *et al* – ***13***

Utilizing "Crittercam" video footage to analyze animal behavior in the forest canopy
Lisa Dabek (Woodland Park Zoo) and Kyler Abernathy (National Geographical Society) – ***14***

Generalization, adoption for small synoptic datasets
Rob Fatland (Microsoft Research) – ***15***

SinBiota 2.0: Displaying Biodiversity and Environmental data from Brazil using Bing Maps
Pedro Feijão, Cleber Mira, Joao Meidanis (Scylla Bioinformatics, Campinas, Brazil), and
Carlos Joly (Institute of Biology, University of Campinas, Brazil) – ***16***

OneBusAway: A Suite of Transit Traveler Information Tools
Brian Ferris, Kari Watkins, and **Alan Borning** (University of Washington) – ***17***

Going Home for Data: In-Home Sensing for e-Science
Jon Froehlich (University of Washington) – ***18***

Green-City Transport Architecture: Closing the loop between sensors, navigators, and routes.
Mario Gerla, Giovanni Pau, Eugenio Giordano (University of California, Los Anglous) – ***19***

Water Resource Model Interoperability
Jonathan Goodall (University of South Carolina) – ***21***

© 2010 Microsoft Corporation. All rights reserved.

TeraPixel: A Spherical Image of the Sky
Dean Guo, Christophe Poulain (Microsoft Research) – ***22***

SEAMONSTER: A Testbed Sensor Web in a Partially Glaciated Watershed in Southeast Alaska
Matt Heavner (Los Alamos National Lab, University of Alaska) – ***23***

SQL is Dead; Long Live SQL: Cloud Query Services for Ad Hoc Research Data
Bill Howe (University of Washington) – ***24***

Bridging the Gaps: Satellites to Science and Desktop to the Cloud
Catherine Van Ingen (Microsoft Research) – ***25***

SQL Server Spatial – A Powerful Platform for Storing and Querying Geospatial Data
Ed Katibah (Microsoft SQL Server) – ***26***

Capturing the Long Tail of Sensor Web
Steve Liang (University of Calgary, Canada) – ***27***

Forecasting Changes in Urban Land Use with Bing Maps
Yi Liu (Tsinghua University, China) and Yan Xu (Microsoft Research) – ***28***

Towards GeoS3Web-based Virtual Environmental Observatories
Yong Liu (University of Illinois at Urbana-Champaign) – ***29***

Easily Disseminating High Volumes of Model Output
John McGee (Renaissance Computing Institute, University of North Carolina at Chapel Hill) – ***30***

Bing Maps
Chris Pendleton (Microsoft Corporation) – ***31***

Towards Proxy Workflow Execution in Environmental Research: a Case Study
Beth Plale (Indiana University Bloomington) – ***32***

Network of environmental sensors in tropical rain forests
Celso von Randow (National Institute for Space Research, Brazil) – ***33***

Citizen Environmental Science and the Geospatial Web 2.0
Renee Sieber (McGill University, Canada) – ***34***

Protecting Watersheds with Web Based Characterization
Stephen Stanley (Washington Department of Ecology) – ***35***

Tsunami forecast challenge: development, implementation, information dissemination
Vasily Titov (National Oceanic and Atmospheric Administration) – ***37***

Open and easy access to geophysical data
Chat Trabant & Tim Ahern (Incorporated Institutions for Seismology) – ***38***

Semantically-enabled Discovery
Evelyne Viegas (Microsoft Research) – ***39***

Quantifying the Environmental benefits of Electronic Software Distribution …
Dan Williams (University of Reading, UK) – ***40***

Sensors, Events, Decisions, and Actions
Roy Williams et al (California Institute of Technology) – ***41***

Seamless Integration and Visualization of Heterogeneous and Distributed Earth Science Data Using Bing Map
Huayi Wu & Chaowei (George Mason University) and Yan Xu (Microsoft Research) – ***42***

Global Geo-Referenced Field Photo Library
Xiangming Xiao (University of Oklahoma) and Pavel Dorovskoy (University of New Hampshire) – ***43***

Spatial Computing: Utilizing spatial principles and distributed computing to support environmental sciences
Chaowei Yang et al (Western Michigan University) and Yan Xu (Microsoft Research) – ***44***

CUAHSI Hydrologic Information System
Ilya Zaslavsky (University of California at San Diego) – ***46***

Exploring the Cross-Cultural Effects of Feedback for Pro-Environmental Behavior Change
Chen Zhao (Microsoft Research), Jon Froehlich & James Landay (University of Washington) – ***47***

© 2010 Microsoft Corporation. All rights reserved.

Introduction

As environmental sciences are changing and becoming more computational and data intensive, moving into the Fourth Paradigm, we at Microsoft Research believe that environmental research can be greatly accelerated through the use of many off the shelf products and technologies. To help with this effort of utilizing technologies to improve our planet, we would like to collaborate with scientific experts to investigate our research technologies and products in a way that significantly leads to new scientific insights.

To solve today's environmental problems, many researchers are dealing with massive datasets, broadly heterogeneous data types, complex computing algorithms and modeling, etc. These are not environmental problems alone but computational and data challenges as well. Collaborating with computing experts will enable these researchers to better focus on their science.

The Microsoft Environmental Research Workshop 2010 is designed to be a one-day intensive mind-swap event between environmental researchers and Microsoft technologists. As a result of this event, we look forward to possibly initiating collaboration projects using Microsoft technologies to solve environmental research problems.

To prepare all of us for the discussions at the workshop, we have asked all attendees to provide a one-page position paper to address the topics including, but not limited to, the following:

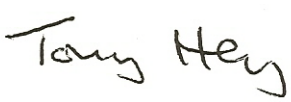
- Position papers on computational challenges in environmental research
- Case studies of environmental research that define software innovation in the field
- Technologies that enable or advance environmental public awareness or citizen science
- Use scenarios of applications and devices designed specifically for environmental research
- Demonstration of a Microsoft technology used in accelerating environmental research

We are thrilled with the position papers being collected here, and we are looking forward to carrying on the discussions further at the event and in the future…

Yan Xu
Event Chair, Microsoft ERW2010
Sr. Research Program Manager
Earth, Energy, and Environment
Microsoft Research

Dan Fay
Director
Earth, Energy, and Environment
Microsoft Research

Daron Green
Sr. Director of eScience
Microsoft Research



Tony Hey
Corporate Vice President
Microsoft External Research

© 2010 Microsoft Corporation. All rights reserved

ERW2010 Agenda

08:00 AM - 09:00 AM	*Continental Breakfast*
09:00 AM - 09:05 AM	**Welcome** Yan Xu (Microsoft Research)
09:05 AM - 09:20 AM	**Opening Speech: Environmental Research at Microsoft** Dan Fay (Microsoft Research)
09:20 AM - 09:40 AM	**Bridging the gaps: satellites to sensors and desktop to cloud** Catharine Van Ingen (Microsoft Research)
09:40 AM - 10:00 AM	**Virtual Environmental Observatories** Yong Liu (National Center for Supercomputing Applications, University of Illinois, Urbana-Champaign)
10:00 AM - 10:10 AM	**A Global Tree Database** You-Wei Cheah (Indiana University)
10:10 AM - 10:20 AM	**Tsunami forecast challenge: development, implementation, information dissemination** Vasily Titov (National Oceanic and Atmospheric Administration (NOAA))
10:20 AM - 10:30 AM	**SQL Server Spatial – A Powerful Platform for Storing and Querying Geospatial Data** Ed Katibah (Microsoft SQL)
10:30 AM - 10:40 AM	*Break*
10:40 AM - 11:50 AM	**Roundtable Discussions** All / Chairs: A.J. Brush, Rob Fatland, Chris Pendleton, Evelyne Viegas
11:50 AM - 01:30 PM	**Group Picture, Lunch, and group presentations** (representatives from each group)
01:30 PM - 01:50 PM	**Open Data Protocol** Pablo Castro (Data Framework, Microsoft)
01:50 PM - 02:10 PM	**Wireless Sensor Networks for Environmental Research: A Case Study from a Field Campaign in a Brazilian Rain Forest** Andreas Terzis (Johns Hopkins University)
02:10 PM - 02:20 PM	**The development of GeoCENS: finding the long tail in sensor web** Steve Liang (Calgary University, Canada)
02:20 PM - 02:30 PM	**Project Dallas - a data marketplace in the Windows Azure** Christian Liensberger (Project Dallas, Microsoft)
02:30 PM - 02:40 PM	**Citizen Environmental Science and the Geospatial Web 2.0** Renee Sieber (McGill University, Canada)
02:40 PM - 02:50 PM	*Break*
02:50 PM - 03:10 PM	**Bing Maps** Chris Pendleton (Microsoft Bing Maps)
03:10 PM - 03:20 PM	**Spatial Computing: Utilizing spatial principles and distributed computing to support environmental sciences** Phil Yang (George Mason University)
03:20 PM - 03:30 PM	**Seamless Integration and Visualization of Heterogeneous and Distributed Earth Science Data Using Bing Map** Hannes Wu (George Mason University)
03:30 PM - 03:40 PM	**Forecasting Crop Disease Risk Using Mesoscale Weather Forecasts** Kathleen M. Baker (Western Michigan University)
03:40 PM - 04:40 PM	**Roundtable Discussions** All / Chairs: James Bellingham, Matt Heavner, Beth Plale, Chaowei Yang
04:40 PM - 05:40 PM	**Group Presentations** (representatives from each group)
05:40 PM - 06:00 PM	**Closing Remarks** All
06:00 PM	*End*

© 2010 Microsoft Corporation. All rights reserved

U.S. Geoscience Information Network (GIN) and Convergence towards Global Data Integration in the Geosciences

Lee Allison[1], Stephen Richard[1], Linda Gundersen[2], & Ian Jackson[3]

[1]Arizona Geological Survey, [2]U.S. Geological Survey, [3]British Geological Survey

Contributing scientific solutions to the challenges of landslides and earthquakes, minerals and mining, water supply and flooding, pollution and erosion, and not least, climate change and energy supply, depends absolutely on geological data. Few of these challenges respect state, national, or science domain frontiers and if we want to assess and address these environmental challenges holistically then we need access to holistic data too. Rich environmental data are extensive, but when they are available, they are often exceptionally difficult to discover, exist in different formats and via different services, with different access conditions.

Efforts around the world are converging towards creation of an integrated global digital data network for the geosciences based on common standards and protocols for data discovery and access, and a shared vision of a distributed, web-based, open source, interoperable data integration. Commonalities include use of OGC-compliant standards and GeoSciML as the data interchange mechanism. The Geosciences Information Network (GIN; http://usgin.org, http://lab.usgin.org) is linking databases and web services in the U.S. state geological surveys (Association of American State Geologists - AASG) and the U.S. Geological Survey (USGS). The USGS Council on Data Integration has validated GIN as a step towards resolving the challenges in linking data resources from all of its four branches – geology, geography, water, and biology. GIN has been adopted as the data exchange mechanism for the U.S. Department of Energy-funded National Geothermal Data System (NGDS) which will establish data nodes in every state in the U.S. and populate the system with large volumes of state-specific data, creating the infrastructure for a national network.

Critical system components of GIN include standardized catalog services to register and discover resources, web map services to display georeferenced images, and feature services to transport data. Open Geospatial Consortium service components are being used to meet all of these requirements. The Catalog Service for the Web (CSW) ISO 19115 profile provides services to search metadata registries and obtain results in a standard format. The GIN project is participating in the Energistics' petroleum consortium's Energy Industry Metadata Standards Working Group, with representatives from the upstream petroleum industry, to develop an industry metadata profile that is compatible with metadata services for other geoscience domains.

The OneGeology (www.onegeology.org) initiative to make accessible online digital geologic map data for the world has 116 participating countries, providing over 120 map services from 46 nations using OGC WMS (Web Map Service), and with four nations now using OGC WFS (Web Feature Service) through a dynamic web portal. These map services demonstrate the maturity of the WMS for production-level data publication. OneGeology–Europe (www.onegeology-europe.eu/) is a European Union project in which 26 national geological surveys and organizations are collaborating to build a continent-wide geoscience data network. Emerging practices from the global project, OneGeology, and the regional initiatives OneGeology-Europe, and GIN, provide a foundation for the potential to create a comprehensive global digital data network of geoscience (and geospatial) information. The next step is providing structured data for geoscience features using OGC Web Feature Services utilizing GeoSciML as the data transport schema.

© 2010 Microsoft Corporation. All rights reserved

2011 GSA Annual Meeting in Minneapolis (9–12 October 2011)
Paper No. 9-9

Presentation Time: 10:30 AM-10:45 AM

NATIONAL GEOTHERMAL DATA SYSTEM

ANDERSON, Arlene F.[1], CUYLER, David[1], SNYDER, Walter S.[2], ALLISON, M. Lee[3], BLACKWELL, David D.[4], and WILLIAMS, Colin F.[5], (1) U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy, Geothermal Technologies Program, 1000 Independence Ave, Washington, DC 20585, Arlene.Anderson@ee.doe.gov, (2) Department of Geosciences, Boise State University, 1910 University Drive, Boise, ID 83725, (3) Arizona Geological Survey, 416 W. Congress, #100, Tucson, AZ 85701-1381, (4) Huffington Department of Earth Sciences, Southern Methodist Univeristy, PO Box 750395, Dallas, TX 75275, (5) U.S. Geological Survey, MS 977, 345 Middlefield Road, Menlo Park, CA 94025

The goal of the U.S. Department of Energy's National Geothermal Data System is to design, build, implement, deploy and populate a national, sustainable, distributed, interoperable network of data and service (application) providers. These providers will develop, collect, serve, and maintain geothermal-relevant data that operates as an integral component of NGDS. As a result the geothermal industry, the public, and policy makers will have access to consistent and reliable data, which in turn, reduces the amount of staff time devoted to finding, retrieving, integrating, and verifying information. With easier access to information, the high cost and risk of geothermal power projects (especially exploration drilling) is reduced.

Five separate NGDS projects provide the data support, acquisition, and access to cyber infrastructure necessary to reduce cost and risk of the nation's geothermal energy strategy and US DOE program goals focused on the production and utilization of geothermal energy. The U.S DOE Office of Energy Efficiency and Renewable Energy Geothermal Technologies Program is developing the knowledge and data foundation necessary for discovery and development of large-scale energy production while the Buildings Technology Program is focused on other practical applications such as direct use and residential/commercial ground source heat pumps.

The NGDS provides expanded reference and resource data for research and development activities (a subset of the US DOE goals) and includes data from across all fifty states and the nation's leading geothermal research centers. Thus, the project incorporates not only high-temperature potential but also moderate and low-temperature locations incorporating US DOE's goal of adding more geothermal electricity to the grid. The program, through its cyberinfrastructure, will help lead to innovative exploration technologies through increased data availability on geothermal energy capacity. Finally, the project will contribute new data from previously unexplored locations.

NGDS is being built using the US Geoscience Information Network (US GIN) data integration framework to promote interoperability across the Earth sciences community and with other emerging data integration and networking efforts.

2011 GSA Annual Meeting in Minneapolis (9–12 October 2011)
General Information for this Meeting

Session No. 9
Advancement in the Exploration Practices Used to Characterize Geothermal Resources
Minneapolis Convention Center: Room 101H-J

8:00 AM-12:00 PM, Sunday, 9 October 2011

Geological Society of America *Abstracts with Programs*, Vol. 43, No. 5, p. 41

2011 GSA Annual Meeting in Minneapolis (9–12 October 2011)
Paper No. 170-9

Presentation Time: 9:00 AM-6:00 PM

U.S. GEOSCIENCE INFORMATION NETWORK: A CRITICAL PATH FOR DATA INTEGRATION IN THE U.S. EARTH SCIENCES

ALLISON, M. Lee, Arizona Geological Survey, 416 W. Congress, #100, Tucson, AZ 85701-1381, lee.allison@azgs.az.gov and GALLAGHER, Kevin T., Core Science Systems, U.S. Geological Curvey, 12201 Sunrise Valley Drive, 108 National Center, Reston, VA 20192

Development efforts for the U.S. Geoscience Information Network (US GIN) over the last two years have crystallized around the Community for Data Integration (CDI) at the USGS, and the 50-state AASG State Geothermal Data project. The next step in developing a USGS-AASG community is to bring these two efforts into closer alignment through greater participation in CDI activities by geoinformatics practitioners from state geological surveys, and implementation of test bed activities by the USGIN partners, within two years.

Test bed activities in the geological survey community will define a scope and provide a foundation to promote the use of specifications developed by the larger geoinformatics community. Adoption of some of these specifications as ‘standards’ by USGS and AASG for use by those organizations will lend authority and motivate wider adoption.

The arc from use case to test bed to production deployments to agreement on ‘standard’ specifications for data discovery and access must be propelled by active interest from the user communities who have a stake in the outcome. The specifications developed will benefit the organizations involved in development, testing and deployment, which motivates participation -- a model that has worked successfully for standards organizations such as OGC, ISO and OASIS.

The governance structure to support such a community process should promote grass root nucleation of interest groups that are the core of development efforts. Some mechanism for community agreement on priorities is desirable because geological survey agencies will need to allocate resources to support development. Loosely knit organizations such as ESIP and the current CDI provide models for this kind of structure.

Because many geological surveys have data archive and dissemination functions as part of their portfolio, some support for the system can be built into the operating expenses and overhead. Sharing of resources and reuse of components can reduce the cost. Wide adoption of similar software, protocols and practices increases the number of stake holders with an interest in supporting the system.

2011 GSA Annual Meeting in Minneapolis (9–12 October 2011)
General Information for this Meeting

Session No. 170--Booth# 41
Geology and Health (Posters)
Minneapolis Convention Center: Hall C
9:00 AM-6:00 PM, Tuesday, 11 October 2011

Geological Society of America *Abstracts with Programs*, Vol. 43, No. 5, p. 419

2011 GSA Annual Meeting in Minneapolis (9–12 October 2011)
Paper No. 223-10

Presentation Time: 11:00 AM-11:15 AM

DELIVERING GEOSCIENCE KNOWLEDGE IN FEDERAL SYSTEMS

JACKSON, Ian, Information Directorate, British Geological Survey, Keyworth, Nottingham, NG12 5GG, United Kingdom, ij@bgs.ac.uk, BROOME, John, Earth Sciences Sector, 615 Booth Street, Ottawa, ON K1A0E8, Canada, and ALLISON, M. Lee, Arizona Geological Survey, 416 W. Congress, #100, Tucson, AZ 85701-1381

Across the globe, geological communities are facing the same four challenges: put simply, how do we best make data discoverable, shareable, viewable and downloadable, so that the user also has access to consistent data at a national and continental level? The principle of managing scientific data and knowledge where it is generated and is best understood is well established in the science community. The distributed nature of most data sources means the complementary delivery mechanism of web map services has become equally prevalent in the spatial data community. Together these two factors are driving a world-wide revolution in the way spatial geoscience information is being disseminated to its users. The outcome is that data are being managed and delivered from multiple component sources - a federated system - ie the individual states within a union. These systems exist in the USA, in Canada, in Australia, and progressively, also in Europe, where the European Union can be regarded as a federal analogue, and where new regulation is placing the force of law behind spatial data infrastructures. In these "systems" addressing the four challenges are however, far from simple. To address them means finding solutions to adequate but workable metadata description, data specifications which encompass the richness of the data but deliver continuity, web map interfaces which allow flexible access but are easy to use, and last but not least intellectual property rules that protect the originator but provide the data the users need. The models for collaboration emerging in each of the federated systems are moving towards consensus on a global digital integration framework in the geosciences. This paper will introduce a session which will draw on the rich experience in North America and Europe, and explore the way the challenges have been articulated and addressed with a strong emphasis on gaining future benefit by sharing the lessons learned

2011 GSA Annual Meeting in Minneapolis (9–12 October 2011)
General Information for this Meeting

Session No. 223
Geoscience Information/Communication: Printed Past, Digital Future—We Hold the Key
Minneapolis Convention Center: Room 101DE
8:00 AM-12:00 PM, Wednesday, 12 October 2011

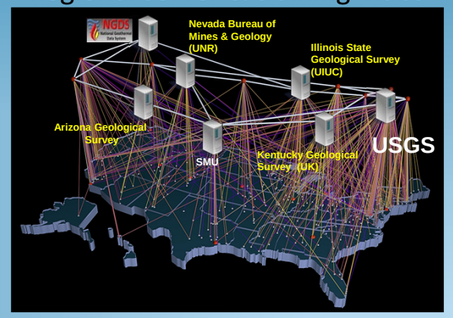
Geological Society of America *Abstracts with Programs*, Vol. 43, No. 5, p. 538

Talks/ Presentations

Application of the U.S. Geoscience Information Network to Deploying a National Geothermal Data System

M. Lee Allison, Stephen M. Richard, Ryan J. Clark, Wolfgang Grunberg, Arizona Geological Survey, Tucson, AZ

Data nodes for all 50 states
Regional server & training hubs

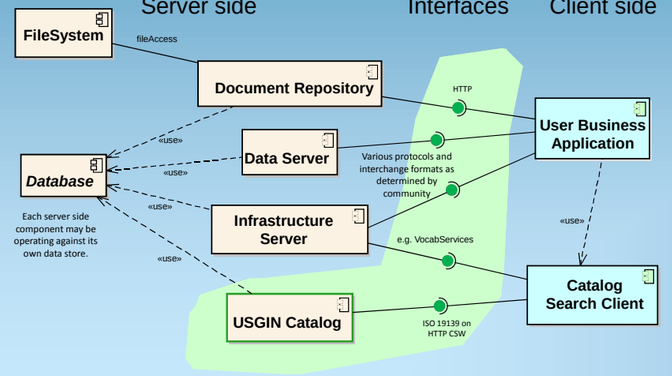


System Data Types

DisplayName	Implementation
Active fault	Feature class
Well log	1-D coverage
Borehole	Feature class
Borehole temperature measurement	Observation class
Crustal strain rate	Observation class
Dike, vein, or marker bed	Feature class
Document object	Metadata record
Drill stem test	Observation class
Earthquake epicenter	Observation class
Flooded mine feature	Feature class
Fluid chemistry	Observation class
Geothermometer observation	Observation class
Contact, fault, fold	Feature class
Lithology log interval	Feature class
Map unit polygon	Feature class
Geothermal facility	Feature class
Geothermal system	Feature class
Geologic unit	Feature class
Heat flow measurement	Observation class
Hot spring	Feature class
Metadata	Metadata record
Physical property (soft typed)	Observation class
Co-produced fluid feature	Feature class
Whole rock radiogenic element analysis	Observation class
Resource suitability	Feature class
Sample	Feature class
Area coverage	2-D raster coverage
Soil classification polygons	Feature class
Structure data point	Observation class
Thermal well	Feature class
Trace element chemistry	Observation class
Volcanic vent area	Feature class
Map layer	WMS layer
Chemical analysis	Observation class

What is USGIN: the Geoscience Information Network

Visit http://usgin.org for more overview information or
http://lab.usgin.org for technical information, documents and discussion

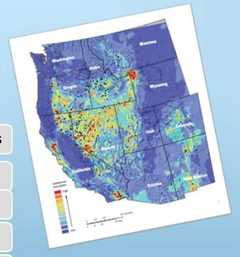


The network is defined by a collection of shared web services and interchange formats that facilitate finding, accessing, and using geoscientific information, and a registry (catalog) of resources that use the services and formats.

The USGIN framework:

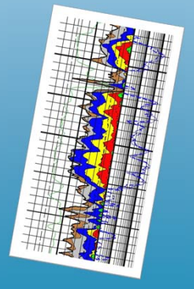
- ✓Data publication using open specifications
- ✓Community participation in development of specifications
- ✓Registration of resources in interoperable catalog system
- ✓Respect for data ownership, tracking of data provenance

- **A coalition of State Geological Surveys is creating a national, sustainable, distributed, interoperable network of data providers representing all 50 states that will develop, collect, serve, and maintain geothermal relevant data as an integral component of the National Geothermal Data System (NGDS, www.geothermaldata.org).**
- U.S. Geoscience Information Network (USGIN) protocols and specifications provide the data integration framework.
- Goals:
  - enhance states' abilities to preserve and disseminate geothermal data
  - facilitate geothermal resource characterization and development
  - expand the scope of data available to the geothermal community
  - foster new services and applications built by third parties to take advantage of the system's capabilities and content
  - contribute materially to creation of a national geoinformatics system through implementation and deployment of NGDS
  - increase operational support for geoinformatics infrastructure through a broader user base.
- Data are made accessible by
  - digitizing at risk legacy, geothermal relevant data (paper records, samples, etc)
  - publishing existing digital data using standard web and data services
  - limited collection of new data in areas lacking critical information.

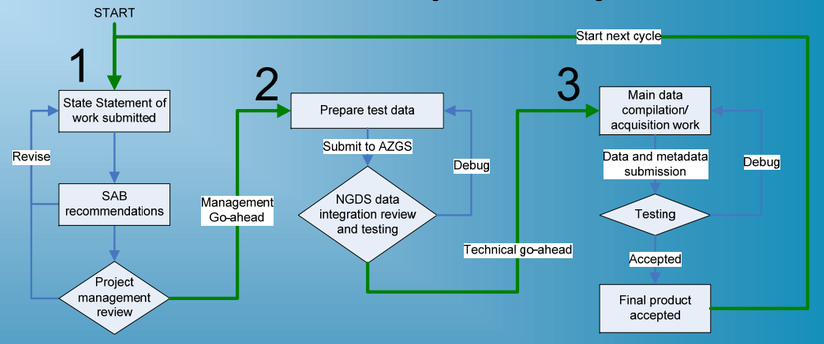


Selected data resources in State Geological Surveys

3.5 million wells – oil & gas, water
195,000 well logs
50,000 geothermal wells
750,000 BHT's
6 Tb existing digital data
>75,000 scanned publications & maps
2.5 million feet of core
600,000 sample logs



Data Development Cycle



http://stategeothermaldata.org

Steve.richard@azgs.az.gov

This material is based upon work supported by the **Department of Energy** under Award Numbers DE-EE0002850 and DE-EE0001120 and by **National Science Foundation** grant EAR-0753154

New Initiatives in the Development of a National Geoinformatics Community (*Invited*)

Abstract

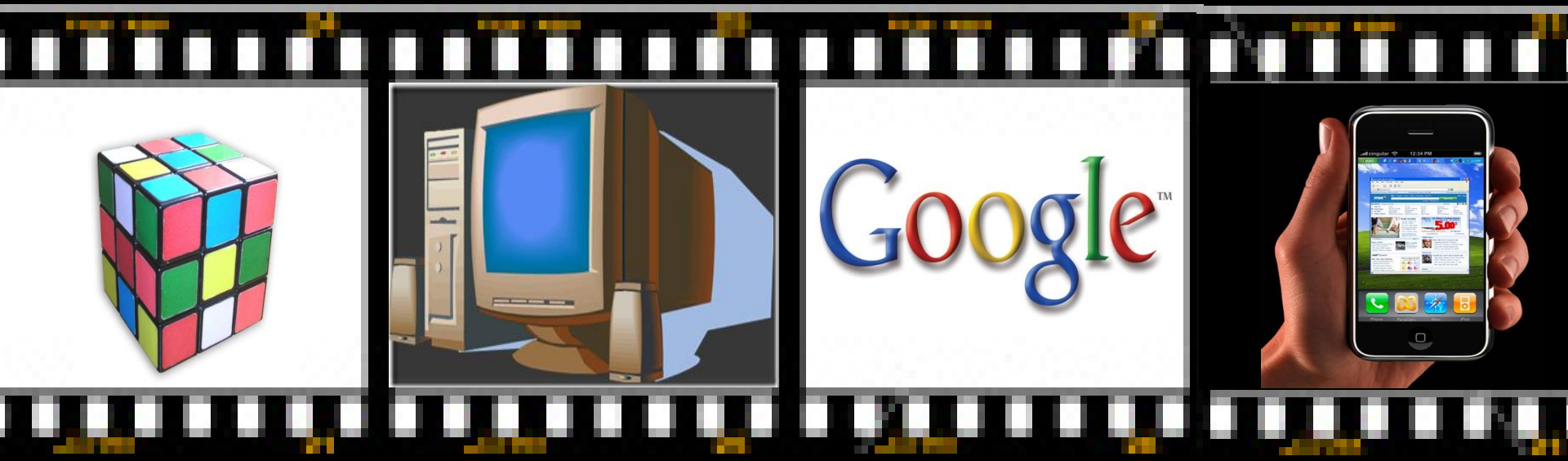
Geoinformatics has emerged as an important component of research and educational initiatives within the geoscience community. As such, it is transforming data collection and delivery, model development, knowledge representation and management, and education & outreach activities. In the past, the community has struggled to develop and maintain digital portals and data & knowledge management infrastructures that can interface with each other and have utility across geoscience disciplines and with the broader community. Recent integrative efforts have focused on the establishment of a National Geoinformatics Community (NGC) that will advance geoinformatics at all levels via outreach, advocacy, and fostering communities of practice. The NGC intends to encourage collaboration across geoscience communities through recording and disseminating community efforts, identifying and facilitating the development of community standards and needs, and exploring social aspects of data & knowledge sharing and ownership. Recent NGC initiatives include an NSF-funded workshop in Fall 2010 that gathered more than 50 members of the geoinformatics and geoscience communities to discuss and advance technology and education & outreach issues. Salient themes of the workshop include identification and discussion of: 1) existing resources and knowledgeable practitioners, 2) impediments to progress, and 3) the necessary future resources for data & knowledge and interoperability. Progress on all of these themes will be presented at this AGU session, along with future initiatives and a call to the geoscience community to become more aware of geoinformatics issues and to get involved in charting future directions.

Cite as: Whitmeyer, S J; Gundersen, L C; Walker, J D; Allison, M L; Babaie, H A; Cervato, C; Fils, D; Richard, S M; and Arrowsmith, R (2010), New Initiatives in the Development of a National Geoinformatics Community (*Invited*), Abstract ED22A-04 presented at 2010 Fall Meeting, AGU, San Francisco, Calif., 13-17 Dec.

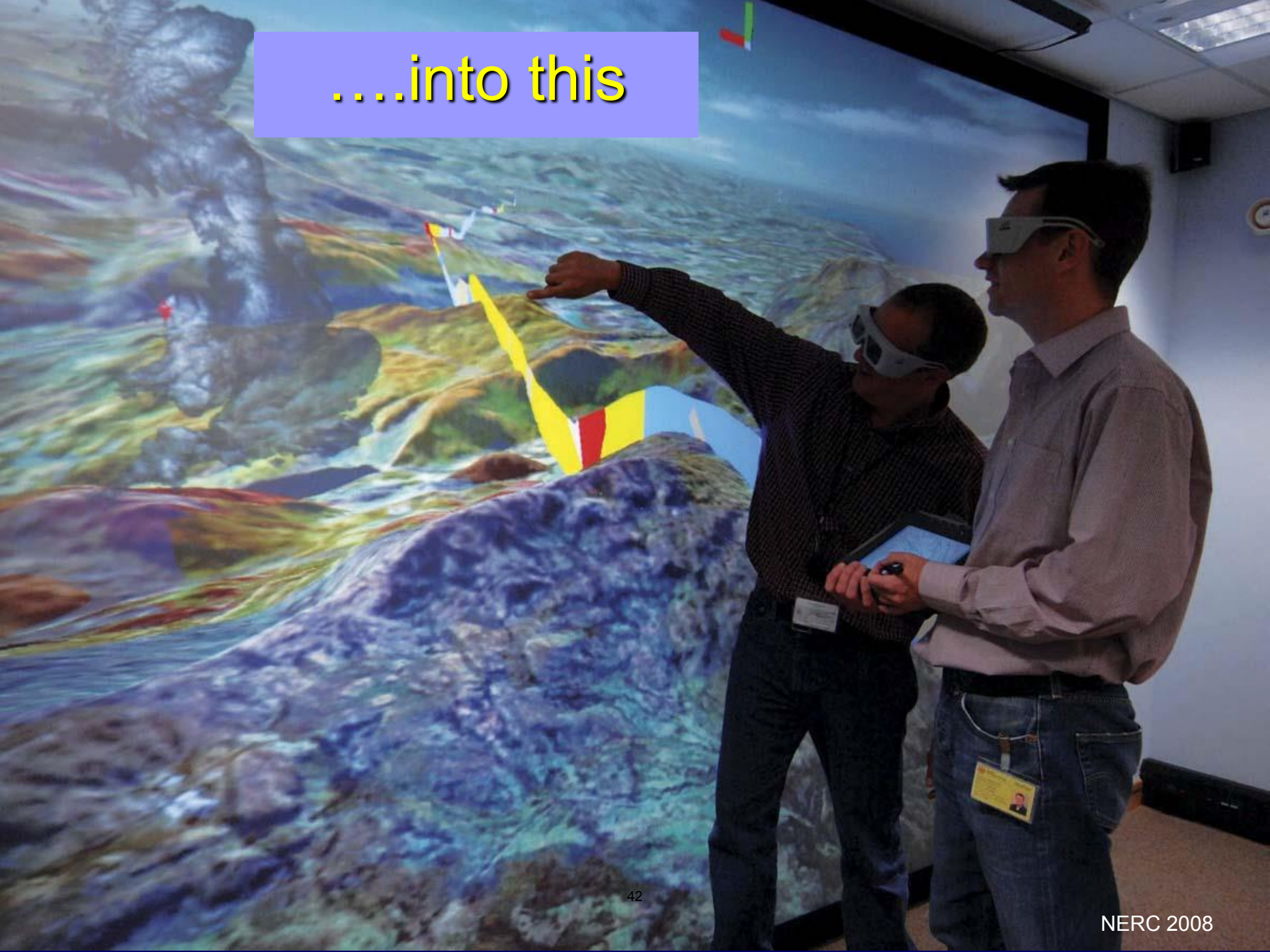
National Geothermal Data System and Global Geosciences Data Integration

Lee Allison, AZGS
Stephen Richard, AZGS

Arlene Anderson, US DOE
David Cuyler, US DOE & Sandia Natl Lab



how to turn this….

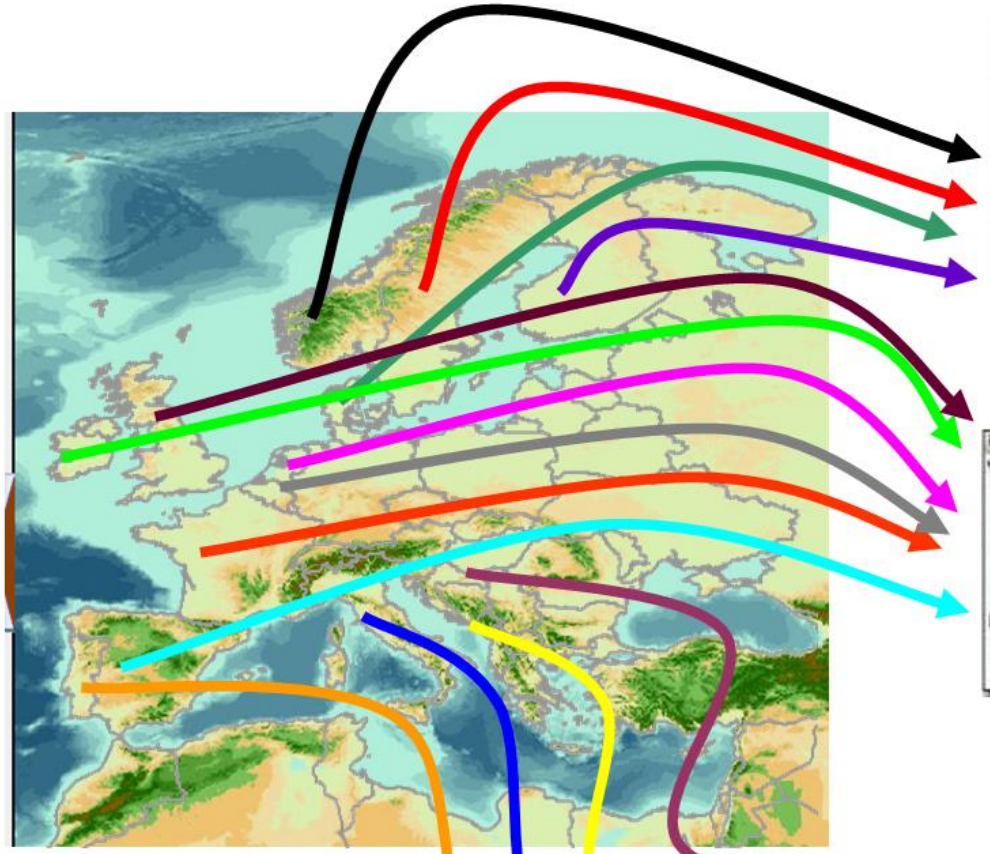

….into this

200 million+ websites –
if you don't have a website, you don't exist.

Prediction -

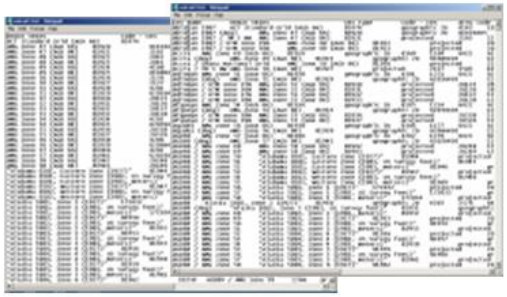
In 5-10 years, if your data are not online in an integrated, interoperable network, you won't exist.

Current electronic delivery



Proprietary Software

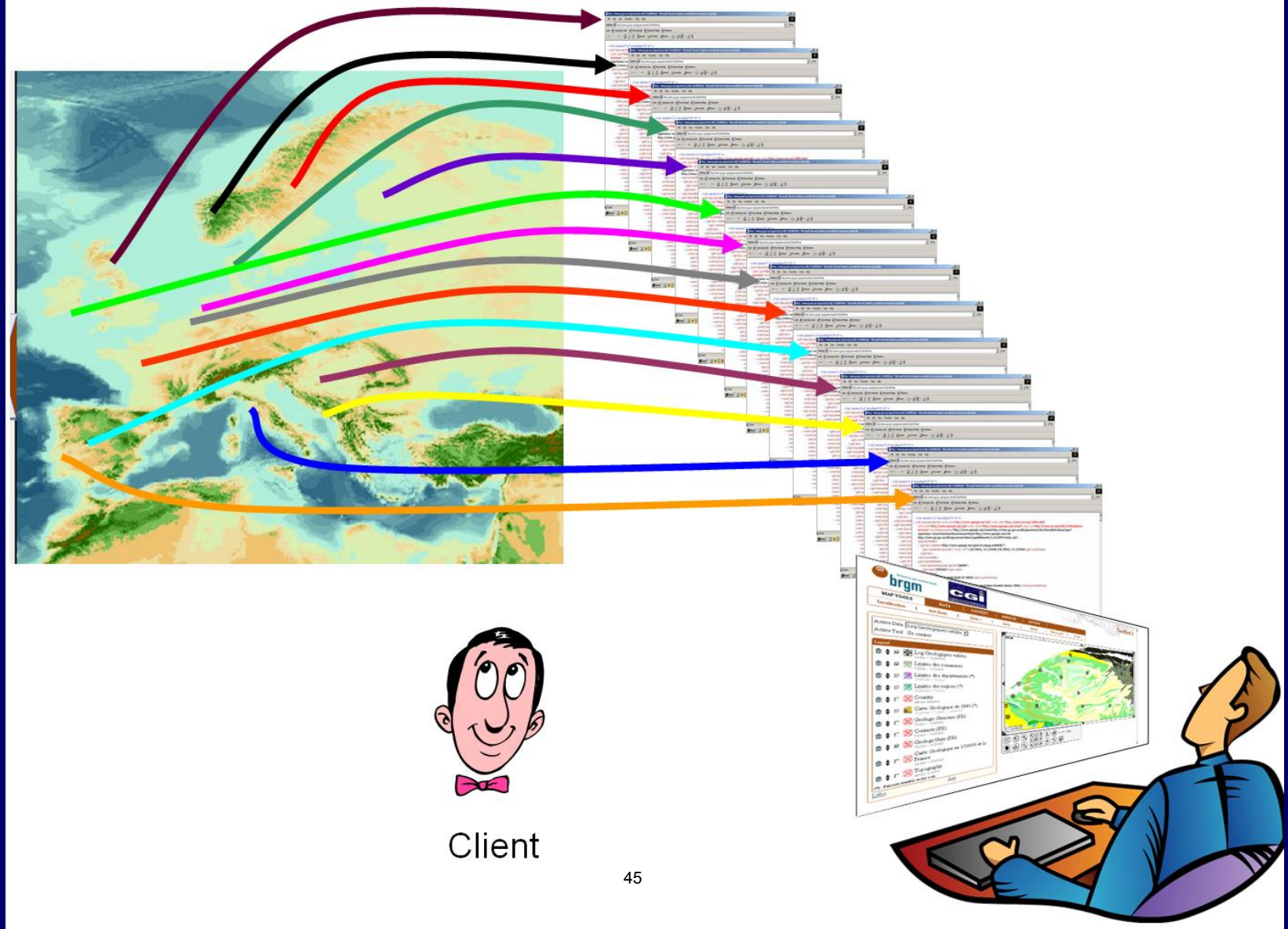
Versions of Software

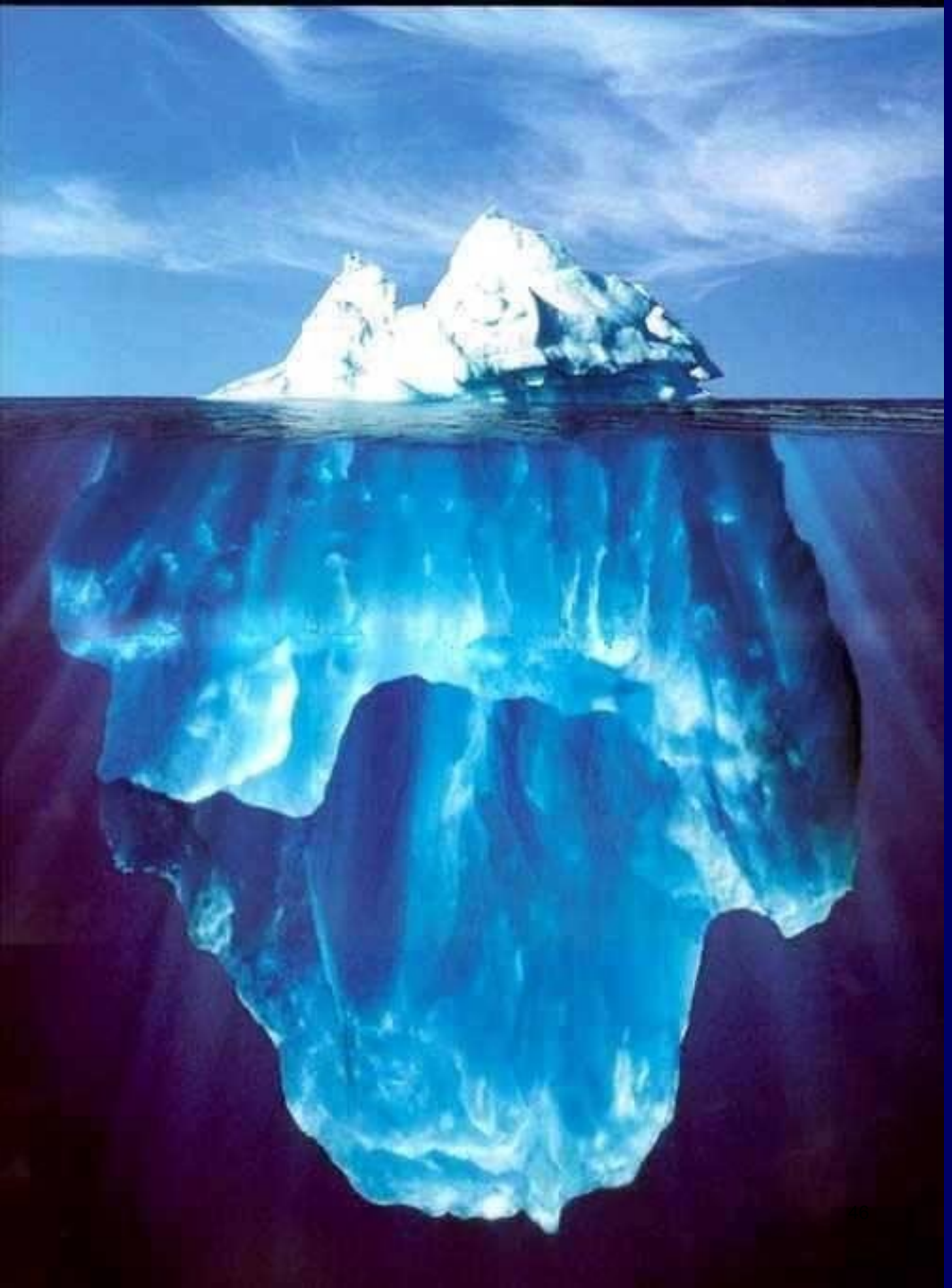


Data Structures

Client

The Goal





Most of the technology exists

Challenges are cultural and organizational

Challenges to building community

Who sets the standards?
Who controls the system?
Who makes the decisions?

One system to rule them all,
one system to find them,
one system to bring them all,
and in the darkness bind them.

Data Integration Framework

USGS
science for a changing world

AASG
Association of American State Geologists

Distributed
Web-based
Interoperable
Open source

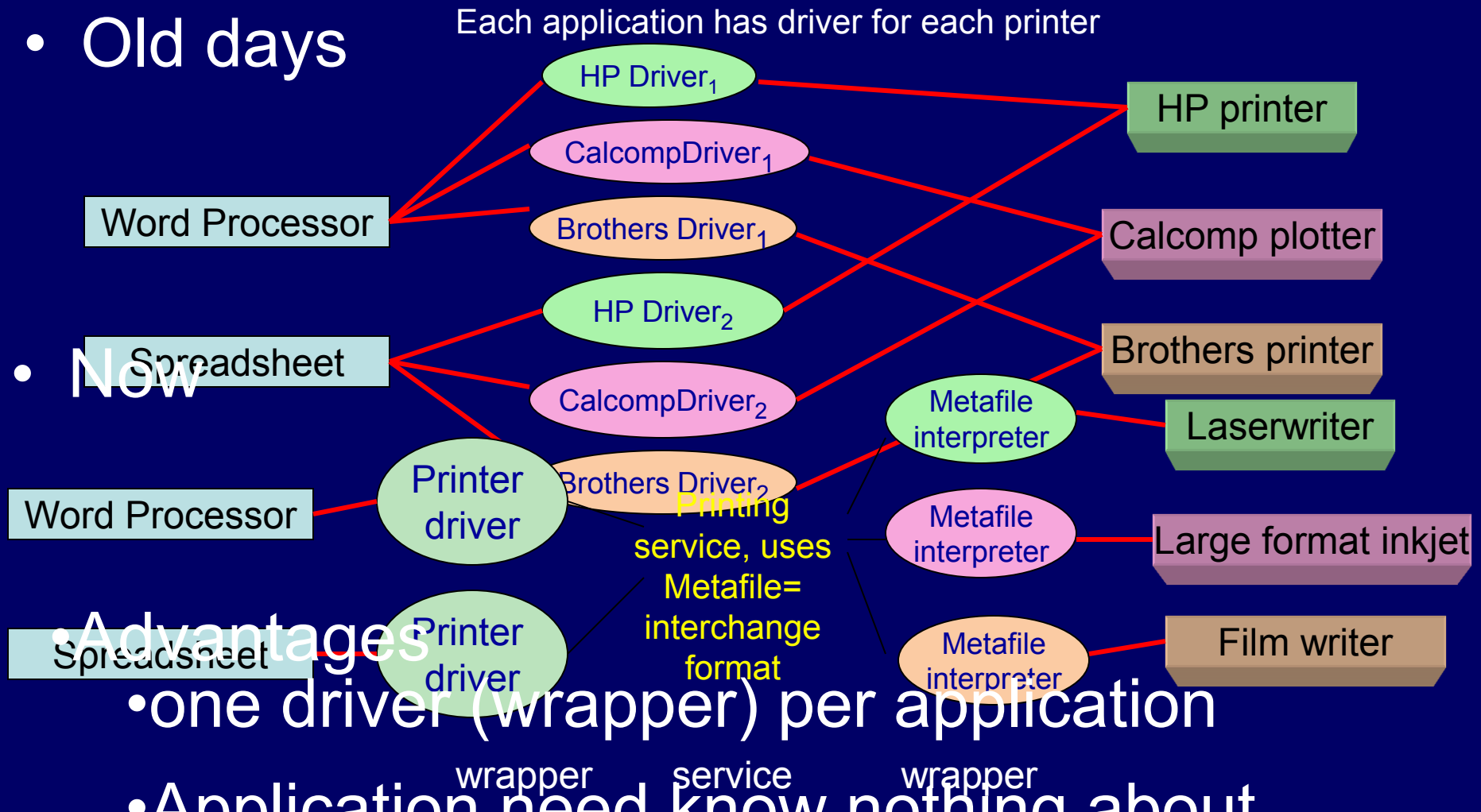
s

aps

Geoscience Information Network (GIN)

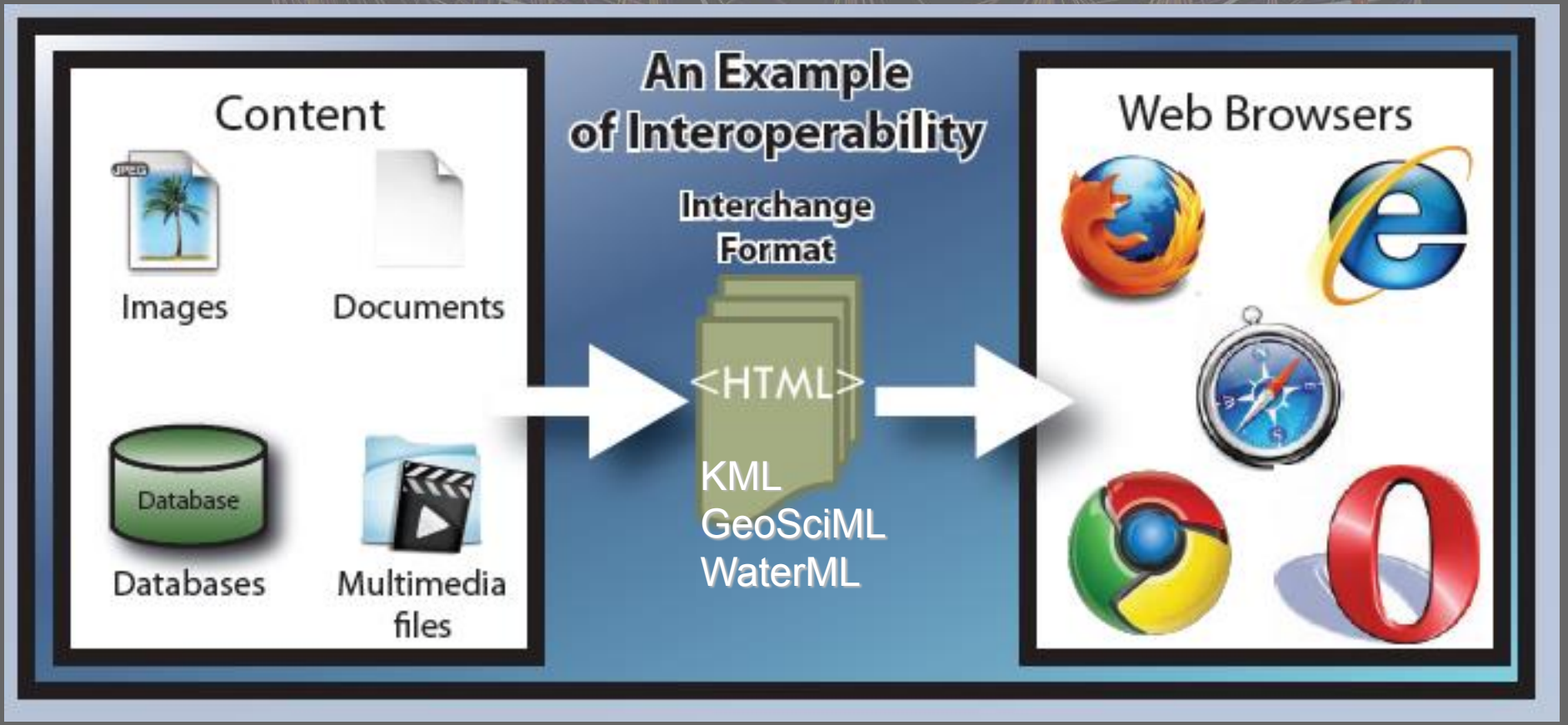
Computer printer services

- Old days



- Now
- Advantages
  - one driver (wrapper) per application
  - Application need know nothing about printer—separation of concerns

Interoperable: “Loosely-coupled”



a decade of advances in cyberinfrastructure and community practice

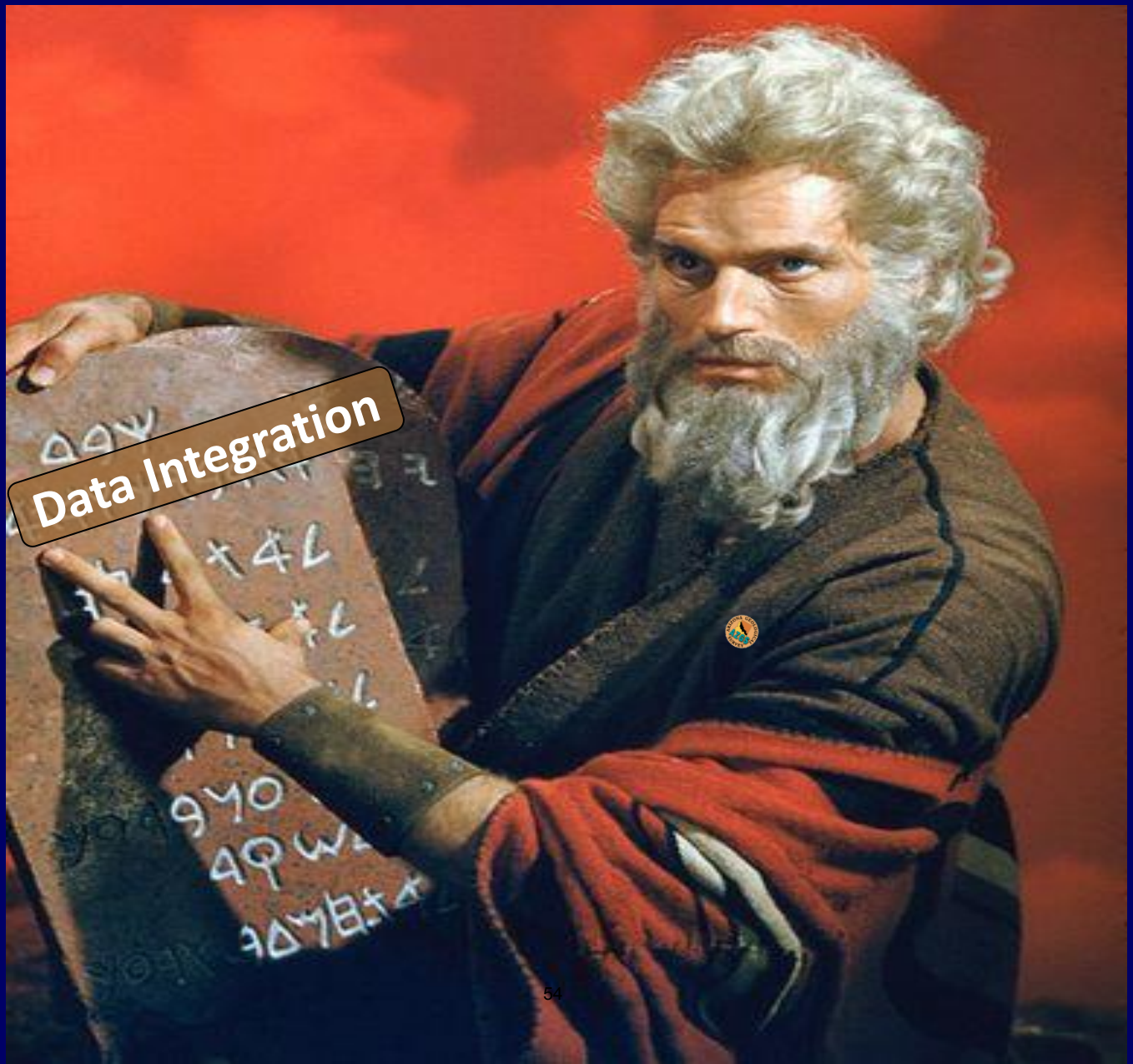
- Modular approach – build a network by adopting and linking existing capabilities

You already use Web services and apps

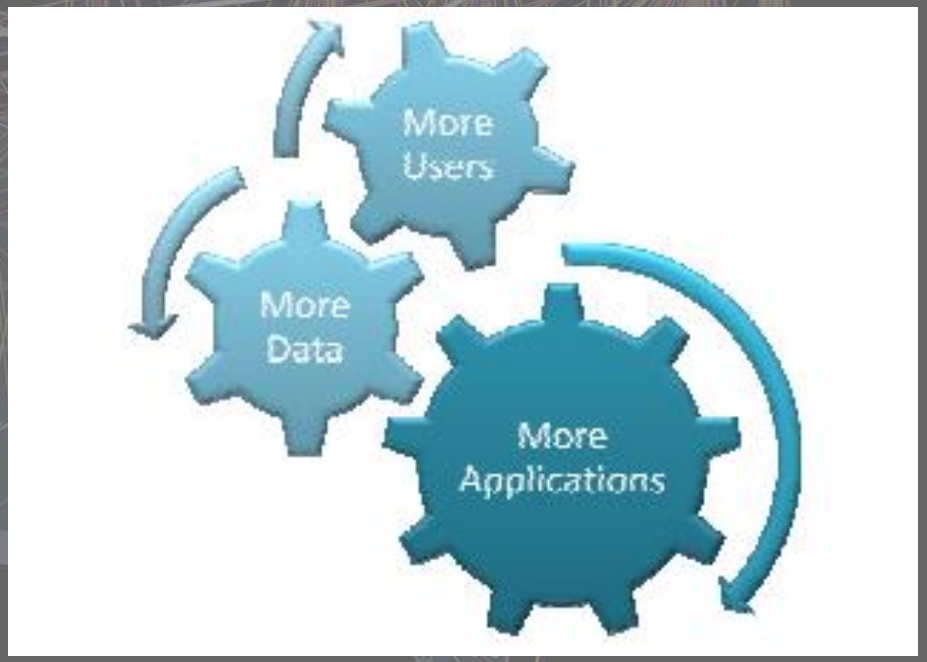
Make your data available to a wide variety of applications: Make it easy for developers to create custom applications, or mash-up your data with any other data sources.

300,000 apps for the iPhone

The network is voluntary, not imposed from above

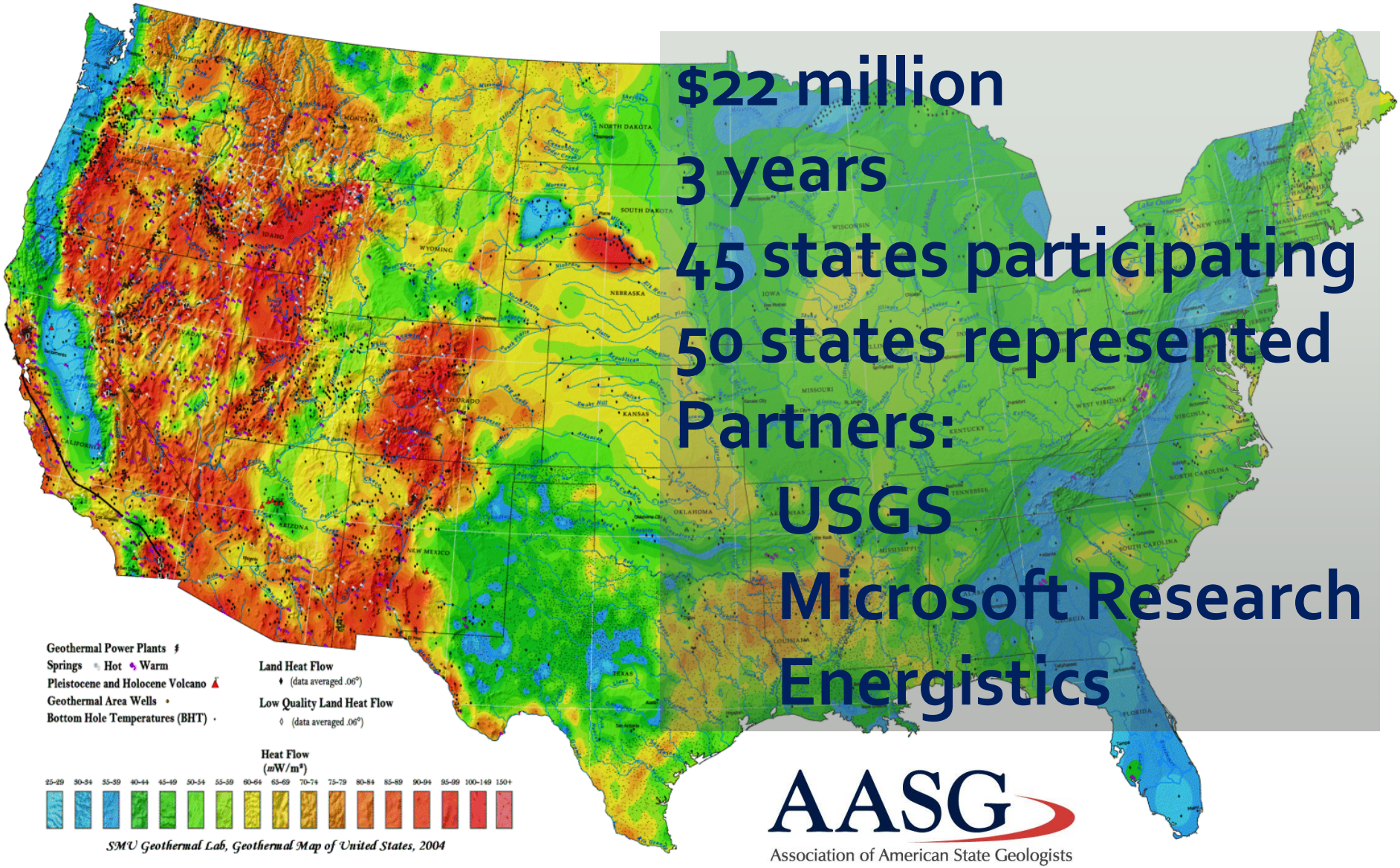


Sustainable business model



Geothermal community's #1 request: ready access to relevant data

State Geothermal Data Project



$22 million

3 years

45 states participating

50 states represented

Partners:

USGS

Microsoft Research

Energistics

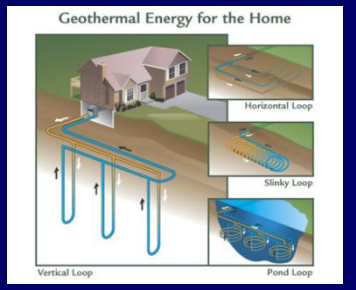
AASG
Association of American State Geologists

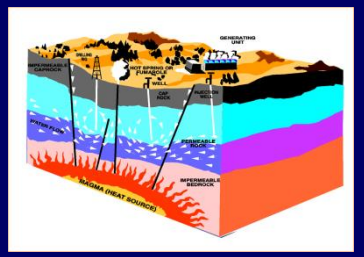
Heat flow - Southern Methodist University

State Geothermal Data

Selected data resources in State Geological Surveys

- 3.5 million wells – oil & gas, water
- 195,000 well logs
- 50,000 geothermal wells
- 750,000 BHT's
- 6 Tb existing digital data
- >75,000 scanned publications & maps
- 2.5 million feet of core
- 600,000 sample logs

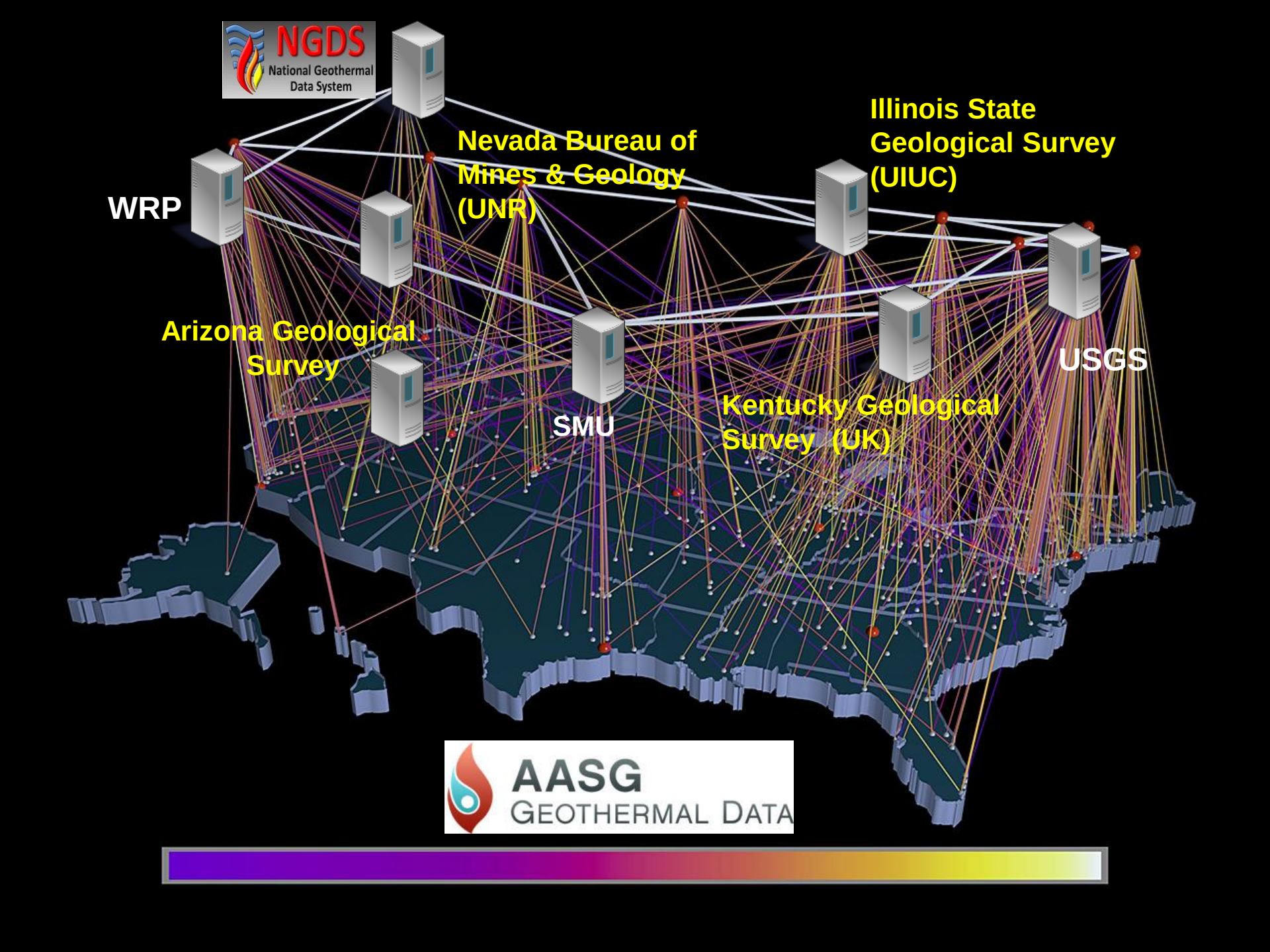




AASG
Association of American State Geologists

The "killer app" for geoscience data integration in the U.S.

DOE & USGS Data ***Boise State University***	**National Assessment** ***USGS***	**University Data** ***Southern Methodist University***	**"State Geological Survey Contributions to the NGDS"** ***AASG - AZGS***	**DOE GTP-funded projects**


NGDS
National Geothermal Data System
WRP
Nevada Bureau of Mines & Geology (UNR)
Illinois State Geological Survey (UIUC)
Arizona Geological Survey
SMU
Kentucky Geological Survey (UK)
USGS
AASG
GEOTHERMAL DATA

WRP Web Mapping Application
WESTERN REGIONAL PARTNERSHIP
WEB MAPPING APPLICATION
LAYERS
SEARCH
GROUP BY
SOURCE
CATEGORY
NONE
SHOW
METADATA
DOWNLOAD
Argonne National Laboratory
Bureau of Land Management
California Energy Commission
Department of Defense
Military Installations
Military Ranges
Military Training Route Centerline
Military Training Route Corridor
Special Use Airspace
Alert Areas
Military Operations Areas
Restricted Areas
Warning Areas
Environmental Systems Research Institute
Global Energy Decisions
National Renewable Energy Labora
The Nature Conservancy
U.S. Census Bureau
U.S. Fish and Wildlife Service
United States Department of Agric
SELECTED LAYERS
CLEAR
LAYERS TO DOWNLOAD
BACK
ZOOM
PAN
LAYER INFO
MEASURE
Las Vegas
Kingman
Flagstaff
Los Angeles
San Diego
Phoenix
Tucson
ARIZONA
100 mi
200 km
10,000 GIS layers from
15 federal agencies, 5 states

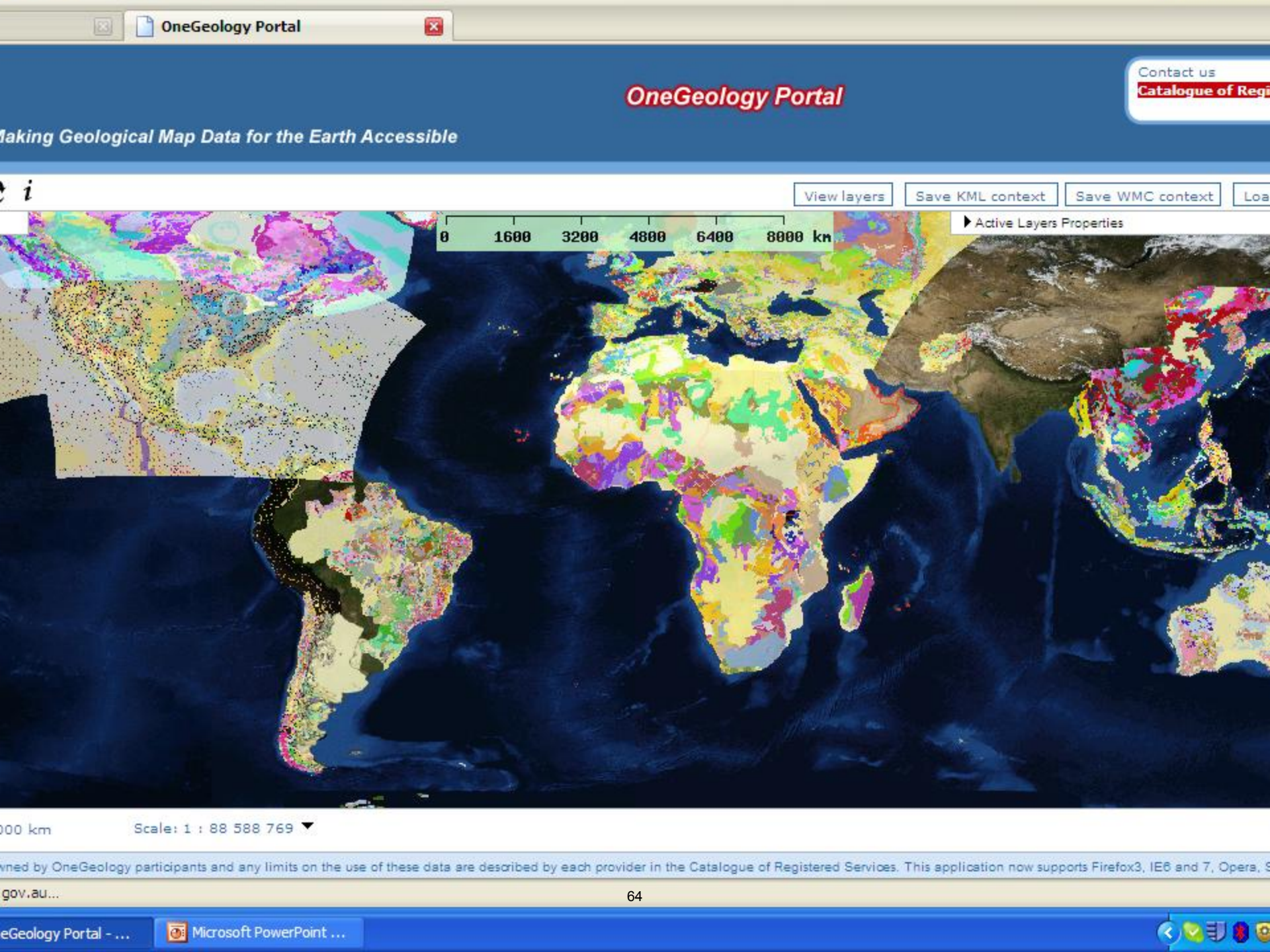
ONE EUROPE ONE GEOLOGY
APPLYING GEOSCIENCE FOR SOCIETY
It works!
ONE
Geology Europe

Energy Industry Metadata Standards Working Group

- End-to-end discovery, access, and exchange of upstream petroleum data

energistics
The Energy Standards Resource Centre

GIN (Geosciences Information Network) Architecture Example – *The NGDS Pilot Project*


OneGeology Portal
OneGeology Portal
Making Geological Map Data for the Earth Accessible
Contact us
Catalogue of Regi
View layers
Save KML context
Save WMC context
Active Layers Properties
0 1600 3200 4800 6400 8000 km
Scale: 1 : 88 588 769
owned by OneGeology participants and any limits on the use of these data are described by each provider in the Catalogue of Registered Services. This application now supports Firefox3, IE6 and 7, Opera,
gov.au...
eGeology Portal - ...
Microsoft PowerPoint ...

Towards a Global Data Network for the Geosciences

AASG Association of American State Geologists

USGS science for a changing world

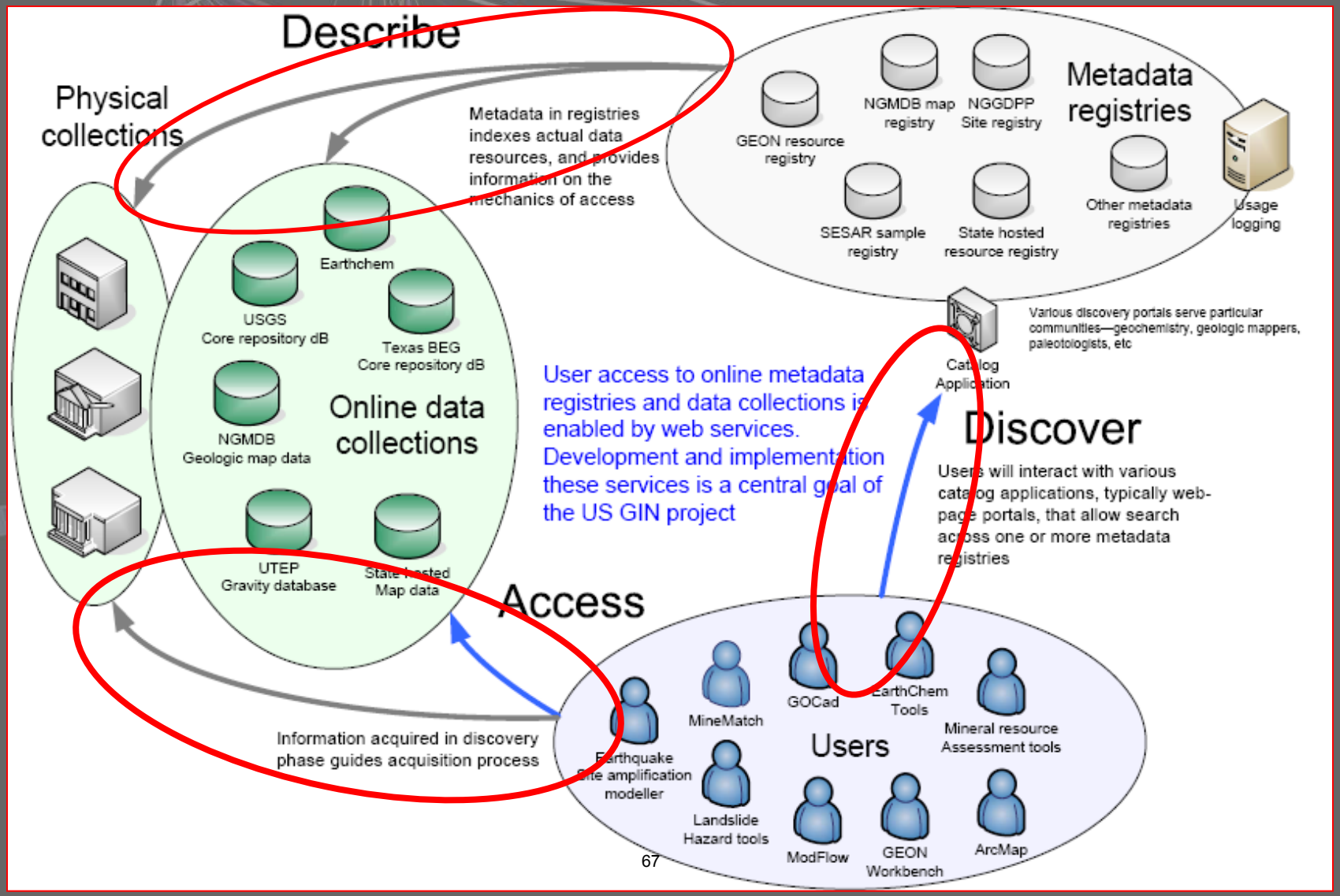
NE Geology

NE Geology Europe

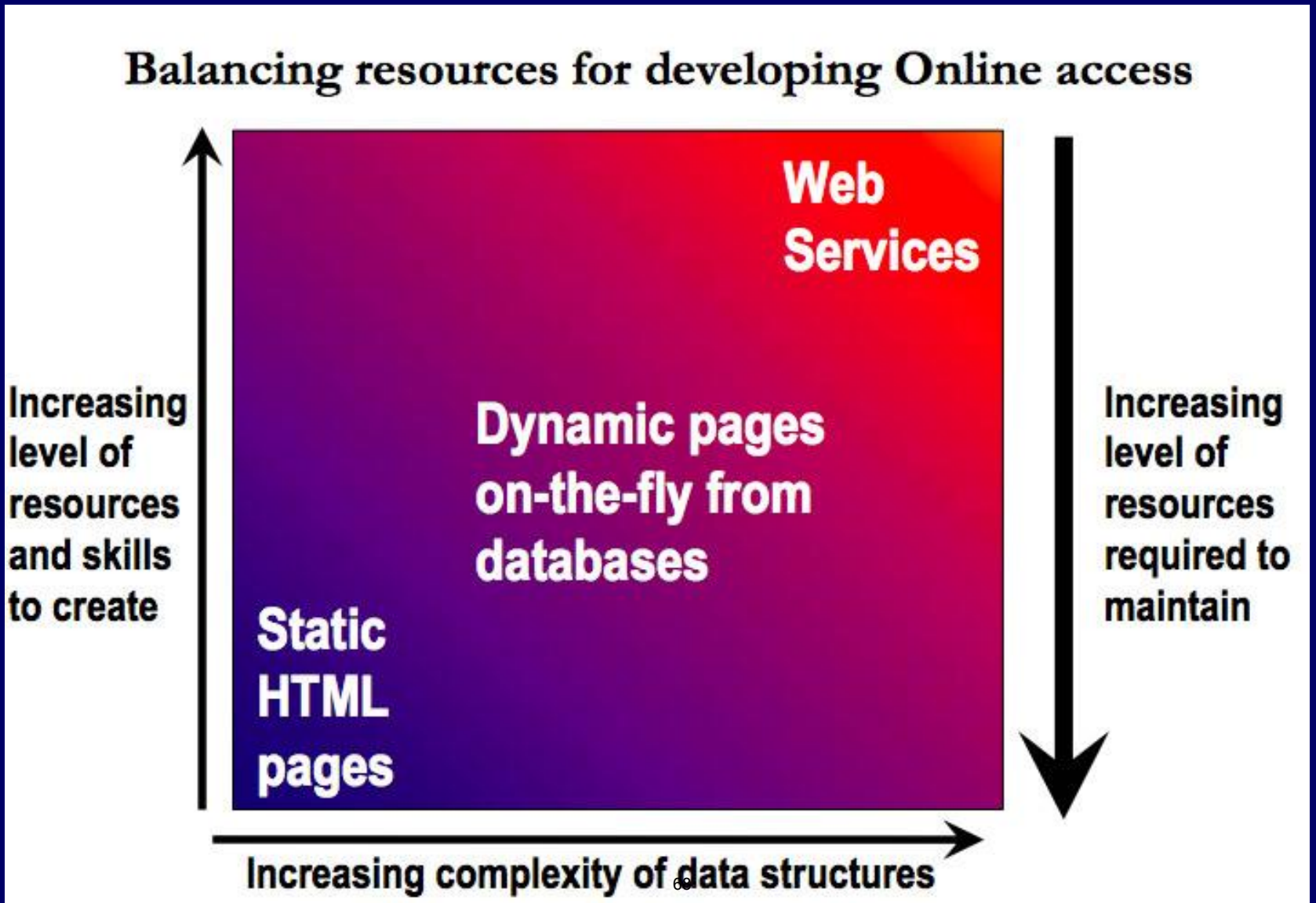
energistics The Energy Standards Resource Centre

ARIZONA GEOLOGICAL SURVEY
AZGS
Live demo

System overview



Shift the Burden from the User to the Provider



Peter Fox, RPI, 2010

Core Science Systems

Core System Informatics

- data integration services
- capacity
- framework for science programs

National Geothermal Data System

NGDS
National Geothermal Data System

Distributed data sources
(*AASG, USGS, universities, labs*)

Network communications
& controls

Legacy data
repository

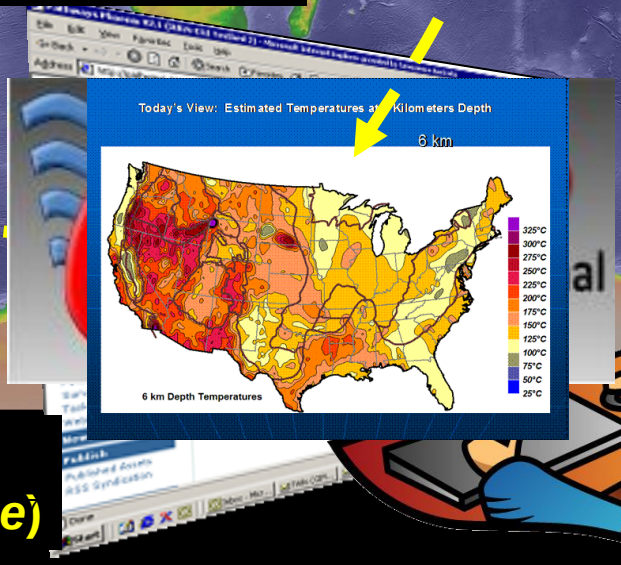
Desktop applications
(*Geothermal Desktop*)

Data models

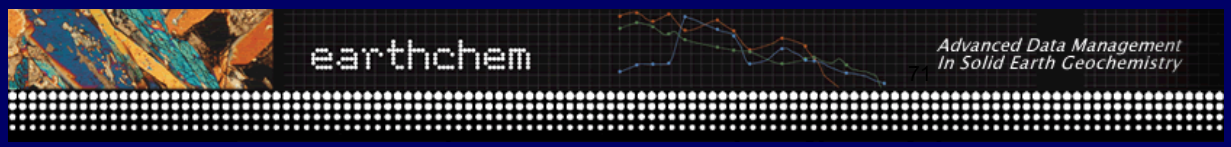
Ontologies, vocabularies

Discovery, access,
exchange (*GIN*)

Portals
(*GeoStratSys, SciScope*)



Partners & Collaborators:



29 countries and European organizations are committed to create a geological map at 1:1.000.000 scale, integrated with metadata initially available in the following languages: English, French, Italian, Spanish, Swedish, Czech and Norwegian.

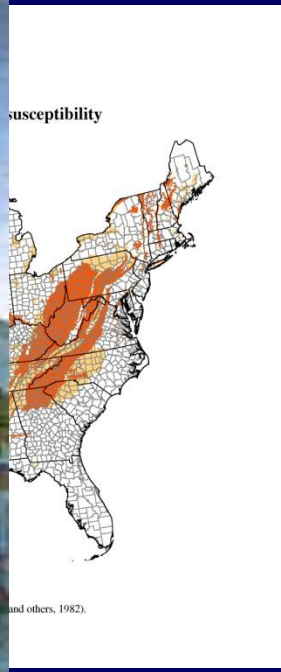
Co-funded by the European Union

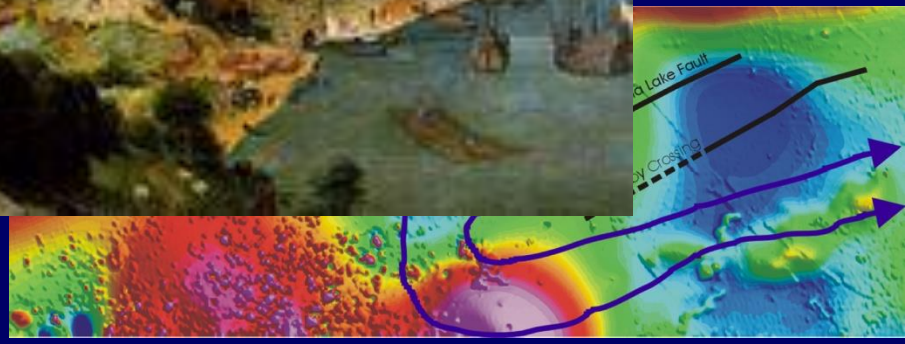
Everything Digital, Online, & Interoperable

Lee Allison, Arizona Geological Survey

1000's of National and Regional Databases

- topograph
  hydrograp
- mineral re
- water
- geochemi
- geophysic
  gravity, ae
- earthquak
- biological
- vegetation/speciation maps





USGS science for a changing world

AASG Association of American State Geologists

Example: the electrical utility

Simple interface– put plug in wall, get electricity

Afghanistan 220 V 50 Hz
Andorra 230 V 50 Hz
Anguilla 110 V 60 Hz
Antigua 230 V* 60 Hz
Cayman Islands 120 V 60 Hz
Cyprus 240 V 50 Hz
Czech Republic 230 V 50 Hz
.....

Online, Interactive Assessment of Geothermal Energy Potential in the U.S

Abstract

Geothermal-relevant geosciences data from all 50 states (www.stategeothermaldata.org), federal agencies, national labs, and academic centers are being digitized and linked in a distributed network via the U.S. Department of Energy-funded National Geothermal Data System (NGDS) to foster geothermal energy exploration and development through use of interactive online 'mashups,' data integration, and applications. Emphasis is first to make as much information as possible accessible, with a long range goal to make data interoperable through standardized services and interchange formats. Resources may be made available as documents (files) in whatever format they are currently in, converted to tabular files using standard content models, or published as Open Geospatial Consortium or ESRI Web services using the standard xml schema. An initial set of thirty geoscience data content models are in use or under development to define standardized interchange format: aqueous chemistry, borehole temperature data, direct use feature, drill stem test, earthquake hypocenter, fault feature, geologic contact feature, geologic unit feature, thermal/hot spring description, metadata, quaternary fault, volcanic vent description, well header feature, borehole lithology log, crustal stress, gravity, heat flow/temperature gradient, permeability, and feature description data like developed geothermal systems, geologic unit geothermal properties, permeability, production data, rock alteration description, rock chemistry, and thermal conductivity. Map services are also being developed for isopach maps (depth to bedrock), aquifer temperature maps, and several states are working on geothermal resource overview maps. Content models are developed preferentially from existing community use in order to encourage widespread adoption and promulgate minimum metadata quality standards. Geoscience data and maps from NGDS participating institutions (USGS, Southern Methodist University, Boise State University Geothermal Data Coalition) are being supplemented with extensive land management and land use resources from the Western Regional Partnership (15 federal agencies and 5 Western states) to provide access to a comprehensive, holistic set of data critical to geothermal energy development. As of August 2011, over 33,000 data resources have been registered in the system catalog, along with scores of Web services to deliver integrated data to the desktop for free downloading or online use. The data exchange mechanism is built on the U.S. Geoscience Information Network (USGIN, http://lab.usgin.org) protocols and standards developed in partnership with the U.S. Geological Survey.

Cite as: Allison, M L; Richard, S M; Clark, R; Coleman, C; Love, D; Pape, E; and Musil, L (2011), Online, Interactive Assessment of Geothermal Energy Potential in the U.S, Abstract IN31D-04 presented at 2011 Fall Meeting, AGU, San Francisco, Calif., 5-9 Dec.

U.S. Geoscience Information Network: A Critical Path for Data Integration in the U.S. Earth Sciences

Kevin T. Gallagher, U.S. Geological Survey, Reston VA and M. Lee Allison, Arizona Geological Survey, Tucson, AZ

U.S. Geoscience Information Network: A Critical Path for Data Integration in the U.S. Earth Sciences

Abstract:

Development efforts for the U.S. Geoscience Information Network (USGIN) over the last two years have crystallized around the Community for Data Integration (CDI) at the USGS, and the 50-state AASG State Geothermal Data project. The next step in developing a USGS-AASG community is to bring these two efforts into closer alignment through greater participation in CDI activities by geoinformatics practitioners from state geological surveys, and implementation of test bed activities by the USGIN partners, within two years.

Test bed activities in the geological survey community will define a scope and provide a foundation to promote the use of specifications developed by the larger geoinformatics community. Adoption of some of these specifications as 'standards' by USGS and AASG for use by those organizations will lend authority and motivate wider adoption.
The arc from use case to test bed to production deployments to agreement on 'standard' specifications for data discovery and access must be propelled by active interest from the user communities who have a stake in the outcome. The specifications developed will benefit the organizations involved in development, testing and deployment, which motivates participation -- a model that has worked successfully for standards organizations such as OGC, ISO and OASIS.

The governance structure to support such a community process should promote grass root nucleation of interest groups that are the core of development efforts. Some mechanism for community agreement on priorities is desirable because geological survey agencies will need to allocate resources to support development. Loosely knit organizations such as ESIP and the current CDI provide models for this kind of structure. Because many geological surveys have data archive and dissemination functions as part of their portfolio, some support for the system can be built into the operating expenses and overhead. Sharing of resources and reuse of components can reduce the cost. Wide adoption of similar software, protocols and practices increases the number of stake holders with an interest in supporting the system.

USGIN – the roadmap

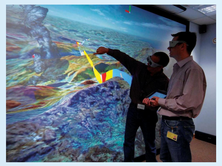
Evolution

- Conceived in 2007
- Meeting of representatives from
  - Association of American State Geologists (AASG)
  - US Geological Survey (USGS)
- Outcome
  - Develop a distributed national infrastructure of digital Earth Science data
  - Use common standards and protocols
  - Preserve data ownership, credit, and control
  - Build on existing systems

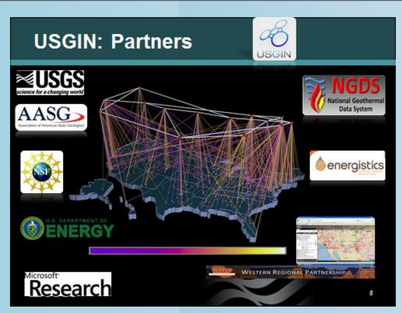
Purpose

- Use standardized services to make State and Federal geological survey data accessible online
- Employ a distributed network
- Use common standards and protocols
- Work with data providers to implement services

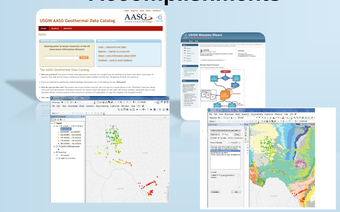
The Challenge of Data



Converting this…… to this



Accomplishments



USGIN Catalog –
- Collection of consistent metadata describing network resources

USGIN Repository –
- Generates and manages metadata for files and Web tools
- Upload capability

Interoperable, Simple Content Models -
Widely utilized client platforms are available for service utilization (ArcGIS, Microsoft Oce, OpenLayers and other free, open-source software)

Web Services Deployment -
OGC Web Feature Services deployed by the Arizona, Illinois, Kentucky, Nevada geological surveys as server hubs for the NGDS.

Partnership Test Bed Activities

Test Bed Proposal 1: OneGeology – USGS

- USGS and AASG collaborate to implement Web Map and Web Feature Services according to the OneGeology profile
  - –USGS: responsible for national maps (surficial deposits, glacial deposits, bedrock geology of US, etc.)
  - –State Geological Surveys: deploy services for state-scale geologic maps
  - –Community: agrees on an integrated portrayal scheme to attain interoperability between map services

Test Bed Proposal 2: Observation Services

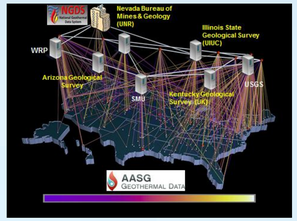
- States and USGS deploy observation services for geochronologic data, geochemical data, gravity stations, or similar commonly available site or sample-based data
- Community agrees on
  - Service profiles and interchange formats
  - Procedures to avoid data duplication
- This Test Bed project could build on and enhance the National Geothermal Data System and EarthChem

Test Bed Proposal 3: Integrated Catalog Capability

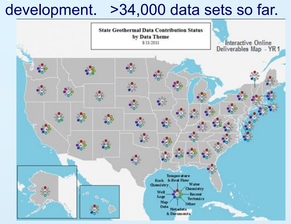
- Achieve interoperable metadata
- Demonstrate ability to harvest between catalog nodes hosted by state surveys and USGS
- Execute search against multiple nodes
- Edit tools to include user authentication and access control

Parallel tracks to building a national federated data integration framework

Design, build, deploy, and populate contributions to the National Geothermal Data System using USGIN framework



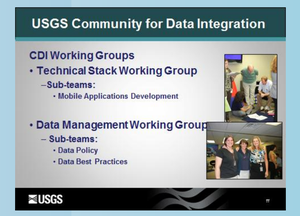
Digital data from 50 states relevant to geothermal energy exploration and development. >34,000 data sets so far.



USGS Community for Data Integration

CDI is a community of practice established in 2009
- –Facilitate discovery of data and tools
- –Improve use of scientific computing capacity
- –Develop, implement scientific data products and services
- –Enable data integration

CDI members
- –USGS data providers/practitioners/scientists /consumers
- –External partners in government, academia, industry



USGS
USGS Data Management
Releases and Documentation
Publications
Publication Policy
Publication Templates
Publication Process
Publication Publishing Info
Publication Publishing Info
Publication Publishing Info

Online, Interactive Assessment of Geothermal Energy Potential in the U.S.A.*

Lee Allison[1], Stephen M. Richard[1], Ryan J. Clark[1], Kim Patten[1], Diane Love[1], Celia Coleman[1], Genhan Chen[1], Jordan Matti[1], Janel Day[1], Esty Pape[1], and Leah Musil[1]

Search and Discovery Article #80226 (2012)**
Posted June 18, 2012

*Adapted from oral presentation at AAPG Annual Convention and Exhibition, Long Beach, California, April 22-25, 2012. Please refer to companion article, "National Geothermal Data System," by Arlene F. Anderson et al., Search and Discovery Article #80225 (2012).
**AAPG©2012 Serial rights given by author. For all other rights contact author directly.

[1]Arizona Geological Survey, Tucson, AZ (lee.allison@azgs.az.gov)

Abstract

Geothermal-relevant geosciences data from all 50 states (www.stategeothermaldata.org), federal agencies, national labs, and academic centers are being digitized and linked in a distributed network via the U.S. Department of Energy-funded National Geothermal Data System (NGDS) to foster geothermal energy exploration and development through use of interactive online 'mashups,' data integration, and applications. Emphasis is first to make as much information as possible accessible, with a long range goal to make data interoperable through standardized services and interchange formats.

An initial set of thirty geoscience data content models are in use or under development to define standardized interchange format: aqueous chemistry, borehole temperature data, direct use feature, drill stem test, earthquake hypocenter, fault feature, geologic contact feature, geologic unit feature, thermal/hot spring description, metadata, quaternary fault, volcanic vent description, well header feature, borehole lithology log, crustal stress, gravity, heat flow/temperature gradient, permeability, and feature description data like developed geothermal systems, geologic unit geothermal properties, permeability, production data, rock alteration description, rock chemistry, and thermal conductivity.

Map services are also being developed for isopach maps, aquifer temperature maps, and several states are working on geothermal resource overview maps. Content models are developed preferentially from existing community use in order to encourage widespread adoption and promulgate minimum metadata quality standards. Geoscience data and maps from NGDS participating institutions (USGS, Southern Methodist University, Boise State University Geothermal Data Coalition) are being supplemented with extensive land management and land use resources from the Western Regional Partnership (15 federal agencies and 5 Western states) to provide access to a comprehensive, holistic set of data critical to geothermal energy

development. As of September 2011, we have over 34,000 records registered in the system catalog, and 234,942 data resources online, along with scores of Web services to deliver integrated data to the desktop for free downloading or online use. The data exchange mechanism is built on the U.S. Geoscience Information Network (USGIN, http://lab.usgin.org) protocols and standards developed as a partnership of the Association of American State Geologists and U.S. Geological Survey.

Online, interactive assessment of geothermal energy potential in the U.S.

www.stategeothermaldata.org

Lee Allison, Stephen M. Richard, Ryan J. Clark, Kim Patten, Diane Love, Celia Coleman, Genhan Chen, Jordan Matti, Janel Day, Esty Pape, and Leah Musil

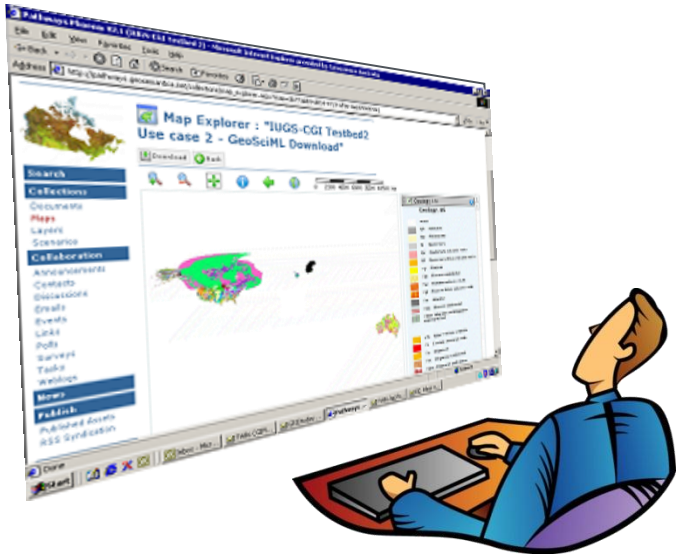
Arizona Geological Survey

AASG GEOTHERMAL DATA

NGDS: The Challenge of Data

1,000s of databases
1,000s of collections
80,000+ geologic maps

Selected data resources in State Geological Surveys
3.5 million wells – oil & gas, water
195,000 well logs
50,000 geothermal wells
750,000 BHT's
6 Tb existing digital data
>75,000 scanned publications & maps
2.5 million feet of core
600,000 sample logs



STATE GEOTHERMAL DATA

USGIN: data integration framework

- **Distributed**
- **Web-based**
- **Interoperable**
- **Open source**
- **Community of Practice**

Separation of roles:

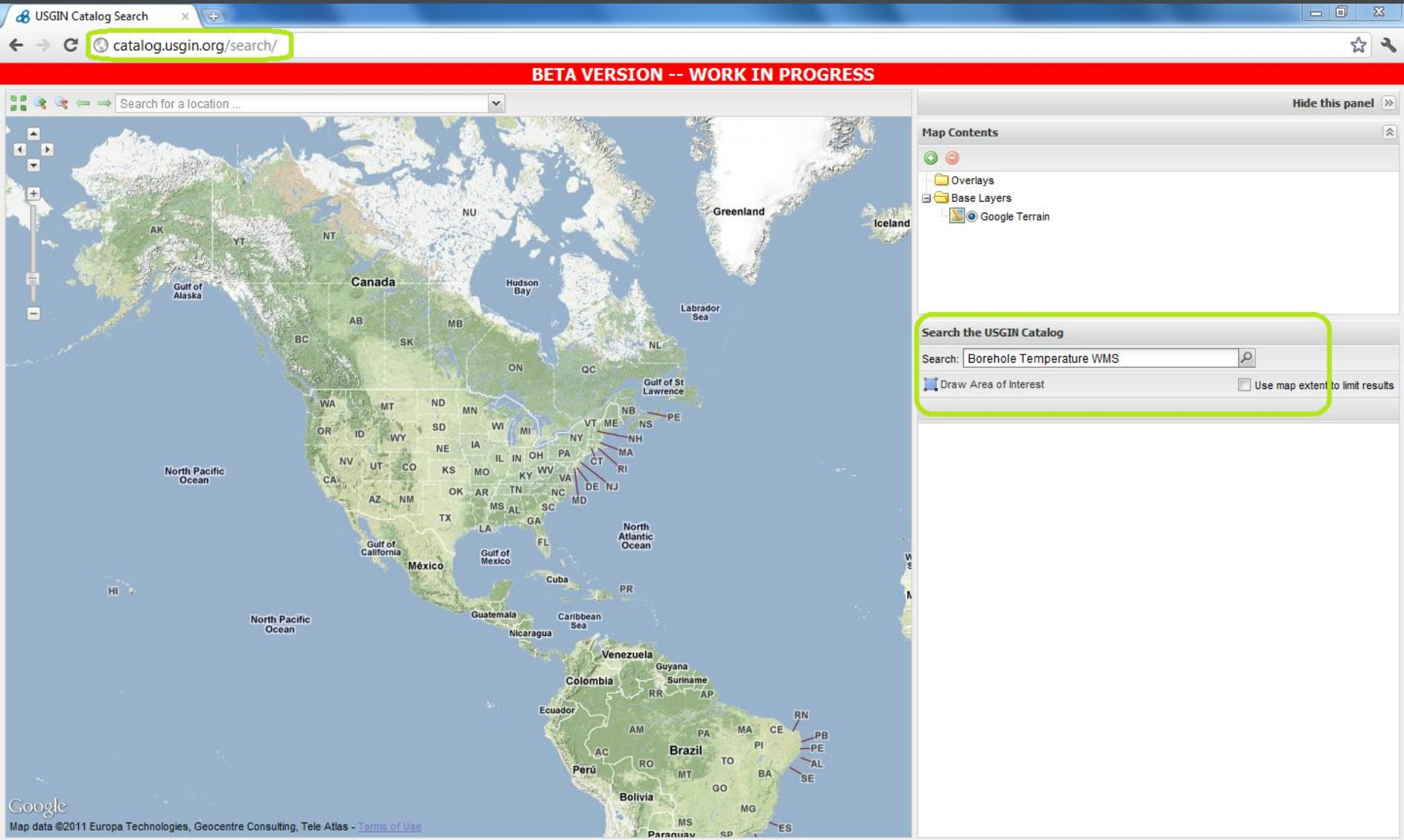
- Data Producer
  - Manage data content
- Data Provider
  - Server performance
  - Archive and backup
  - Service implementation, schema mapping
- Data Consumer
  - Find resources
  - Extract useful information
  - User interface

USGIN

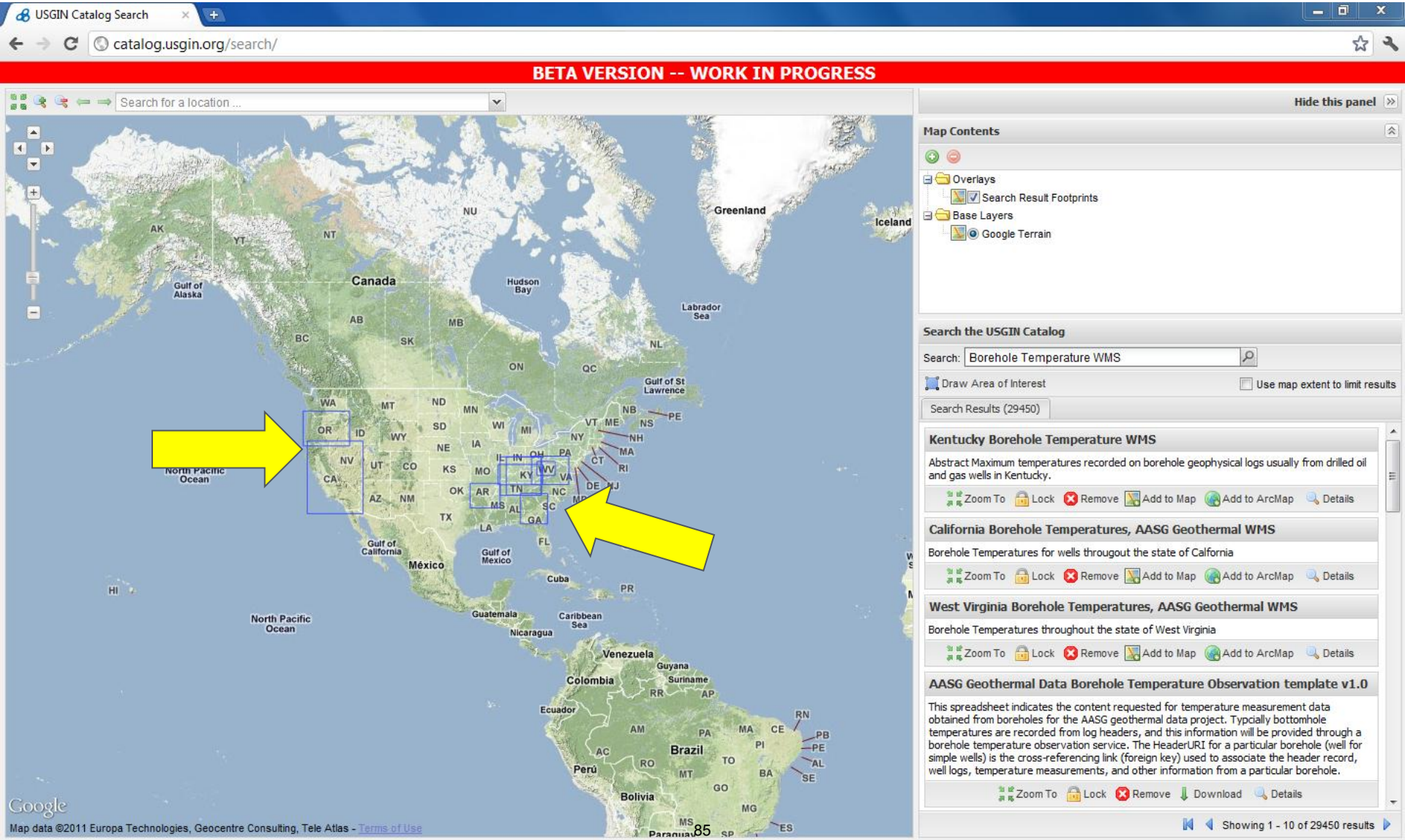
Geothermal Data System

- **Discover** via the Catalog
- **Access** via the Web
- **Explore** using Applications

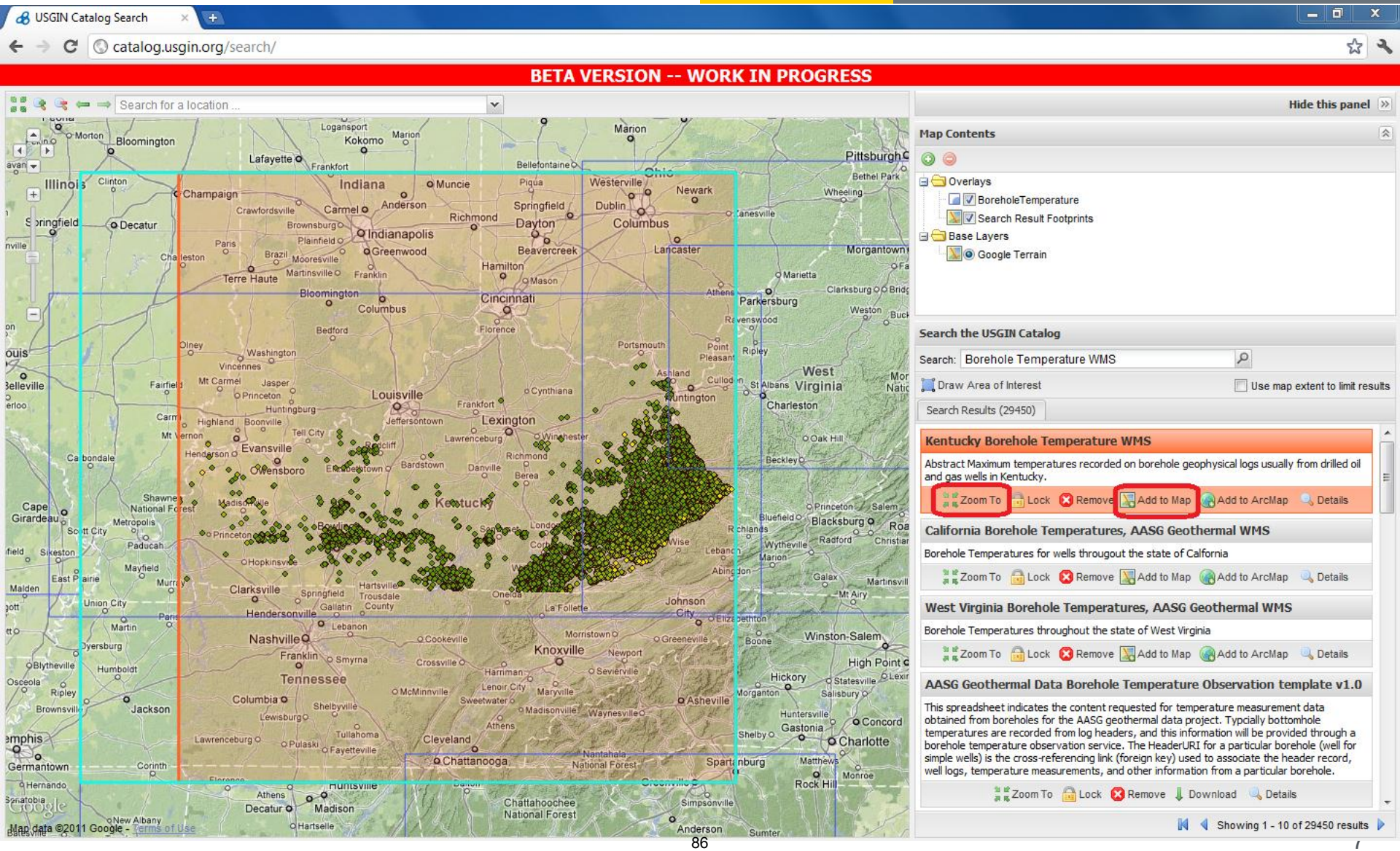
Catalog search – Borehole Temperatures



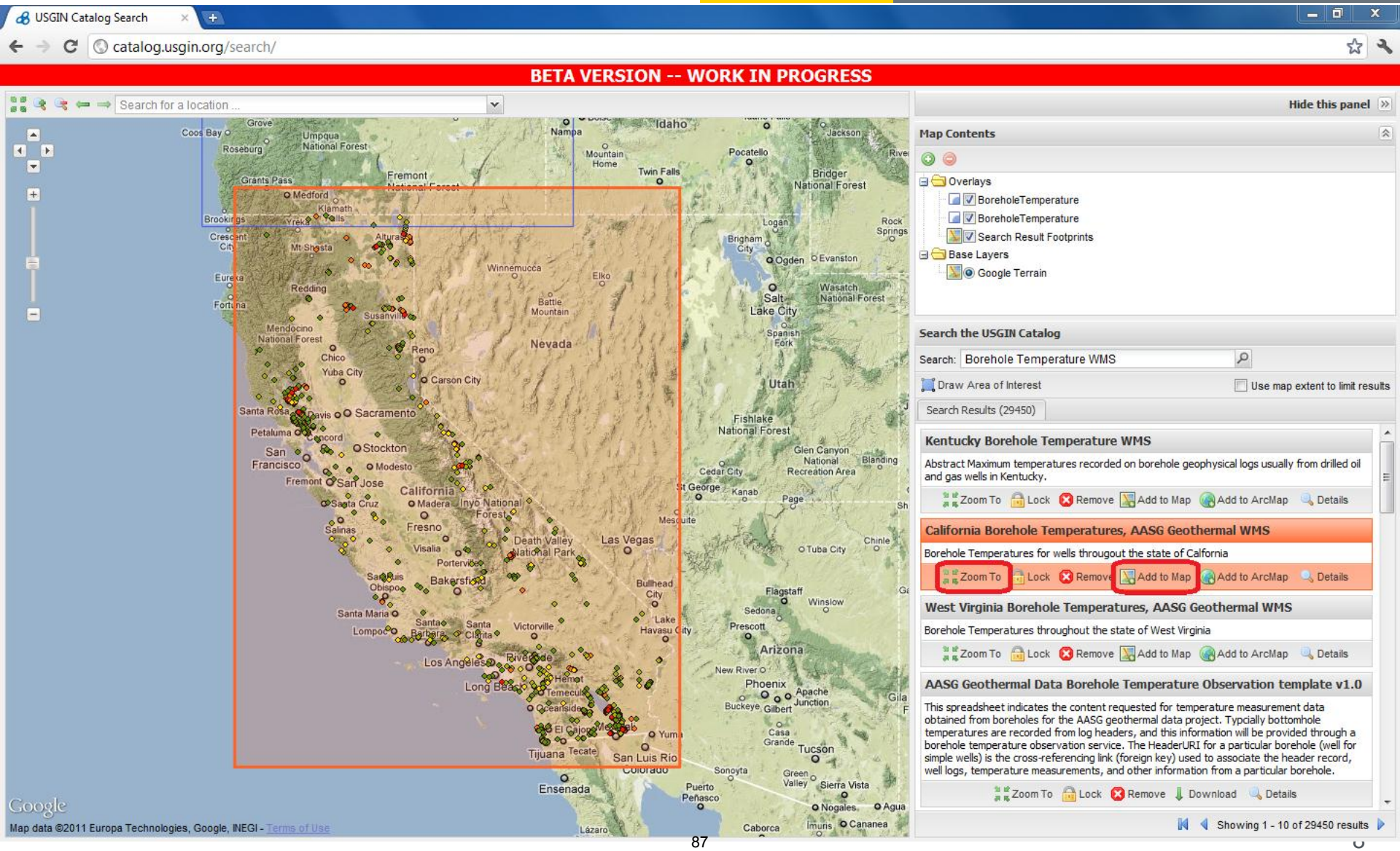
Borehole Temperature search results



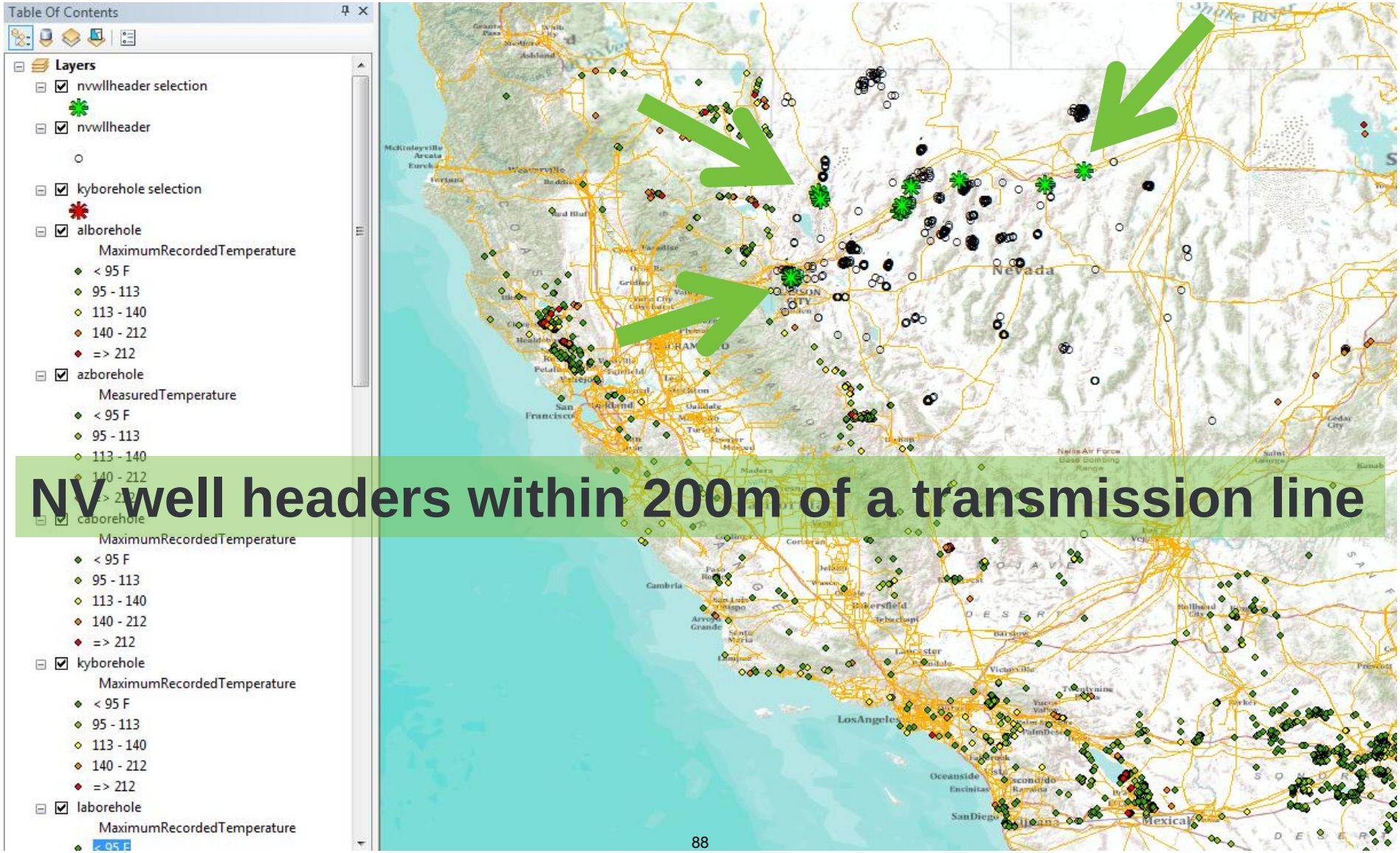
Zoom to and add to map – KY BHTs



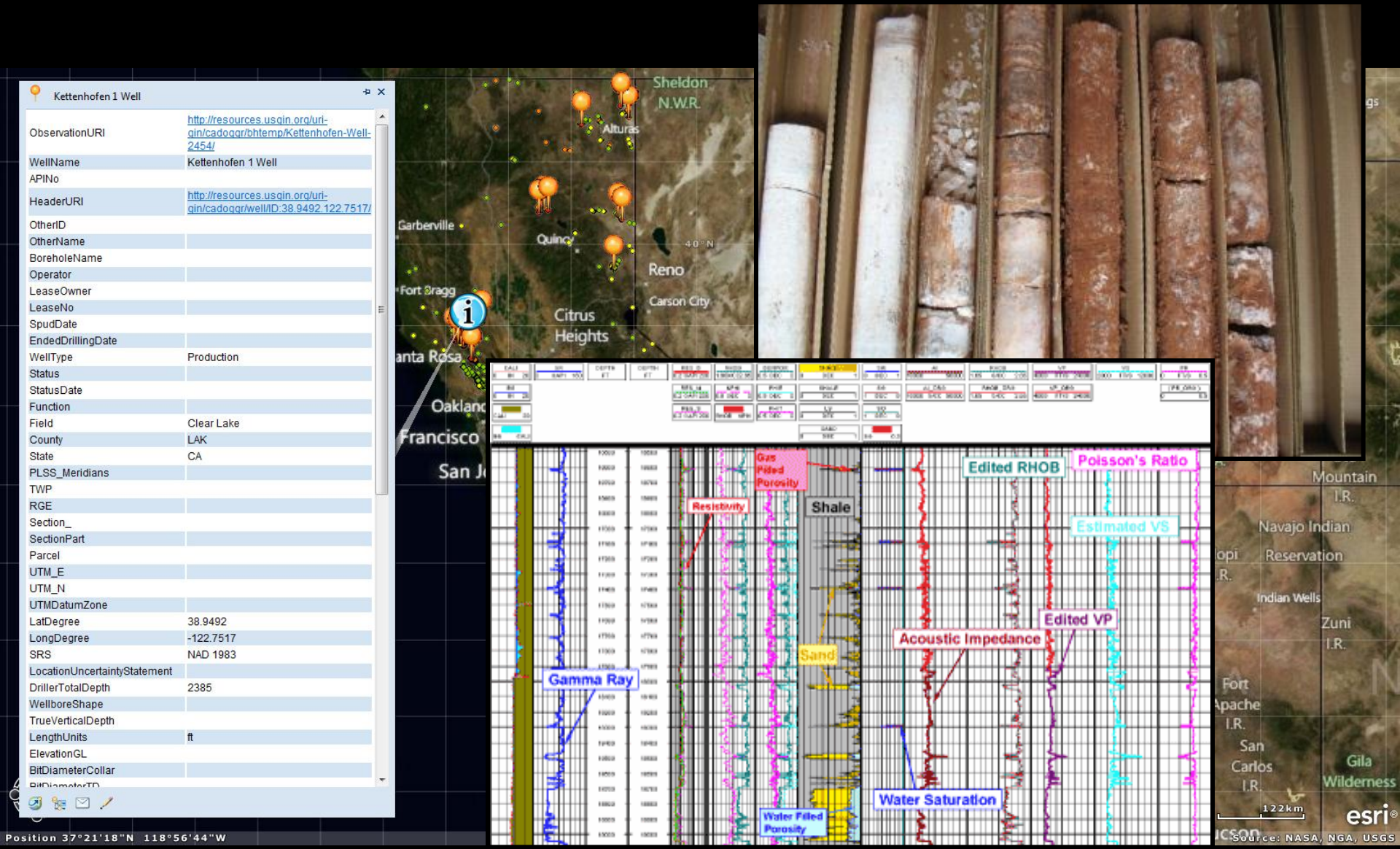
Zoom to and add to map – CA BHTs



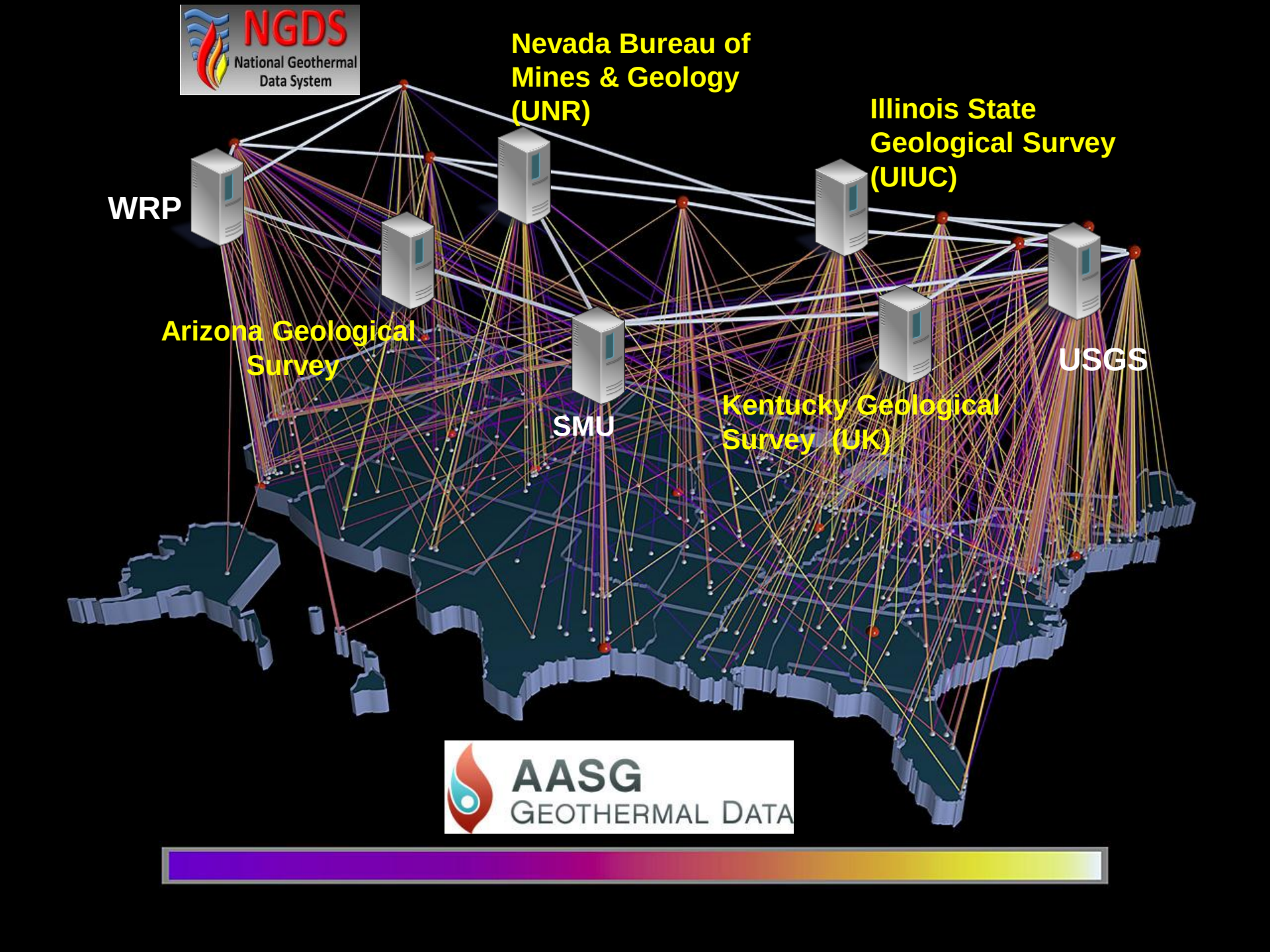
Exploring the Data in our Application – Well Headers in Nevada using ArcGIS



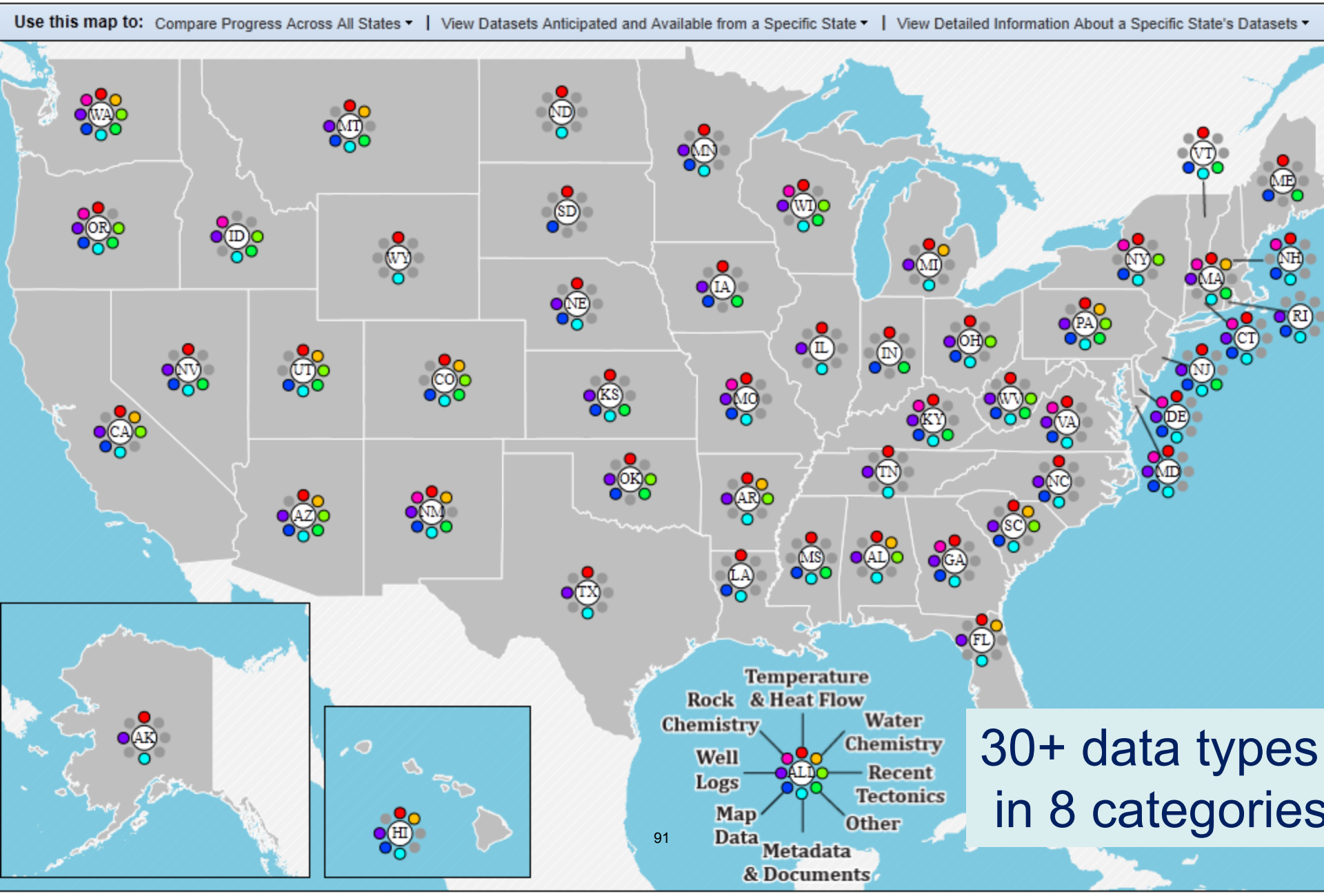
NV well headers within 200m of a transmission line



Individual features in live datasets may be queried for attributes

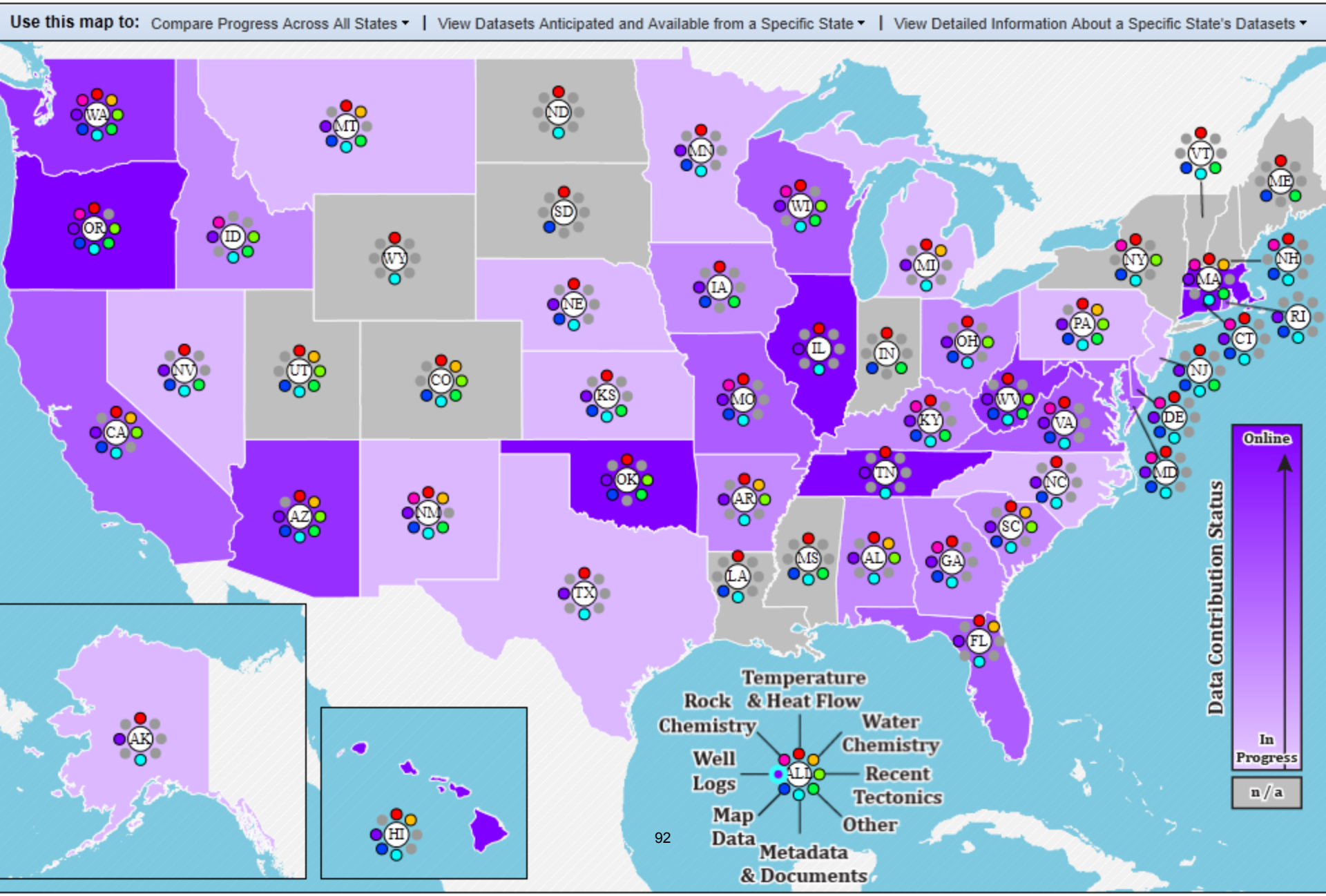

NGDS
National Geothermal Data System
Nevada Bureau of Mines & Geology (UNR)
Illinois State Geological Survey (UIUC)
WRP
Arizona Geological Survey
SMU
Kentucky Geological Survey (UK)
USGS
AASG
GEOTHERMAL DATA

Data sets from states

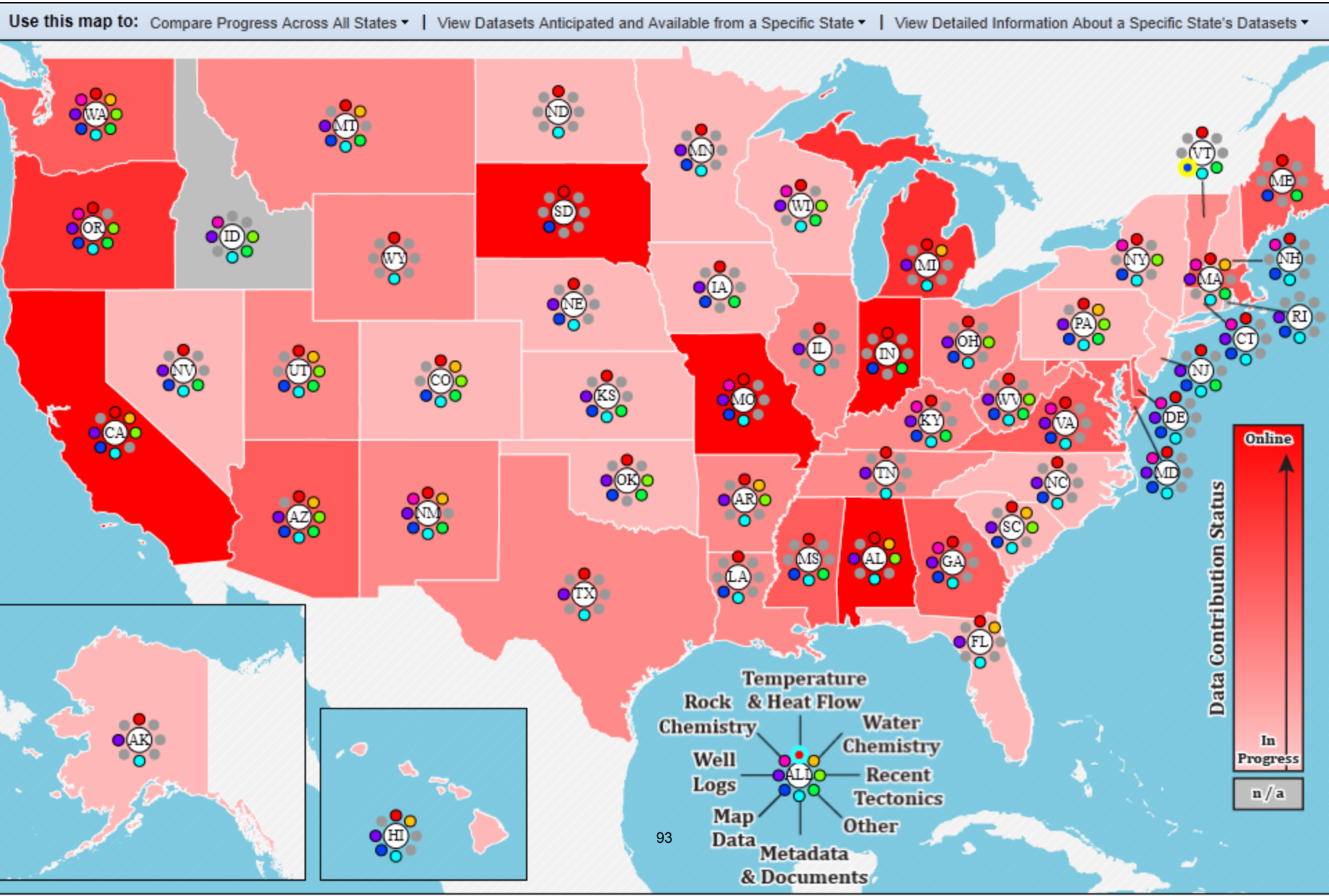


30+ data types in 8 categories

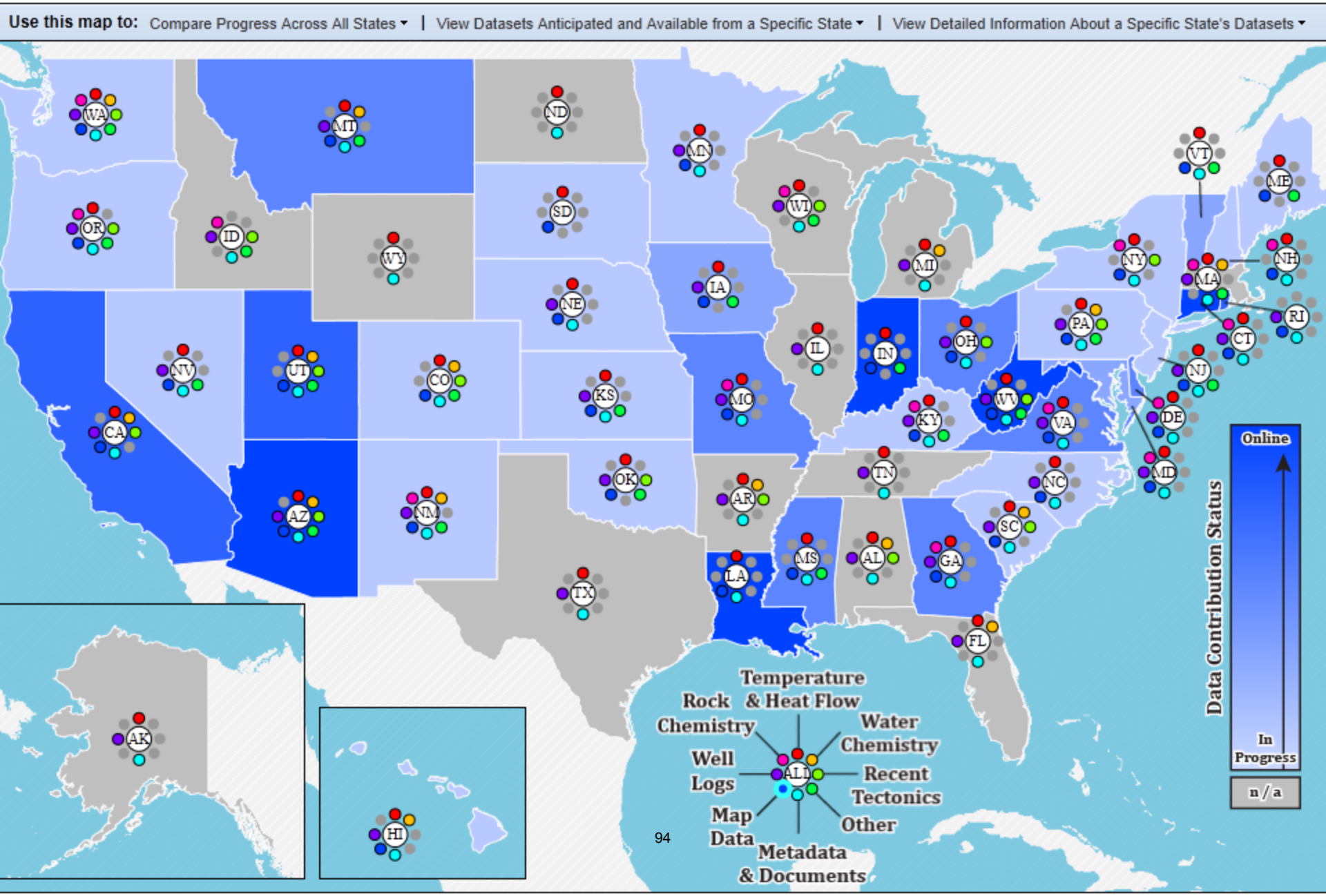
Well Logs


Use this map to: Compare Progress Across All States | View Datasets Anticipated and Available from a Specific State | View Detailed Information About a Specific State's Datasets
WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MS
MI
IN
KY
TN
AL
OH
WV
GA
FL
SC
NC
VA
PA
NY
VT
ME
NH
MA
RI
CT
NJ
DE
MD
AK
HI
Temperature & Heat Flow
Water Chemistry
Recent Tectonics
Other
Metadata & Documents
Map Data
Well Logs
Rock Chemistry
ALL
Data Contribution Status
Online
In Progress
n/a

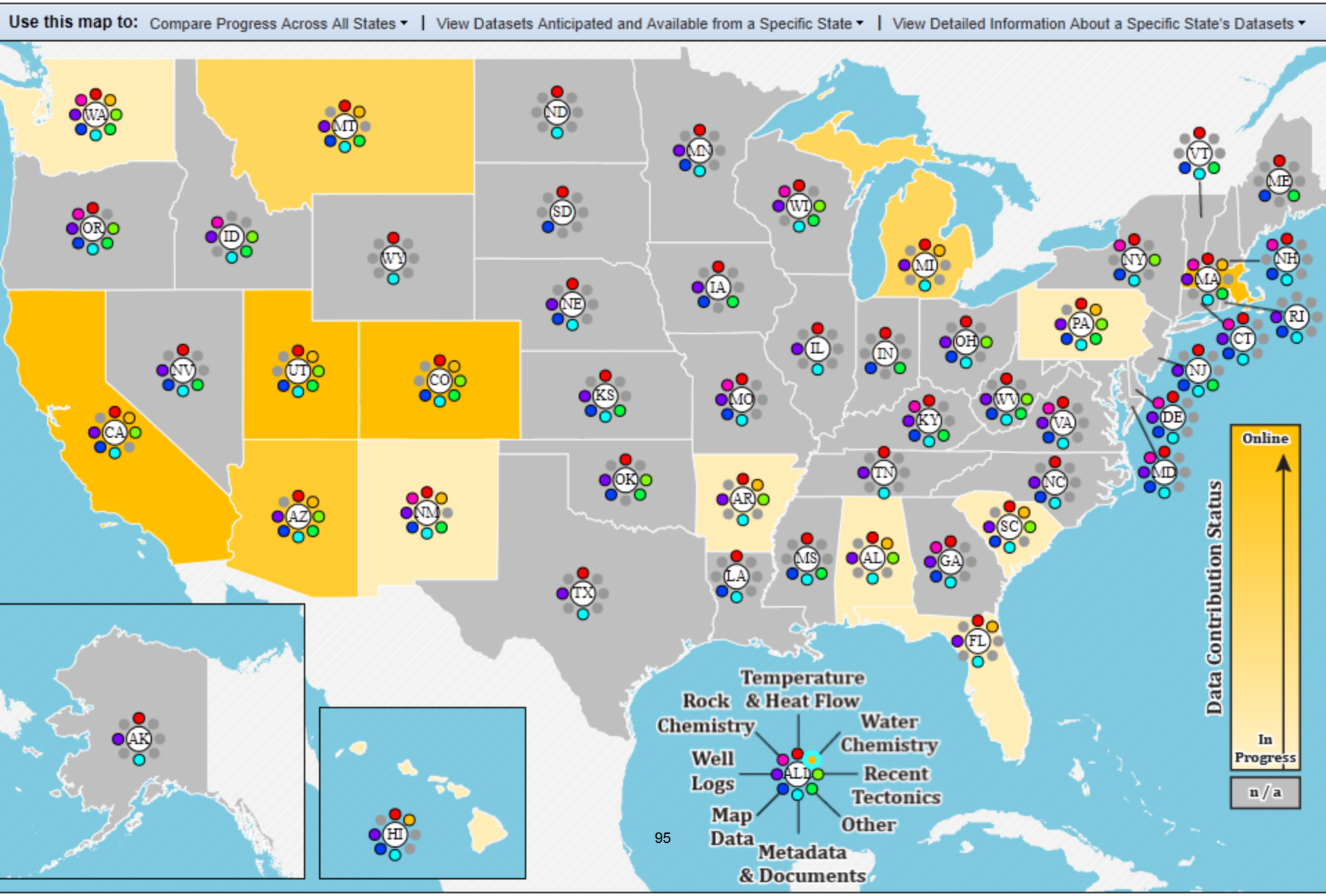
Temperature & Heat Flow



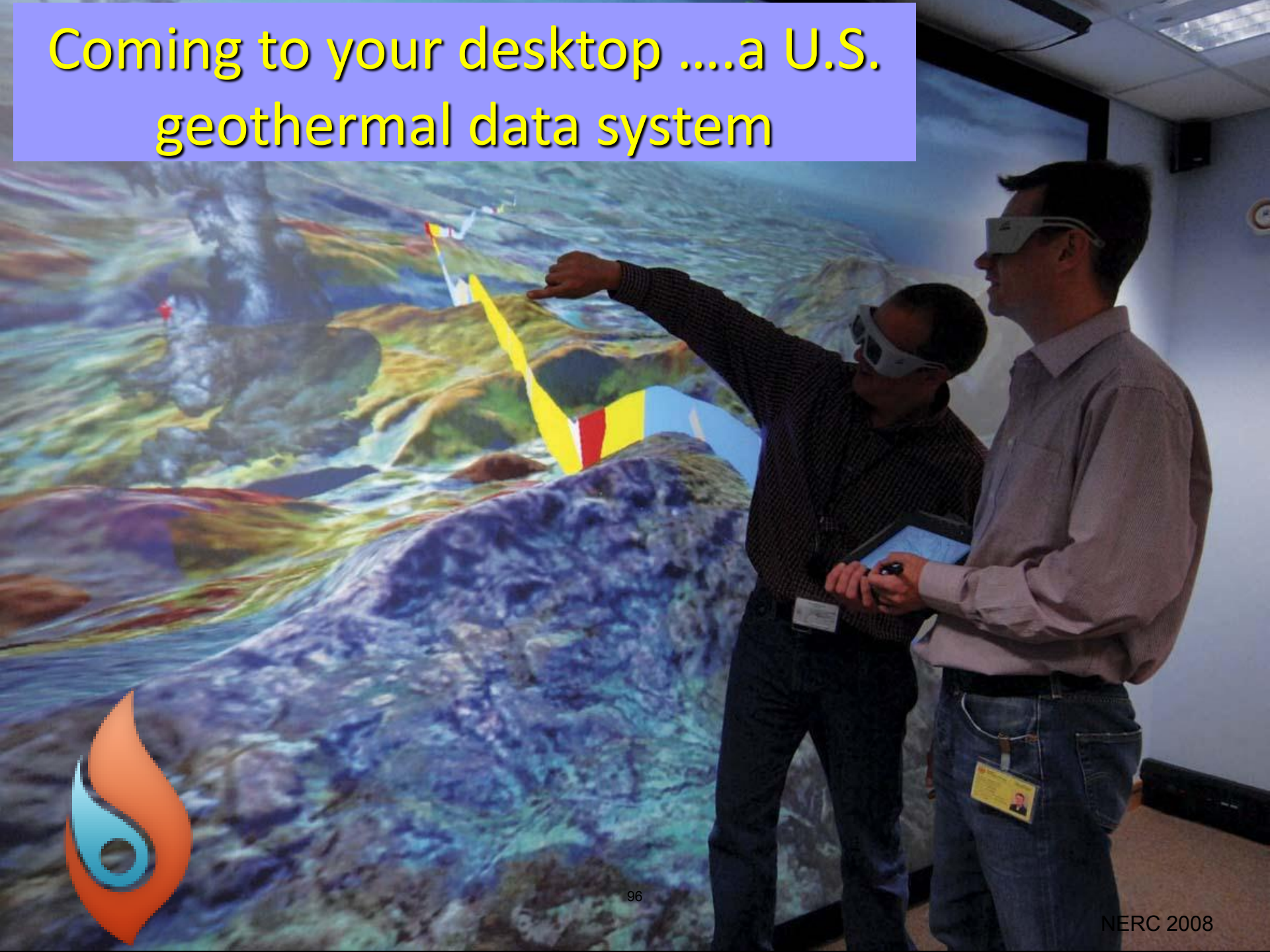
Map Data


Use this map to: Compare Progress Across All States | View Datasets Anticipated and Available from a Specific State | View Detailed Information About a Specific State's Datasets
WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MS
MI
IN
KY
TN
AL
OH
WV
GA
FL
SC
NC
VA
PA
NY
VT
ME
NH
MA
RI
CT
NJ
DE
MD
AK
HI
Temperature & Heat Flow
Water Chemistry
Recent Tectonics
Other
Metadata & Documents
Map Data
Well Logs
Rock Chemistry
ALL
Data Contribution Status
Online
In Progress
n / a

Water Chemistry



Coming to your desktop ….a U.S. geothermal data system



[PS]National Geothermal Data System*

Arlene F. Anderson[2], David Cuyler[2], Walter S. Snyder[3], M. L. Allison[1], David D. Blackwell[4], and Colin F. Williams[5]

Search and Discovery Article #80225 (2012)**
Posted June 18, 2012

*Adapted from extended abstract prepared in conjunction with poster presentation at AAPG Annual Convention and Exhibition, Long Beach, California, April 22-25, 2012. Please refer to companion article, "Online, Interactive Assessment of Geothermal Energy Potential in the U.S.A.," by Lee Allison et al., Search and Discovery Article #80226 (2012).
**AAPG©2012 Serial rights given by author. For all other rights contact author directly.

[1]Arizona Geological Survey, Tucson, AZ (lee.allison@azgs.az.gov)
[2]Office of Energy Efficiency & Renewable Energy, Geothermal Technologies Program, US Department of Energy, Washington, DC
[3]Department of Geosciences, Boise State University, Boise, ID
[4]Huffington Department of Earth Sciences, Southern Methodist University, Dallas, TX
[5]U.S. Geological Survey, Menlo Park, CA

Abstract

The goal of the U.S. Department of Energy's National Geothermal Data System is to design, build, implement, deploy and populate a national, sustainable, distributed, interoperable network of data and service (application) providers. These providers will develop, collect, serve, and maintain geothermal-relevant data that operates as an integral component of NGDS. As a result, the geothermal industry, the public, and policy makers will have access to consistent and reliable data, which in turn, reduces the amount of staff time devoted to finding, retrieving, integrating, and verifying information. With easier access to information, the high cost and risk of geothermal power projects (especially exploration drilling) is reduced.

Five separate NGDS projects provide the data support, acquisition, and access to cyber infrastructure necessary to reduce cost and risk of the nation's geothermal energy strategy and US DOE program goals focused on the production and utilization of geothermal energy. The U.S DOE Office of Energy Efficiency and Renewable Energy Geothermal Technologies Program is developing the knowledge and data foundation necessary for discovery and development of large-scale energy production while the Buildings Technology Program is focused on other practical applications such as direct use and residential/commercial ground source heat pumps. The NGDS provides expanded reference and resource data for research and development activities (a subset of the US DOE goals) and includes data from across all fifty states and the nation's leading academic geothermal centers. Thus, the project incorporates not only high-temperature potential but also moderate and low-temperature locations incorporating US DOE's goal of adding more geothermal electricity to the grid. The program, through its development of data

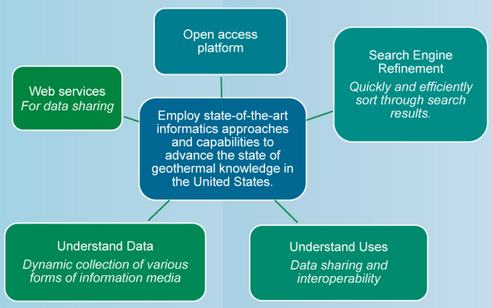
integration cyberinfrastructure, will help lead to innovative exploration technologies through increased data availability on geothermal energy capacity. Finally, the project will contribute new data from previously unexplored locations. NGDS is being built using the US Geoscience Information Network (US GIN) data integration framework to promote interoperability across the Earth sciences community and with other emerging data integration and networking efforts.

National Geothermal Data System

Arlene Anderson (US Department of Energy, Geothermal Technologies Program), David Cuyler (Sandia National Laboratory), Walter Snyder (Boise State University), M. Lee Allison (Arizona Geological Survey), David Blackwell (Southern Methodist University), & Colin F. Williams (USGS)

Abstract: The goal of the U.S. Department of Energy's National Geothermal Data System is to design, build, implement, deploy and populate a national, sustainable, distributed, interoperable network of data and service (application) providers. These providers will develop, collect, serve, and maintain geothermal-relevant data that operates as an integral component of NGDS. As a result the geothermal industry, the public, and policy makers will have access to consistent and reliable data, which in turn, reduces the amount of staff time devoted to finding, retrieving, integrating, and verifying information. With easier access to information, the high cost and risk of geothermal power projects (especially exploration drilling) is reduced.
Five separate NGDS projects provide the data support, acquisition, and access to cyber infrastructure necessary to reduce cost and risk of the nation's geothermal energy strategy and US DOE program goals focused on the production and utilization of geothermal energy. The U.S DOE Office of Energy Efficiency and Renewable Energy Geothermal Technologies Program is developing the knowledge and data foundation necessary for discovery and development of large-scale energy production while the Buildings Technology Program is focused on other practical applications such as direct use and residential/commercial ground source heat pumps.
The NGDS provides expanded reference and resource data for research and development activities (a subset of the US DOE goals) and includes data from across all fifty states and the nation's leading academic geothermal centers. Thus, the project incorporates not only high-temperature potential but also moderate and low-temperature locations incorporating US DOE's goal of adding more geothermal electricity to the grid. The program, through its development of data integration cyberinfrastructure, will help lead to innovative exploration technologies through increased data availability on geothermal energy capacity. Finally, the project will contribute new data from previously unexplored locations.
NGDS is being built using the US Geoscience Information Network (US GIN) data integration framework to promote interoperability across the Earth sciences community and with other emerging data integration and networking efforts.

NGDS Design Objectives



Multiple projects in collaboration

- DOE Geothermal Data Repository — *Boise State University*
- National Resource Assessment & Classification — *USGS*
- National Temperature and Heat Flow Data and Maps — *Southern Methodist University*
- State Geological Survey Contributions — *Association of American State Geologists*
- Energy and Geosciences Institute — Oregon Institute of Technology; Stanford Reservoir Engineering

NGDS is a nationwide network of data repositories hosting geothermal-relevant data. Each participating data "node" receives, manages and provides content independently. Data are made publicly available through a distributed framework and a unified catalog search interface provided by the NGDS.

U.S. DEPARTMENT OF ENERGY This material is based upon work supported by the **Department of Energy** and by **National Science Foundation** grant EAR-0753154

The "killer app" for data integration in the U.S.

goal

•To design, build, implement, deploy and populate a national, sustainable, distributed, interoperable network of data and service (application) providers.

strategy

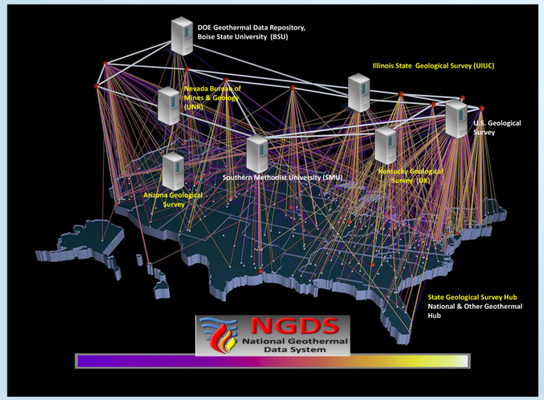
- Data providers will develop, collect, serve, and maintain geothermal-relevant data that operates as an integral component of NGDS.

success

•The geothermal industry, the public, and policy makers will have access to consistent and reliable data.

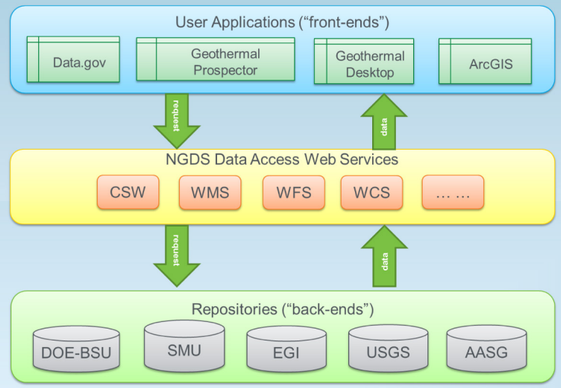
•The high cost and risk of geothermal exploration & development is reduced.

•Increase geothermal base load power production



Locations of "Charter" NGDS Hubs, where data integration, curation and development are underway. NGDS is built on the US Geoscience Information Network (USGIN) data integration framework

Basic NGDS Structure



SMU NGDS Publication Keywords

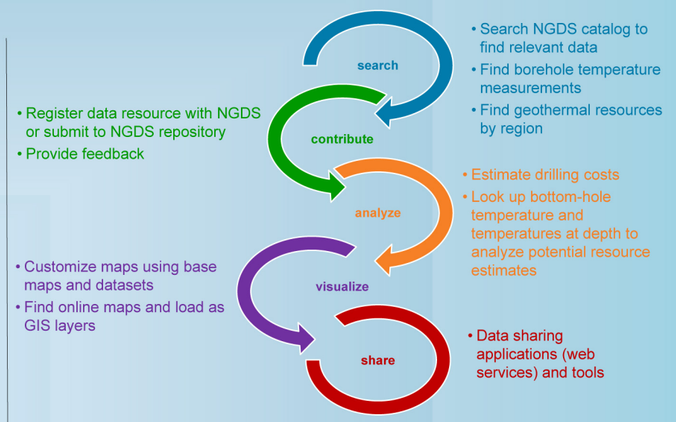
Keyword (primary)	Synonyms (or highly related)
well	wells; borehole; boreholes; downhole
geothermal exploration	energy source development; geophysical prospecting; geothermal exploration and development
seismic surveys	seismic survey; seismic refraction; seismic reflection; reflection seismic; reflection; seismic noise
resistivity surveys	electrical resistivity; resistivity; electric conductivity; electrical conductivity
magnetotelluric surveys	magnetotelluric; magnetotellurics; MT survey; electromagnetic surveys; telluric
heat flow	heat-flow; heat flows, heat flow density
well log	well logs; well logging; logs; lgoging; borehole logs; drillers' logs; open hole

Keywords will allow users to search the NGDS. Associating synonyms or highly-related words with "primary" keywords will ensure users are led to reliable and accurate search results

SMU NGDS Publication Keywords

Keywords will allow users to search the NGDS. Associating synonyms or highly-related words with "primary" keywords will ensure users are led to reliable and accurate search results.

What can you do with the NGDS?

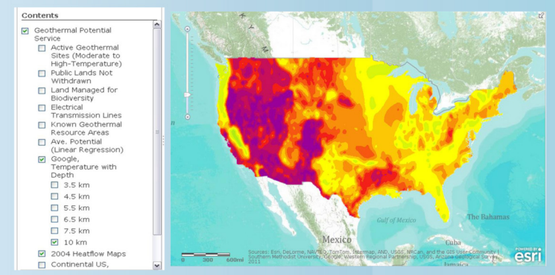


Analysis & Decision Making

Assessing and categorizing the nation's geothermal resources and consolidating all geothermal data through a publicly accessible data system will support research, stimulate public interest, promote market acceptance and investment, and in turn the growth of the geothermal industry.

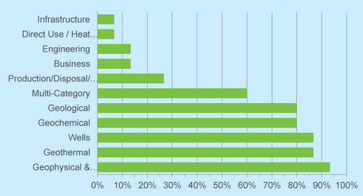
- How are geothermal resources defined?
- How are new classes of larger energy resources characterized?
- How is the management and expansion of known geothermal fields optimized?
- How can a path for technology growth into the future be ensured—providing the science and engineering basis for conventional and enhanced geothermal systems (EGS)?

Analytical Integration

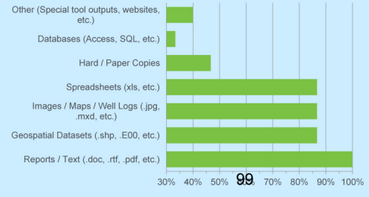


An example of NGDS Analytical Integration is shown here. Where an ESRI ArcGis application is used to PULL IN Google temperature and depth maps as well as SMU 2004 heatflow maps.

Intended Data Categories



Intended Data Formats



State	Funding Amount
Washington	$1,249,146
Nevada	$1,070,639
Utah	$966,834
Oregon	$958,847
Idaho	$873,562
Texas	$743,481
Colorado	$617,021
New Mexico	$605,483
Kentucky	$585,977
Massachusetts	$515,901
Illinois	$507,809
Hawaii	$499,951
Montana	$401,009
Indiana	$378,499
Wisconsin	$329,135

Top 15 Funded States

Drilling Project s	Funding Received
Idaho*	$457,662.80
Nevada*	$504,201.80
Oregon*	$526,803.80
Utah*	$516,294.80
Washington	$648,878.80

Supplemental funding for new data acquisition

Non Drilling Projects	Funding Received
Arizona	$179,976.00
Colorado	$174,763.00
Indiana	$69,975.00
Maine	$49,912.00
Massachusetts	$74,839.00
New Jersey	$49,989.00
New Mexico	$200,000.00
Oklahoma	$20,000.00
Pennsylvania	$83,425.00
Vermont	$78,870.00
West Virginia	$42,858.00

Geophysical Research Abstracts
Vol. 14, EGU2012-10066, 2012
EGU General Assembly 2012
© Author(s) 2012

Strategic Roadmap for the U.S. Geoscience Information Network

M. L. Allison (1), K. T. Gallagher (2), S.M. Richard (1), and V. B. Hutchison (2)

(1) Arizona Geological Survey, Tucson, United States (lee.allison@azgs.az.gov), (2) U.S. Geological Survey, Reston, United States (kgallagher@usgs.gov)

An external advisory working group has prepared a 5-year strategic roadmap for the U.S. Geoscience Information Network (USGIN). USGIN is a partnership of the Association of American State Geologists (AASG) and the U.S. Geological Survey (USGS), who formally agreed in 2007 to develop a national geoscience information framework that is distributed, interoperable, uses open source standards and common protocols, respects and acknowledges data ownership, fosters communities of practice to grow, and develops new Web services and clients.

The intention of the USGIN is to benefit the geological surveys by reducing the cost of online data publication and access provision, and to benefit society through easier (lower cost) access to public domain geoscience data. This information supports environmental planning, resource-development, hazard mitigation design, and decision-making. USGIN supposes that sharing resources for system development and maintenance, standardizing data discovery and creating better access mechanisms, causes cost of data access and maintenance to be reduced. Standardization in a wide variety of business domains provides economic benefits that range between 0.2 and 0.9% of the gross national product. We suggest that the economic benefits of standardization also apply in the informatics domain. Standardized access to rich data resources will create collaborative opportunities in science and business. Development and use of shared protocols and interchange formats for data publication will create a market for user applications, facilitating geoscience data discovery and utility for the benefit of society.

The USGIN Working Group envisions further development of tools and capabilities, in addition to extending the community of practice that currently involves geoinformatics practitioners from the USGS and AASG. Promoting engagement and participation of the state geological surveys, and increasing communication between the states, USGS, and other stakeholders are prerequisites for community development. A key element of community building is personal interaction. The USGIN community can establish an identity for geological survey informatics practitioners, can assist in prioritizing technical development that is specific to the geological survey community, and can leverage development taking place in the larger community. Policies, protocols, and procedures for developing, reviewing, and distributing specifications can be adopted from established practices developed by existing organizations, such as the OGC. Documenting and promoting best practices through demonstrations, education, and outreach within the geological survey community is paramount for fostering deployment of interoperable services for data discovery and distribution.

Evolution of the current Balkanized geoinformatics practice into a more cohesive and effective community has been and will continue to be an incremental process. The role of USGIN as an entity in this larger community requires organization, planning, promotion, and funding. As a member of a community activity, the role of USGIN as a leader in the community must be organic and emergent. Essential implementation activities include:
• Establish a long-term governance model
• Develop a business model
• Explore testbed opportunities
• Develop marketing strategy

Geophysical Research Abstracts
Vol. 14, EGU2012-13196, 2012
EGU General Assembly 2012
© Author(s) 2012

GET21: Geoinformatics Training and Education for the 21st Century Geoscience Workforce

C. Baru (1), L. Allison (2), P. Fox (3), C. Keane (4), R. Keller (5), and S. Richard (2)

(1) University of California, San Diego, San Diego Supercomputer Center, La Jolla, CA, United States (baru@sdsc.edu), (2) Arizona Geological Survey, Tucson, AZ, (3) Dept. of Computer Science, Renssalaer Polytechnic Institute, Troy, NY, (4) American Geological Institute, Alexandria, VA, (5) ConocoPhillips School of Geology and Geophysics, University of Oklahoma, Norman, OK

The integration of advanced information technologies (referred to as cyberinfrastructure) into scientific research and education creates a synergistic situation. On the one hand, science begins to move at the speed of information technology, with science applications having to move rapidly to keep apace with the latest innovations in hardware and software. On the other hand, information technology moves at the pace of science, requiring rapid prototyping and rapid development of software and systems to serve the immediate needs of the application. The 21st century geoscience workforce must be adept at both sides of this equation to be able to make the best use of the available cyber-tools for their science and education endeavors.

To reach different segments of the broad geosciences community, an education program in geoinformatics must be multi-faceted, ranging from areas dealing with modeling, computational science, and high performance computing, to those dealing with data collection, data science, and data-intensive computing. Based on our experience in geoinformatics and data science education, we propose a multi-pronged approach with a number of different components, including summer institutes typically aimed at graduate students, postdocs and researchers; graduate and undergraduate curriculum development in geoinformatics; development of online course materials to facilitate asynchronous learning, especially for geoscience professionals in the field; provision of internship at geoinformatics-related facilities for graduate students, so that they can observe and participate in geoinformatics "in action"; creation of online communities and networks to facilitate planned as well as serendipitous collaborations and for linking users with experts in the different areas of geoscience and geoinformatics.

We will describe some of our experiences and the lessons learned over the years from the Cyberinfrastructure Summer Institute for Geoscientists (CSIG), which is a 1-week institute that has been held each summer (August) at the San Diego Supercomputer Center, University of California, San Diego, since 2005. We will also discuss these opportunities for GET21 and geoinformatics education in the context of the newly launched EarthCube initiative at the US National Science Foundation.

U.S. Geothermal Data System

M. Lee Allison and Stephen M. Richard: Arizona Geological Survey, Tucson, Arizona, USA

STATE GEOTHERMAL DATA

The "killer app" for data integration in the U.S.

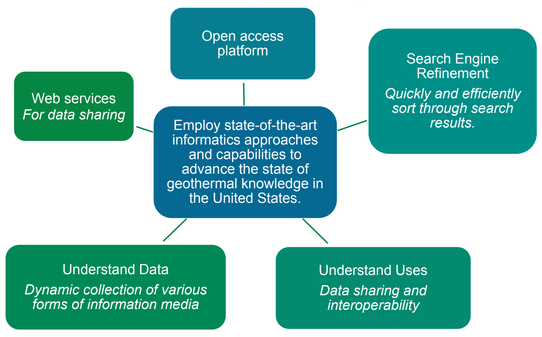
Abstract: The goal of the U.S. Department of Energy's National Geothermal Data System is to design, build, implement, deploy and populate a national, sustainable, distributed, interoperable network of data and service (application) providers. These providers will develop, collect, serve, and maintain geothermal-relevant data that operates as an integral component of NGDS. As a result the geothermal industry, the public, and policy makers will have access to consistent and reliable data, which in turn, reduces the amount of staff time devoted to finding, retrieving, integrating, and verifying information. With easier access to information, the high cost and risk of geothermal power projects (especially exploration drilling) is reduced.

Five separate NGDS projects provide the data support, acquisition, and access to cyber infrastructure necessary to reduce cost and risk of the nation's geothermal energy strategy and US DOE program goals focused on the production and utilization of geothermal energy. The U.S DOE Office of Energy Efficiency and Renewable Energy Geothermal Technologies Program is developing the knowledge and data foundation necessary for discovery and development of large-scale energy production while the Buildings Technology Program is focused on other practical applications such as direct use and residential/commercial ground source heat pumps.

The NGDS provides expanded reference and resource data for research and development activities (a subset of the US DOE goals) and includes data from across all fifty states and the nation's leading academic geothermal centers. Thus, the project incorporates not only high-temperature potential but also moderate and low-temperature locations incorporating US DOE's goal of adding more geothermal electricity to the grid. The program, through its development of data integration cyberinfrastructure, will help lead to innovative exploration technologies through increased data availability on geothermal energy capacity. Finally, the project will contribute new data from previously unexplored locations.

NGDS is being built using the US Geoscience Information Network (US GIN) data integration framework to promote interoperability across the Earth sciences community and with other emerging data integration and networking efforts.

NGDS Design Objectives



Multiple projects in collaboration

DOE Geothermal Data Repository — *Boise State University*

National Resource Assessment & Classification — *USGS*

National Temperature and Heat Flow Data and Maps — *Southern Methodist University*

State Geological Survey Contributions — *Association of American State Geologists*

Energy and Geosciences Institute; Oregon Institute of Technology; Stanford Reservoir Engineering

NGDS is a nationwide network of data repositories hosting geothermal-relevant data. Each participating data "node" receives, manages and provides content independently. Data are made publicly available through a distributed framework and a unified catalog search interface provided by the NGDS

goal

To design, build, implement, deploy and populate a national, sustainable, distributed, interoperable network of data and service (application) providers.

strategy

Data providers will develop, collect, serve, and maintain geothermal-relevant data that operates as an integral component of NGDS.

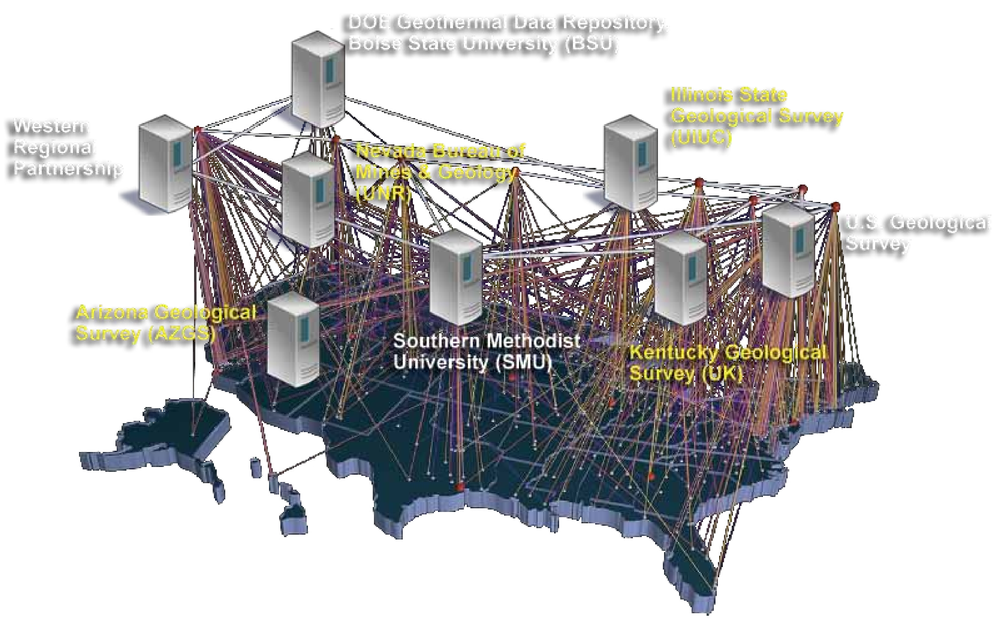
success

The geothermal industry, the public, and policy makers will have access to consistent and reliable data.

- The high cost and risk of geothermal exploration & development is reduced.

➡ **Increase geothermal base load power production!**

Distributed, federated, interoperable, Web-based

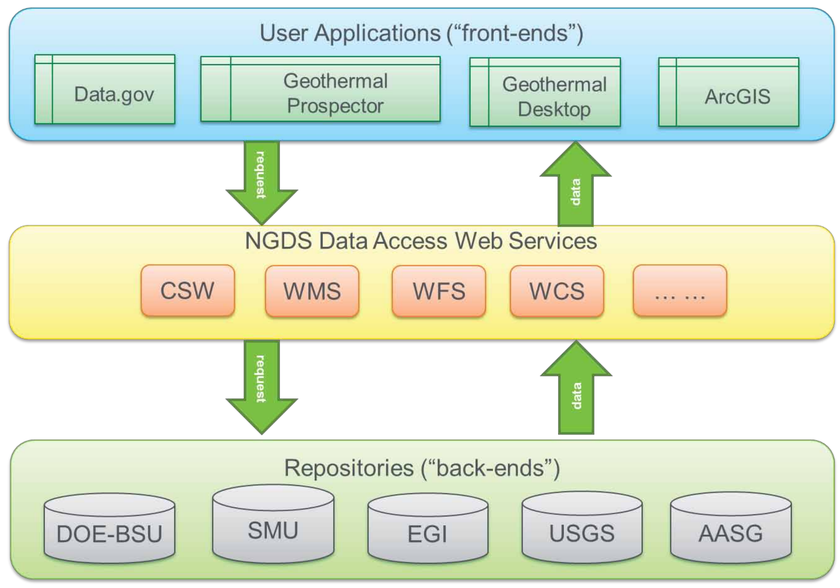


Locations of "Charter" NGDS Hubs, where data integration, curation and development are underway. NGDS is built on the US Geoscience Information Network (USGIN) data integration framework

Built on the US Geoscience Information Network

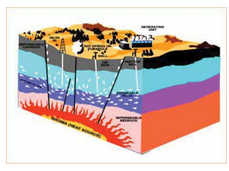
USGS science for a changing world · USGIN · AASG Association of American State Geologists

Basic NGDS Structure



Content Models Templates Interchange Formats

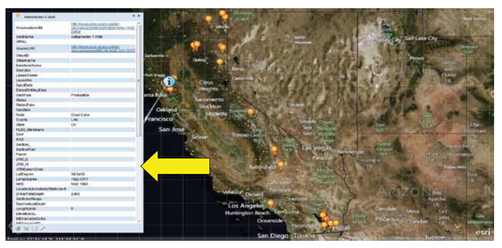
- Active Fault/Quaternary Fault
- Aqueous Chemistry
- Borehole Temp. Observation Feature
- Direct Use Feature
- Drill Stem Test Observations
- Earthquake Hypocenter
- Fault Feature
- Geologic Contact Feature
- Geologic Unit Feature
- Geothermal Area
- Geothermal Fluid Production
- Geothermal Power Plant
- Heat Flow
- Heat Pump Facility
- Lithology Interval Log Feature
- Metadata
- Thermal/Hot Spring Feature
- Volcanic Vents
- Well Header
- Well Log Data Compilation Workbook
- Well Log Observation
- Isopach map
- Permeability
- Rock chemistry
- Thermal conductivity measurement
- Gravity Data



What can you do with the NGDS?

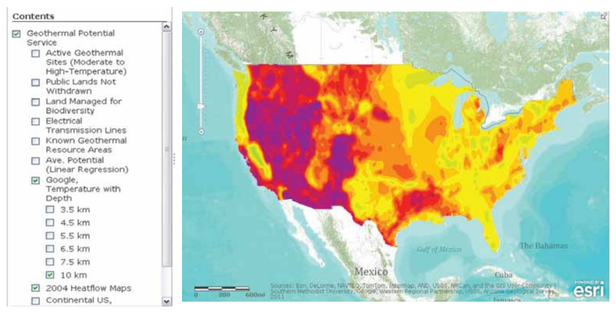
Example: Borehole temperature search results

Zoom and add to map (California borehole data)



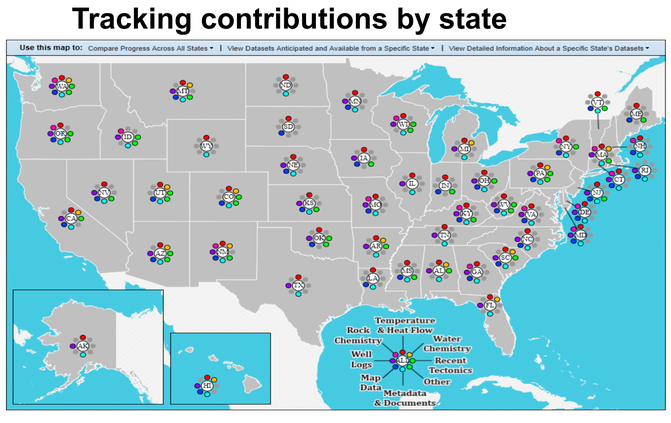
Individual features in live datasets may be queried for attributes

Analytical Integration



An example of NGDS Analytical Integration is shown here. Where an ESRI ArcGis application is used to PULL IN Google temperature and depth maps as well as SMU 2004 heatflow maps.

Tracking contributions by state



Data resources in State Geological Surveys

Selected data resources in State Geological Surveys

- 3.5 million wells – oil & gas, water
- 195,000 well logs
- 50,000 geothermal wells
- 750,000 BHT's
- 6 Tb existing digital data
- >75,000 scanned publications & maps
- 2.5 million feet of core
- 600,000 sample logs

>2 million records digital, online and growing fast!

- Metadata: >34,000 records
- Map Data: 589 records
- Well Log Data: 572,645 records
- Rock Chemistry: 458 records
- Temperature and Heat-flow: 84,086 records
- Water Chemistry: 4,141 records
- Recent Tectonics: 9,031 records

U.S. DEPARTMENT OF ENERGY

This material is based upon work supported by the Department of Energy award DE-EE000285D & DE-EE0001120 and by National Science Foundation grant EAR-0753154

NSF

Geophysical Research Abstracts
Vol. 14, EGU2012-13563-1, 2012
EGU General Assembly 2012
© Author(s) 2012

International Convergence on Geoscience Cyberinfrastructure

M. L. Allison (1), R. Atkinson (2), D.K. Arctur (3), S. Cox (4), I. Jackson (5), S. Nativi (6), and L.A. Wyborn (7)

(1) Arizona Geological Survey, Tucson, United States (lee.allison@azgs.az.gov), (2) CSIRO, Lucas Heights, Australia (Rob.atkinson@csiro.au), (3) Open Geospatial Consortium, Austin, United States (darctur@opengeospatial.org), (4) CSIRO, Perth, Australia (Simon.Cox@csiro.au), (5) OneGeology Secretariat, Keyworth, United Kingdom (ianjackson2011@gmail.com), (6) Consiglio Nazionale delle Ricerche, Roma, Italy (stefano.nativi@cnr.it), (7) Geoscience Australia, Canberra, Australia (lesley.wyborn@ga.gov.au)

There is growing international consensus on addressing the challenges to cyber(e)-infrastructure for the geosciences. These challenges include: Creating common standards and protocols; Engaging the vast number of distributed data resources; Establishing practices for recognition of and respect for intellectual property; Developing simple data and resource discovery and access systems; Building mechanisms to encourage development of web service tools and workflows for data analysis; Brokering the diverse disciplinary service buses; Creating sustainable business models for maintenance and evolution of information resources; Integrating the data management life-cycle into the practice of science.

Efforts around the world are converging towards de facto creation of an integrated global digital data network for the geosciences based on common standards and protocols for data discovery and access, and a shared vision of distributed, web-based, open source interoperable data access and integration. Commonalities include use of Open Geospatial Consortium (OGC) and ISO specifications and standardized data interchange mechanisms. For multidisciplinarity, mediation, adaptation, and profiling services have been successfully introduced to leverage the geosciences standards which are commonly used by the different geoscience communities -introducing a brokering approach which extends the basic SOA archetype.

Principal challenges are less technical than cultural, social, and organizational. Before we can make data interoperable, we must make people interoperable. These challenges are being met by increased coordination of development activities (technical, organizational, social) among leaders and practitioners in national and international efforts across the geosciences to foster commonalities across disparate networks. In doing so, we will 1) leverage and share resources, and developments, 2) facilitate and enhance emerging technical and structural advances, 3) promote interoperability across scientific domains, 4) support the promulgation and institutionalization of agreed-upon standards, protocols, and practice, and 5) enhance knowledge transfer not only across the community, but into the domain sciences, 6) lower existing entry barriers for users and data producers, 7) build on the existing disciplinary infrastructures leveraging their service buses. . All of these objectives are required for establishing a permanent and sustainable cyber(e)-infrastructure for the geosciences.

The rationale for this approach is well articulated in the AuScope mission statement: “Many of these problems can only be solved on a national, if not global scale. No single researcher, research institution, discipline or jurisdiction can provide the solutions. We increasingly need to embrace e-Research techniques and use the internet not only to access nationally distributed datasets, instruments and compute infrastructure, but also to build online, ‘virtual’ communities of globally dispersed researchers."

Multidisciplinary interoperability can be successfully pursued by adopting a "system of systems" or a "Network of Networks" philosophy. This approach aims to: (a) supplement but not supplant systems mandates and governance arrangements; (b) keep the existing capacities as autonomous as possible; (c) lower entry barriers; (d) Build incrementally on existing infrastructures (information systems); (e) incorporate heterogeneous resources by introducing distribution and mediation functionalities. This approach has been adopted by the European INSPIRE (Infrastructure for Spatial Information in the European Community) initiative and by the international GEOSS (Global Earth Observation System of Systems) programme.

Access to State Geological Survey data for geothermal energy development in the United States

Stephen M. RICHARD and M. Lee ALLISON, Arizona Geological Survey,
Tucson, AZ, USA; steve.richard@azgs.az.gov

The US Department of Energy is investing in an information management, discovery, and access platform to facilitate the use of geoscientific information to reduce the costs and risk in geothermal energy exploration development. State geological surveys are compiling and digitizing data in their holdings to create web- accessible resources. Resources are registered in the geothermal data system catalog for public discovery.

Data resources remain under the stewardship of their originating organizations, and are accessed by users through links provided in metadata records.

Interchange formats for feature-services are developed based on the existing data holdings and review by domain practitioners.

Content models for 18 interchange formats are in use or in review, and more than 129, web services have been deployed from four regional hubs.

Search interface for resources: http://search.stategeothermaldata.org/

Data Access:

Various levels of data publication are supported.

First tier: Data online and can be discovered. Existing file-based, web-accessible resources in their current data format and schema registered with metadata in the catalog.

--Format: downloadable file or a service of some sort
--Requirement: catalogued in a meaningful way
--What you get: no interoperability, only discoverability

Second Tier: Data discoverable with basic web service interoperability

--Format: data tabular, and available in OGC web service. Metadata are in a catalog
--What you get: Tier One discoverability and a basic level of interoperability. Software can be written to ingest data directly.
--Interoperability = software can ingest anyone else's data shared as an OGC web service, but user has to figure out semantics

Third Tier: Analyze across data sets. For widely used and high-value datasets, additional work is justified to make the data directly accessible through web services using standard protocols and interchange formats

--Format: data are transformed into a standard interchange format, are accessible via OGC web service, and metadata are in a catalog for them
--What you get:
*data are discoverable
*software can ingest data through standard OGC service calls
*software can know what that data means
*directly compare that data to other datasets in the same format
*Software can do analyses that cross dataset boundaries

Content models for interchange:

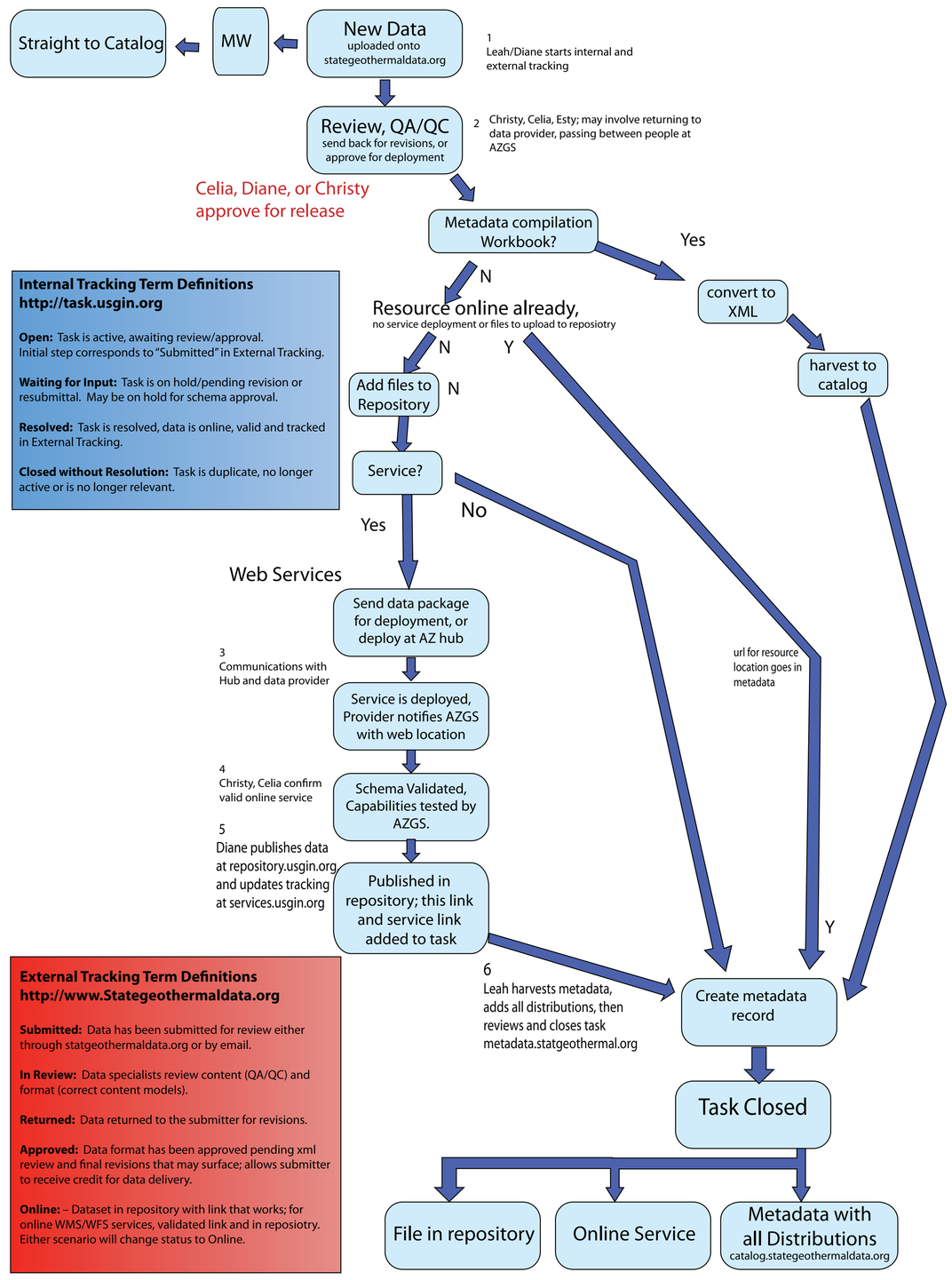
GML simple features for interchange: denormalized, property values are strings or numbers, simple point, line, or polygon geometry

OGC WFS implementation, works with existing clients


```
<wfs:FeatureCollection xsi:schemaLocation='http://stategeothermaldata.org/uri-gin/aasg/xmlschema/drillstemtest/1.8
        http://schemas.usgin.org/schemas/DrillStemTest1.8.xsd  http://www.opengis.net/wfs http://schemas.opengis.net/wfs/1.1.0/wfs.xsd'
        xmlns:aasg='http://stategeothermaldata.org/uri-gin/aasg/xmlschema/drillstemtest/1.8' xmlns:gml='http://www.opengis.net/gml'
        xmlns:wfs='http://www.opengis.net/wfs' xmlns:xlink='http://www.w3.org/1999/xlink'
        xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'>
  <gml:boundedBy>
            <gml:Envelope srsName='urn:ogc:def:crs:EPSG:6.9:4326'>
                <gml:lowerCorner>31.587935559999998 -114.79352706</gml:lowerCorner>
                <gml:upperCorner>36.997509219999998 -109.06292003999999</gml:upperCorner>
            </gml:Envelope>
  </gml:boundedBy>
  <gml:featureMember>
            <aasg:DrillStemTest gml:id='F831__1'>
                <aasg:OBJECTID>1</aasg:OBJECTID>
                <aasg:ObservationURI>http://resources.usgin.org/uri-gin/azgs/drillstemtest/AZGS:DST1api200505019/
                    </aasg:ObservationURI>
                <aasg:WellHeaderURI>http://resources.usgin.org/uri-gin/azgs/well/api:0200505019/</aasg:WellHeaderURI>
                <aasg:WellName>Sinclair Oil and Gas 1 Santa Fe Pacific</aasg:WellName>
                <aasg:APINo>200505019</aasg:APINo>
                <aasg:OtherID>06</aasg:OtherID>
                <aasg:DSTName>DST1</aasg:DSTName>
                <aasg:DSTNumber>1</aasg:DSTNumber>
                <aasg:Field>Wildcat</aasg:Field>
                <aasg:County>COCONINO</aasg:County>
                <aasg:State>Arizona</aasg:State>
                <aasg:PLSS_Meridians>Gila and Salt River</aasg:PLSS_Meridians>
                <aasg:TWP>28N</aasg:TWP>
                <aasg:RGE>01W</aasg:RGE>
                <aasg:Section_>35</aasg:Section_>
                <aasg:SectionPart>SW NE</aasg:SectionPart>
                <aasg:UTM_E>377302.29999999999</aasg:UTM_E>
                <aasg:UTM_N>3958890</aasg:UTM_N>
                <aasg:UTMDatumZone>12_NAD27</aasg:UTMDatumZone>
                <aasg:LatDegree>35.768207050000001</aasg:LatDegree>
                <aasg:LongDegree>-112.3573439</aasg:LongDegree>
                <aasg:SRS>WGS 84</aasg:SRS>
                <aasg:DSTTargetFormation>Naco | Martin</aasg:DSTTargetFormation>
                <aasg:DepthTopOpenZone>2827</aasg:DepthTopOpenZone>
                <aasg:DepthBottomOpenZone>2882</aasg:DepthBottomOpenZone>
                <aasg:UnitsOfMeasure>ft</aasg:UnitsOfMeasure>
                <aasg:FormationTD>Tapeats</aasg:FormationTD>
                <aasg:FlowHistory>3/4&amp;quot; bottom hole choke, 1&amp;quot; surface choke, open with slight blow, decreased, dead in 10
        minutes, open 75 minutes, recovered 10&amp;apos; drilling mud, hydrostatic pressure 1300-1300#, all other pressures 0#.</aasg:FlowHistory>
                <aasg:PressureInitialShutIn>0</aasg:PressureInitialShutIn>
                <aasg:HydrostaticPressure>1600</aasg:HydrostaticPressure>
                <aasg:HydrostaticPressureMethod>initial | final</aasg:HydrostaticPressureMethod>
                <aasg:FlowPressureMin>0</aasg:FlowPressureMin>
                <aasg:FlowPressureMax>0</aasg:FlowPressureMax>
                <aasg:RelatedResource>Graphic Lithologic Log, http://repository.azgs.az.gov/resources/og/OGPwelllogs/0006AMST.tif |
        Graphic Lithologic Log, http://repository.azgs.az.gov/resources/og/OGPwelllogs/0006AMSTD2.tif | Electrical Log,
        http://repository.azgs.az.gov/resources/og/OGPwelllogs/0006EL.tif | cuttings, metadata fileIdentifier= metadata
        fileIdentifier=258D6616-17C2-46F5-9000-E4CE2BB48489 | samples, metadata fileIdentifier=E5A43576-5FE6-4726-8897-16357AADBF1F |
        permit files, http://resources.usgin.org/uri-gin/azgs/doc/permitfile/0006/</aasg:RelatedResource>
                <aasg:InformationSource>2010-11-23 Information transferred from paper logs to spreadsheet by Leah Musil,  QA
        checked by Esty Pape, imported to Access table by S. M. Richard; updates to fields by join with S.J. Rauzi API6
        table.</aasg:InformationSource>
                <aasg:Shape>
                    <gml:Point>
                        <gml:pos>35.768207050000001 -112.3573439</gml:pos>
                    </gml:Point>
                </aasg:Shape>
            </aasg:DrillStemTest>
  </gml:featureMember>
</wfs:FeatureCollection>
```


Example WFS response XML schema for interchange formats are at:
http://schemas.usgin.org/schemas/

Data acquisition workflow



Metadata: content requirements

title **description** **resource_url**	Essential-- reject from catalog if missing
publication_date author west_bounding_longitude south_bounding_longitude resource contact point resource type	Minimum to be considered basic metadata
resource_Language metadata_contact_name metadata_contact point metadata_date metadata_language	Additional information required for metadata record maintenance; can mostly be automatically included in records

Technical framework for the system

- ISO 19139 XML metadata profile for metadata interchange
  http://repository.usgin.org/uri_gin/usgin/dlio/499,
- Open Geospatial Consortium (OGC) CSW 2.0.2 for catalog search
  http://catalog.usgin.org/geoportal/csw?request=GetCapabilities&service=CSW,
- OGC web map services (WMS) for delivery of georeferenced map images,
- OGC web feature service (WFS) simple features for delivering datasets
- NetCDF OpenDAP for grid-based data delivery.
- Features are assigned http URIs that dereference by default to XML WFS features, for some wells html pages are also available representations
  http://lab.usgin.org/groups/using-django-usgin/uri-redirection-engine-pictures

http://lab.usgin.org http://stategeothermaldata.org
This material is based upon work supported by the Department of Energy under Award Numbers DE-EE0002850 and DE-EE0001120 and by National Science Foundation grant EAR-0753154

National Geothermal Data System: State Geological Survey Contributions to Date

Abstract

In collaboration with the Association of American State Geologists the Arizona Geological Survey is leading the effort to bring legacy geothermal data to the U.S. Department of Energy's National Geothermal Data System (NGDS). NGDS is a national, sustainable, distributed, interoperable network of data and service (application) providers entering its final stages of development. Once completed the geothermal industry, the public, and policy makers will have access to consistent and reliable data, which in turn, reduces the amount of staff time devoted to finding, retrieving, integrating, and verifying information. With easier access to information, the high cost and risk of geothermal power projects (especially exploration drilling) is reduced. This presentation focuses on the scientific and data integration methodology as well as State Geological Survey contributions to date. The NGDS is built using the U.S. Geoscience Information Network (USGIN) data integration framework to promote interoperability across the Earth sciences community and with other emerging data integration and networking efforts. Core to the USGIN concept is that of data provenance; by allowing data providers to maintain and house their data. After concluding the second year of the project, we have nearly 800 datasets representing over 2 million data points from the state geological surveys. A new AASG specific search catalog based on popular internet search formats enables end users to more easily find and identify geothermal resources in a specific region. Sixteen states, including a consortium of Great Basin states, have initiated new field data collection for submission to the NGDS. The new field data includes data from at least 21 newly drilled thermal gradient holes in previously unexplored areas. Most of the datasets provided to the NGDS are being portrayed as Open Geospatial Consortium (OGC) Web Map Services (WMS) and Web Feature Services (WFS), meaning that the data is compatible with a variety of visualization software. Web services are ideal for the NGDS data for a number of reasons including that they preserve data ownership in that they are read only and new services can be deployed to meet new requirements without modifying existing applications.

Cite as: Patten, K; Allison, M L; Richard, S M; Clark, R; Love, D; Coleman, C; Caudill, C; Matti. J; Musli, L; Day, J; and Chen, G (2012), National Geothermal Data System: State Geological Survey Contributions to Date, Abstract H12F-07 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3-7 Dec.

OneGeology - a Geoscience Exemplar for Worldwide Cyberinfrastructure Capacity-building and Scientific Innovation

Abstract

OneGeology is a trail-blazing global initiative that has helped propel the geosciences into the forefront of cyberinfrastructure development with potentially transformative impacts on scientific and technical innovation across broad areas of society. In the five years since its launch, 117 nations, through their Geological Surveys have signed the OneGeology protocols and nearly half are serving up national geological maps as Web services at varying scales, with the remainder developing those capabilities. In federal systems, states and provinces are increasingly adding higher resolution spatial data to the national contributions to the global system. The OneGeology concept of a distributed, open-source, Web-service based network has become the archetype for transforming data into knowledge and innovation. This is not only revolutionizing the geosciences but offering opportunities for governments to use these cutting-edge capabilities for broad innovation and capacity building. Across the globe, communities are facing the same four challenges: put simply, how do we best make data discoverable, shareable, viewable and downloadable, so that the user also has access to consistent data at a national and continental level? The principle of managing scientific and societal data and knowledge where they are generated and are best understood is well established in the geoscience community and can be scaled up and transferred to other domains and sectors of society. The distributed nature of most data sources means the complementary delivery mechanism of Web map services has become equally prevalent in the spatial data community. Together these factors are driving a world-wide revolution in the way spatial information is being disseminated to its users. Industry, academia, and governments are quickly adopting and adapting to this new paradigm and discovering that very modest investments in this emerging field are reaping tremendous returns in national capacity and triggering a wave of innovation and economic development symptomatic of previous deployment of new infrastructures, from transportation networks to the electrical grid to the Internet. OneGeology continues to implement and deploy critical cyberinfrastructure capabilities in best practices, definitions, and standards on data management. The global adoption of OneGeology is also lowering the barriers to accessing the world's digital resources.

Cite as: van Daalen, T and Allison M L (2012), OneGeology - a Geoscience Exemplar for Worldwide Cyberinfrastructure Capacity-building and Scientific Innovation, Abstract IN21A-1469 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3-7 Dec.

GeoSciML and EarthResourceML Update, 2012

Abstract

CGI Interoperability Working Group activities during 2012 include deployment of services using the GeoSciML-Portrayal schema, addition of new vocabularies to support properties added in version 3.0, improvements to server software for deploying services, introduction of EarthResourceML v.2 for mineral resources, and collaboration with the IUSS on a markup language for soils information. GeoSciML and EarthResourceML have been used as the basis for the INSPIRE Geology and Mineral Resources specifications respectively. GeoSciML-Portrayal is an OGC GML simple-feature application schema for presentation of geologic map unit, contact, and shear displacement structure (fault and ductile shear zone) descriptions in web map services. Use of standard vocabularies for geologic age and lithology enables map services using shared legends to achieve visual harmonization of maps provided by different services. New vocabularies have been added to the collection of CGI vocabularies provided to support interoperable GeoSciML services, and can be accessed through http://resource.geosciml.org. Concept URIs can be dereferenced to obtain SKOS rdf or html representations using the SISSVoc vocabulary service. New releases of the FOSS GeoServer application greatly improve support for complex XML feature schemas like GeoSciML, and the ArcGIS for INSPIRE extension implements similar complex feature support for ArcGIS Server. These improved server implementations greatly facilitate deploying GeoSciML services. EarthResourceML v2 adds features for information related to mining activities. SoilML provides an interchange format for soil material, soil profile, and terrain information. Work is underway to add GeoSciML to the portfolio of Open Geospatial Consortium (OGC) specifications.

Cite as: Richard S M and Commission for the Management and Application International Union of Geological Sciences Interoperability Working Group (2012), GeoSciML and EarthResourceML Update, 2012, Abstract IN31A-1490 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3-7 Dec.

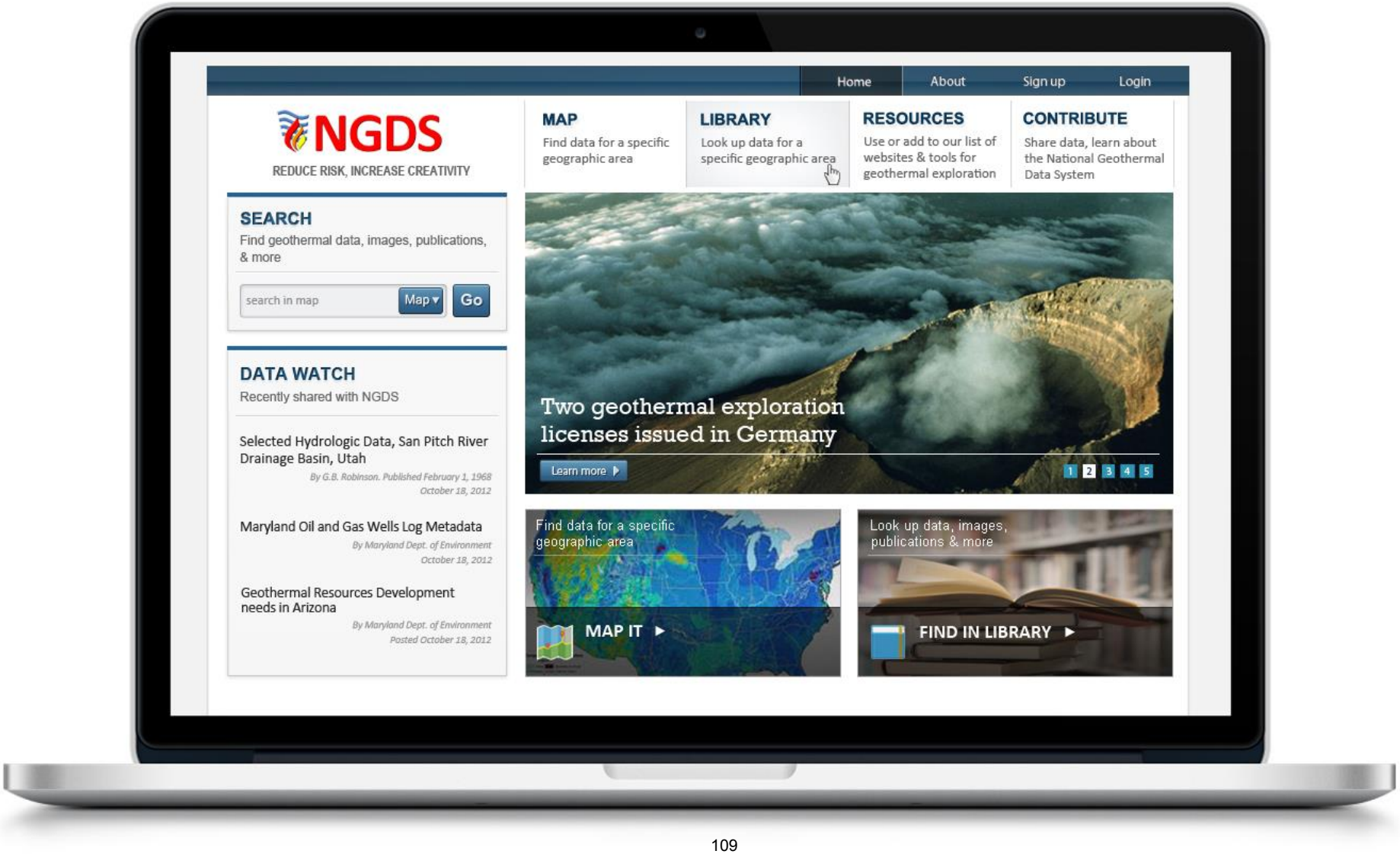
A Lithology Based Map Unit Schema For OneGeology Regional Geologic Map Integration

Abstract

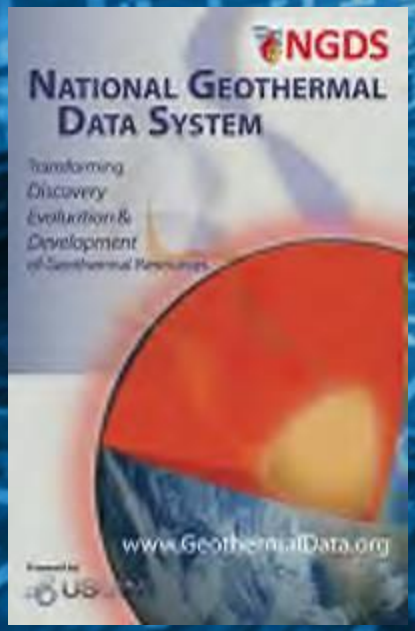
A system of lithogenetic categories for a global lithological map (GLiM, http://www.ifbm.zmaw.de/index.php?id=6460&L=3) has been compiled based on analysis of lithology/genesis categories for regional geologic maps for the entire globe. The scheme is presented for discussion and comment. Analysis of units on a variety of regional geologic maps indicates that units are defined based on assemblages of rock types, as well as their genetic type. In this compilation of continental geology, outcropping surface materials are dominantly sediment/sedimentary rock; major subdivisions of the sedimentary category include clastic sediment, carbonate sedimentary rocks, clastic sedimentary rocks, mixed carbonate and clastic sedimentary rock, colluvium and residuum. Significant areas of mixed igneous and metamorphic rock are also present. A system of global categories to characterize the lithology of regional geologic units is important for Earth System models of matter fluxes to soils, ecosystems, rivers and oceans, and for regional analysis of Earth surface processes at global scale. Because different applications of the classification scheme will focus on different lithologic constituents in mixed units, an ontology-type representation of the scheme that assigns properties to the units in an analyzable manner will be pursued. The OneGeology project is promoting deployment of geologic map services at million scale for all nations. Although initial efforts are commonly simple scanned map WMS services, the intention is to move towards data-based map services that categorize map units with standard vocabularies to allow use of a common map legend for better visual integration of the maps (e.g. see OneGeology Europe, http://onegeology-europe.brgm.fr/ geoportal/ viewer.jsp). Current categorization of regional units with a single lithology from the CGI SimpleLithology (http://resource.geosciml.org/201202/ Vocab2012html/ SimpleLithology201012.html) vocabulary poorly captures the lithologic character of such units in a meaningful way. A lithogenetic unit category scheme accessible as a GeoSciML-portrayal-based OGC Styled Layer Description resource is key to enabling OneGeology (http://oneGeology.org) geologic map services to achieve a high degree of visual harmonization.

Cite as: Moosdorf, N and Richard S M (2012),: A Lithology Based Map Unit Schema For OneGeology Regional Geologic Map Integration, Abstract IN51C-1706 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3-7 Dec.

NGDS – transforming geothermal exploration

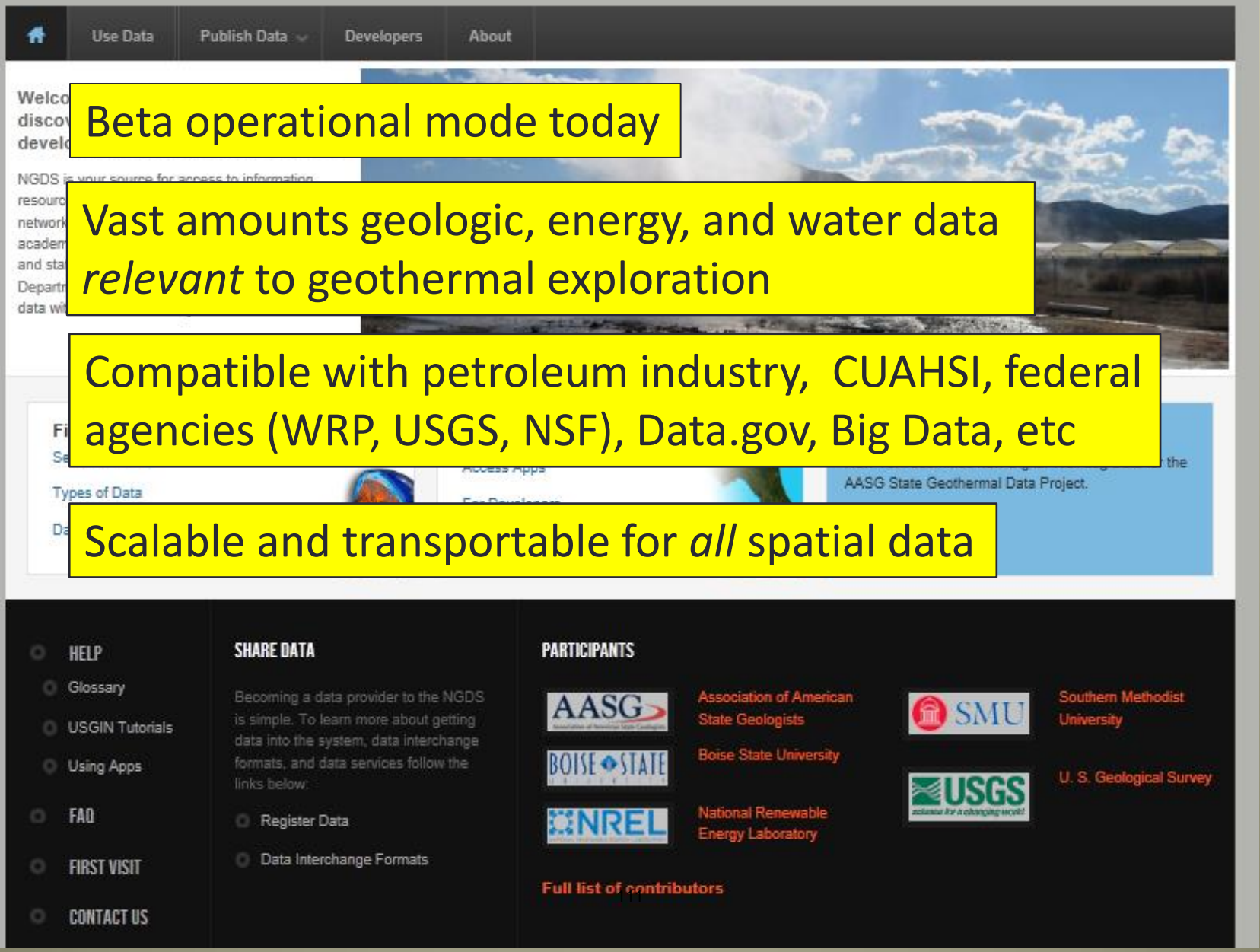


How does it work?



NGDS National Geothermal Data System

www.geothermaldata.org



Data resources online in NGDS

- > 10,000 data sets from all 50 states
- > 5.2 million records digital, online, and growing fast!

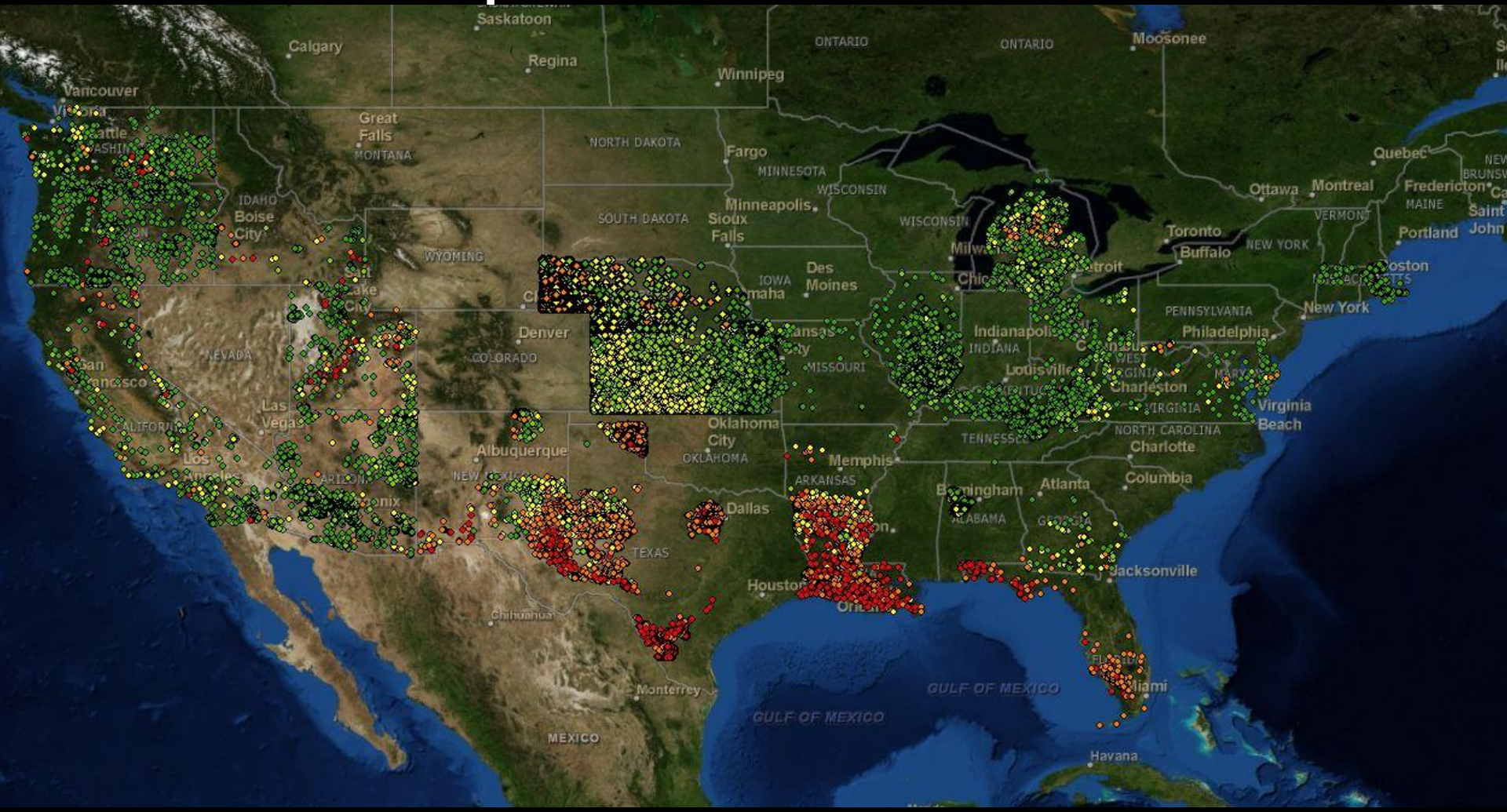
>840,000 oil & gas wells

>410,000 water wells

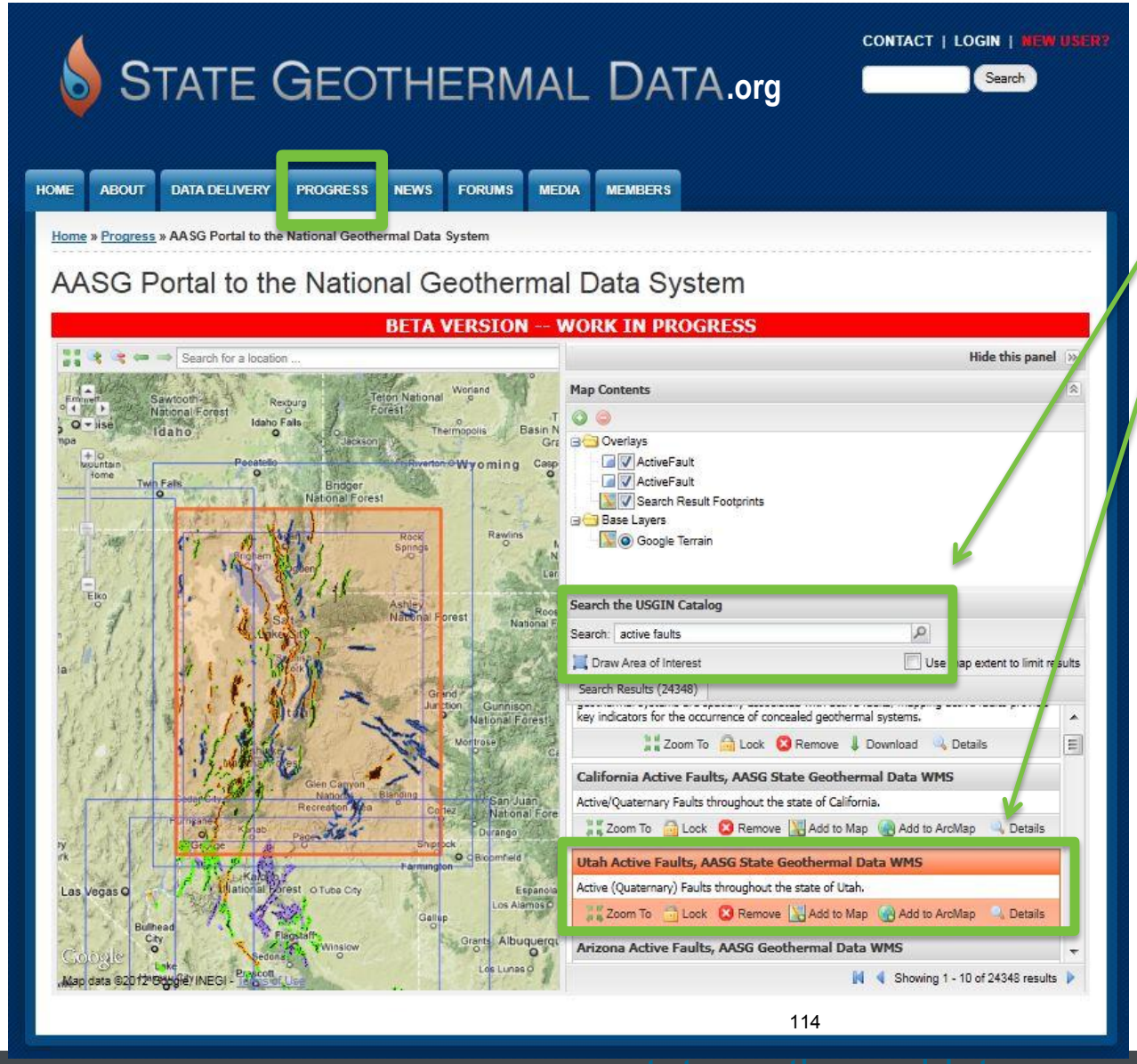
>9,000 geothermal wells

NGDS National Geothermal Data System

Nationwide Borehole Temperature Data



Basic Search Interface



Accessing data through a map based search on www.stategeothermaldata.org:

Search Box

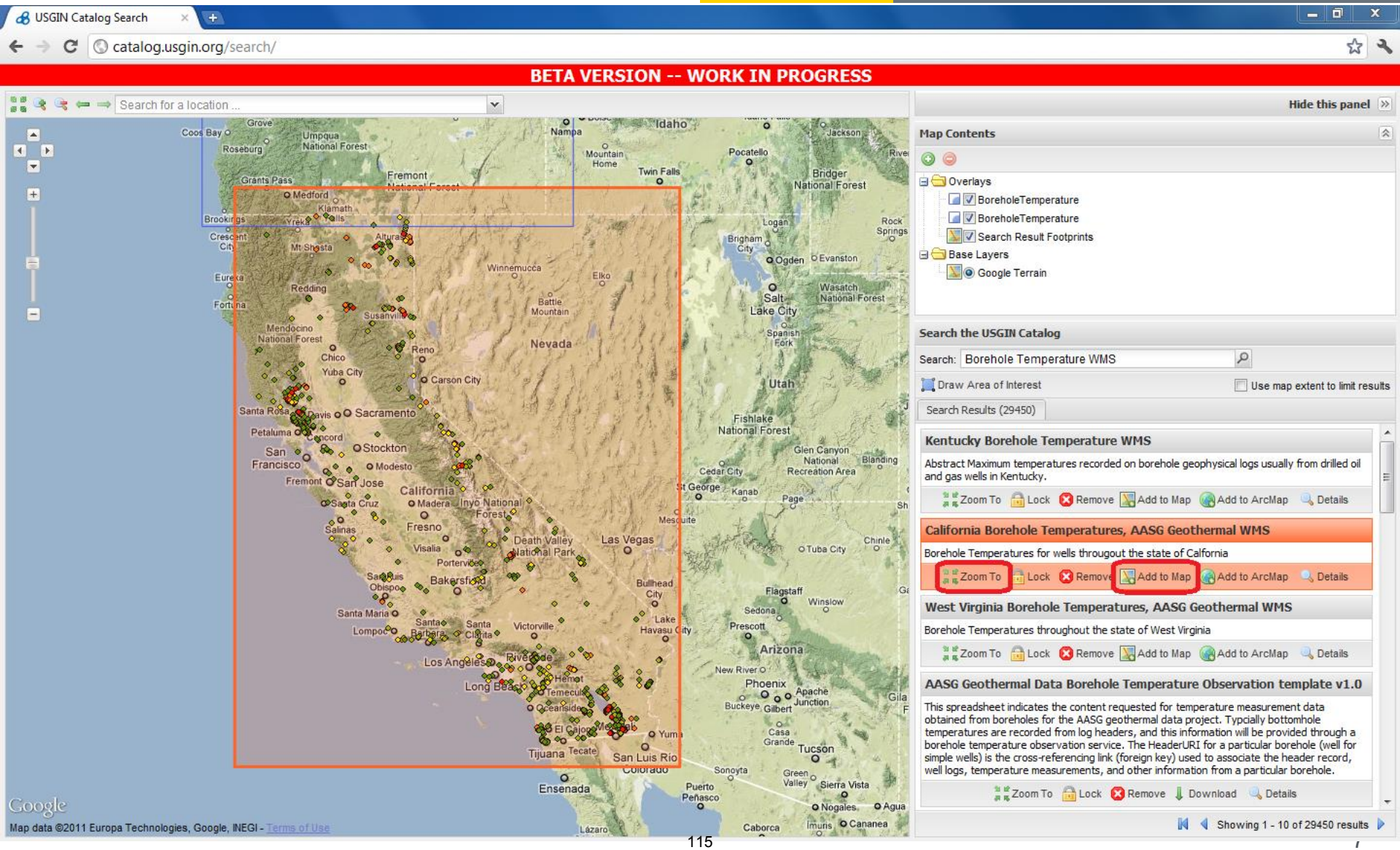
WMS Functions:

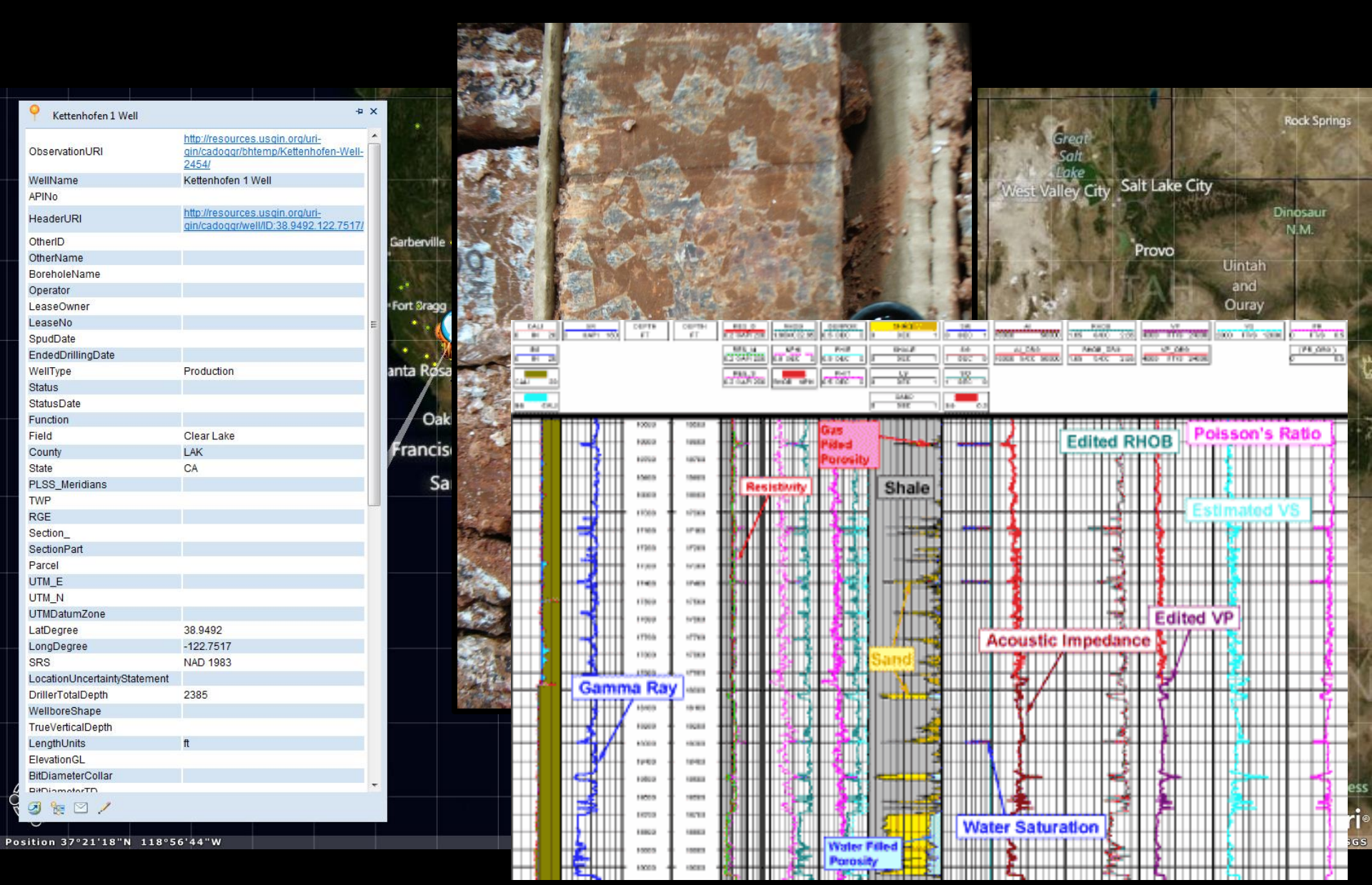
- Zoom To
- Lock (in Search)
- Remove (from Search)
- Add to Map
- Add to ArcMap
- Details

Other Dataset Functions:

- Zoom To
- Lock
- Remove
- Download
- Details

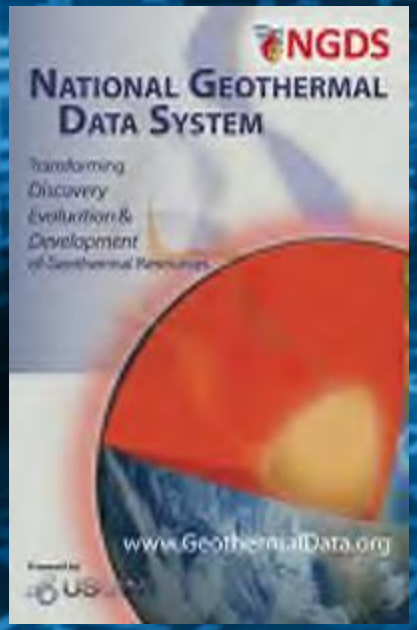
Zoom to and add to map – CA BHTs

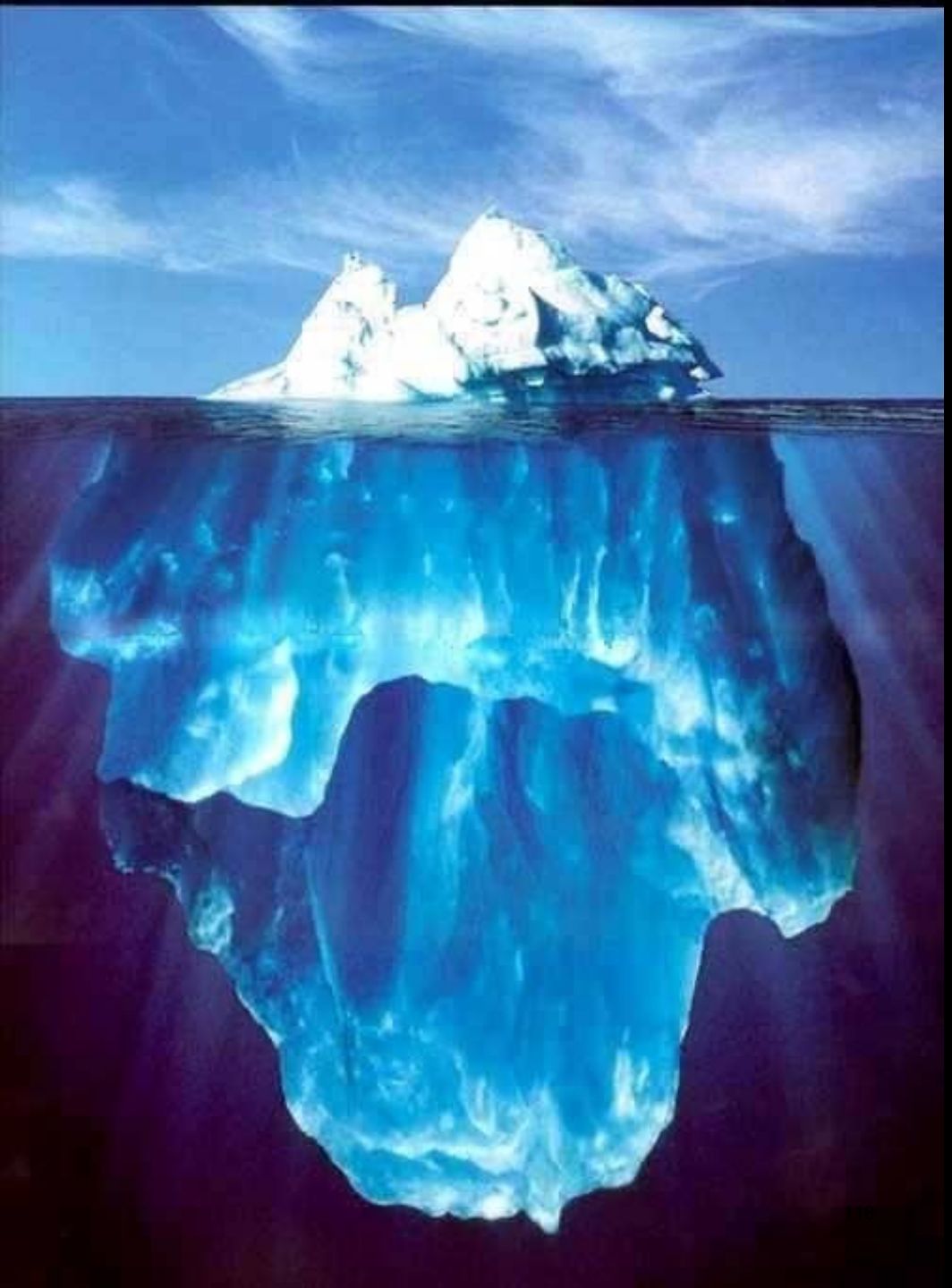




Individual features in live datasets may be queried for attributes

How did we put it together?





Most of the technology exists

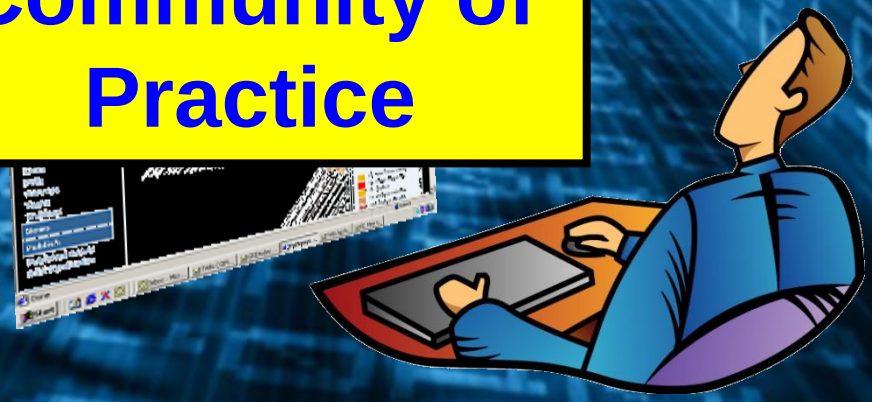
Challenges are cultural and organizational

Geoscience Information Network (USGIN)

USGS
science for a changing world

AASG
Association of American State Geologists

Distributed
Web-based
Interoperable
Open source
Community of Practice



USGS
science for a changing world

USGIN

Community for
Data Integration

ScienceBase

WRP
WESTERN REGIONAL PARTNERSHIP

WESTERN REGIONAL PARTNERSHIP

15+ Federal Agencies and 5 Governors - NM, AZ, UT, NV, CA

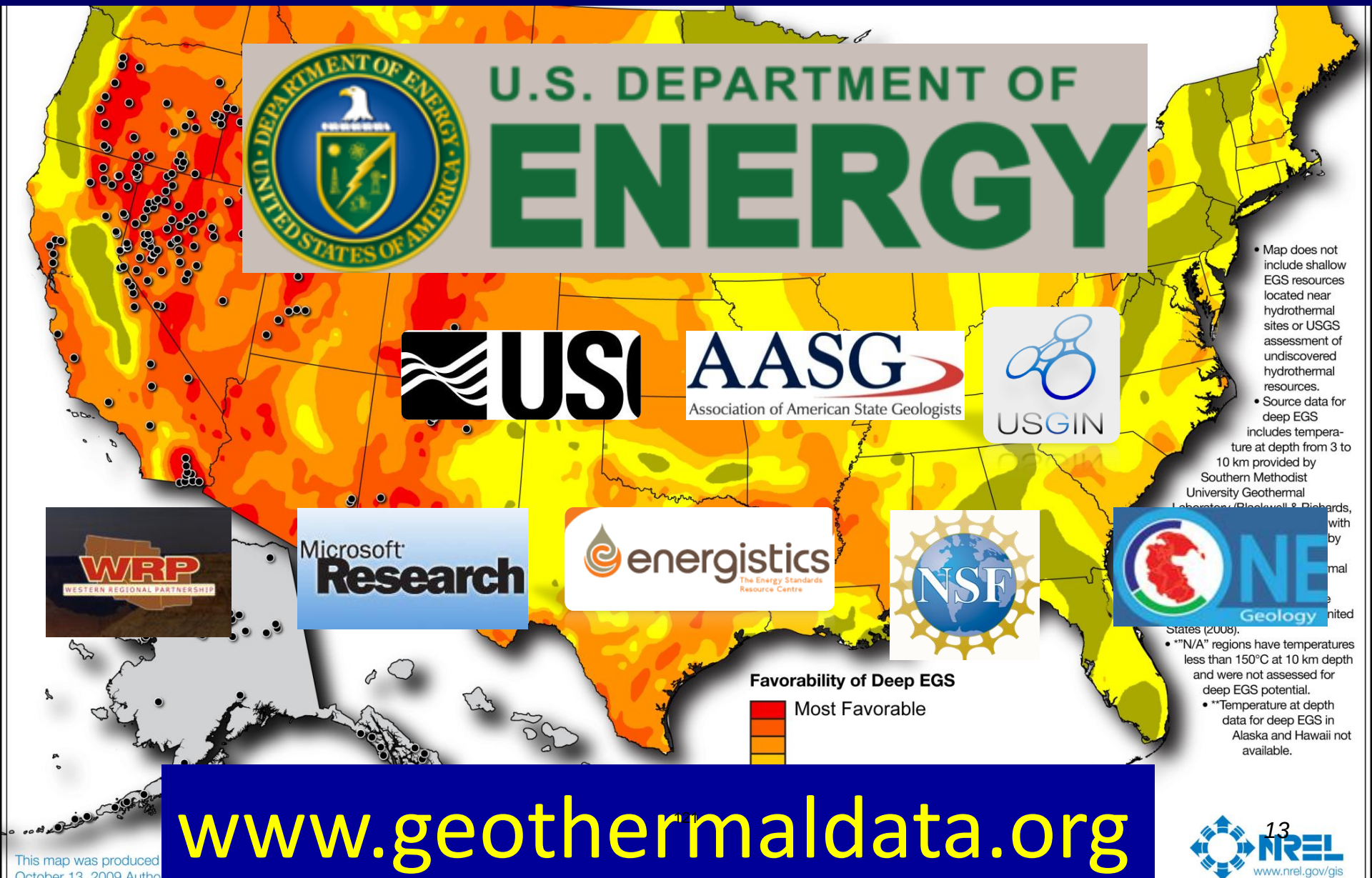
10,000 GIS layers for land use management

energistics
The Energy Standards Resource Centre

Energy Industry Metadata Standards Working Group

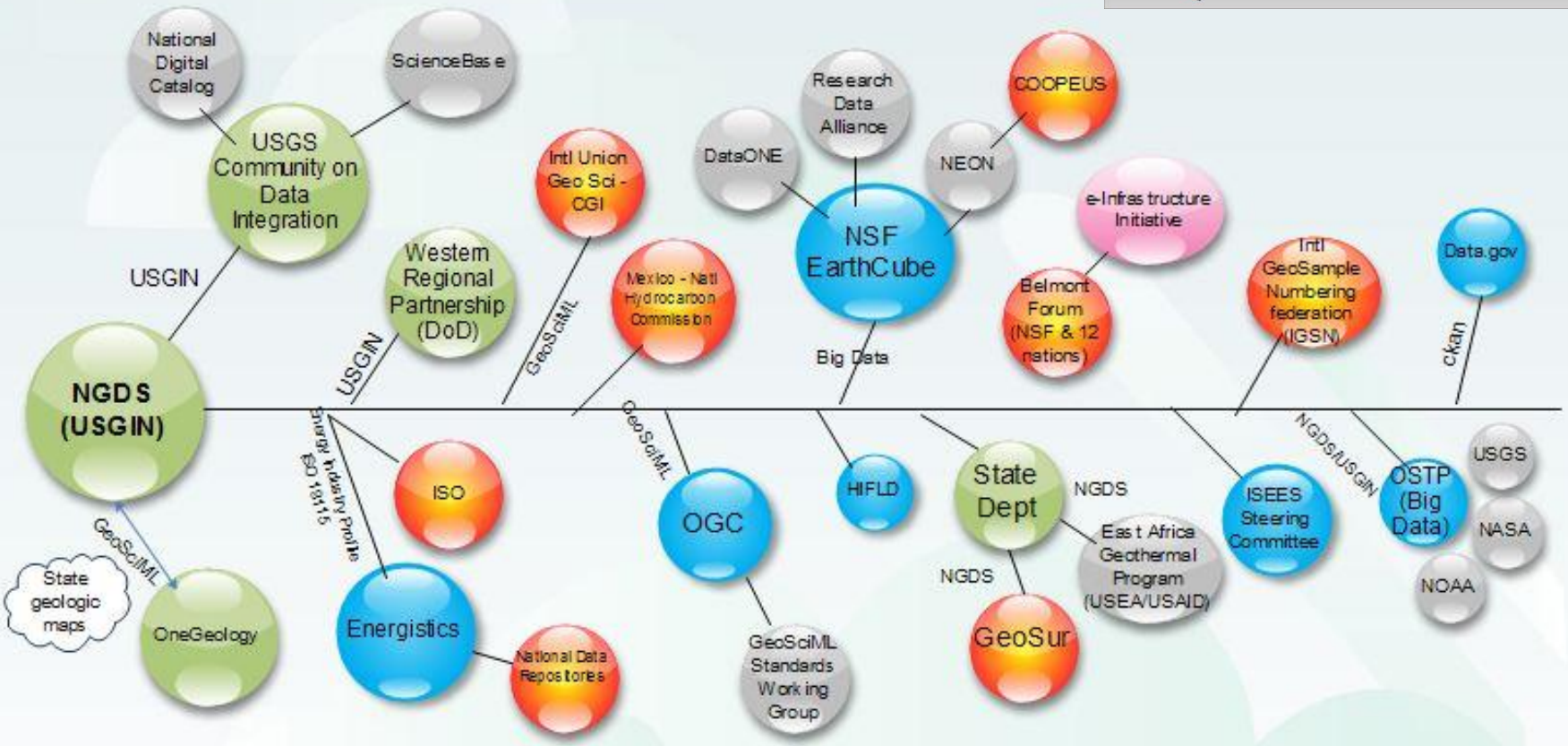
End-to-end discovery, access, and exchange of upstream petroleum data

The *DataWeb*

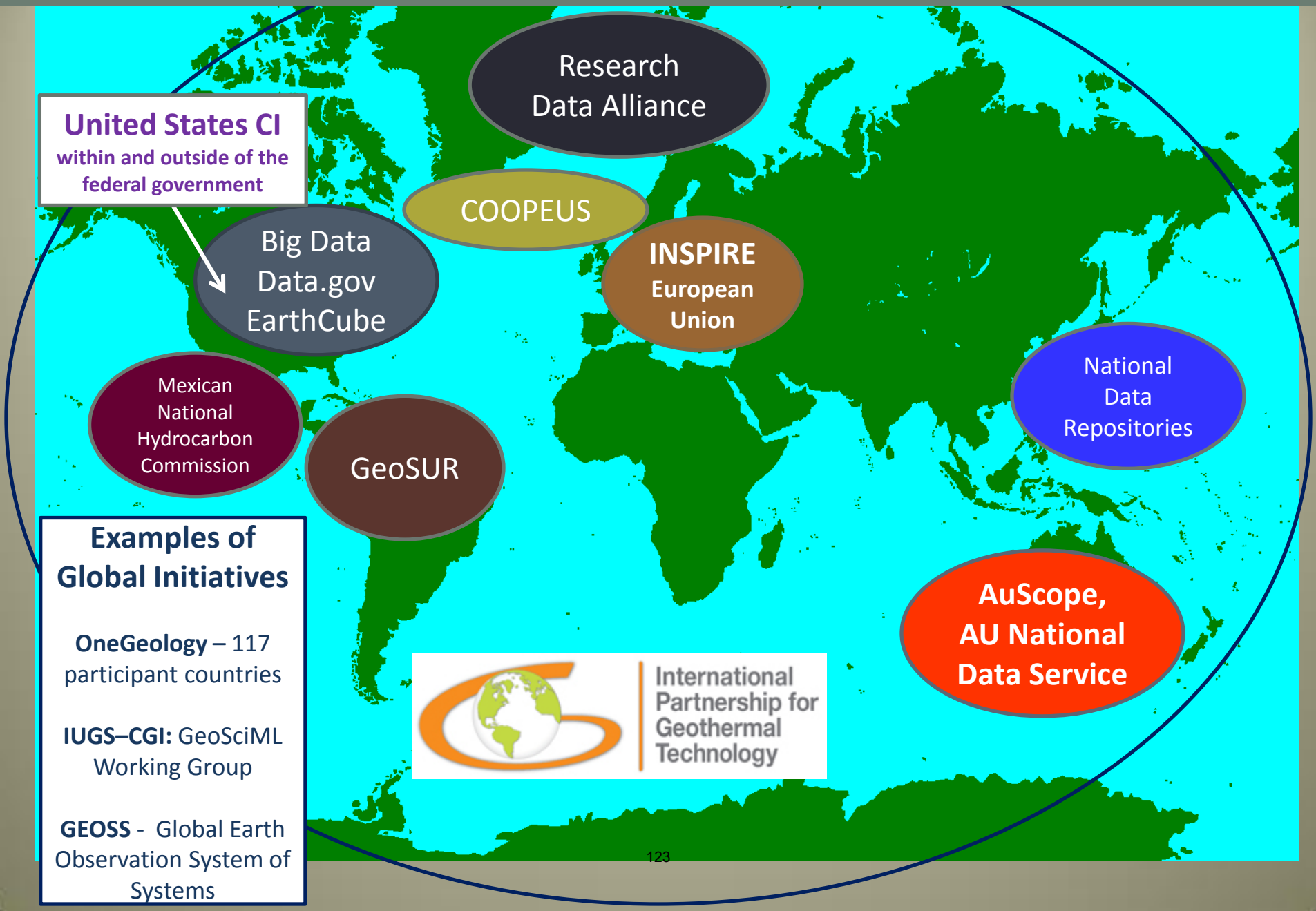


www.geothermaldata.org

National Geothermal Data System



International cyberinfrastructure initiatives



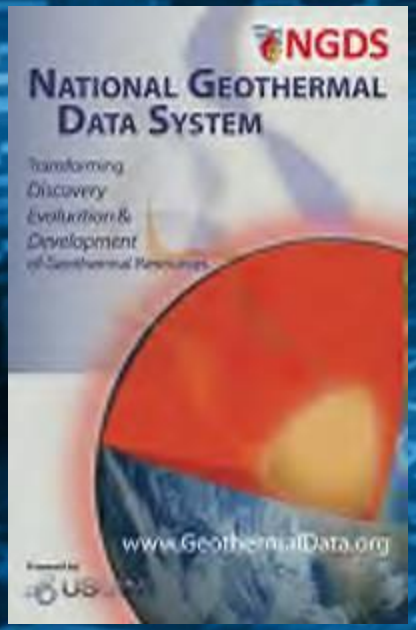
Examples of Global Initiatives

OneGeology – 117 participant countries

IUGS–CGI: GeoSciML Working Group

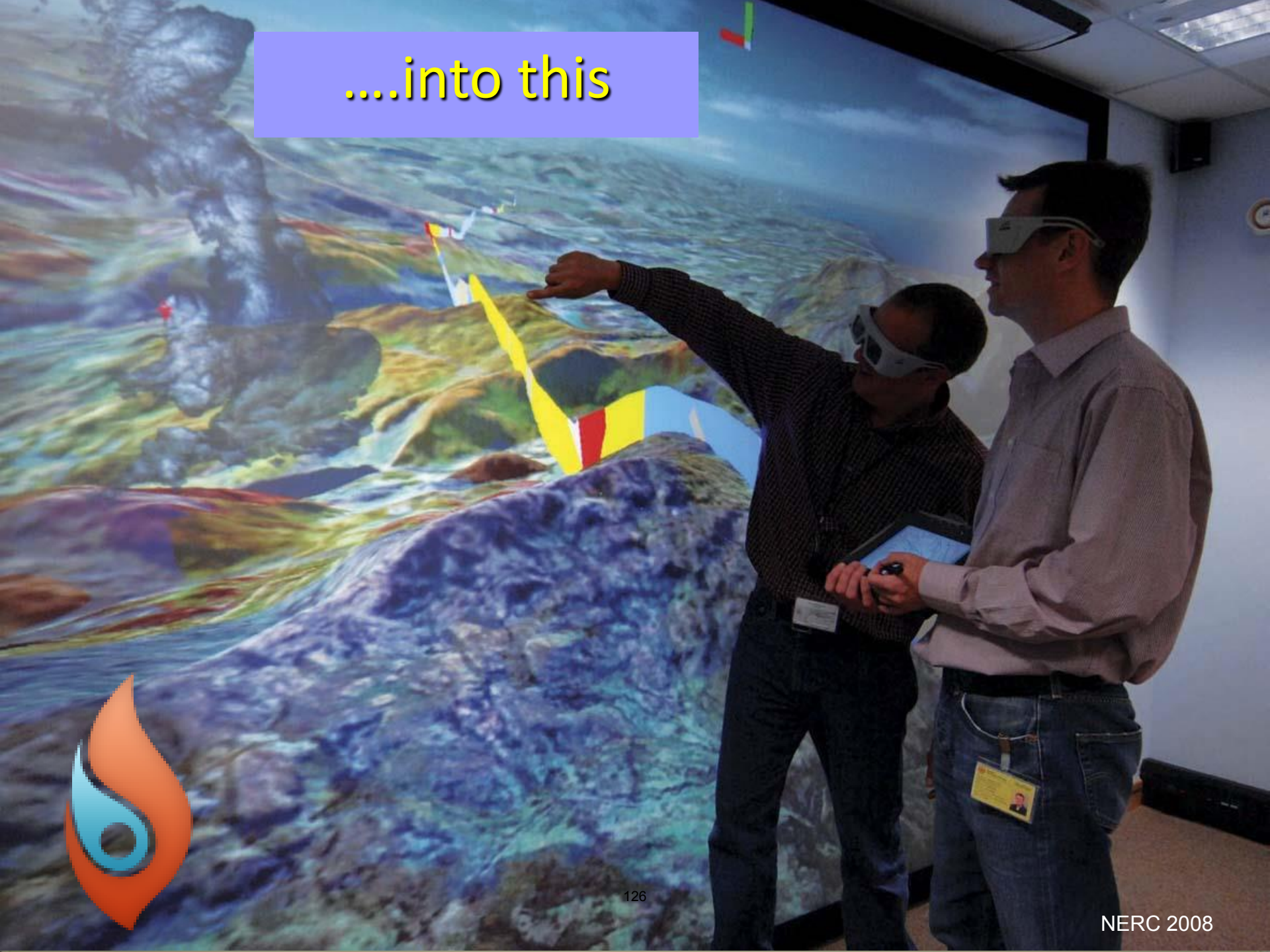
GEOSS - Global Earth Observation System of Systems

How does it help industry?



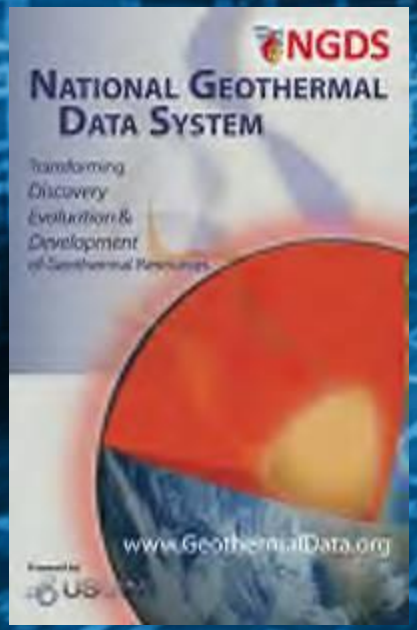
how to turn this….

….into this

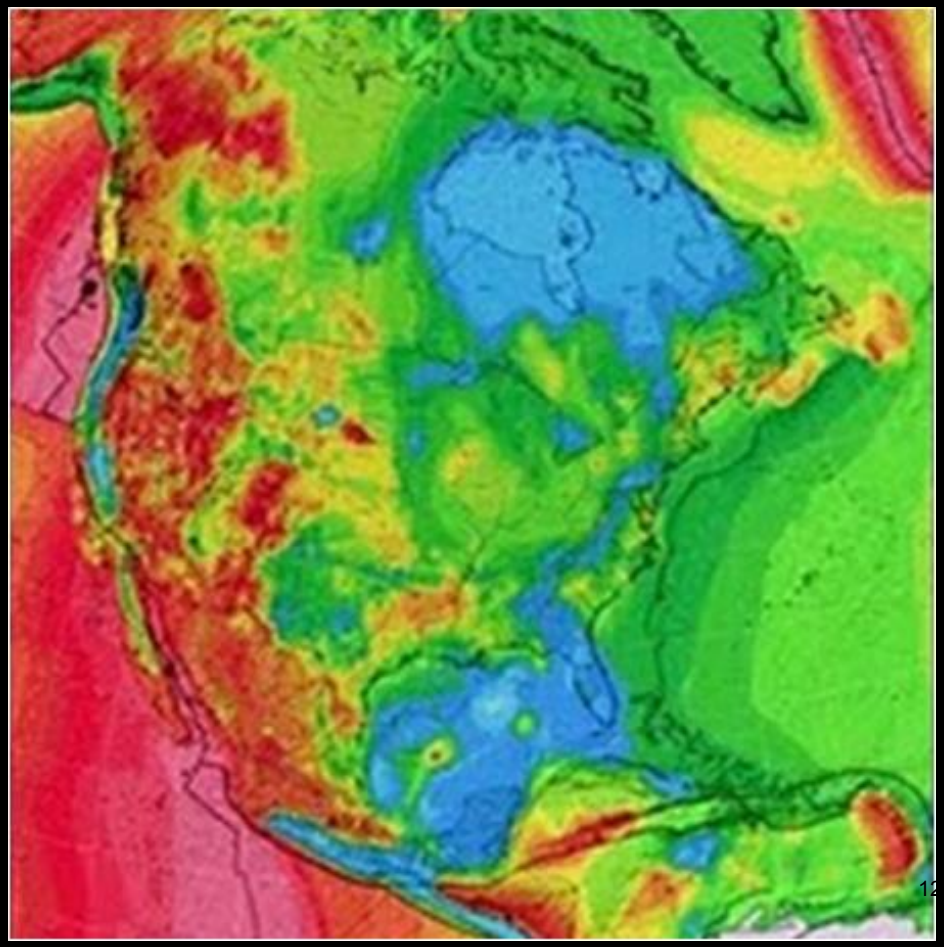


NERC 2008

How does NGDS reduce risk?



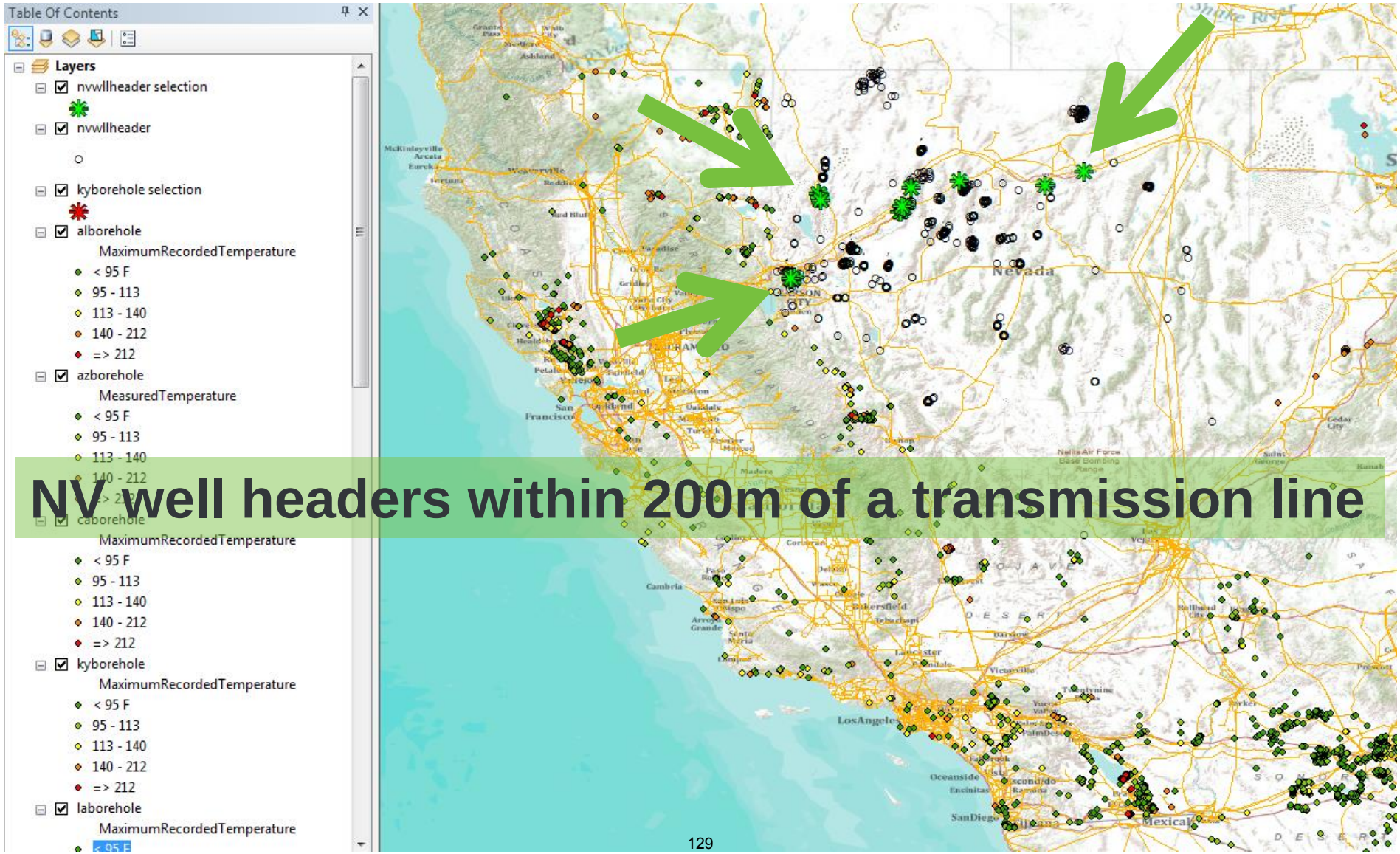
Geothermal Exploration: Everything Digital, Online, and Interoperable



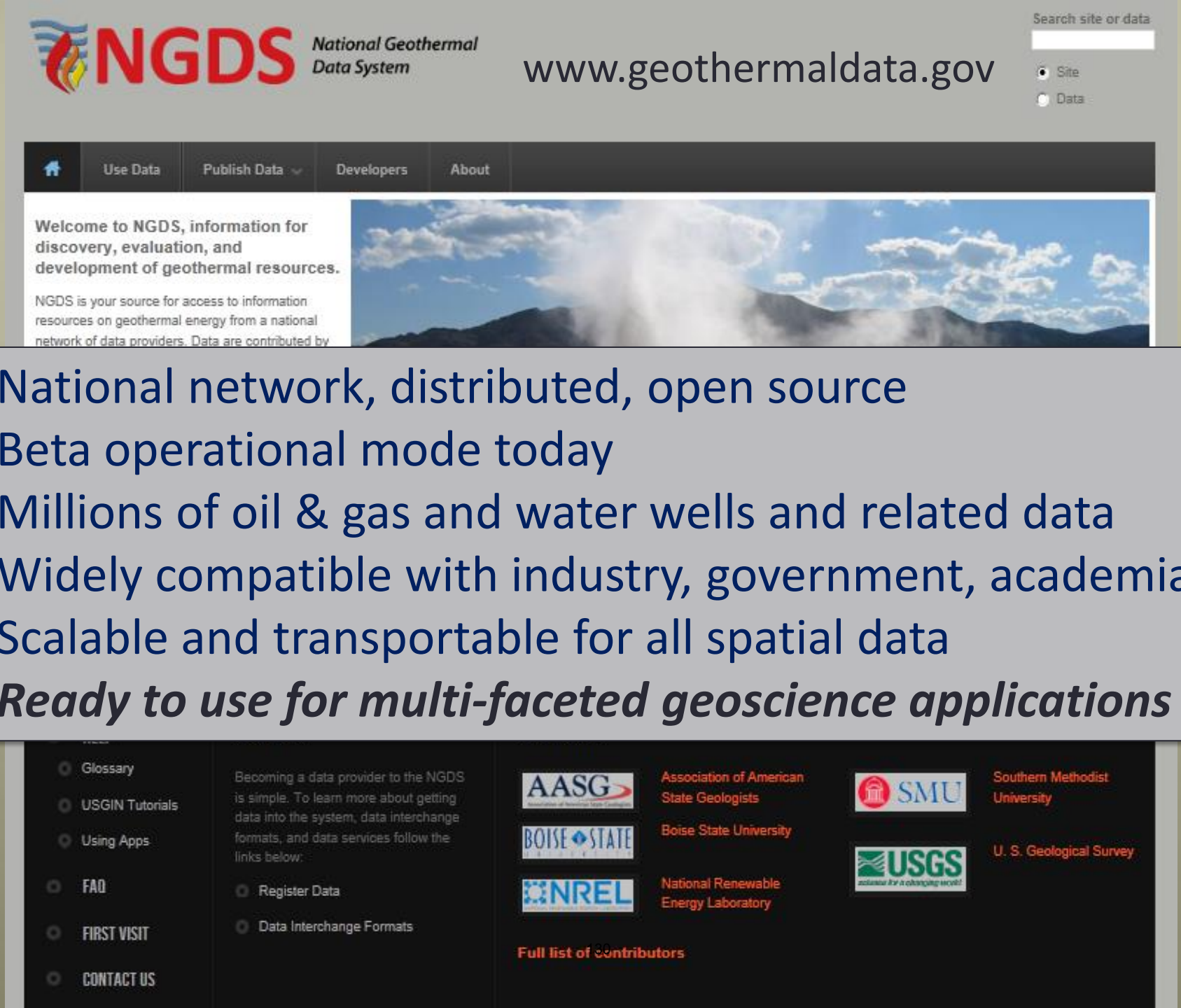
NGDS National Geothermal Data System

www.geothermaldata.org

Exploring the Data in our Application – Well Headers in Nevada using ArcGIS



NV well headers within 200m of a transmission line


NGDS
National Geothermal Data System
www.geothermaldata.gov
Search site or data
Site
Data
Use Data
Publish Data
Developers
About
Welcome to NGDS, information for discovery, evaluation, and development of geothermal resources.
NGDS is your source for access to information resources on geothermal energy from a national network of data providers. Data are contributed by
National network, distributed, open source
Beta operational mode today
Millions of oil & gas and water wells and related data
Widely compatible with industry, government, academia
Scalable and transportable for all spatial data
Ready to use for multi-faceted geoscience applications
Glossary
USGIN Tutorials
Using Apps
FAQ
FIRST VISIT
CONTACT US
Becoming a data provider to the NGDS is simple. To learn more about getting data into the system, data interchange formats, and data services follow the links below:
Register Data
Data Interchange Formats
AASG
Association of American State Geologists
BOISE STATE
Boise State University
NREL
National Renewable Energy Laboratory
SMU
Southern Methodist University
USGS
U. S. Geological Survey
Full list of contributors

Action Items

Spin off USGIN as independent not-for-profit
NGDS as first big client
Diversified client base – geospatial, natural resources, state and federal governments, international
Diversified funding base –

- Grants and contracts
- Services
- Repositories
- Ads
- Subscriptions
- Licenses

NGDS
National Geotherma
Data System

GEOTHERMAL DATA in the US National Geothermal Data System

NGDS Design and Development Team

Stephen Richard (presenting), Arlene Anderson, David Blackwell, Cathy Chickering, Toni Boyd, Roland Horne, Matthew MacKenzie, Joseph Moore, Duane Nickull, Lisa Shevenell

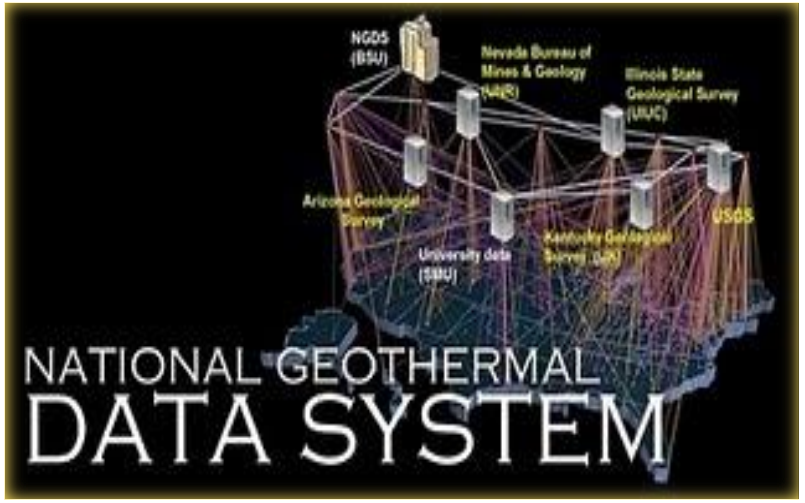
2013 CanGEA Fifth Annual Conference
Calgary Canada, March 21, 2013

Outline

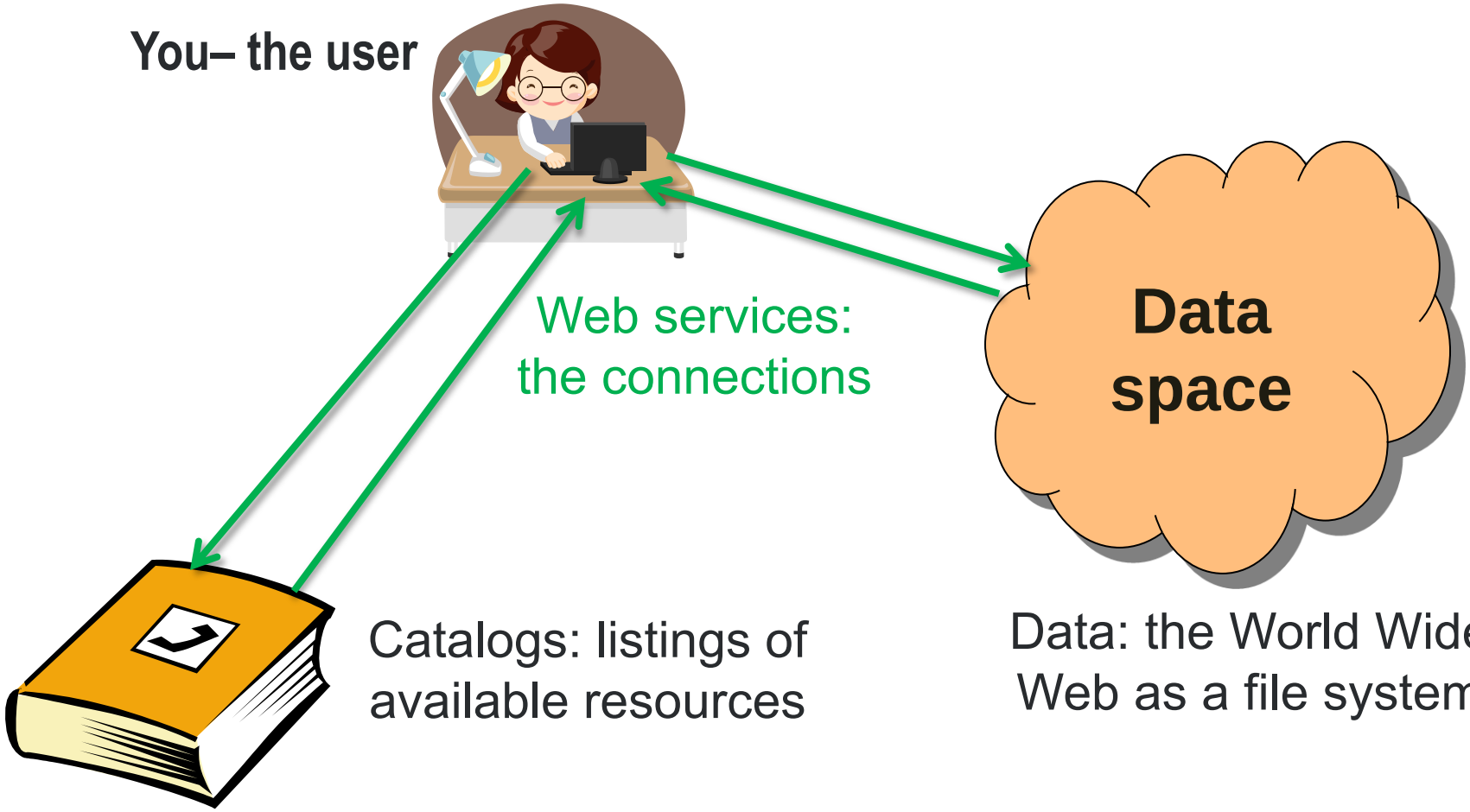
- **How is it put together**
  - Approach to Data Management
  - The Domain Steering Committee
  - Recommended data types
- How does it work
  - Content Models
  - Interchange formats
  - Catalog

Overview

- NGDS is a data system, not a database
  - Discover and access geothermal data from many different sources
- System is unified by:
  - Catalog and standardized metadata
  - Data access protocols and interchange formats (it's the Web for data)



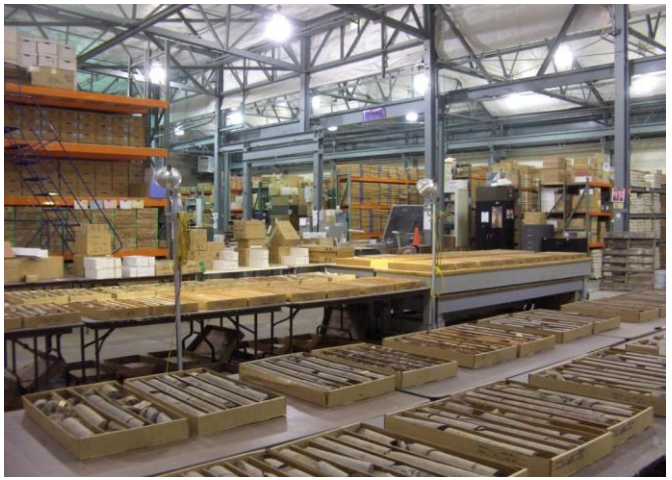
A distributed data system

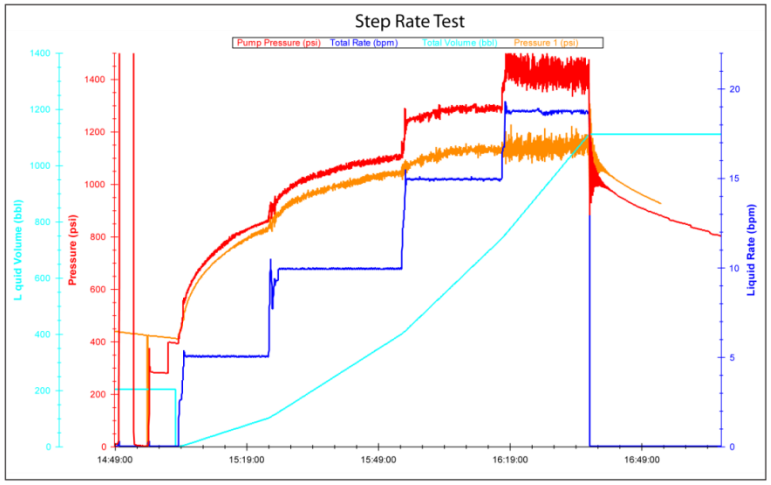


Data Types

- **Three sorts of information:**
  - Data – information about the Earth and geothermal energy systems
  - Metadata – Information about system data or samples
  - Annotation – notes added to by the community (ratings, comments, reviews…)

EGI Geothermal Sample Library





Outline

- How is it put together
  - Approach to Data Management
  - The Domain Steering Committee
  - **Recommended data types**
- How Does it work
  - Content Models
  - Interchange formats
  - Catalog

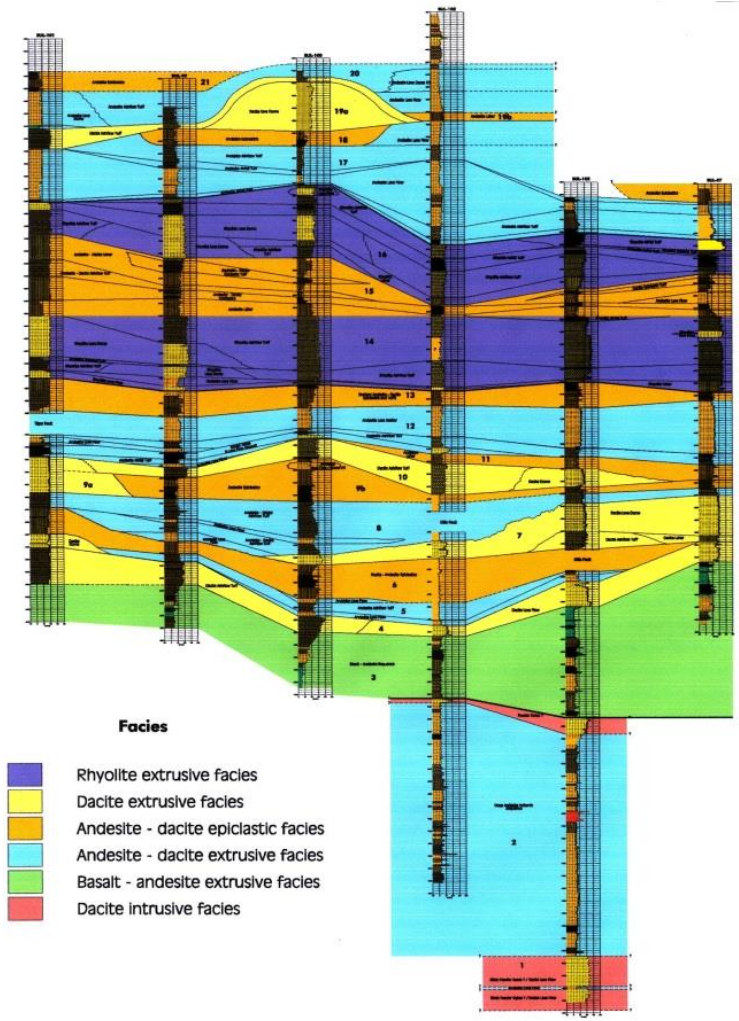
Domain Steering Committee

- What data are important?
- Solution: A ***Domain Steering Committee*** formed to work on data issues!

  - ***Toni Boyd***: Geo-Heat Center Oregon Institute Technology: direct use data - national
  - ***Roland Horne***: Stanford University-Stanford Reservoir Engineering Conference papers and research results
  - ***Joseph Moore***: Energy & Geoscience Institute-legacy resource and research data (**Chair**)
  - ***Lisa Shevenell***: University Nevada, Reno – legacy to recent resource and research data

Domain Steering Committee

- Solicit input from experts
  - Content models
  - User stories – functionality and data output
  - Sustainability – future user input
- Prioritize data types needed by the geothermal community
- Review content models and user stories with DOE invited experts
- Provide recommendations to Project Management

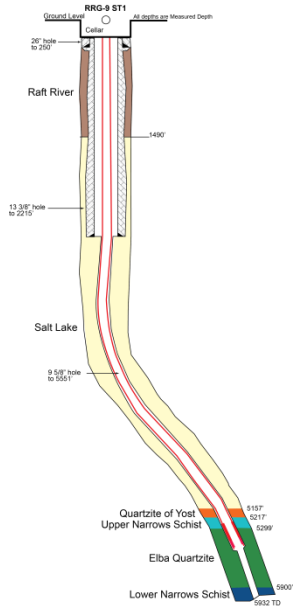


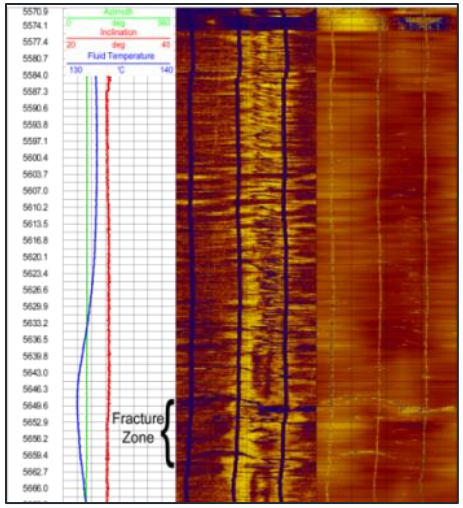
Data type recommendations

- Well Logging
- Geochemistry
- Geophysics
- Land Status and Ownership
- Power Plant History
- Well field Development Costs
- Geothermal system resource models

Wells: data requirements

- Links to related information:
  - Log and well data to geographic location
  - From well to all related data: drilling records, daily reports, well maintenance, costs, sundry notices, regulatory filings, all logs, time lines….
  - Technical papers to well locations
- 3-D geometry of well bore
- History of well (drilling, completion, production, rework, repurpose, abandonment…)
- **Solution**: Interchange formats for well, borehole temperature obs, well log obs, lithology interval obs, fluid production/ injection/disposal, tests…; related resource links

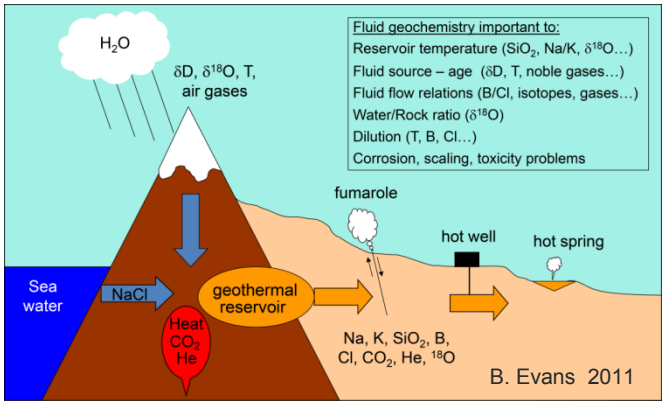




Geothermal fluid characterization

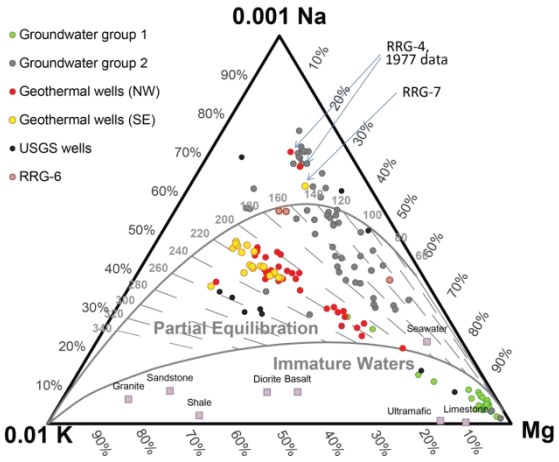
Property:	Approach
Reservoir temperature	Temperature logs
Fluid composition	Aqueous geochemistry obs
Fluid production	Fluid production obs
Fluid origin, age, mixing	No data yet
Direction of fluid movement	No data yet
Steam fraction	Aqueous geochemistry obs

J. Lovekin, GeothermEx, Inc. 2010



Aqueous Geochemistry

- Analytical data for geothermometry, Piper and triangular diagrams:
  - Need time of sampling information
  - Content model based on Powell and Cummings (rev 2010) spreadsheet
- **Solution**: Geochemistry interchange formats
- Access to relational database with exhaustive data compilation: Tier 2 files



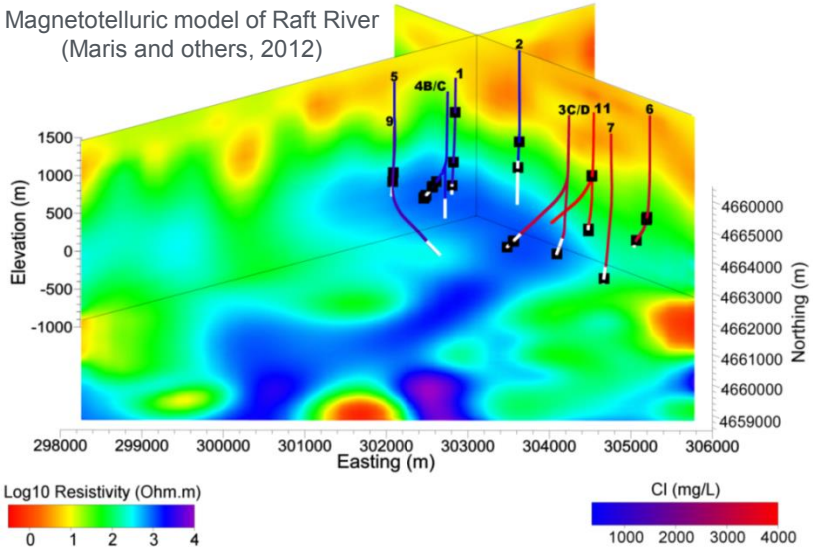
Raft River Chemistry

Cathy Janik with a miniseparator

Geophysical data types

- Gravity, magnetic, resistivity… Continuously varying coverage
- Data model is simple: X,Y,Z,value
- Data represented using numeric arrays (grids).
  Solution: Web Coverage Service, OpenDAP/NetCDF
- Commonly only accessible as an image of a data visualization
- Need to georeference these data visualizations to put in context:
  Solution: Web Map Services

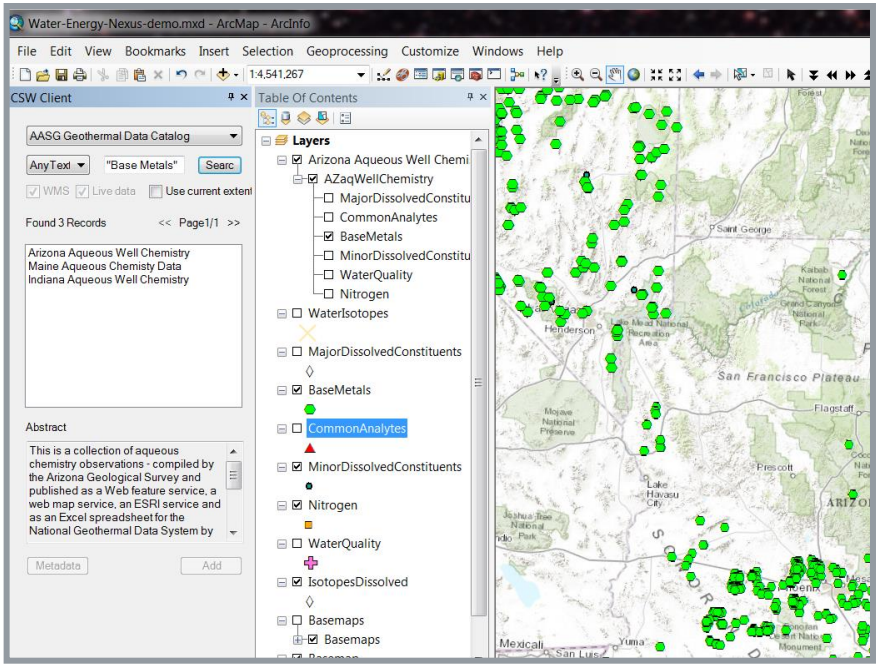
Magnetotelluric model of Raft River (Maris and others, 2012)



Magnetotelluric (MT) Surveys

Other Data

- Geologic Maps
  - Rock unit distribution
  - Location, orientation of contacts and faults
  - **Solution**: GeoSciML feature services and map services
- Infrastructure
  - Land ownership: Parcel maps; under stewardship of various land agencies
  - Power lines, transportation, water…
  - **Solution**: services from other agencies (Western Regional partnership, USGS and others)



Outline

- How is it put together
  - Approach to Data Management
  - The Domain Steering Committee
  - Recommended data types
- **How does it work**
  - Content Models
  - Interchange formats
  - Catalog

What does Interoperability mean?

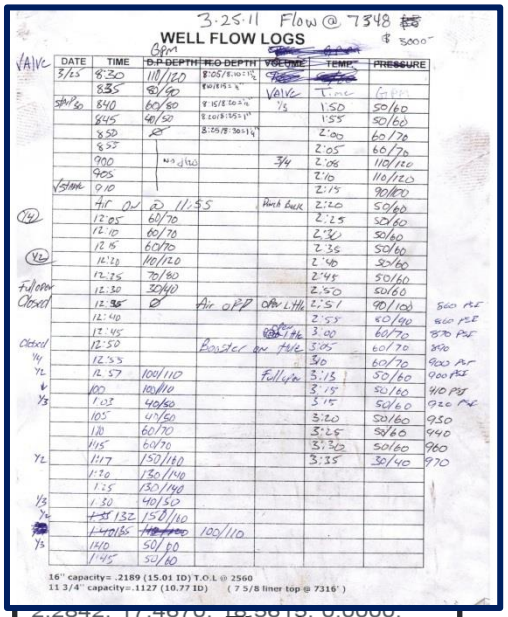
… the ability of diverse systems and organizations to exchange information and use the information that has been exchanged…

To me this means that when I want data:

- I always know how to ask for it
- I always know the format in which I'll receive it
- I always know what it means

Data Interoperability Tiers

- Tier 1: unstructured content in files
  - Text, images, graphics files
  - MS Word, Adobe Acrobat, Illustrator, TIFF….
- Tier 2: Structured data in custom formats
  - May be accessed in files or via services
  - Excel, Microsoft Access, dbf, xml…
- Tier 3: Structured data in standard schema and encoding
  - Delivered via standard service, may also be accessed in files.
  - Excel, Microsoft Access, dbf, xml…



Tier 1

Tier 2

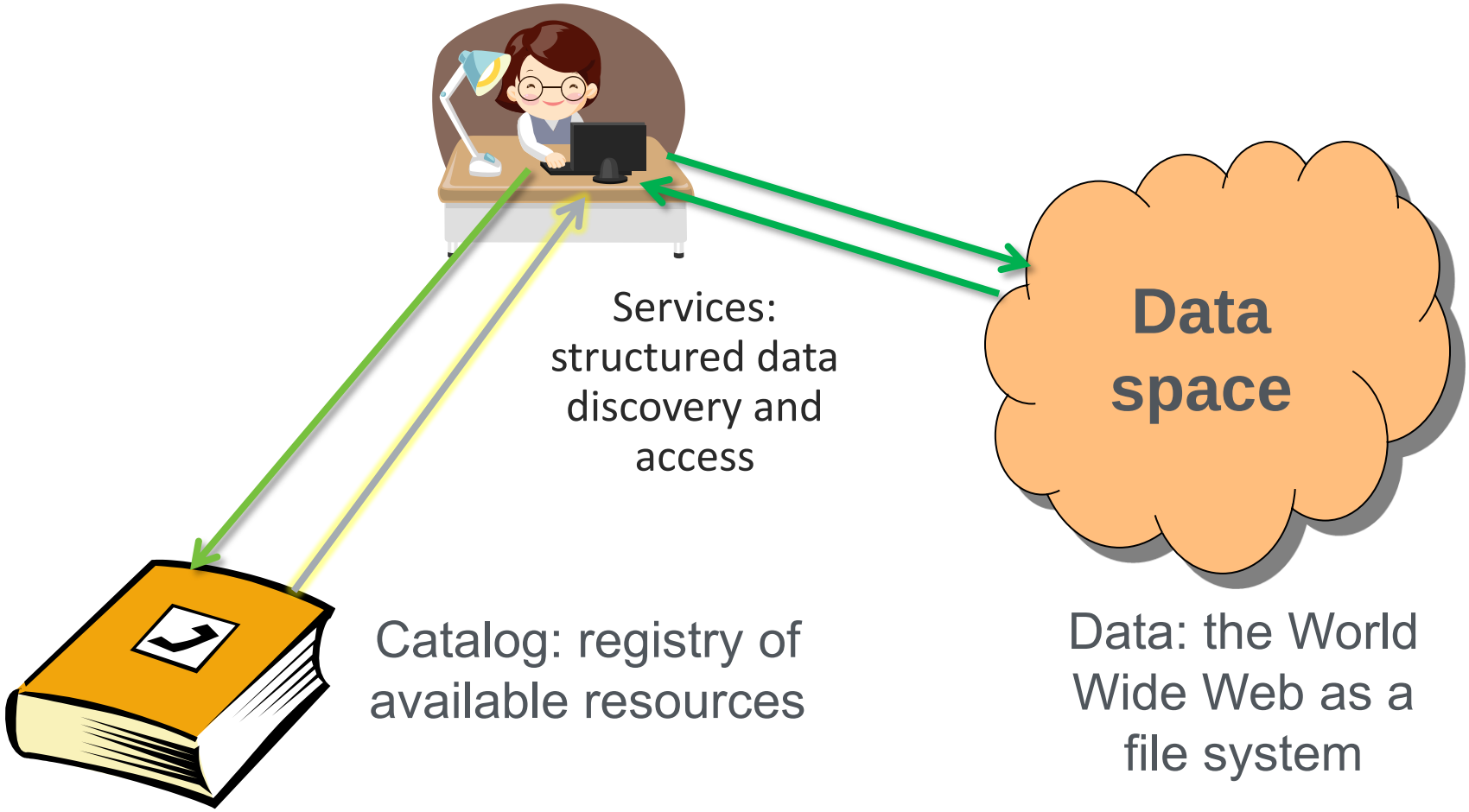
Content Models

- Active Fault/Quaternary Fault
- Aqueous Chemistry
- Borehole Temperature Observation Feature
- Direct Use Feature
- Fault Feature
- Fluid Flux Injection and Disposal
- Geologic Contact Feature
- Geologic Unit Feature
- Geothermal Area
- Geothermal Fluid Production
- Geothermal Power Plant
- Heat Flow
- Heat Pump Facility
- Lithology Interval Log Feature
- Metadata
- Physical Sample
- Powell Cummings Geothermometry
- Power Plant Production
- Radiogenic Heat Production
- Seismic Event Hypocenter
- Thermal Conductivity
- Thermal/Hot Spring Feature
- Volcanic Vents
- Well Fluid Production
- Well Header
- Well Log Observation
- Well Test Observations

Being part of the system

- 1. Any Tier: Resource is described by **metadata** in the system catalog
- Tier 1– file is accessible on the Web
- Tier 2– Ditto, but need more documentation (data format)
- Tier 3 – Accessible on Web via service: provider has to validate format

The Catalog


Services: structured data discovery and access
Data space
Catalog: registry of available resources
Data: the World Wide Web as a file system

Metadata: Strong Conventions

- Minimum metadata content recommendations

Title (1 entry): Succinct (preferably <250 characters) name of the resource.
Description (1 entry): Inform the reader about the resource's content as well as its context.
Originators (1 to many entries): Authors, editors, or corporate authors/curators of the resource.
Publication Date (1 entry): Publication, origination, or update date (not temporal extent) for the resource. Use a "year" or ISO 8601 date and time format. Alternative date formatting must be machine readable and consistent across all datasets. If no publication date is known, estimate the publication date range, enter the oldest year as the publication date, and include the estimated date range in the Description field.
Geographic Extent - *Horizontal* (1 entry, point or minimum bounding rectangle): Values given in decimal degrees using the WGS 84 datum. Required if resource has location Some resources may not be usefully described by an extent; if no extent is specified the default is Earth.
Access Instructions (1 entry): Text description of how to access the resource.
Distribution Contact Party (1 entry): The party (name of organization or person, etc.) to contact about accessing the resource.
Distribution Contact Email (1 entry): How to contact the party responsible for distribution
Metadata Date (1 entry): Last metadata update/creation date-time stamp in ISO 8601 date and time format. This may be automatically updated on metadata import if a metadata format conversion is necessary.
Metadata Contact Party (1 entry): The party (name of organization or person, etc.) to contact with questions about the metadata itself
Metadata Contact Email (1 entry): How to contact the party responsible for metadata content and accuracy
Metadata Specification (1 entry): Identifier for metadata specification used to create a metadata record encoding this content.

- Standard interchange format: USGIN Profile for ISO 19139 metadata (XML)

```
▼<gmd:MD_Metadata xmlns:gmd="http://www.isotc211.org/2005/gmd"
 xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:gml="http://www.opengis.net/gml"
 xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instan
 xsi:schemaLocation="http://www.isotc211.org/2005/gmd
 http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd">
  ▼<gmd:fileIdentifier>
     <gco:CharacterString>3c529816-a0f4-11e0-8768-0024e880c1d2</gco:CharacterString>
   </gmd:fileIdentifier>
  ▼<gmd:language>
     <gco:CharacterString>eng</gco:CharacterString>
   </gmd:language>
  ▼<gmd:characterSet>
     <gmd:MD_CharacterSetCode
     codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resou
     codeListValue="utf8">utf8</gmd:MD_CharacterSetCode>
   </gmd:characterSet>
  ▼<gmd:hierarchyLevel>
     <gmd:MD_ScopeCode
     codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resou
     codeListValue="dataset">dataset</gmd:MD_ScopeCode>
   </gmd:hierarchyLevel>
  ▼<gmd:hierarchyLevelName>
     <gco:CharacterString>Dataset</gco:CharacterString>
   </gmd:hierarchyLevelName>
```

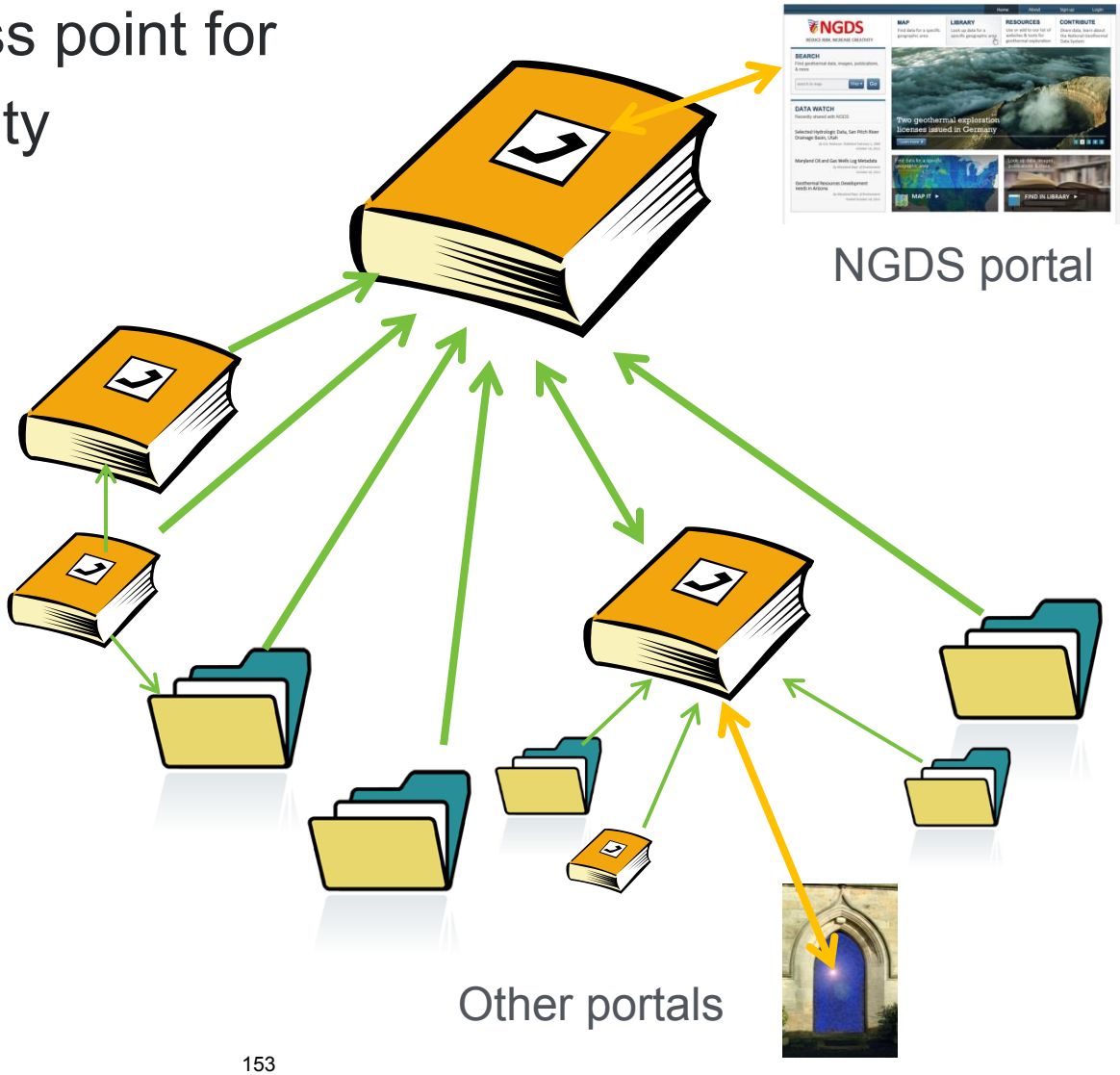
Metadata Content

- Title
- Description
- Extent
  - Geographic
  - Vertical
  - Temporal
- Access instructions
- Keywords
- Originator(s)
- Date
- Resource ID
- Access constraints
- Language
- Quality
- Lineage
- Citation
- Distribution contact
- Metadata
  - Date
  - Contact
  - Specification
  - Identifier

Online document: http://lab.usgin.org/profiles/doc/metadata-content-recommendations

Metadata cloud

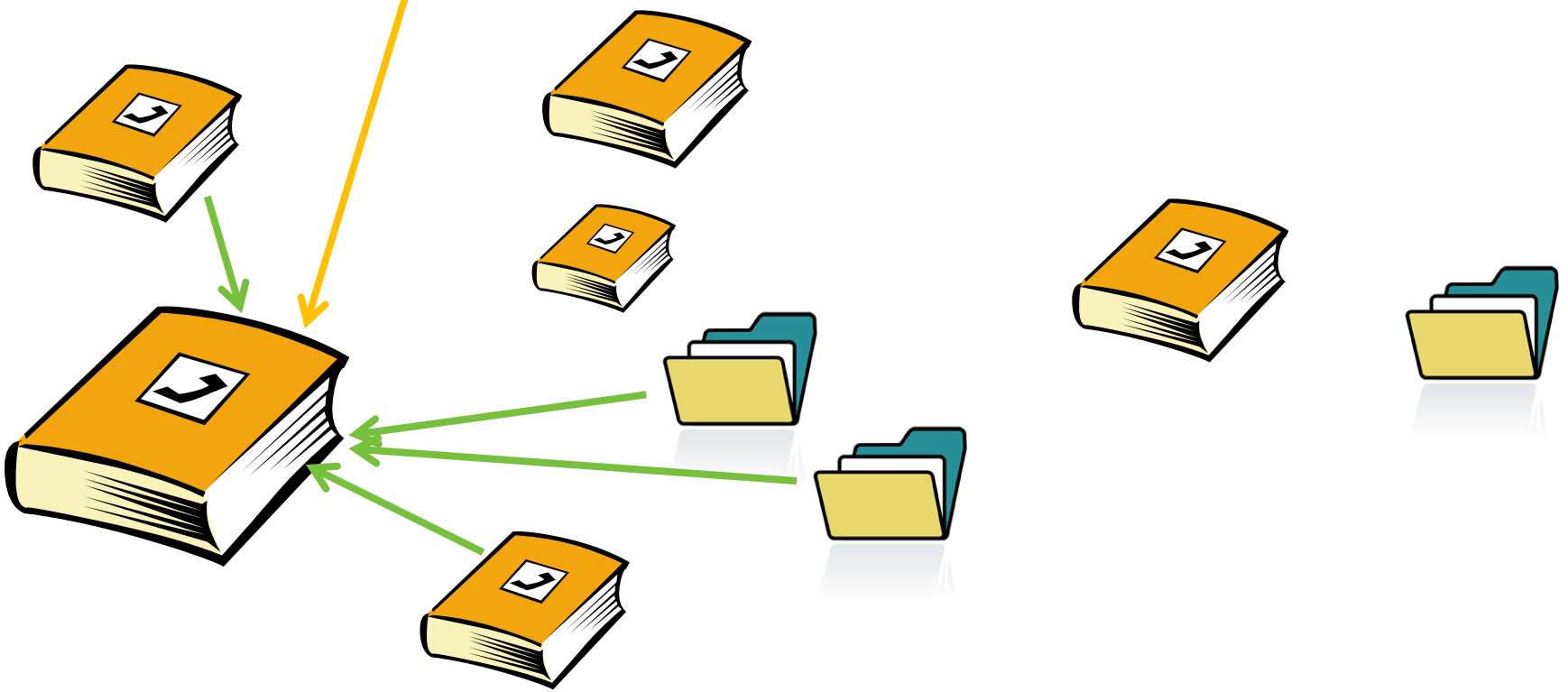
- NGDS portal – access point for geothermal community
- NGDS aggregator: collect metadata for all NGDS resources
- Metadata contributors



Other Portals

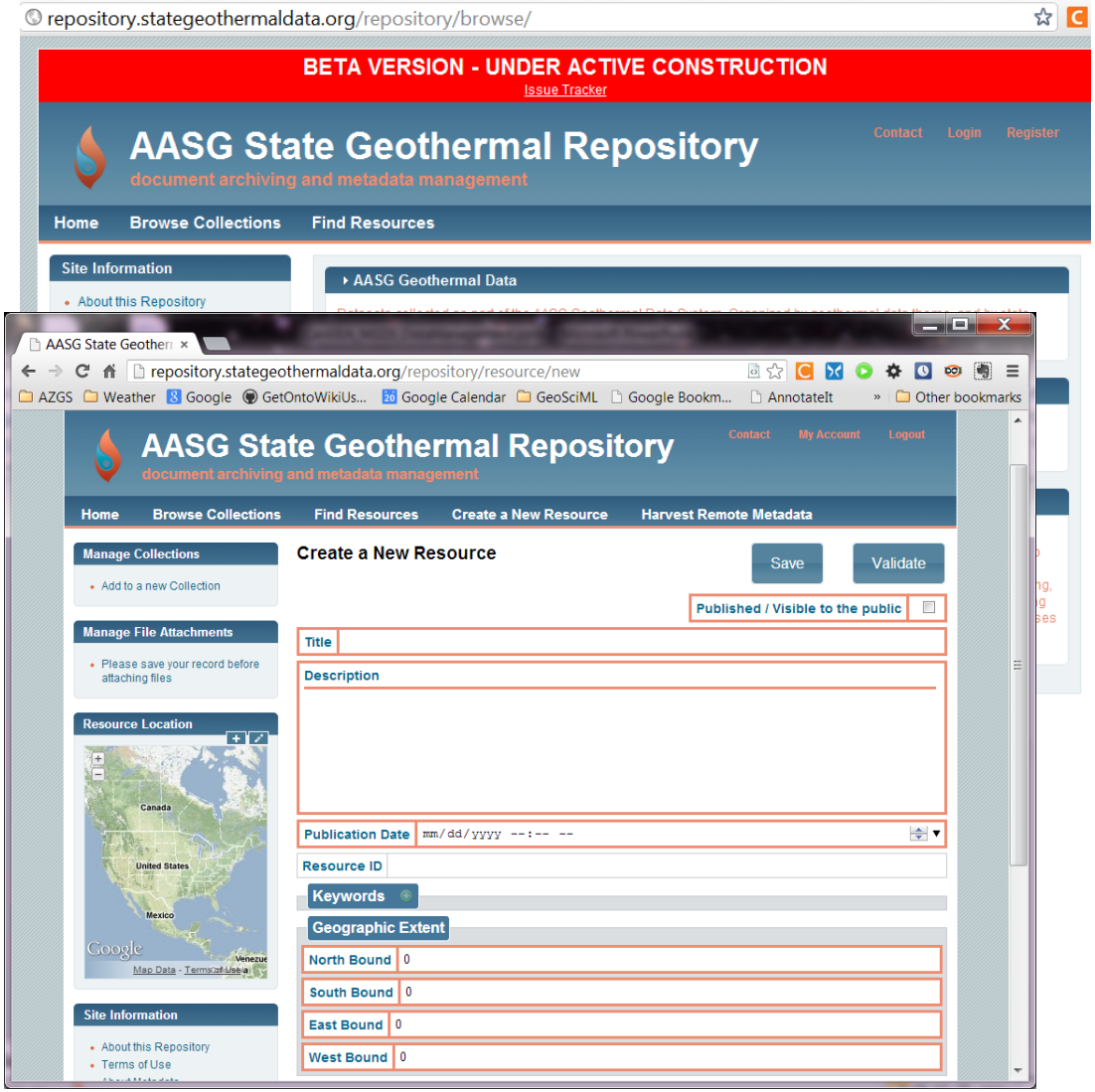
CanGEA
CANADIAN GEOTHERMAL ENERGY ASSOCIATION

CanGEA portal

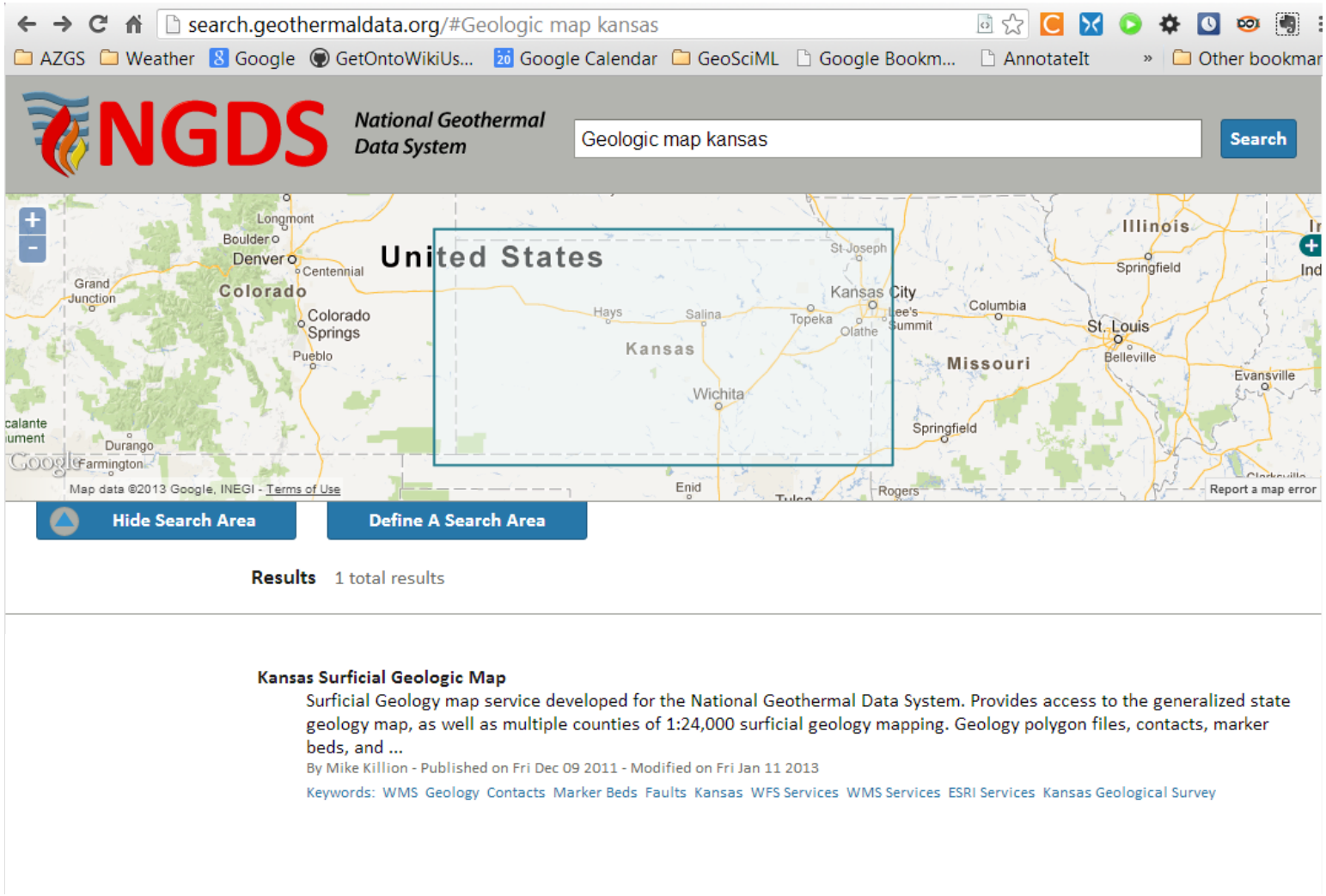


Metadata Repository Tool

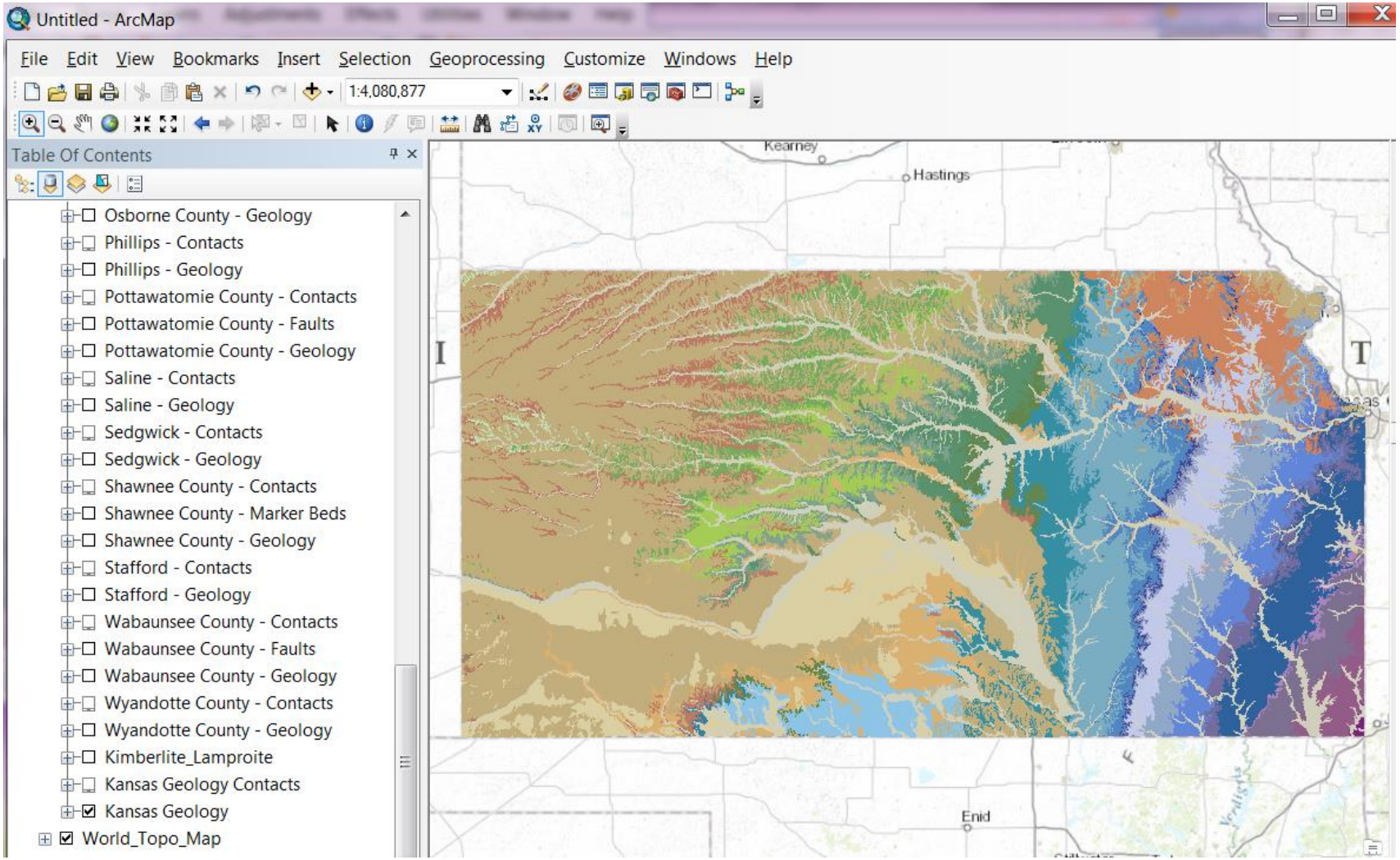
- Generate and manage metadata
- Registry of resources
- Upload capability
- Django/NodeJS implementation



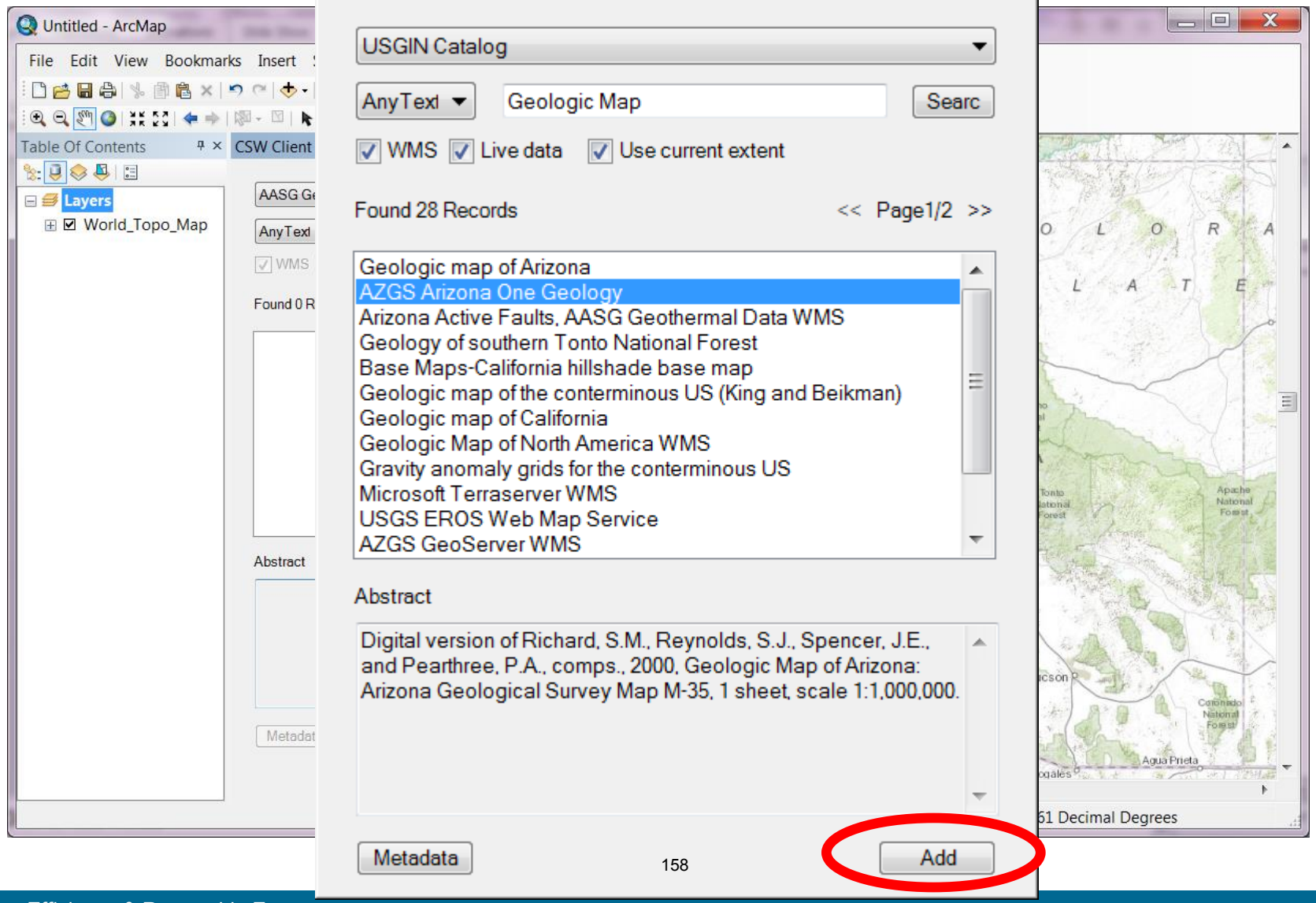
Search Client



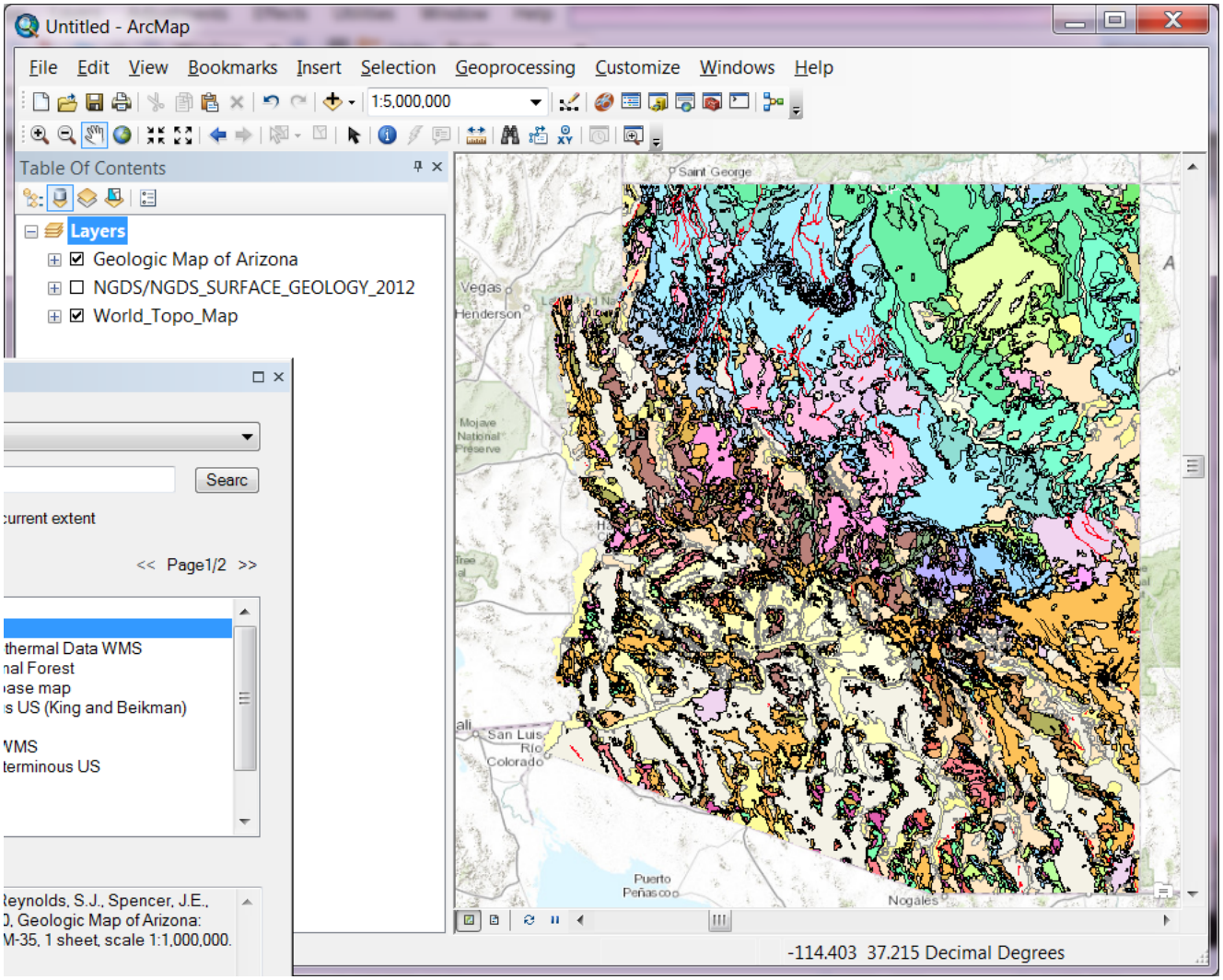
Geologic Map



ArcGIS CSW Client

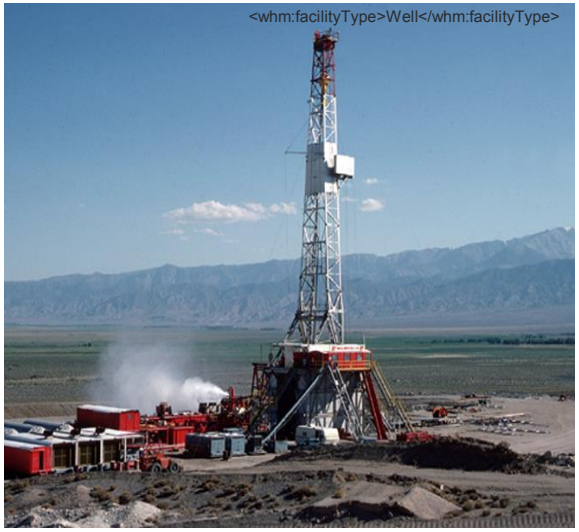


Display map in ArcMap



Tier 3: Interchange Formats

- Conventions for how information will be encoded to send between computers
  - Designed for consumption by computer software
- Encoding formats
  - HTML, XML, JSON, NetCDF, TIFF, JPG…
- Content Model:
  - Defines the information content
  - Data schema
  - May be implemented with various encoding formats

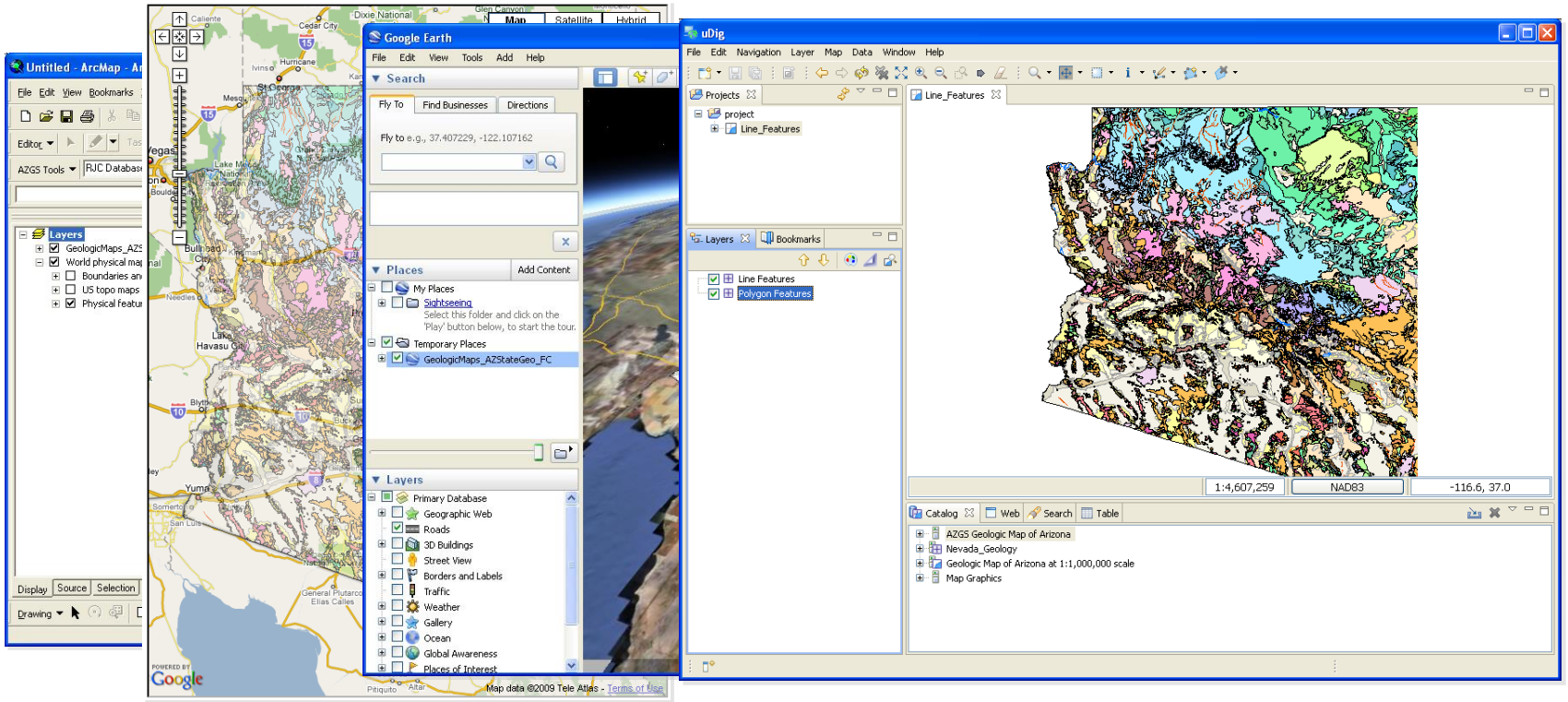


Content Models (Tier 3)

- Active Fault/Quaternary Fault
- Aqueous Chemistry
- Borehole Temperature Observation Feature
- Direct Use Feature
- Fault Feature
- Fluid Flux Injection and Disposal
- Geologic Contact Feature
- Geologic Unit Feature
- Geothermal Area
- Geothermal Fluid Production
- Geothermal Power Plant
- Heat Flow
- Heat Pump Facility
- Lithology Interval Log Feature
- Metadata
- Physical Sample
- Powell Cummings Geothermometry
- Power Plant Production
- Radiogenic Heat Production
- Seismic Event Hypocenter
- Thermal Conductivity
- Thermal/Hot Spring Feature
- Volcanic Vents
- Well Fluid Production
- Well Header
- Well Log Observation
- Well Test Observations

OGC Services -- Useful in Existing Applications

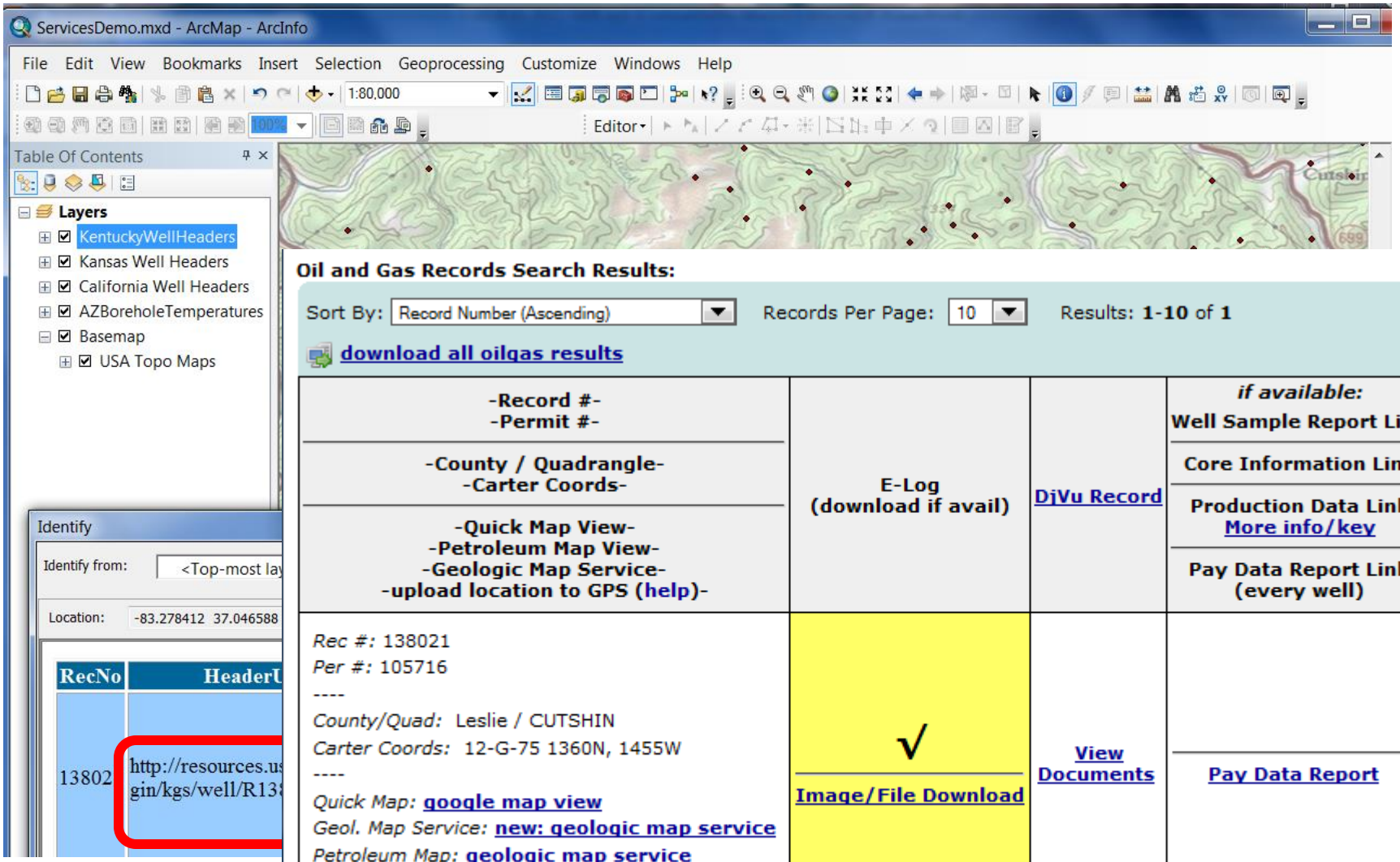
- WMS for providing a symbolized portrayal of vector data
- WFS for full access to attributes and to download vector data
- WCS for access to continuous raster data

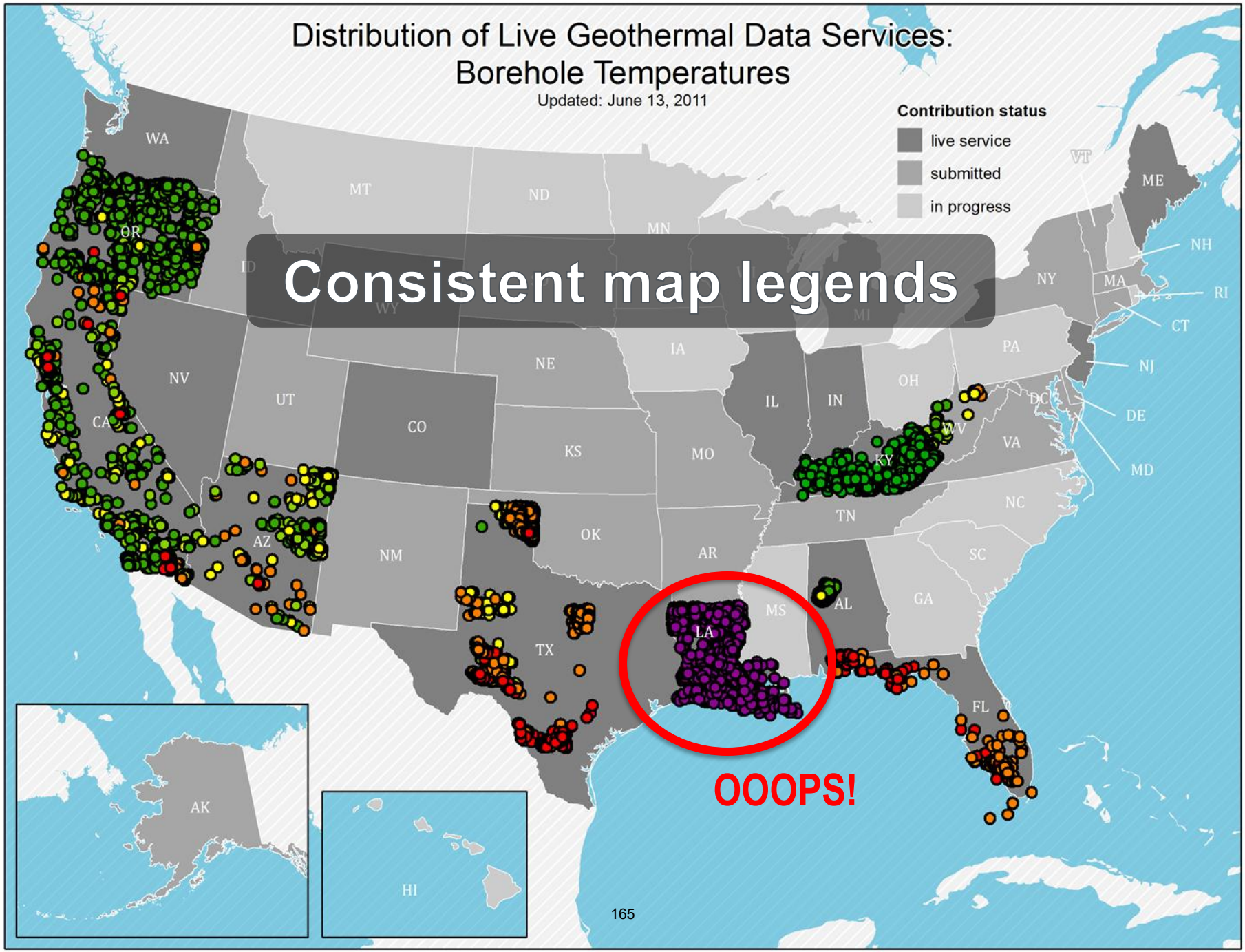


Protocol: Web Map Service (WMS)

- Simple, widely used, many clients
- Access to georeferenced map image
- Servers
  - MapServer, GeoServer, Deegree, ArcGIS
- Data integration = agreeing on standard portrayal schemes (legends)

WMS for well headers



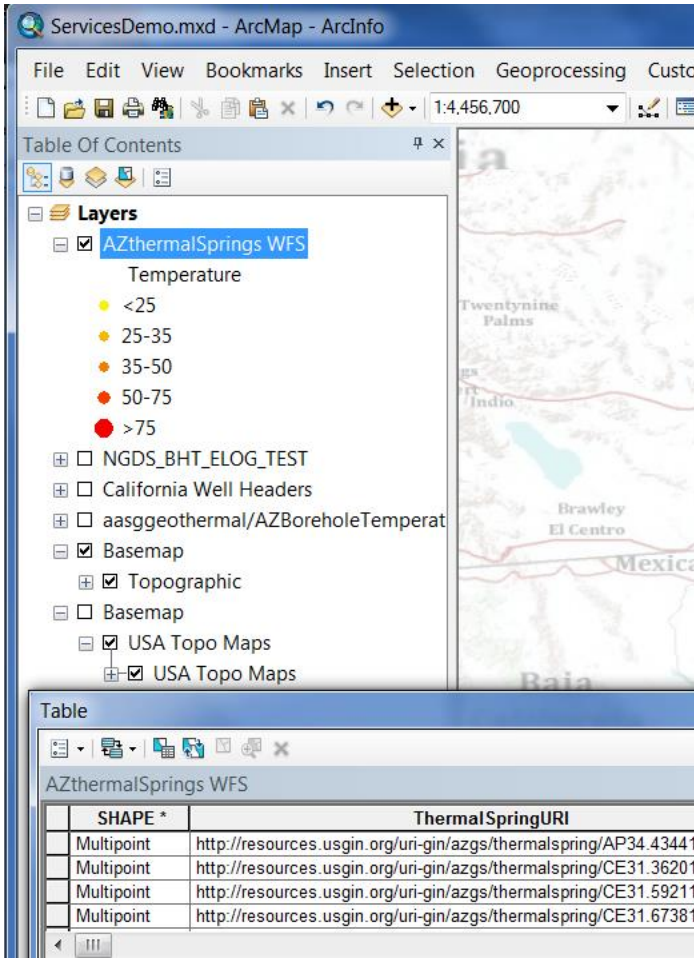


Project sponsored by the U.S. Department of Energy under award DE-EE0002850 to the Arizona Geological Survey acting on behalf of the Association of American State Geologists. Cartography by JCD.

Protocol – Web Feature Service (WFS)

- Open Geospatial Consortium Web Feature Service
  - Standard, structured encoding of content
  - Each feature defines content model (xml schema)
- Feature service –
  - designed for data access based on features (representation of geolocated entity with attributes)

WFS example



services.azgs.az.gov/arcgis/services/aasggeothermal/AZThermalSprings/MapServer/WFS

```
<wfs:FeatureCollection xsi:schemaLocation="http://stategeothermaldata.org/uri-gin/aasg/xmlschema/simple
http://services.azgs.az.gov/arcgis/services/aasggeothermal/AZThermalSprings/MapServer/WFSServer?request
http://www.opengis.net/wfs http://schemas.opengis.net/wfs/1.1.0/wfs.xsd" xmlns:aasg="http://stategeothe
xmlns:gml="http://www.opengis.net/gml" xmlns:wfs="http://www.opengis.net/wfs" xmlns:xsi="http://www.w3.
  <gml:boundedBy>
    <gml:Envelope srsName="urn:ogc:def:crs:EPSG:6.9:4326">
      <gml:lowerCorner>31.361999999999998 -114.745</gml:lowerCorner>
      <gml:upperCorner>36.899500000000003 -109.1289</gml:upperCorner>
    </gml:Envelope>
  </gml:boundedBy>
  <gml:featureMember>
    <aasg:thermalSprings gml:id="F2__1">
      <aasg:OBJECTID>1</aasg:OBJECTID>
      <aasg:ThermalSpringURI>http://resources.usgin.org/uri-gin/azgs/thermalspring/AP34.4344109.3986/
      <aasg:Label>Salado Springs: 22C</aasg:Label>
      <aasg:Name>Salado Springs</aasg:Name>
      <aasg:Source>Waring, Gerald A., 1965. Thermal Springs of the United States and Other Countries o
      y.USGS  Prof Paper 492, 407 pp. | Berry, George W. , Grim, Paul J.  and lkelman, Joy A.  June 19
      d Atmospheric Administration Key to Geophysical Records Documentation No. 12.</aasg:Source>
      <aasg:FeatureType>warm spring</aasg:FeatureType>
      <aasg:SourceURI>http://www.ngdc.noaa.gov/hazard/data/pp492.pdf</aasg:SourceURI>
      <aasg:LandLeaseOwner>unknown</aasg:LandLeaseOwner>
      <aasg:County>Apache</aasg:County>
      <aasg:OtherLocationName>AMS SAINT JOHNS | USGS SALADO 7.5</aasg:OtherLocationName>
      <aasg:PLSS_Meridians>Gila and Salt River Meridian</aasg:PLSS_Meridians>
      <aasg:State>Arizona</aasg:State>
      <aasg:Range>28E</aasg:Range>
      <aasg:Township>12N</aasg:Township>
      <aasg:Section_>17</aasg:Section_>
      <aasg:LatDegree>34.434399999999997</aasg:LatDegree>
      <aasg:LongDegree>-109.3986</aasg:LongDegree>
      <aasg:LocationUncertaintyStatement>Group of springs on topo</aasg:LocationUncertaintyStatement>
      <aasg:Temperature>21.850000000000001</aasg:Temperature>
      <aasg:TemperatureUnits>C</aasg:TemperatureUnits>
      <aasg:TempMeasurementProcedure>unknown</aasg:TempMeasurementProcedure>
      <aasg:Flow>2</aasg:Flow>
      <aasg:FlowMeasurementProcedure>unknown</aasg:FlowMeasurementProcedure>
      <aasg:FlowUnits>GPM</aasg:FlowUnits>
      <aasg:FlowMeasurementDateTime>unknown</aasg:FlowMeasurementDateTime>
      <aasg:MeasurementSource>unknown</aasg:MeasurementSource>
      <aasg:Classification>6</aasg:Classification>
      <aasg:RelatedWaterChemistry>unknown</aasg:RelatedWaterChemistry>
```

Conclusions

- NGDS is a data system, not a database
- The World Wide Web is the file system and data repository
- Resources include:
  - Structured and unstructured data in files (most of the content)
  - Standardized structured data accessible through web services (metadata and some data)
- Standardized information exchanges:
  - Well headers, borehole temperatures, well logs, seismic event hypocenters, borehole lithology intervals, heat pump facilities, aqueous chemistry, thermal conductivity, well flow tests…

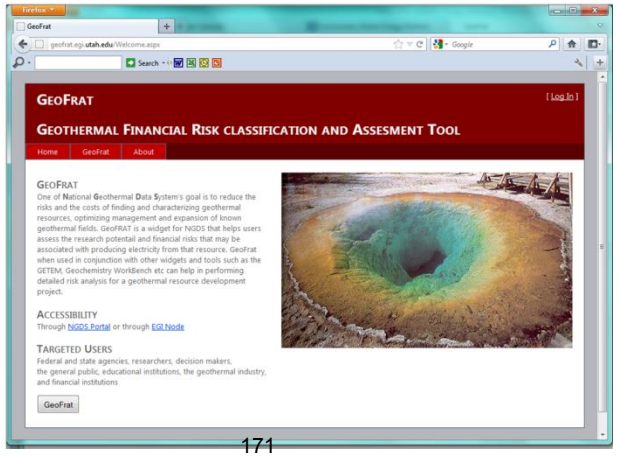
Thank You

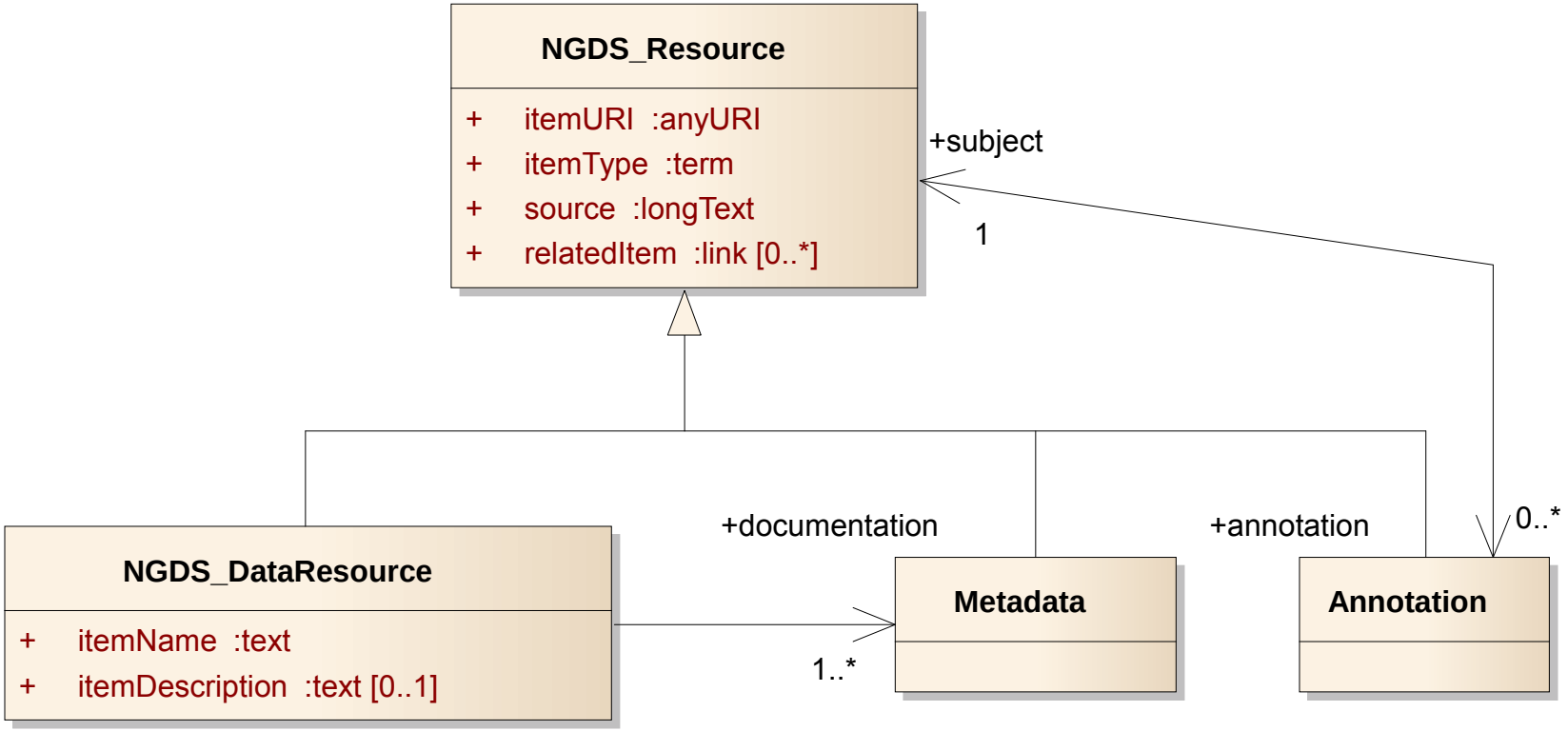
Intentionally blank

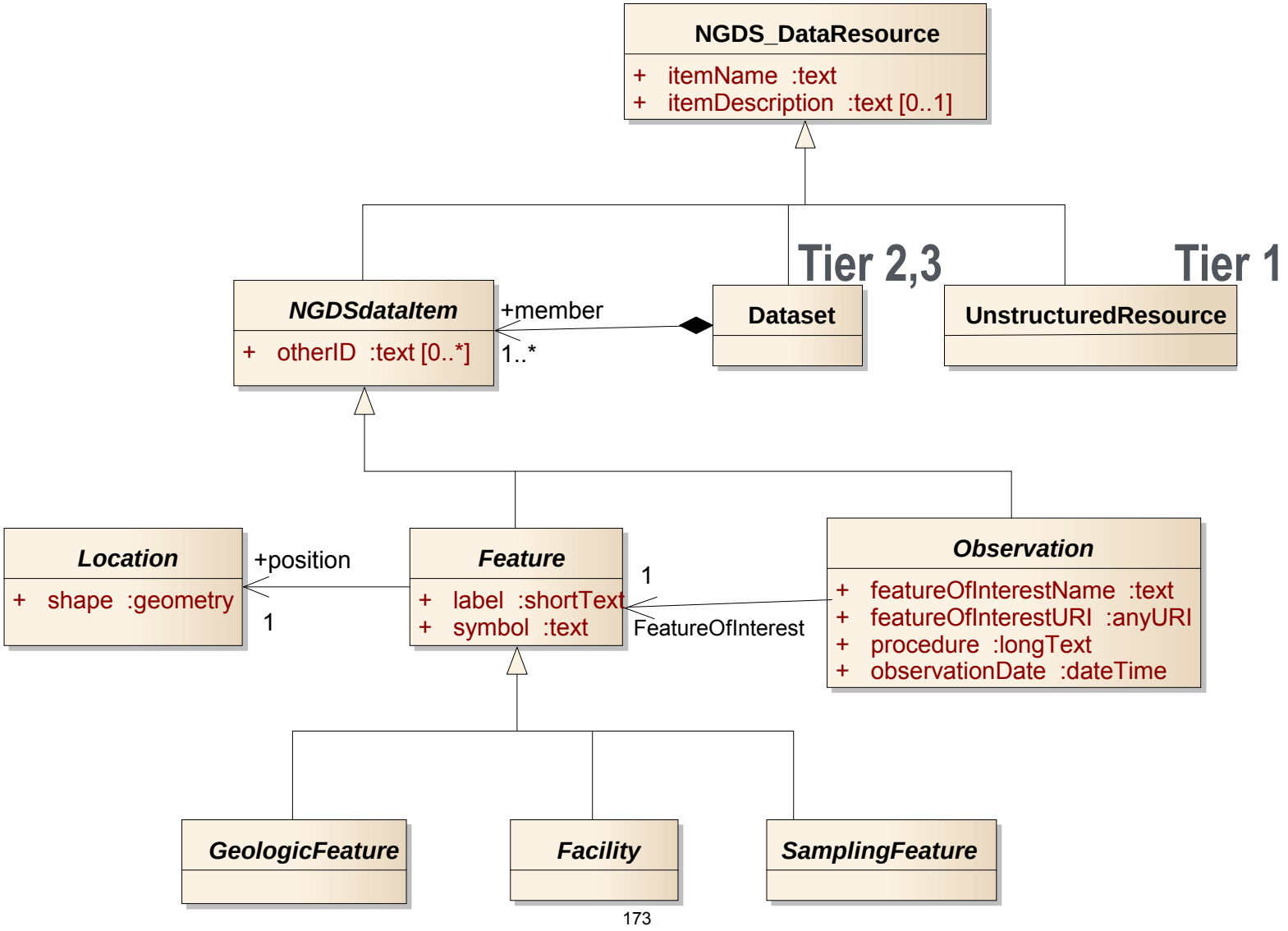
- Begin extra slides

Approach to Data Management

- Use WWW infrastructure, Open Geospatial Consortium (OGC) services, ISO standards
- Simple interchange formats
- Tiers of interoperability

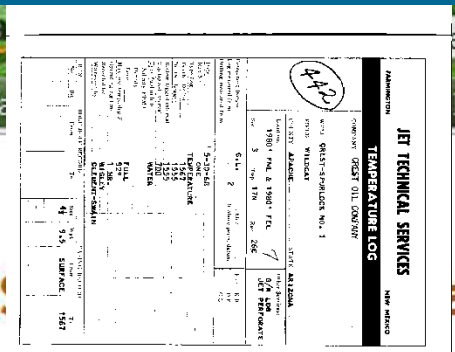


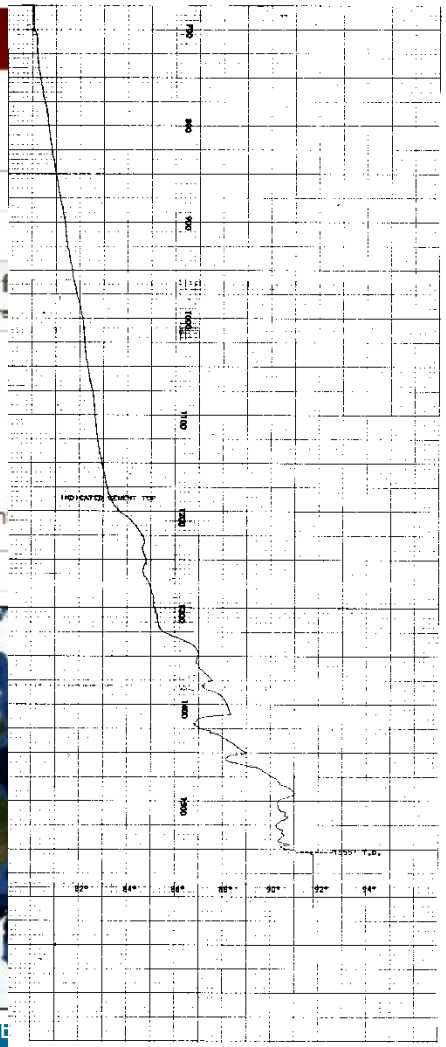

NGDS_Resource
+ itemURI :anyURI
+ itemType :term
+ source :longText
+ relatedItem :link [0..*]
+subject
1
NGDS_DataResource
+ itemName :text
+ itemDescription :text [0..1]
+documentation
1..*
Metadata
+annotation
0..*
Annotation


NGDS_DataResource
+ itemName :text
+ itemDescription :text [0..1]
Tier 2,3
Tier 1
NGDSdataItem
+ otherID :text [0..*]
+member
1..*
Dataset
UnstructuredResource
Location
+ shape :geometry
+position
1
Feature
+ label :shortText
+ symbol :text
1
FeatureOfInterest
Observation
+ featureOfInterestName :text
+ featureOfInterestURI :anyURI
+ procedure :longText
+ observationDate :dateTime
GeologicFeature
Facility
SamplingFeature

Catalog system issues

- How to manage metadata
  - Various streams for importing
  - Management system based on conceptual model
- Metadata for machines– connect user to services inside of their user environment
  - Requires machine-actionable links
- Document and file searching should be done with standard search engines





eoportal/catalog/main/home.page

aasggeo... | EarthSc... | http...WFS | News an... | USG...

Login Register Help About Feedback

USGIN AA ata Catalog

AASG Association of American State Geologists

USGIN

HOME SEARCH

Search

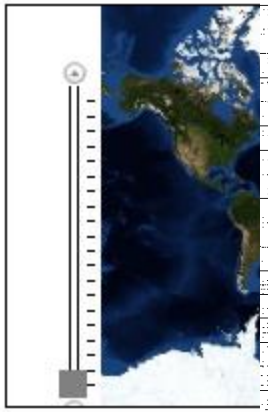
"temperature log"

Records shown f
Click here to select differe

Additional Options
Clear

WHERE

Anywhere In



Results 1-10 of 142 record(s) 1 2 3 4 5 > Last

Expand results Zoom To Results Zoom To Searched Area

Scanned Temperature Log for Hallett Development 1 State

Scanned Temperature Log for Merrion 1 Navajo-E

Scanned Temperature Log for Arkla Exploration 64 State

Scanned Temperature Log for Crest Oil 1 Spurlock

Tiff image file, scanned with NeuraLog scanner. Originator organization is the name of the company that ran the log where that is available. Includes log run 1. Logged by Jet Technical Services. Formation at TD is Precambrian granite.

Open Preview Details Metadata Zoom To

Scanned Temperature Log for AmeriGas 3 Roach-Baker

Scanned Temperature Log for AmeriGas 3 Roach-Baker

Scanned Temperature Log for AmeriGas 3 Roach-Baker

Scanned Temperature Log for AmeriGas 2 Roach-Baker

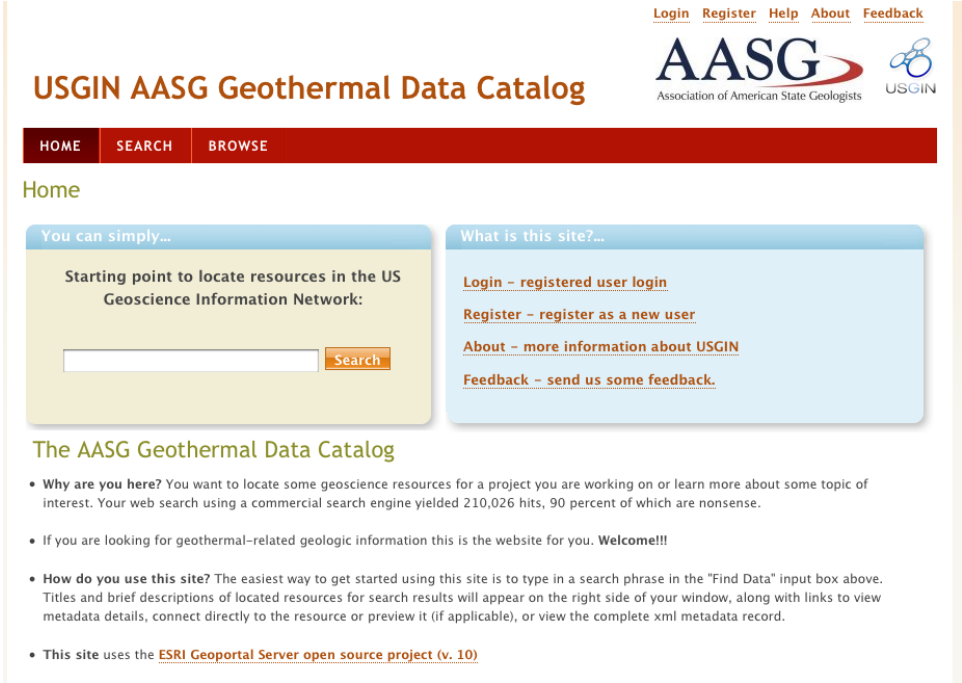
Scanned Temperature Log for Phillips Petroleum A1 State

Domain Steering Committee Members

- ***Toni Boyd***: Geo-Heat Center Oregon Institute Technology: direct use data - national
- ***Roland Horne***: Stanford University-Stanford Reservoir Engineering Conference papers and research results
- ***Joseph Moore***: Energy & Geoscience Institute-legacy resource and research data (**Chair**)
- ***Lisa Shevenell***: University Nevada, Reno – legacy to recent resource and research data

Search interfaces

- Catalog -- for data discovery
- Collection of consistent metadata describing network resources
- Tested various catalog service implementations;
- Current interface is ESRI Geoportal



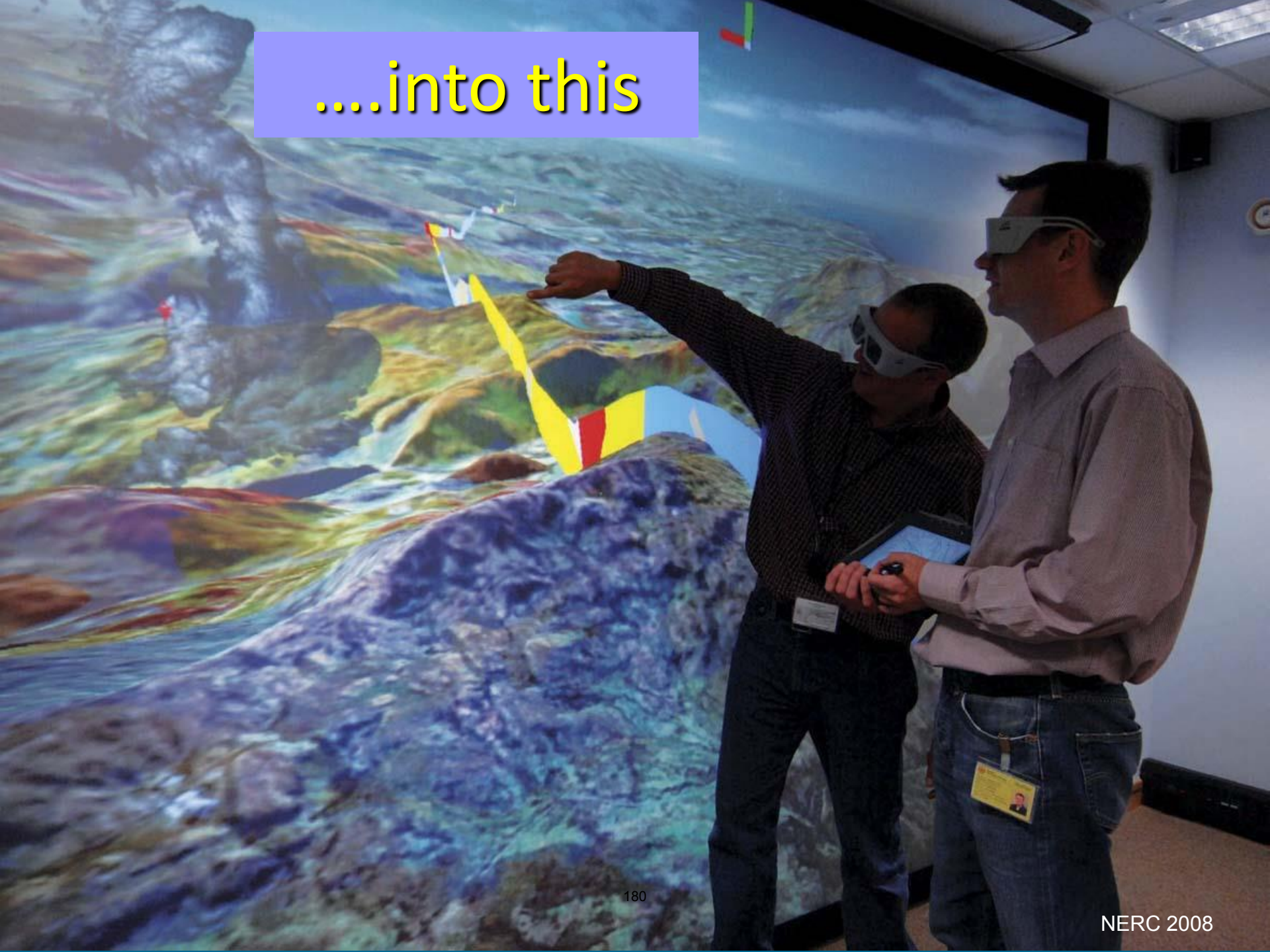
(http://catalog.geothermaldata.org/geoportal)

AASG Geothermal data

- Catalog file-based resources in repositories
- Simple feature (flat file) schema
  - 27 content models
- More on the way –
- http://geothermaldata.org

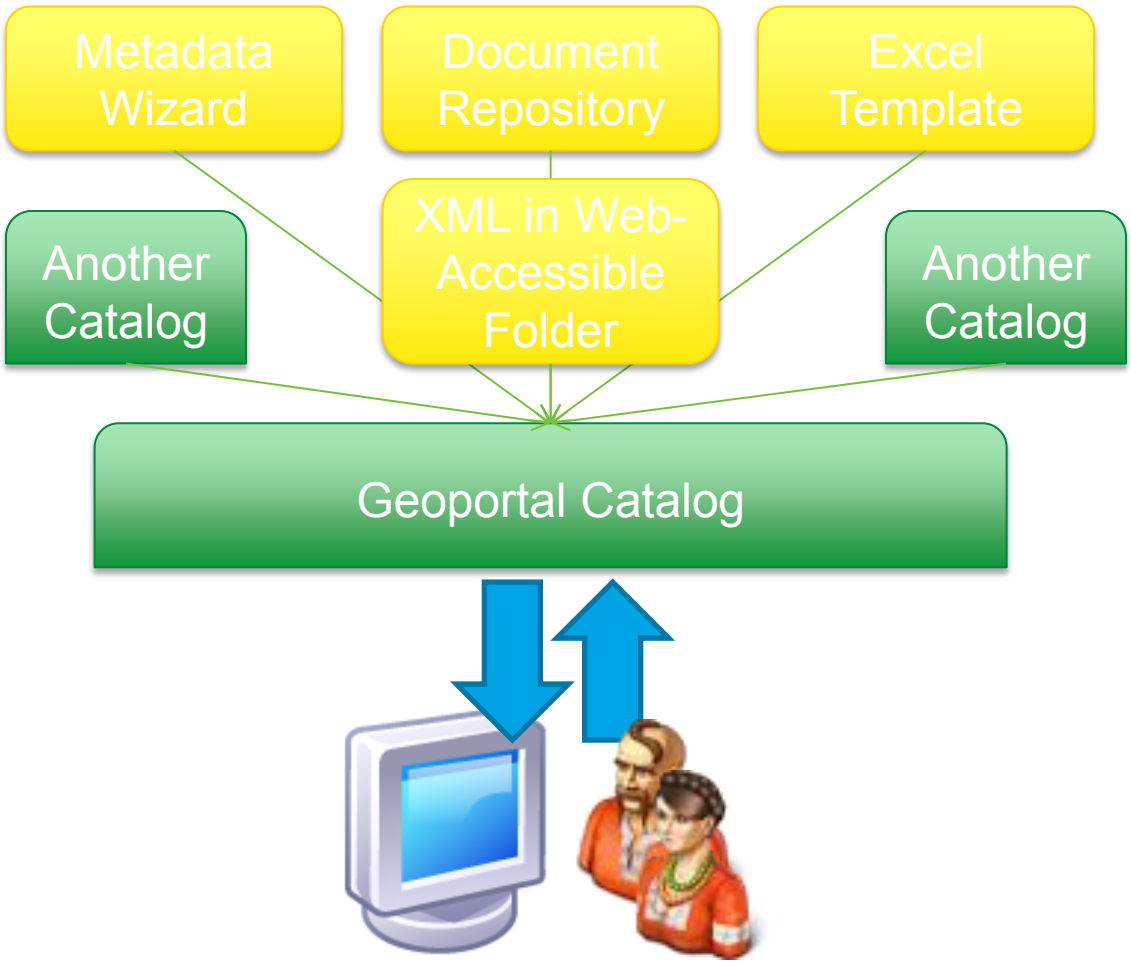
The goal is simple, really….

turn this….


….into this

Why a Geoportal Catalog?

- Allow a variety of metadata creation methods to be aggregated
- Allow distributed metadata records to be aggregated: harvesting between catalogs
- Provide a consistent interface for searching and retrieving metadata records: CSW
- Why ESRI's Geoportal Server?
  - Geonetwork was too hard
  - They made it open source!



Geophysical Research Abstracts
Vol. 16, EGU2014-15742, 2014
EGU General Assembly 2014
© Author(s) 2014. CC Attribution 3.0 License.

National Geothermal Data System (USA): an Exemplar of Open Access to Data

M Lee Allison (1), Stephen Richard (1), Harold Blackman (2), Arlene Anderson (3), and Kim Patten (1)

(1) Arizona Geological Survey, Tucson, United States (lee.allison@azgs.az.gov), (2) Boise State University, Boise, United States (haroldblackman@boisestate.edu, (3) U.S. Department of Energy, Washington, United States (arlene.anderson@ee.doe.gov

The National Geothermal Data System's (NGDS – www.geothermaldata.org) formal launch in April, 2014 will provide open access to millions of data records, sharing -relevant geoscience and longer term to land use data to propel geothermal development and production. NGDS serves information from all of the U.S. Department of Energy's sponsored development and research projects and geologic data from all 50 states, using free and open source software. This interactive online system is opening new exploration opportunities and potentially shortening project development by making data easily discoverable, accessible, and interoperable.

We continue to populate our prototype functional data system with multiple data nodes and nationwide data online and available to the public. Data from state geological surveys and partners includes more than 6 million records online, including 1.72 million well headers (oil and gas, water, geothermal), 670,000 well logs, and 497,000 borehole temperatures and is growing rapidly. There are over 312 interoperable Web services and another 106 WMS (Web Map Services) registered in the system as of January, 2014. Companion projects run by Southern Methodist University and U.S. Geological Survey (USGS) are adding millions of additional data records. The DOE Geothermal Data Repository, currently hosted on OpenEI, is a system node and clearinghouse for data from hundreds of U.S. DOE-funded geothermal projects.

NGDS is built on the US Geoscience Information Network (USGIN) data integration framework, which is a joint undertaking of the USGS and the Association of American State Geologists (AASG). NGDS complies with the White House Executive Order of May 2013, requiring all federal agencies to make their data holdings publicly accessible online in open source, interoperable formats with common core and extensible metadata.

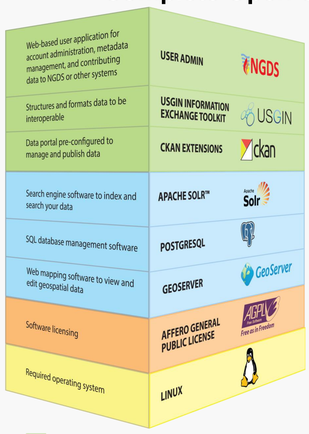
The National Geothermal Data System is being designed, built, deployed, and populated primarily with support from the US Department of Energy, Geothermal Technologies Office. To keep this system operational after the original implementation will require four core elements: continued serving of data and applications by providers; maintenance of system operations; a governance structure; and an effective business model. Each of these presents a number of challenges currently under consideration.

National Geothermal Data System:
Case Studies on Exploration and Development of Potential Geothermal Sites Through Distributed Data Sharing

Arlene Anderson, Lee Allison, Steve Richard, Christy Caudill Daugherty, and Kim Patten

The NGDS rolled out version 1 of the system on April 30, 2014 using the US Geoscience Information Network (USGIN) as its data integration platform. Currently, the NGDS features a comprehensive user interface for searching and accessing nearly 39,000 documents and more than 9 million data points shared by scores of data providers across the U.S. The NGDS supports distributed data sharing, permitting the data owners to maintain the raw data that is made available to the consumer.

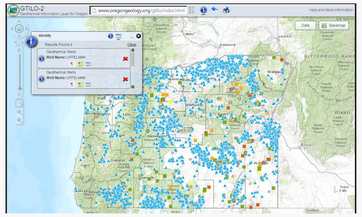
USGIN Stack for NGDS: Complete Open-Source Solution

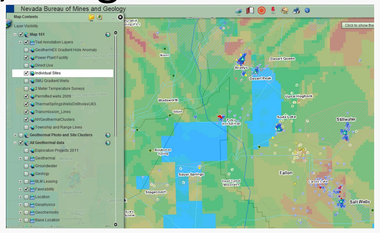


- Data and metadata management tools
- Out-of-the-box functionality
- User interfaces for:
  - Publishing documents and reports on the web
  - Creating and publishing GIS data and map web services on the web

USGIN
Open & Interoperable Data Solutions
usgin.org

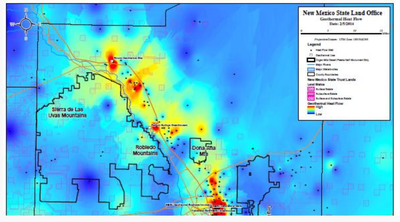
Nevada, Oregon



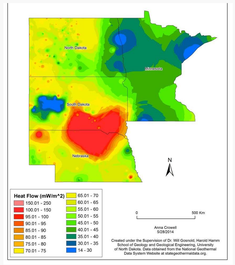


Map applications for data access

Updated heat flow maps generated using NGDS data



Doña Ana County, New Mexico

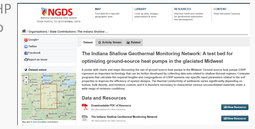


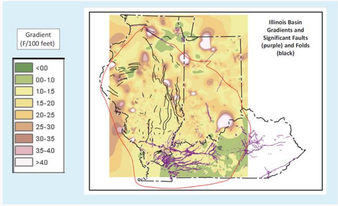
North Dakota, Minnesota, and Nebraska

Indiana

Monitoring in-situ measurements for six sites
- shallow subsurface thermal conductivity
- temperature gradients, 0 - 6 feet depth
- soil moisture

Thermal data from the NGDS will support the design of more efficient GSHP systems tailored to specific geological conditions

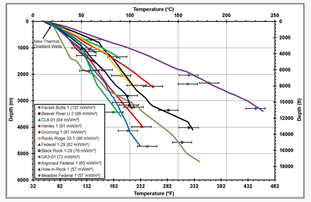




Geothermal gradient analysis from 26,000+ wells used to parameterize reservoir temperatures

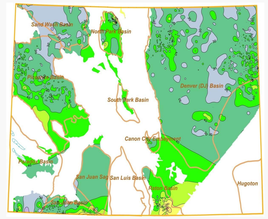
Utah

Black Rock Basin:
- Confirmation of high temperature geothermal resource at depth in a deep sedimentary basin
- Seven of the nine drill holes were funded by the U.S. Department of Energy as part of the NGDS project.
- Greentech Media listed this finding as one of the top geothermal events of 2012.



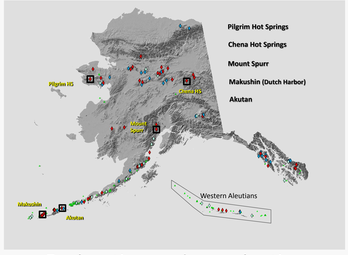
Computed temperature-depth profiles in Black Rock Desert based on borehole temperature data

Colorado



Temperature gradients derived from borehole temperature data

Alaska



Exploration and reevaluation underway at five sites

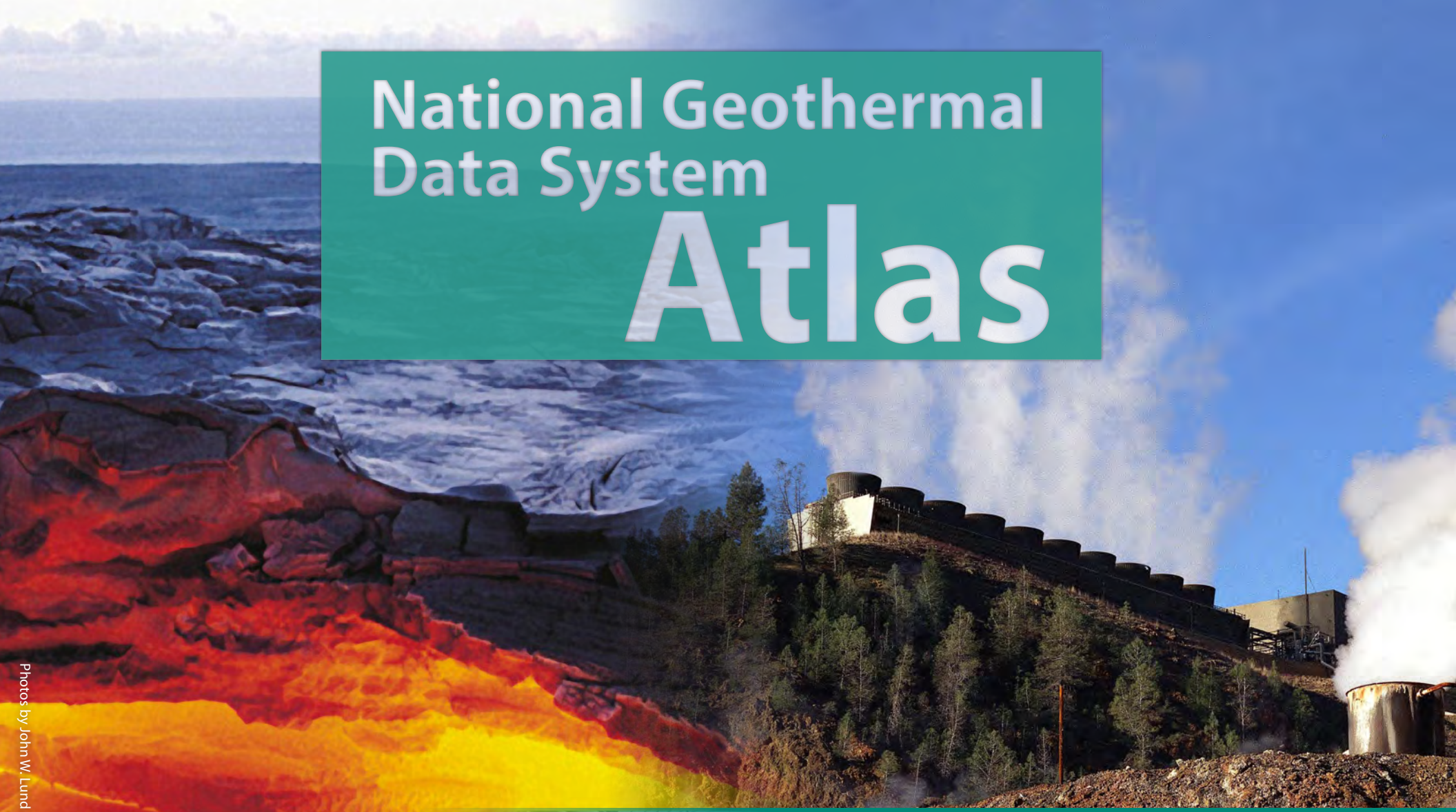
Hawaii

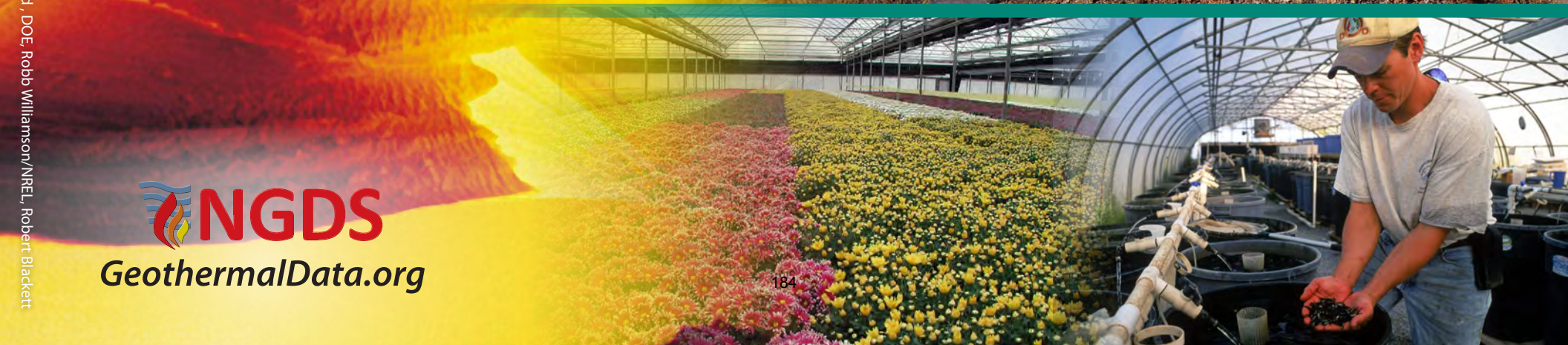
- NGDS instigated statewide assessment of geothermal resources using magnetotelluric surveys.
- Digitized records from approximately 5,700 wells will be used to interpret and analyze the MT data.
- NGDS project has stimulated new geothermal project development and cooperation between key players

This material is based upon work supported by the Department of Energy under Award Numbers DE-EE0002850 and DE-EE0001120 and NSF grant EAR-0753154

U.S. DEPARTMENT OF ENERGY

National Geothermal Data System Atlas





NGDS

GeothermalData.org

Photos by John W. Lund , DOE, Robb Williamson/NREL, Robert Blackett

USGIN
powers NGDS

USGIN is free and open source

Create your own node on an internal network
Stream content via your website or portal

Contact us
Randi.Bellassai@usgin.org
USGIN.org

U.S. Geoscience Information Network

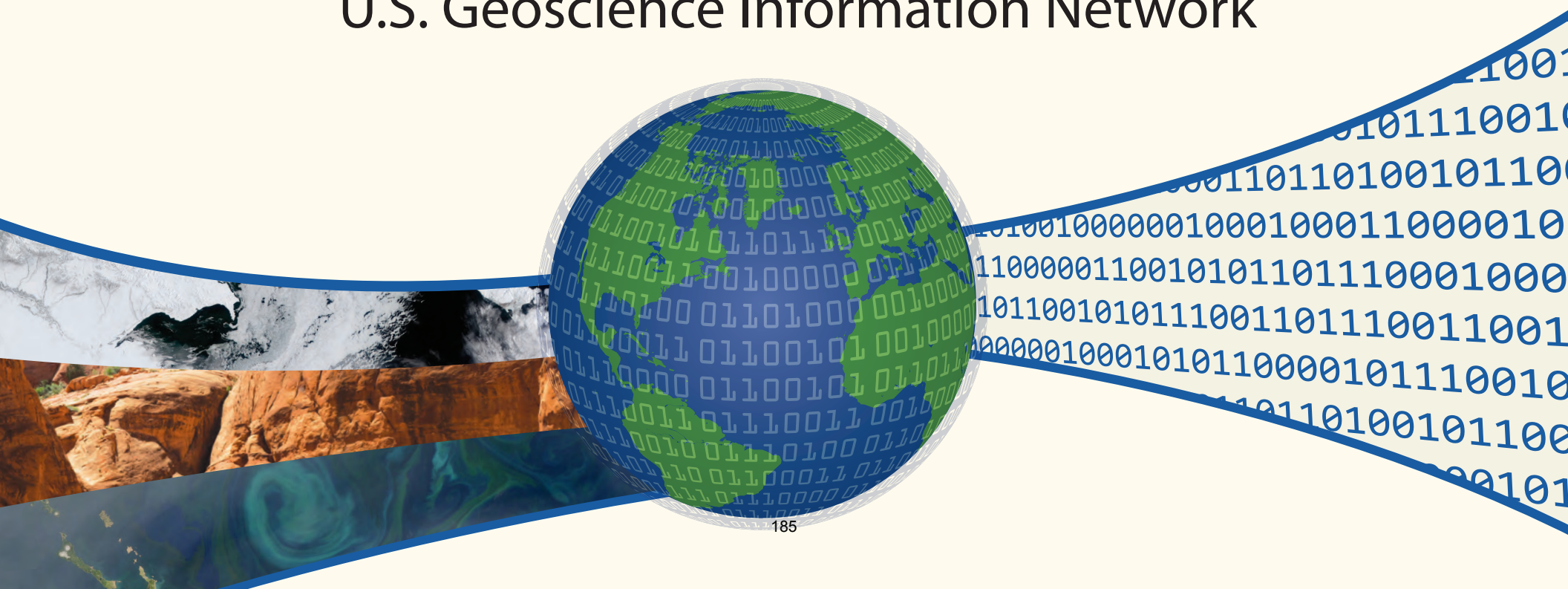


Table of Contents

National Geothermal Data System Overview 4
USGIN Powers NGDS 5
New Geothermal Data Exposed at NGDS 6
NGDS Serving Data via Distributed Hubs 8
About the NGDS Atlas 9
Alabama 10
Alaska 11
Arizona 13
Arkansas 15
California 16
Colorado 17
Connecticut and Massachusetts 19
Delaware 20
Florida 21
Hawaii 22
Idaho 23
Illinois 25
Indiana 27
Iowa 28
Kansas 29
Kentucky 30
Louisiana 31
Maine 32
Michigan 33
Minnesota 34
Mississippi 35
Missouri 36
Montana 37
Nebraska 38
Nevada 39
New Hampshire 40
New Jersey 41
New Mexico 42
New York 43
North Carolina, Georgia, and South Carolina 44
Tennessee and Maryland 45
North Dakota 46
Ohio 47
Oklahoma 48
Oregon 49
Pennsylvania 50
Rhode Island 51
South Dakota 52
Texas 53
Utah 54
Vermont 55
Virginia 56
Washington 57
West Virginia 58
Wisconsin 59
Wyoming 60

Acknowledgments

Major coordination provided by Association of American State Geologists (AASG).

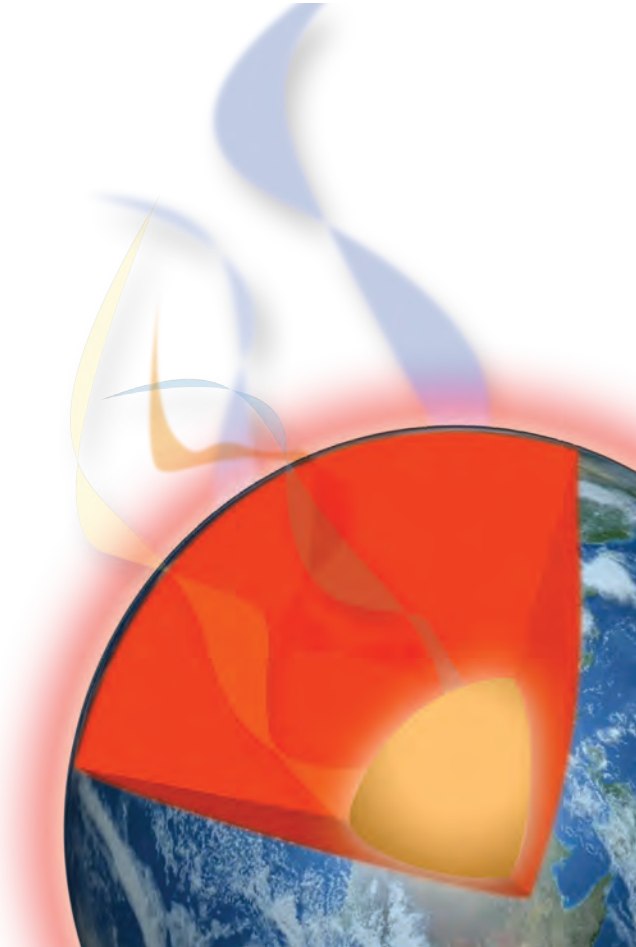
The U.S. Department of Energy's Geothermal Technologies Program, an Office of Energy Efficiency & Renewable Energy program, funded NGDS development under Award No. DE-E0002850.

AASG

U.S. DEPARTMENT OF ENERGY

The editors thank Diane Love for invaluable assistance in building this atlas.

Atlas design and layout by Stephanie Mar, AZGS.



National Geothermal Data System Overview

In May, 2014, U.S. Department of Energy Secretary Ernest Moniz formally launched the National Geothermal Data System (NGDS) at the White House Energy Datapalooza.

The $22 million NGDS project, funded by the U.S. Department of Energy Geothermal Technologies Office, coordinated the nationwide digitization, submission and exposure of geothermal-relevant information.

The NGDS is an online, open-source platform that facilitates the discovery and access of geothermal-related data. The overarching goal is to provide the geothermal energy exploration industry, academia and business with ready access to digital data to break down one of the geothermal energy industry's greatest barriers to development and deployment of this promising clean energy source; the high cost of exploration.

The NGDS was created under two grants. The first was a design and build project administered by Boise State University (more information on the design and build project is available at *http://www.osti.gov/scitech/biblio/1163354*). The second grant coordinated state submissions from more than 50 collaborators, including state geological surveys and sub-awardees, led by the Arizona Geological Survey. The NGDS is now bringing unprecedented volumes of new and legacy data online.

The maps, data products, and information in the State Geothermal Atlas are just a sample of derived data contributed to NGDS under the "State Contributions to the National Geothermal Data System" project. A comprehensive online catalog of state products is available at *geothermaldata.org*.

geothermaldata.org

NGDS

AASG

"One of the most successful [cyberinfrastructure] programs to date is the National Geothermal Data System (NGDS)…"

—Sara E. Pratt, "Earth Magazine" August/September 2013

USGIN Powers NGDS

Access approximately 10 million data records:

- 1.85 million well headers (all wells)
- 520,000 water wells (supply, monitoring, irrigation)
- 666,800 borehole geophysical logs (oil and gas, water, other wells),
- 2 million borehole lithology data records (intervals and intercepts)
- 815,600 digital geologic map features (contacts, units, active faults, volcanic vents, geothermal areas)
- 538,800 borehole temperatures (well sand springs)
- 661,500 well fluid production records
- 387,800 aqueous chemistry records (wells and springs)
- 73,300 drill stem/well test records
- 53,800 geochem records (physical sample, thermal conductivity, and geochemistry)
- 14,000 direct use and heat pump facilities

The records listed above are available as interoperable data that can be seamlessly mapped across all states. Additional data, metadata, and maps are available in structured and unstructured formats (e.g. spreadsheets and scanned images), including:

- 66,600 all other types of data (seismic, gravity, heat flow, hydraulics properties, geothermometry, power plants)
- 59,300 documents, reports, and papers, including 6,300 geological maps
- 3 million data points from Southern Methodist University

GDR Geothermal Data Repository

NREL NATIONAL RENEWABLE ENERGY LABORATORY

USGS science for a changing world

The U.S. Geosciences Information Network (USGIN) is the engine that drives the NGDS.

Developed through the US Department of Energy over five years with American Reinvestment and Recovery Act funding, the NGDS is now managed by the U.S. Geosciences Information Network Foundation, Inc (USGIN).

USGIN

USGIN is a data sharing framework based on conventions and protocols that support a distributed, scalable, digital network specializing in geospatial data. The framework enables the management of interoperable data flawlessly in a number of common formats.

A joint undertaking of the United States Geological Survey and the Association of American State Geologists, USGIN uses the same free, open-source content management system as *Data.gov*, making data fully transparent to researchers and developers. USGIN standards divide data into three tiers: scanned files that are not machine readable (Tier 1), structured data that is machine readable but does not conform to a standard type (Tier 2), and data structured into a standardized schema that allows it to be read in numerous formats and by multiple software programs (Tier 3).

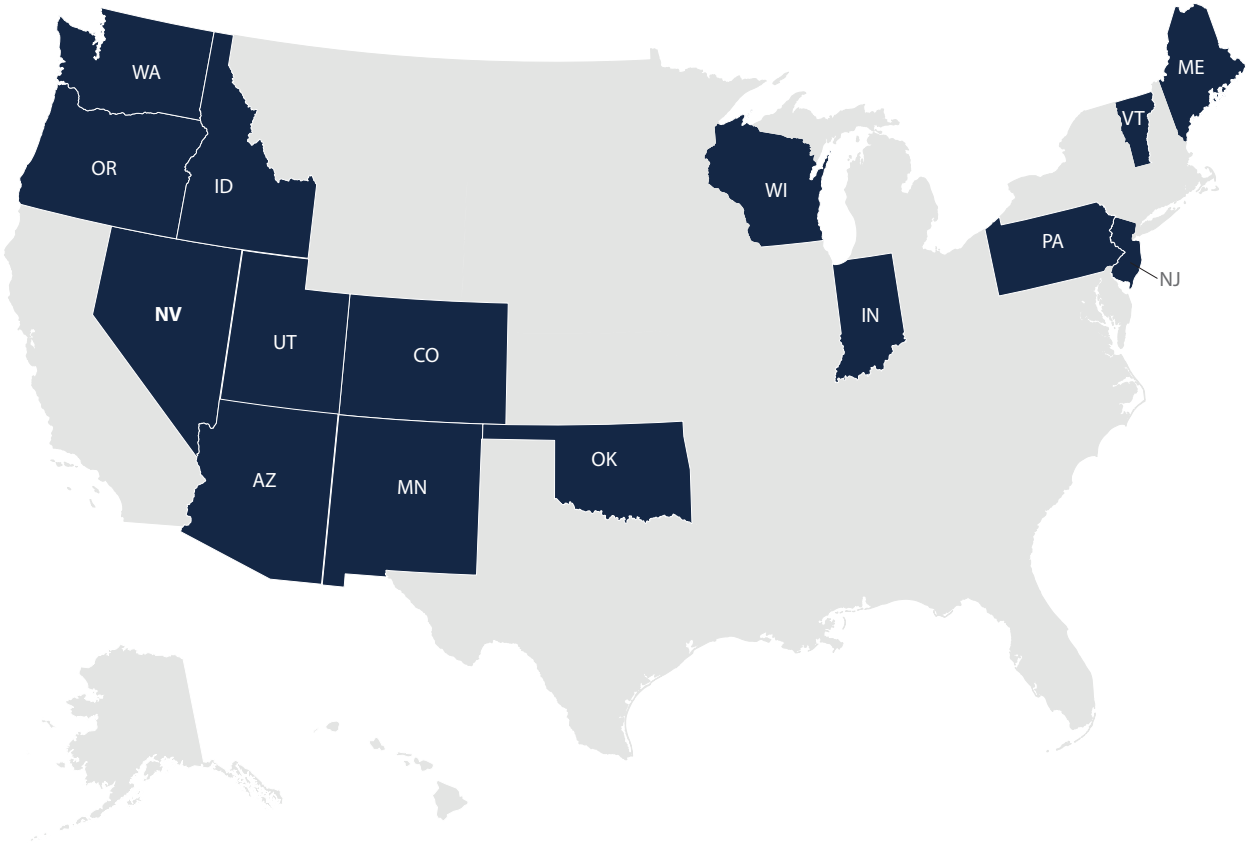
The publicly accessible NGDS data platform draws from millions of newly digitized records, including other DOE geothermal projects, to present information found in the DOE Geothermal Data Repository (GDR) and National Renewable Energy Laboratories (NREL).

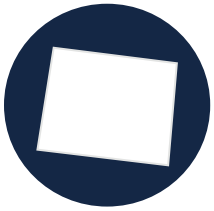
New Geothermal Data Gathered for NGDS

New data gathered in this project, including some from exploration wells, are now available from Arizona, Colorado, Idaho, Indiana, Massachusetts, Maine, New Jersey, New Mexico, Nevada, Oklahoma, Oregon, Pennsylvania, Utah, Vermont, Washington, and West Virginia. Positive results are forthcoming and include the discovery of a new geothermal resource in Utah's Black Rock Desert. Data records for this new exploration are available through the NGDS catalog.
Arizona collected new data as water chemistry from springs and wells, new geologic map data in three locations of known thermal springs.

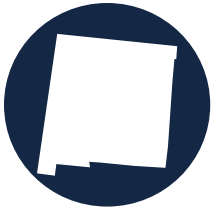
Ongoing Results

The NGDS project provides the data foundation for discovery and development of both large- and moderate-scale energy production and other practical applications, such as direct use and residential/commercial ground source heat pumps. The innovative data management approach adopted by NGDS may well lead to innovations and exploration in other Earth sciences and geospatial fields.





Colorado collected new chemistry data/analyses and obtained new thermal conductivity measurements on existing core.



New Mexico collected new temperature-depth logs from existing wells, new aqueous chemistry samples for analysis, and rock samples for geochronology analysis.

Utah drilled 12 new geothermal exploration wells and obtained well logs, aqueous geochemistry and thermal conductivity data.

Idaho drilled 4 new geothermal exploration wells, obtained lithology logs, thermal conductivity measurements, and heat flow data.

Nevada drilled 2 new geothermal wells and obtained logs, thermal conductivity measurements and aqueous geochemistry data.

Vermont collected new temperature-depth logs and other geophysical logs from a set of existing deep wells without temperature logs. In addition, new rock samples were collected for thermal conductivity measurements.

Massachusetts collected surface rock samples, and measured thermal conductivity and Uranium, Thorium, and Potassium concentrations to calculate radiogenic heat production.

Oklahoma measured thermal conductivity and radiogenic heat production data from existing drill core samples and newly collected surface samples.

Washington collected new aqueous chemistry samples for analyses, completed new temperature-depth logs in existing holes, conducted new geologic mapping related to geothermal activity.

Maine collected new temperature-depth logs from existing deep wells.

Oregon drilled 3 new geothermal exploration wells and obtained logs, aqueous geochemistry and thermal conductivity data.

Wisconsin drilled new wells as part of the primary project (not Supplemental) and obtained well logs and thermal conductivity data.

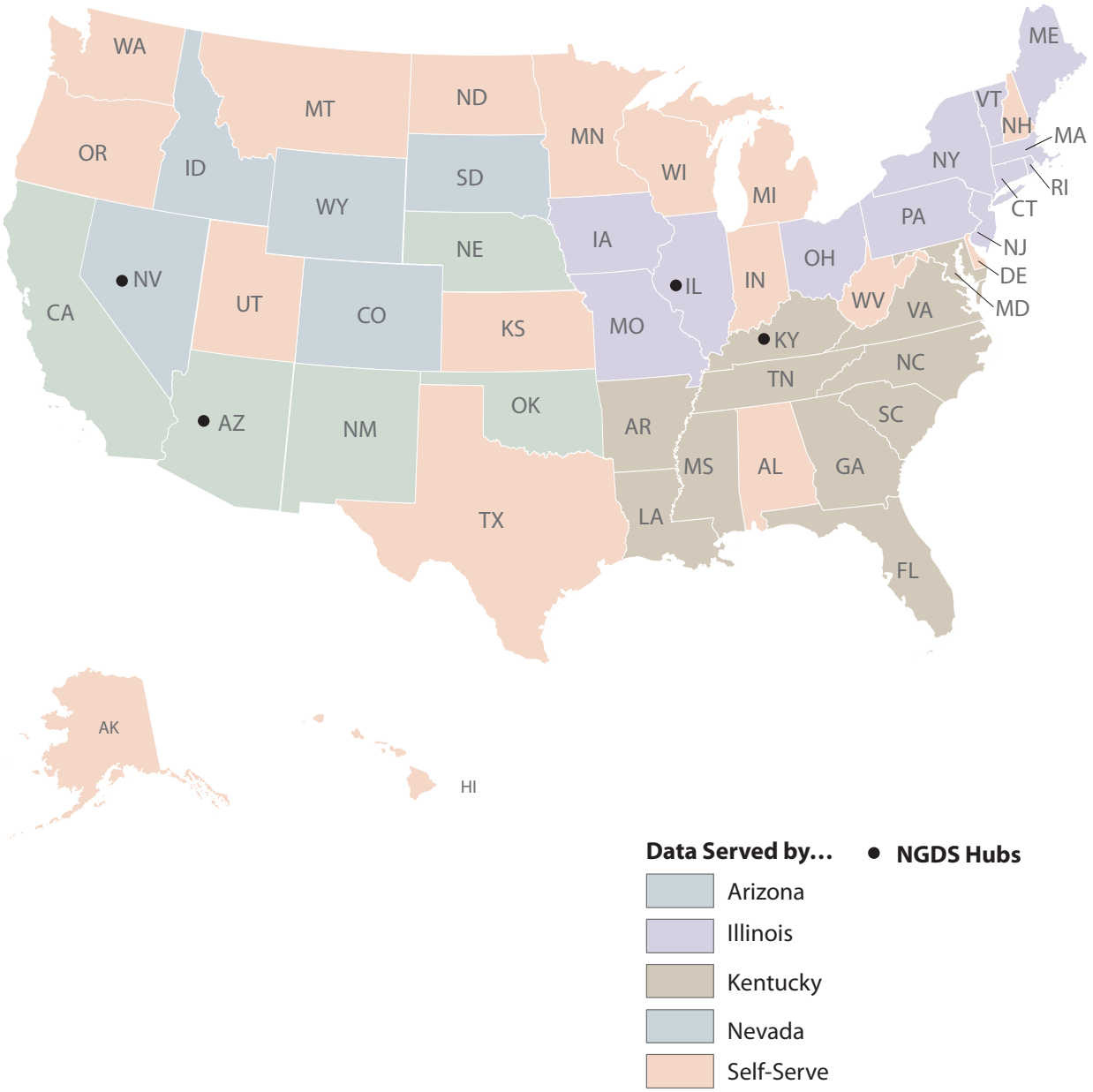
New Jersey collected new temperature-depth logs from existing deep water wells.

Pennsylvania ran well logs in new deep wells drilled by industry in 2014.

West Virginia collected new thermal conductivity measurements from existing core and cuttings samples.

NGDS Serving Data via Distributed Hubs

The NGDS features contributions of geothermal-related data from all 50 U.S. states that is referenced through a central catalog system. Actual data is either independently hosted by states using GeoPortal or GIN-stack software or it is served by a Hub. Currently, four Hubs are responsible for serving data from 34 states. Nevada and Arizona Hubs cover most of the west and westernmost Midwestern states. Illinois and Kentucky Hubs serve data from the Northeast and the Southeast, respectively. States self-hosting their data, known as Nodes, are geographically scattered throughout the country.



About the NGDS Atlas

The chief objective of the NGDS atlas is to inform the reader of the nature, quality and quantity of geothermal-related data available for all 50 U.S. States, free and online through the NGDS catalog.

Currently, over 7.2 million data records are exposed through the NGDS catalog. That number will grow as State Surveys, DOE-funded contributors, industry, and others add fresh data. Potential data contributors and web developers are encouraged to visit our "Data Providers and Web Developers'" webpage to see how to get started (*http://geothermaldata.org/ngds/publish_data*).

Showcasing the breadth and diversity of NGDS holdings—which ranges broadly from PDF documents to Excel data tables, to map services displaying geologic maps, to well log data, to geochemical analyses—is an extraordinary challenge.

NGDS Atlas Data Flower

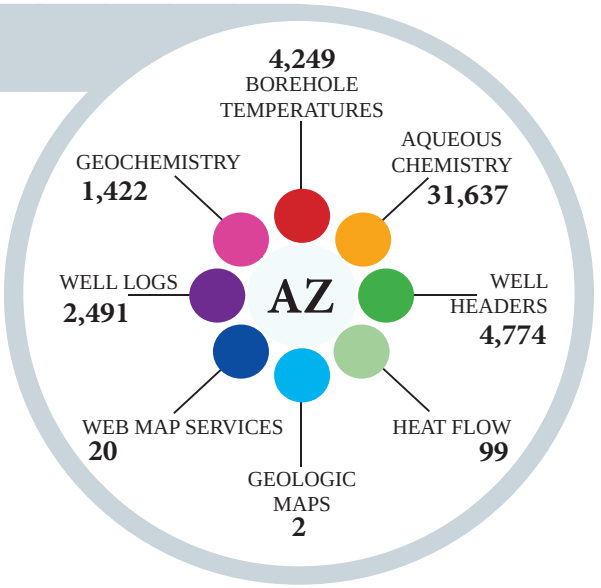
Our solution: to build and display an NGDS Atlas data flower for each state. The data flower displays major data categories and volumes of data available at the NGDS Catalog for each state (*geothermal.org*).

Major Categories

- **Borehole temperatures**—down-hole temperature values
- **Aqueous chemistry**—water chemistry parameters, e.g., pH, alkalinity, conductivity, flow, metals
- **Well headers**—location and elevation of wells, ID
- **Heat flow**—measured or estimated heat flow
- **Geologic maps**—provided by State Surveys
- **Web Map Services**—access to digital maps
- **Well logs**—well report on location and lithology
- **Geochemistry**—major- & trace-element analyses

Graphic Display

We selected graphics to showcase some of the accomplishments that characterize each state's involvement in NGDS. Some states leveraged their participation to build interactive maps and other resources for displaying their data; for others, data compilation has led to the discovery and exploration of previously untapped geothermal resources. The states overwhelmingly report positive outcomes in data management and understanding of subsurface systems as a result of the NGDS project.



The NGDS Atlas data flower shows at a glance the type and volume of data available in NGDS for each state.

The geothermal-geologic data used to construct this map is now available digitally in the NGDS catalog.

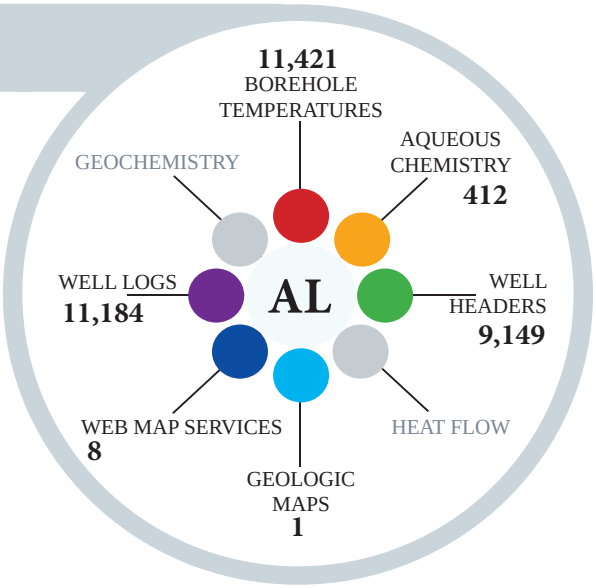
NGDS is powered by USGIN

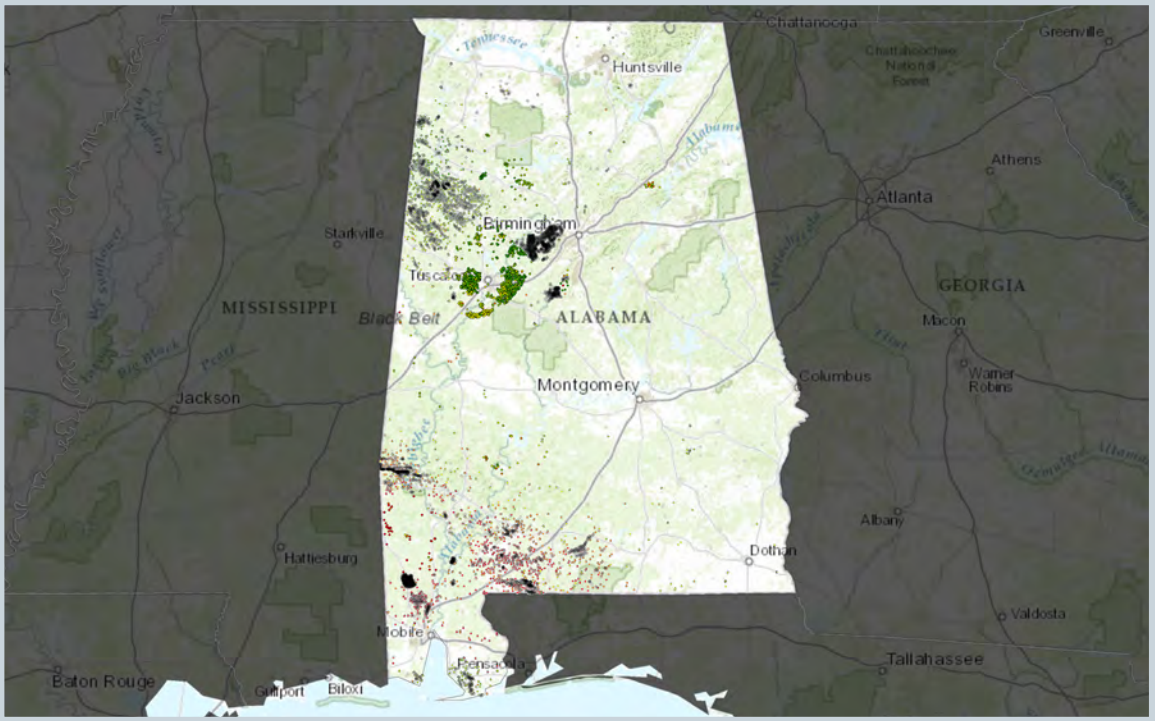
Alabama

Self-hosting NGDS node

In late 2014, Alabama became a self-hosting node for NGDS using the CKAN based software stack called GIN-stack. Alabama is slowly populating their website with available NGDS services. In addition to services for well headers, borehole temperatures, and well logs, anticipated web services will include aqueous chemistry (version 1.10), aqueous well chemistry, borehole lithology intervals, seismic event hypocenters, and well fluid production.

Alabama's ArcGIS map services, including OGB web services portraying their geothermal data submissions, can be located at *http://map.ogb.state.al.us/ogbmaps/rest/services/OGB/map/MapServer*. ■





Graphic representation of Web Map Services conveying borehole temperature measurements and well headers provided by Alabama as part of the National Geothermal Data System, soon to be served through Alabama's independent service.

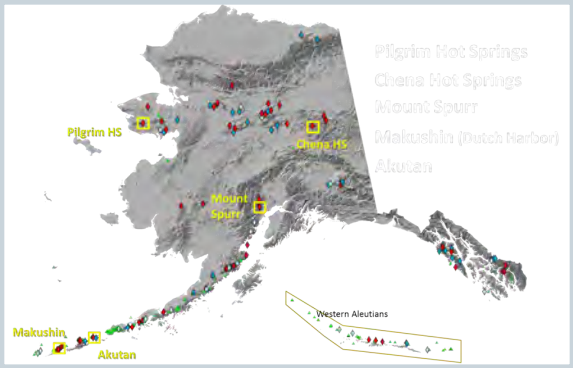
Alaska

Land of Volcanoes

With more than 100 volcanoes and over 50 known hot springs (above 50 degrees C) and 46 warm springs, Alaska is a potential geothermal bonanza. With NGDS support, the Alaska Division of Geological & Geophysical Surveys (DGGS) compiled existing information on exploratory geothermal wells drilled in the 1980s and in 2012 by Alaska Center for Energy and Power.

Data Management & Best Practices

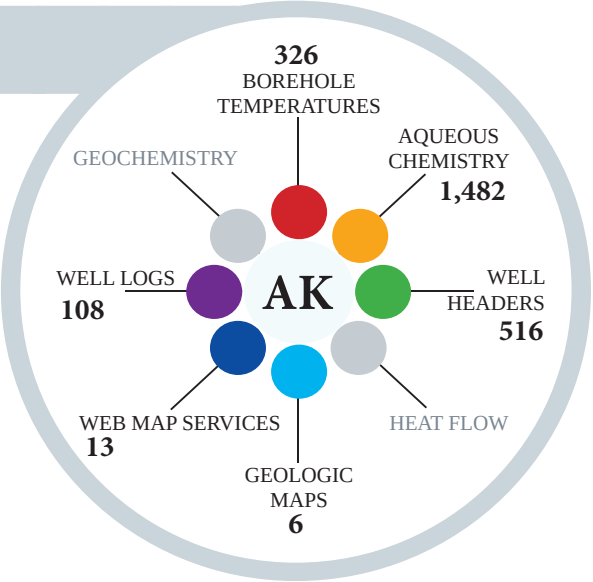
Geothermal-related data are now organized for exposure and easy retrieval—a great improvement over scattered paper records. This structured NGDS data approach breathes new life into legacy data and spotlights existing data gaps, which can drive new exploration projects.

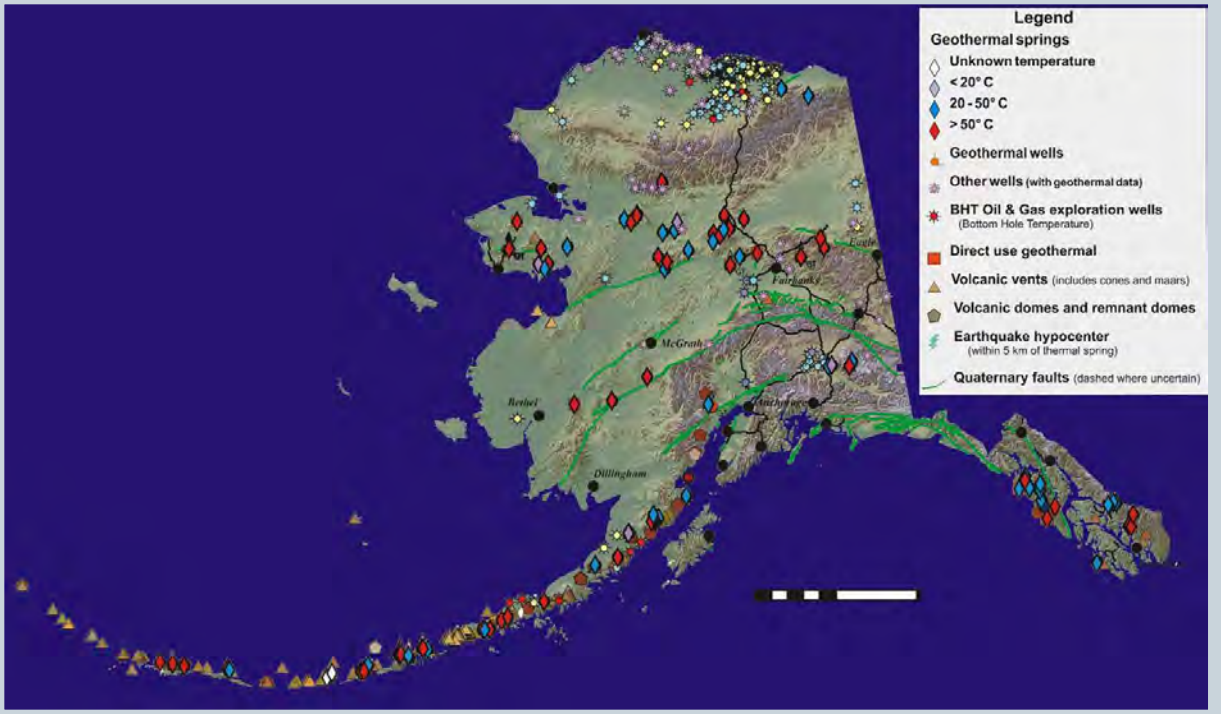


A map of geothermal resources in Alaska; highlighted areas represent recently discovered resources.

As part of their NGDS program, Alaska researchers compiled at-risk legacy and new exploration data for free access at *geothermaldata.org*. The team structured their data to a whopping 30 NGDS-approved schema, creating interoperable Tier 3 data sets that can be read in multiple formats. Tier 3 data is interoperable data, structured and standardized to open in multiple formats and to effortlessly merge with additional data sources in a content model for ease of analysis.

Alaska became a self-serving node of the NGDS using ▸▸





A compilation of digital geothermal data provided by Alaska, including geothermal springs, well data, direct use geothermal sites, and seismic and volcanic features. This is an online, interactive GIS-based, digital map to replace the print-only 1983 analog map.

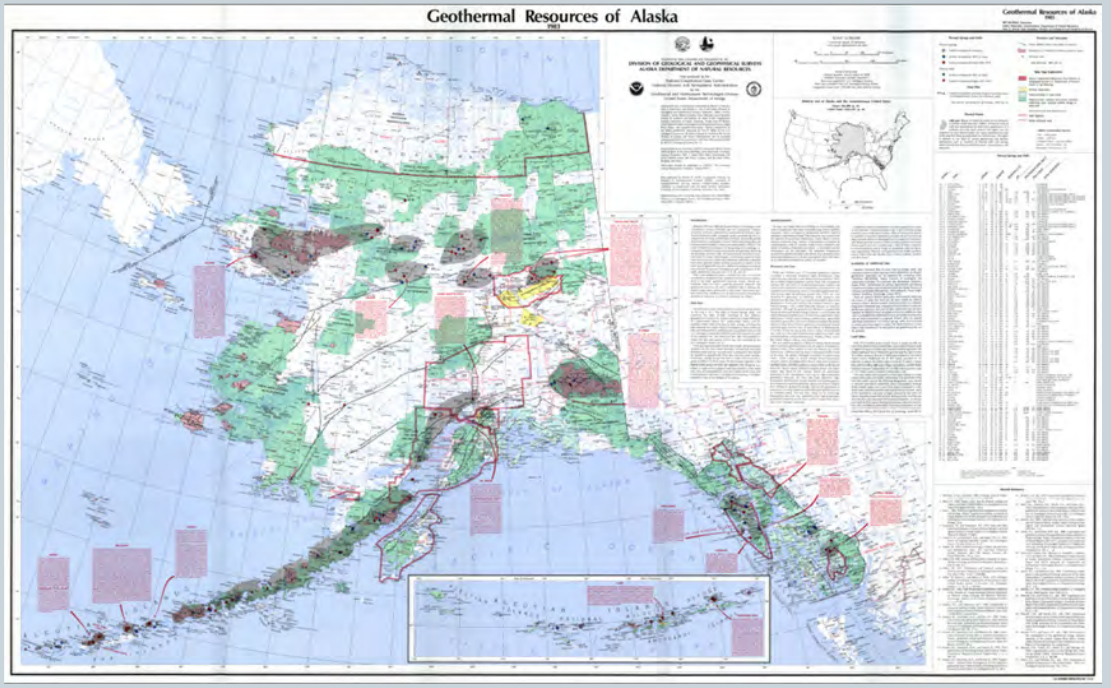
GeoPortal free and open source software in late 2014. The DGGS uses NGDS data to create an online , interactive geothermal map product that will replace paper maps. Look for the Alaska Geothermal Resource Map at *http://www.dggs.alaska.gov/pubs/*.

Geothermal Exploration & New Discoveries

Beginning in 2009, AIDEA-Alaska Energy Authority funded exploration of Hot Springs Valley on Akutan island to evaluate the potential for geothermal development to provide energy to both the City of Akutan and nearby Trident food processing plant.

Newly discovered and currently explored resources include:

- Pilgrim Hot Springs
- Chena Hot Springs
- Mount Spurr
- Makushin (Dutch Harbor)
- Akutan ■



An analog map of geothermal resources in Alaska from 1983; a classic example of legacy data digitized for NGDS.

Arizona

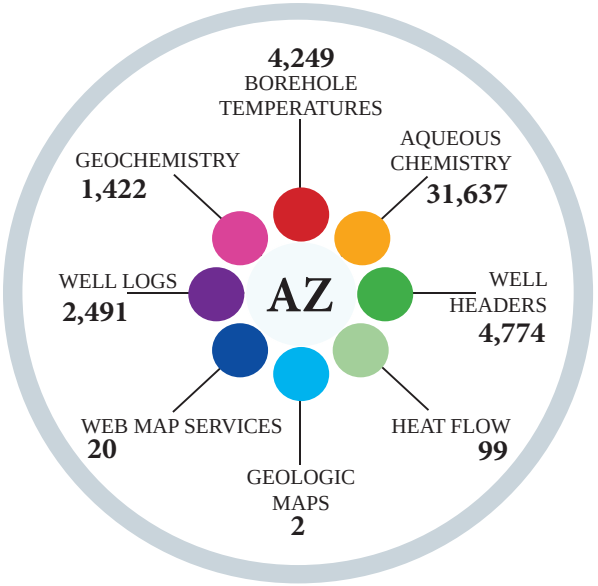
Interoperable Data

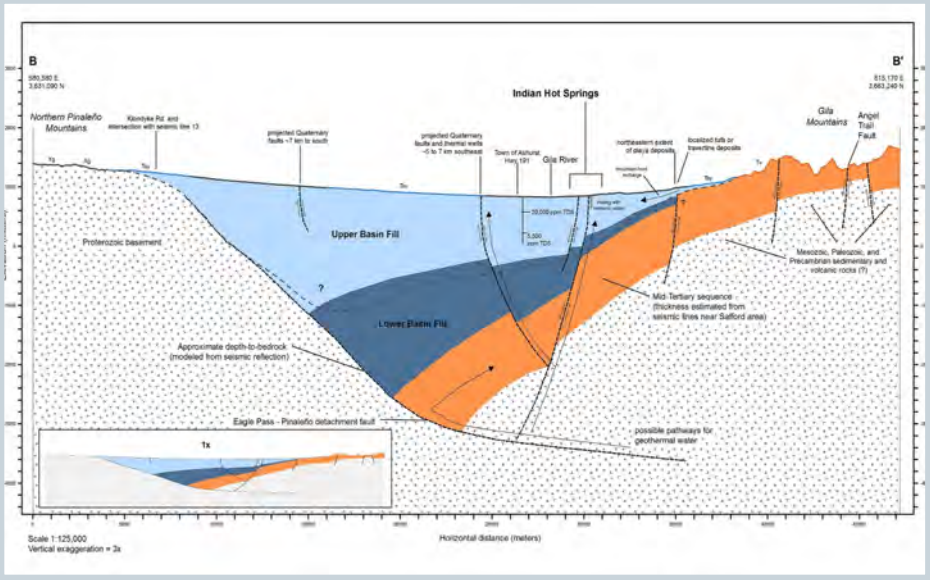
Working with the AASG, team members at the AZGS helped develop the framework for USGIN (US Geoscience Information Network), the collection of standards and protocols that the NGDS project is built upon. Interoperable Arizona data added to the NGDS Catalog includes over 4,000 document metadata for existing and newly scanned documents. Downloadable data is available in interoperable exchange content models or reports and documents. The two submitted geological maps are the geothermal map and the AZGS interactive digital map which covers all the state's geology. This map is available online at *http://data.azgs.az.gov/geologic-map-of-arizona/*.

December 2012 | Verde Hot Spring along the Verde River, southeast of Camp Verde, Arizona. Hot springs throughout Arizona were sampled for isotope and geochemistry analysis, excluding Verde Hot Spring (pictured) which was attempted but inaccessible due to high water flow, to assess the source of thermal groundwater exhibiting unusually high temperature anomalies. The Camp Verde hot springs burble away near the remnants of an old hotel that burnt down in 1962. It was apparently a swank spot in the Jazz Age, but rumors that Al Capone used it as an occasional hideaway are pure fiction.

New Data Collection

The AZGS supplemental 'new data' collection project provided new thermal spring and well aqueous chemistry, rock chemistry, geothermometry data, and temperature data for 17 thermal springs, and geological investigation to determine thermal groundwater sources for three thermally active areas hosting hot springs in southeastern Arizona. A final report detailed Arizona geothermal background, data collection methods and results, and new geologic maps for Aqua Caliente Spring, Hooker Hot Springs, and the Pinaleño Mountains-Safford Basin geothermal area (OFR-14-06 for the *Supplemental Investigation of Thermal Springs throughout Arizona: Geochemical, Isotopic, and Geological Characterization, Arizona Basin and Range Province).*





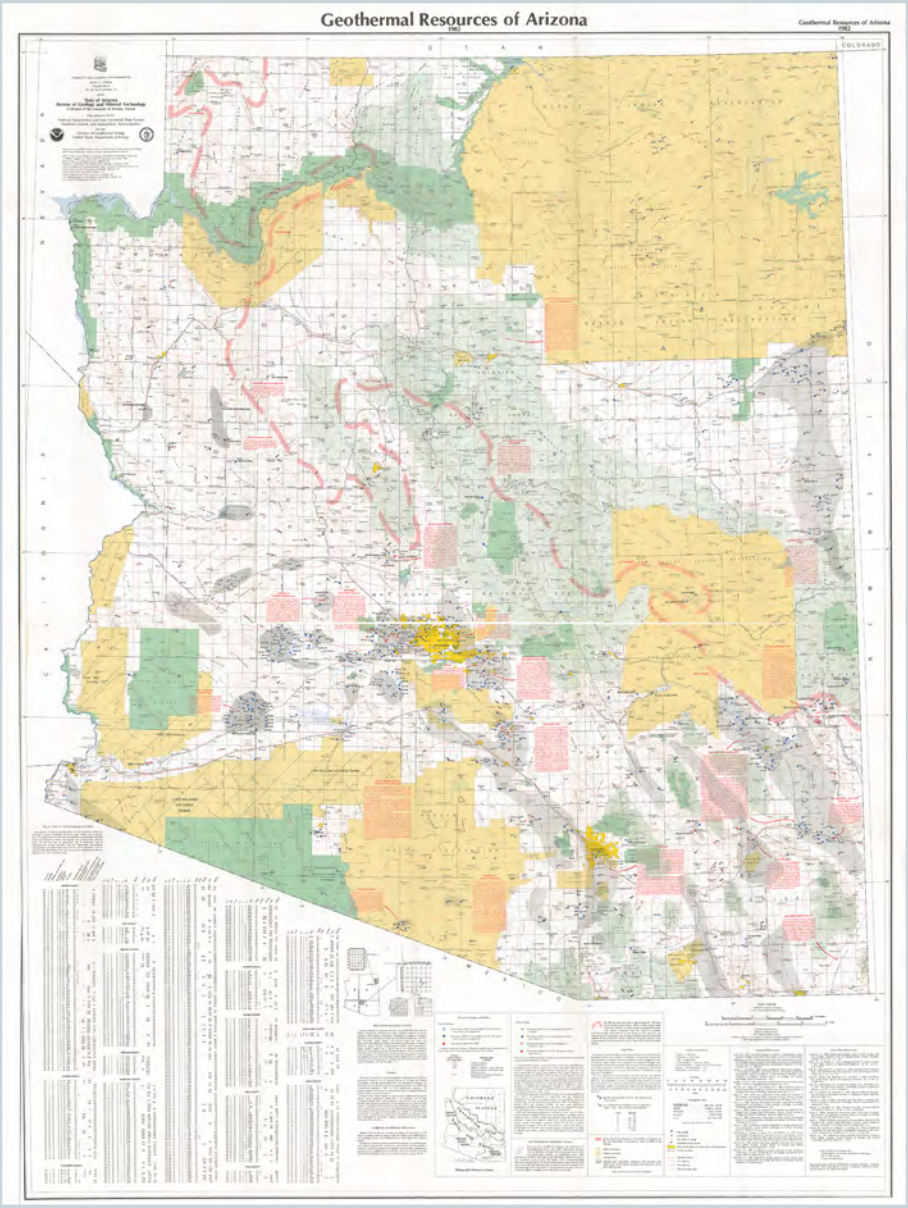
Geologic Cross Section A-A' Across Safford Basin, Cactus Flat to Buena Vista Hot Springs (AZGS 2014) produced for "Investigation of Thermal Springs throughout Arizona: Geochemical, Isotopic, and Geological Characterization, Arizona Basin and Range Province"; authored by Diane S. Love, Brian F. Gootee, Joseph P. Cook, Michael K. Mahan and Jon E. Spencer.

Geothermal Potential

Sedimentary basins in Arizona's Basin and Range Province show potential for high-to-moderate heat flow, moderate to high permeability, especially in fractured and faulted deep basin-fill units, and may be viable targets for large-volume conductive geothermal resources. Temperatures increase with depth below the surface in a near-linear geothermal gradient. Temperatures measured in oil and gas wells greater than 8,000 feet deep may exceed 100 °C. Thermal waters can also materialize at the surface as hot springs with temperatures exceeding 50°C, although most thermal springs temperatures in Arizona range from 20°C to 50°C.

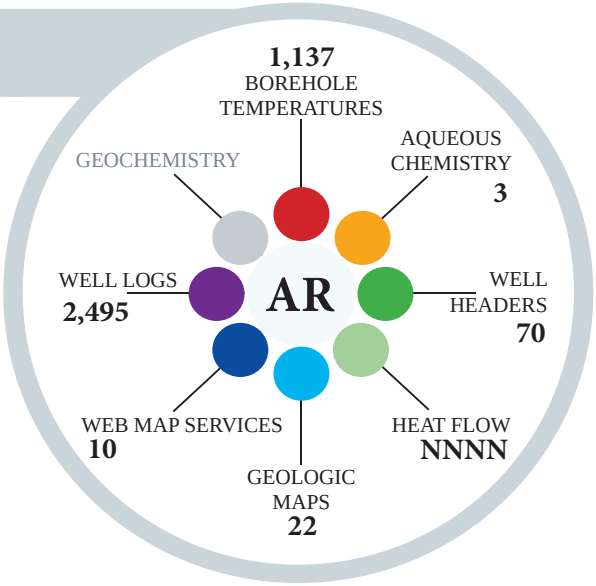
Arid central and southern Arizona supports a large agricultural economic base and several large, rapidly growing metropolitan areas. There is substantial competition between rural agricultural regions and cities over current and future use and control of groundwater, including low-temperature geothermal water sources (Witcher, 1995).

Abundant existing deep water-supply wells and irrigation wells have potential for conversion to private and small commercial geothermal direct-heat uses. Geothermal space heating and district heating are becoming more popular in areas where the resource is co-located with population and feasible for community facilities with large heating/cooling loads, such as schools. Geothermal aquaculture and green-housing are major direct-use applications showing noticeable growth in Arizona. Many geothermal direct-use applications have potential to conserve water while enhancing the value of farm production per acre (Witcher 1995). ■



Map displaying Arizona's geothermal resources (1980). The data comprising this map are now available in the NGDS catalog.

Arkansas



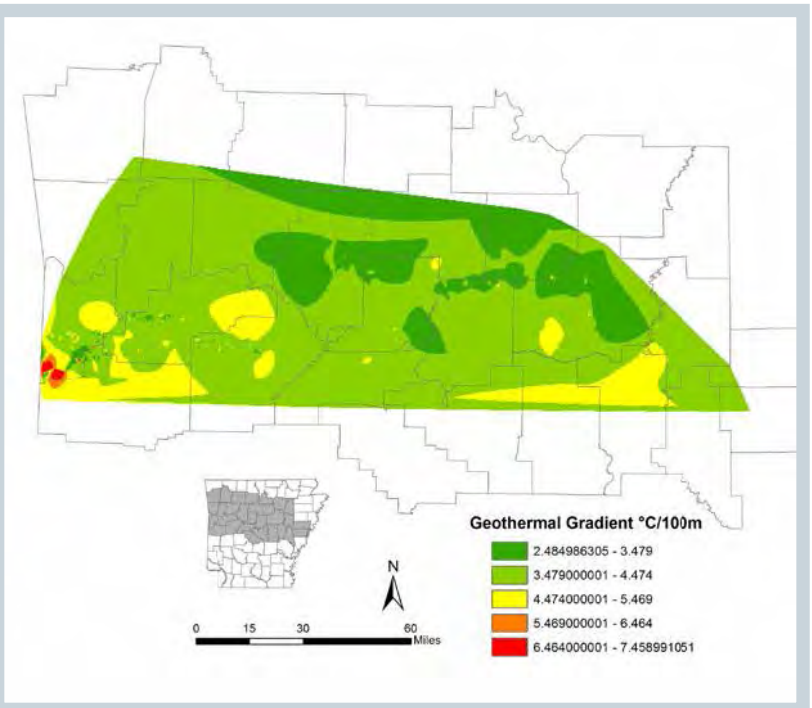
New & Legacy Data

Along with providing reports, maps and cross sections pertaining to basin structure, stratigraphy geothermal gradients, active faults and earthquakes, the Arkansas Geological Survey (AGS) loaded the NGDS project with more than 2,800 digital well logs reporting lithologies and down hole temperatures, as well as bottom-hole temperatures from across the state. Around 120 thermal conductance measurements were performed on core samples from southern Arkansas along with some data provided by the brine industry in southern Arkansas.

Bottom-hole temperatures (BHT) and temperature to depth data from over 1600 well logs from the Arkoma Basin resulted in new geothermal gradient maps. This area of northwestern Arkansas is part of the Fayetteville shale gas play.

Geothermal Potential

Core samples from 18 wells sampling the subsurface Jurassic Smackover Formation in southwest Arkansas were analyzed for thermal conductivity, thermal gradient, and heat flow values by the AGS. Interpretation of the results suggests that areas on northwestern Columbia and northeastern Lafayette Counties exhibit the highest geothermal potential, and further investigations are recommended. ■

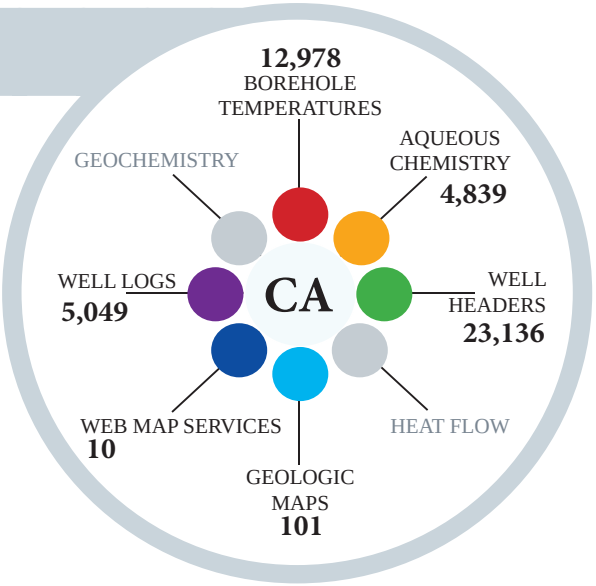


Newly produced geothermal gradient map in degrees Celsius per 100 m for north-central Arkansas.

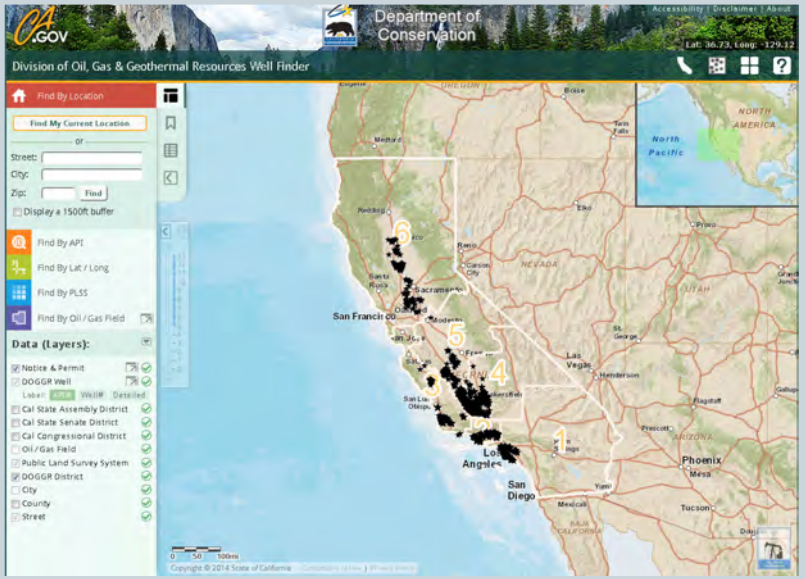
California

New & Legacy Data

California data was compiled by the AZGS with much appreciated cooperation from the California Department of Oil Gas and Geothermal Resources (DOGGR), U.S. Geological Survey (USGS), California Geological Survey, the California Energy Commission, and other entities. For California, 36 deliverable items with 25 Web Map Services were added to the repository, with over 87,000 records for online access. Data was published as an Esri Map Service, web map service, web feature service, and as an Excel workbook for download as appropriate. Metadata for reports and documents were added to the *StateGeothermalData.org* repository and cataloged for the USGIN Catalog.



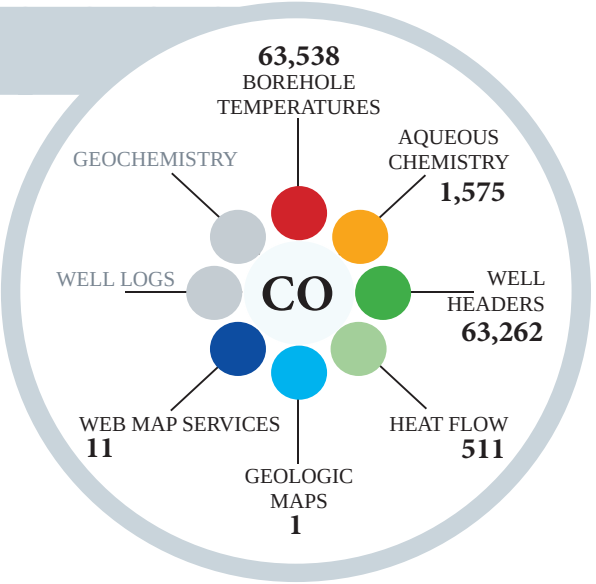
As part of Stanford's participation in NGDS, the Stanford School of Earth Sciences collected and collated about 8,000 technical geothermal papers (including downloadable PDF files), and provided geolocations for those papers specific to particular geothermal fields or sites. Stanford collaborated with the Geothermal Resources Council (GRC) to provide mutual database merges, so that the Stanford collection appears within the GRC database, and vice versa. In addition, they organized and provided some adsorption data (experimental) from rocks of The Geysers geothermal field. ■



Interactive well finder for oil, gas and geothermal wells throughout California.

Published in 2002, these data and more are now digitally available through the NGDS.

Colorado

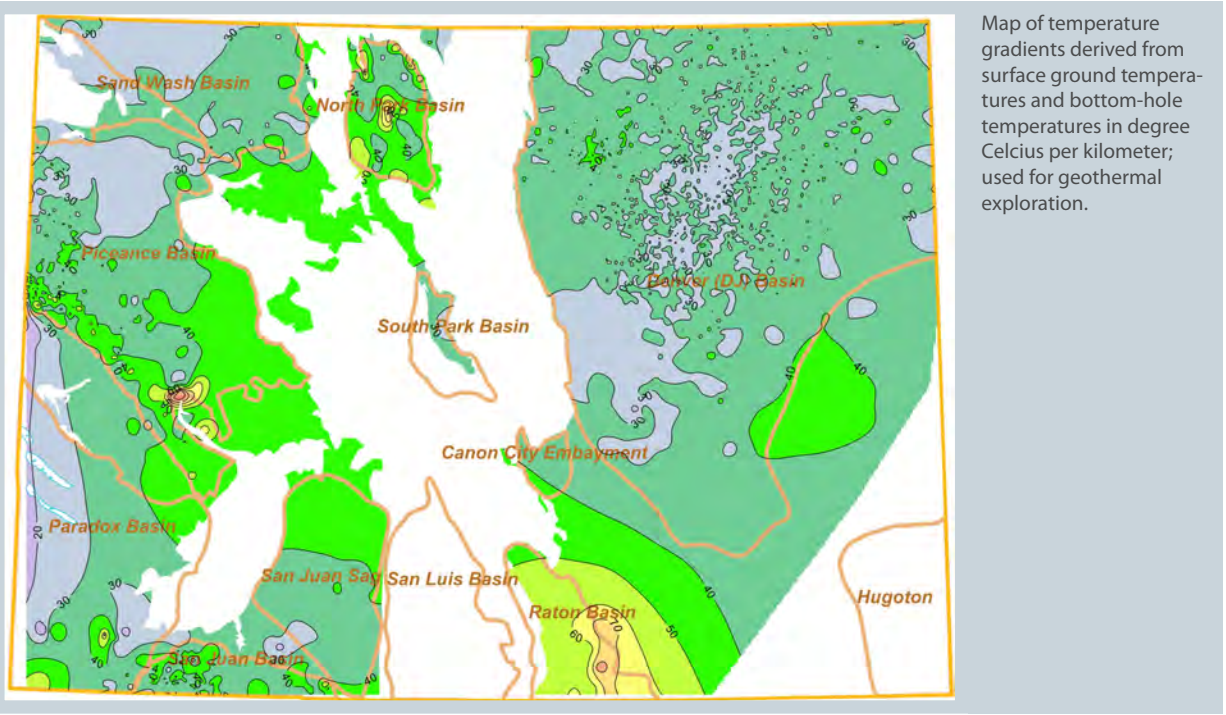


The team at the Colorado Geological Survey focused on digitizing thousands of geothermal-related records to bring legacy data to the NGDS Catalog. The result: better data management practices; out of the box thinking for geothermal assessments; and more industry outreach and involvement in geothermal exploration.

New exploration and assessment maps are proving invaluable in siting locations for electricity production in sedimentary basins, which is considered to be a new horizon. Companies and investors in the energy industry are lining up to break new ground in high yield, renewable, geothermal energy production.

Data Management & Best Practices

The Colorado Geological Survey has been collecting and compiling geothermal-relevant data since the mid-1970s. Most records were preserved as analog data; some data were stored on CD-ROMs. As a result of the NGDS project, however, the analog data has been digitized and converted to machine-readable formats, making them interoperable, searchable, readily cataloged. As new data are collected or legacy data updated, they'll be added to the catalog for public use.



Map of temperature gradients derived from surface ground temperatures and bottom-hole temperatures in degree Celcius per kilometer; used for geothermal exploration.

Industry Benefits

Data of particular importance and of great interest to the energy exploration community have been organized during the NGDS project. For instance, private sector companies are using thermal spring chemistry and geothermometry to help prospect for geothermal resources in Colorado. Bottom-hole temperatures and derived gradients have promoted industry interest in geothermal resources in sedimentary basins, with strong interest from Pioneer Natural Resources and other groups for developing an enhanced geothermal system (EGS) in the Raton Basin. There is additional interest from a major well-logging company to model basin temperatures to better predict optimal temperatures for well-logging, a potentially salable service.

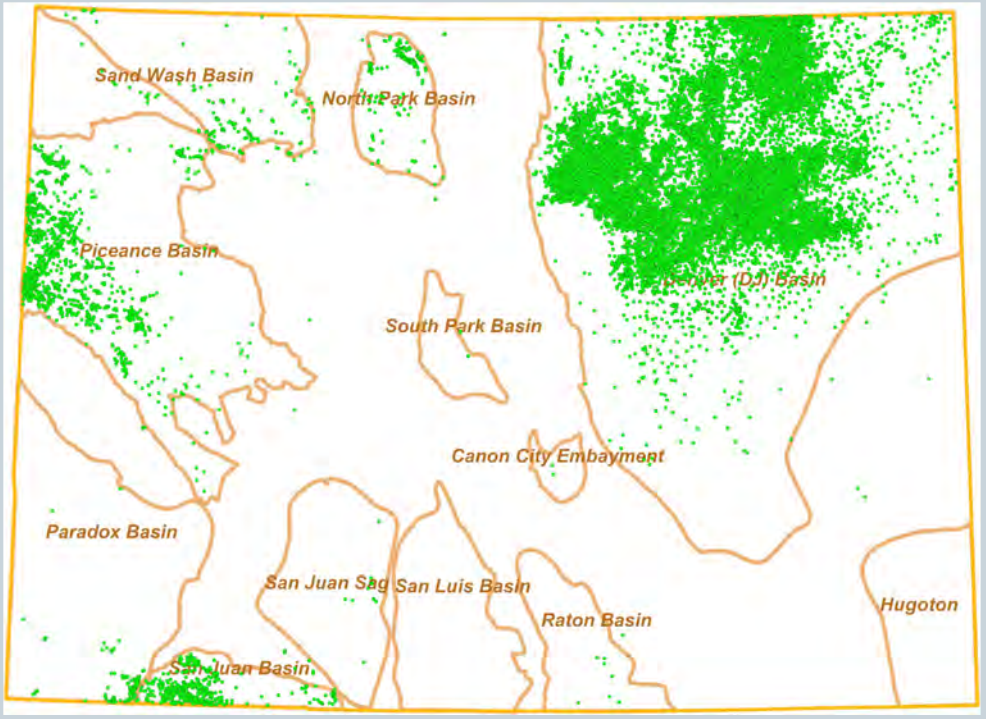
In addition, development groups are using specific temperature depth logs to to target deep geothermal test wells scheduled for drilling in summer 2014.

Geothermal Exploration & New Discoveries

This geothermal gradient map was created using different stratigraphic intervals in which borehole temperatures were selected; gradients were then calculated using ground-surface temperatures, from a model calibrated with climate stations and the BHT/depth data in the stratigraphic interval.

Colorado has potential for some small electricity production from conventional geothermal resource systems. However, these maps reveal much larger potential resources at depth in sedimentary basins that have been previously unexplored at the detail provided by the maps. Exploitation of these resources could bring a new clean, sustainable, renewable energy resource to Colorado, providing revenue and jobs.

These maps are the first maps that account for the effect on the thermal gradients of the low thermal conductivity shales that are widespread in Colorado. Previous maps showed anomalies that were only a function of shallow wells and shallow lithology. These maps may be used with geologic maps and depth sections to reliably estimate temperatures at depth. ■

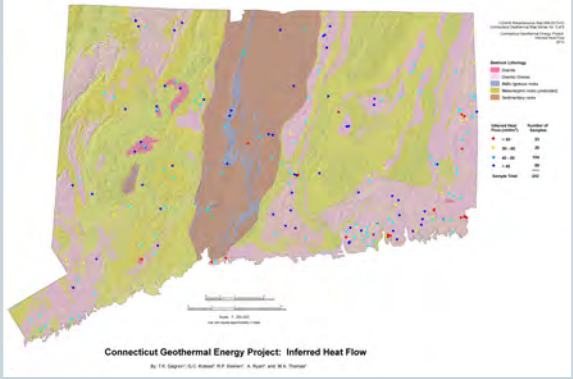


Locations of 26,178 wells used to construct the above temperature gradient map. Data from these wells is available through NGDS.

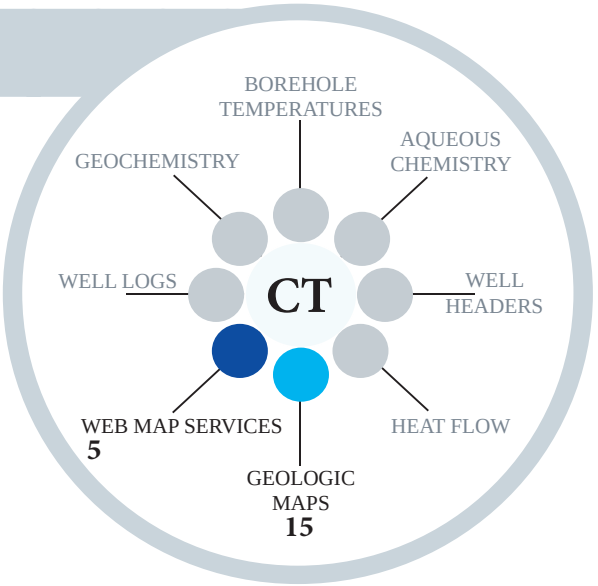
Connecticut and Massachusetts

Geothermal Assessment

The Connecticut and Massachusetts Geological Surveys collaborated on their contribution to the NGDS project. As a result, Connecticut produced the first statewide geothermal potential assessment. The objective: locate state geothermal resources with an eye towards better design of Enhanced Geothermal Systems (EGS) in bedrock and unconsolidated sediments.



Connecticut inferred heat flow map produced in 2013 from data collected as part of CT's contribution to the NGDS project.

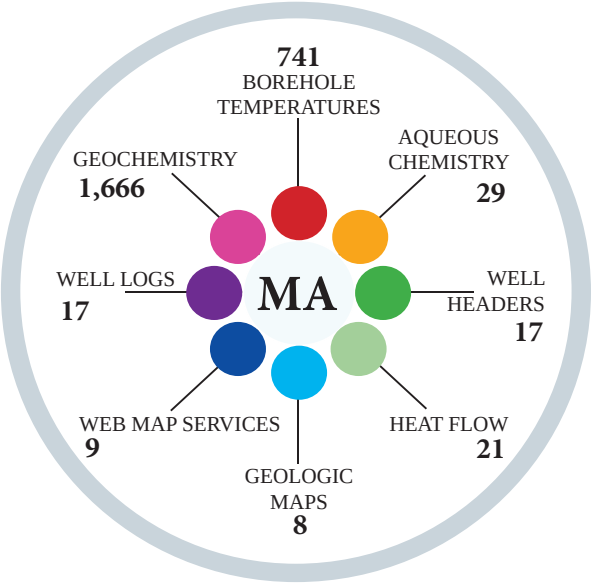


Applied Research

Bedrock units suspected capable of producing radiogenic heat at depth were the primary focus of this study. Other bedrock units were analyzed to provide a regional view of the geothermal potential across the state. A total of 55 bedrock units were targeted and 242 samples collected. Using the surficial and Quaternary maps of Connecticut, 20 unconsolidated sediment units were targeted for the collection of 100 sediment samples.

Rock chemistry, density and thermal conductivity of the bedrock samples were used to calculate heat production, heat flow, and thermal profiles at depth. Thermal conductivity measurements were made and physical profiles of sediment (grain size, sand, silt, clay percent, bulk density, porosity) were created.

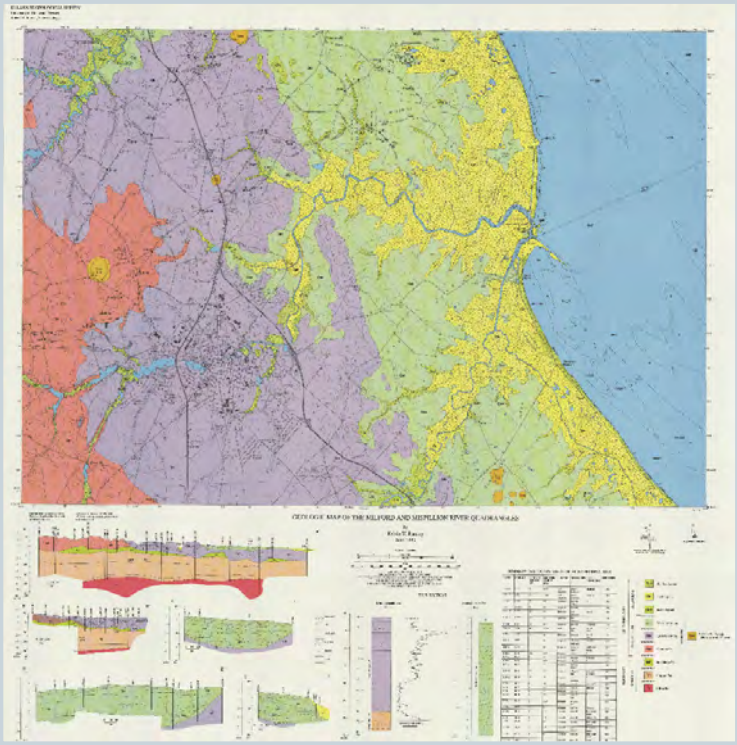
A geothermal resource map series has been compiled from these data. Information depicted in the series includes heat production, inferred heat flow, thermal conductivity, and thermal profile maps for bedrock, and a thermal conductivity map for sediments. This map series will guide further work on deep geothermal resources of Connecticut bedrock in the future. All data and map services and products are available through the NGDS. The map series and summary information can also be accessed from the Connecticut Geological Survey webpage: *http://www.ct.gov/deep/cwp/view.asp?a=2701&Q=534290&deepNav_GID=1641.* ■



Delaware

A Node at USGIN

The Delaware Geological Survey (DGS) has established a node on the United States Geoscience Information Network (USGIN). The DGS USGIN node exposes geoscience data related to the state of Delaware.

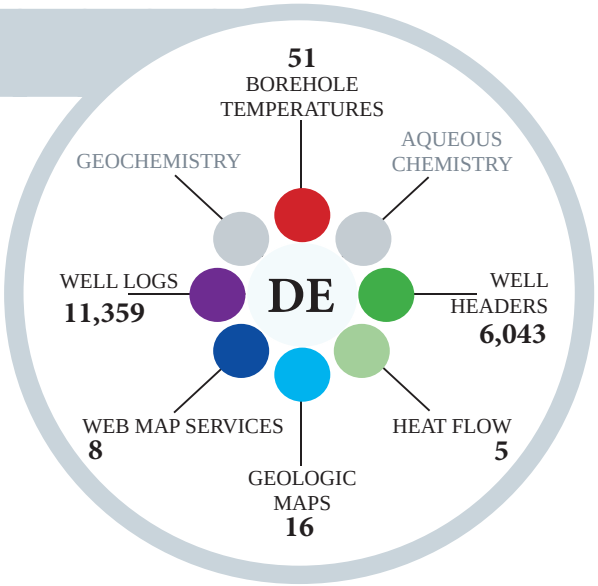


A scanned geologic map of the Milford and Mispillion River Quadrangles. This map represents one of the many products to be digitized for NGDS.

Legacy Data

DGS installed and configured open and interoperable web mapping and metadata services for surface geologic units and contacts within Delaware. These maps were developed at 1:24K and 1:100K scale for the state. All of the data structures match the GeoSciML-Portrayal schema and are distributed as WMS and WFS mapping services. Metadata for all of the data adhere to the USGIN profile standard of the ISO19139 schema, with support of the Catalog Service for the Web (CSW) 2.0.2 service, enabling easy search and retrieval of the metadata records and access to the associated map services. DGS built its IT infrastructure for serving the geoscience data through free and open source software, such as Geoserver (mapping) and GeoNetwork (metadata) applications.

Based on the USGIN node infrastructure, DGS contributed 1:100K scale surface maps to the OneGeology initiative; an international consortium dedicated to exposing and delivering digital geologic map data from around the world. DGS serves map services of data collected for the NGDS. These include datasets of well logs, well headers, thermal conductivity, heatflow and borehole temperatures. Numerous other datasets related to the NGDS are referenced and made available through metadata records. Both OneGeology and NGDS services are hosted through the USGIN node. ■

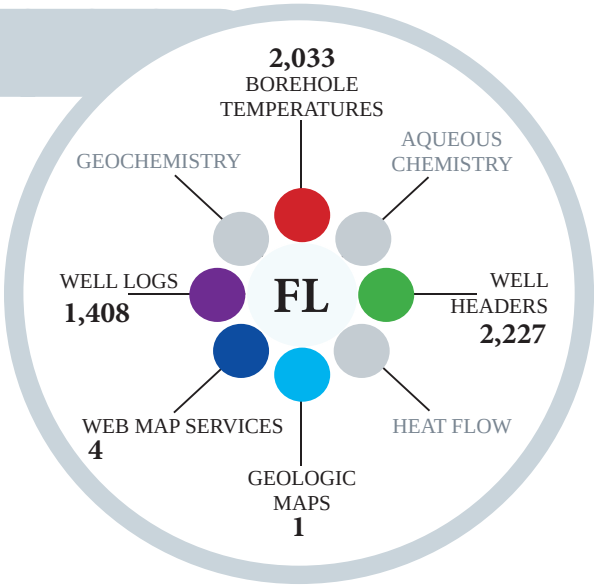


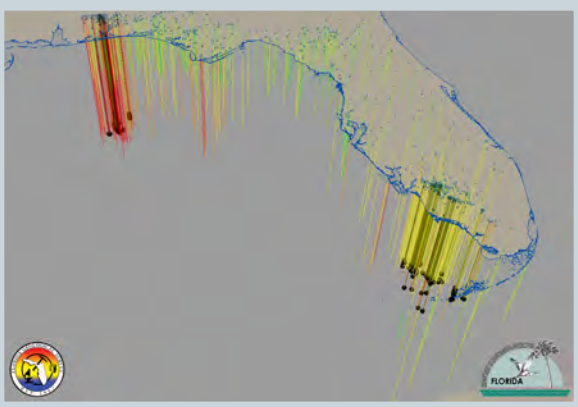
Florida

New & Legacy Data

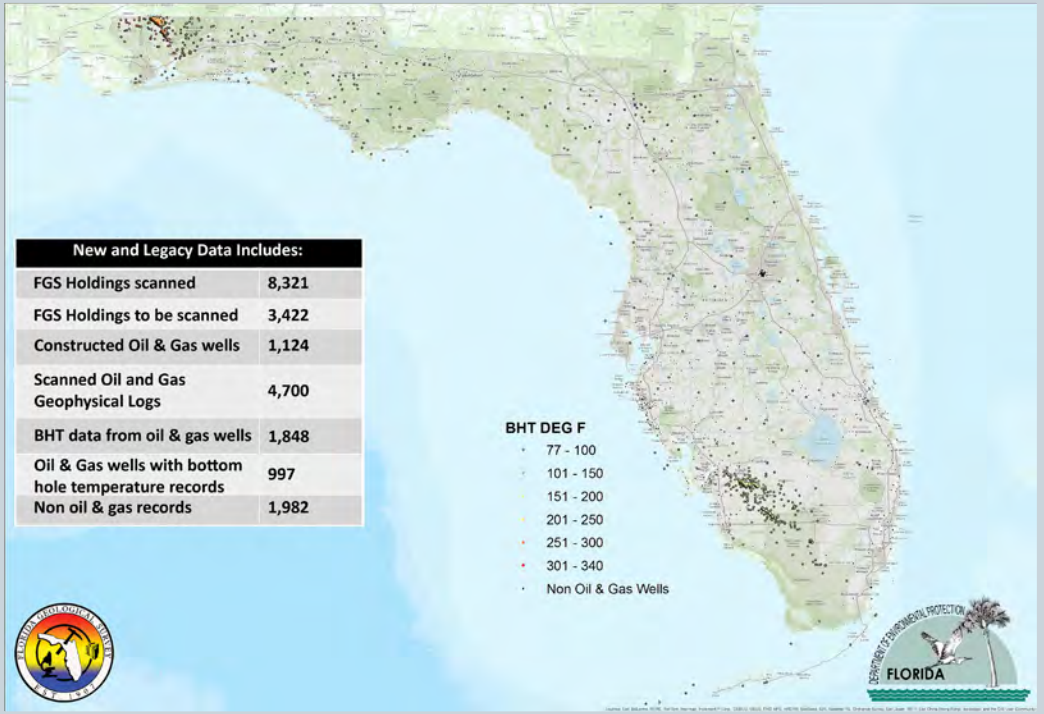
The data collected, digitized and organized through the NGDS has meshed nicely with the Florida Geological Survey's (FGS) efforts to rescue and provide public access to historic and at-risk geological datasets. The FGS scanned more than 4,000 oil and gas geophysical logs, documenting borehole temperatures from oil and gas wells. Supplementary records from other Florida agencies were compiled and the FGS documented nearly 2,000 temperature records from shallow wells throughout the state. The FGS continues to add to its digital collections by organizing and scanning geophysical logs as well as building metadata sheets for that data.

The main usage of geothermal energy in Florida is centered on heating and cooling projects for households and office buildings. An ongoing project in northwest Florida involves a geothermal co-production demonstration unit attached to an oil and gas well. The project produced somewhere between 0.20 and 1.0 MW of energy by skimming the hot water that is co-produced with the oil and gas. ■



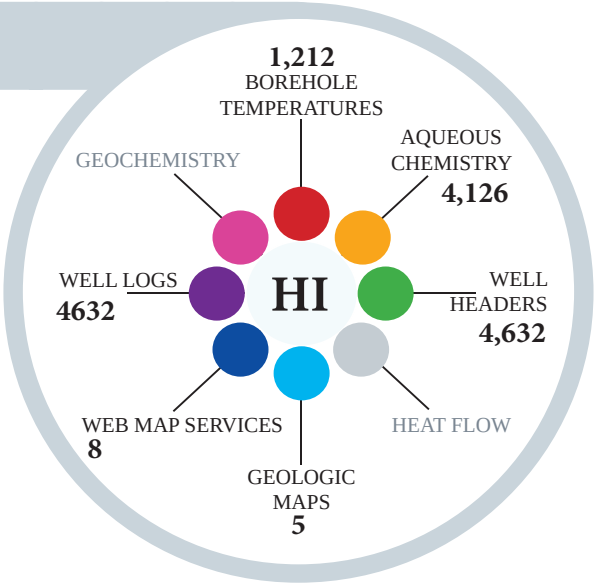


A three-dimensional image demonstrating borehole temperature at depth; longer lines represent deeper wells; green lines represent cooler borehole temperature measurements; and warm colors, like yellow and red, represent hotter borehole temperature measurements.



A compilation of digital geothermal data provided by Florida for NGDS, including FGS Holdings, oil and gas wells, well logs and borehole temperature records.

Hawaii



Industry Benefits

The NGDS project, run through University of Hawaii (UH), fostered unprecedented communication between multiple state agencies. Staff of the Department of Land and Natural Resources (DLNR), the Department of Business and Economic Development and Tourism Energy Office, large landholders, and the Department of Hawaiian Home Lands have cooperated with the UH team in compiling data and reports to incorporate into NGDS. Work with these organizations has led to further collaborative efforts and additional funding resources to develop new information on geothermal resources in Hawaii.

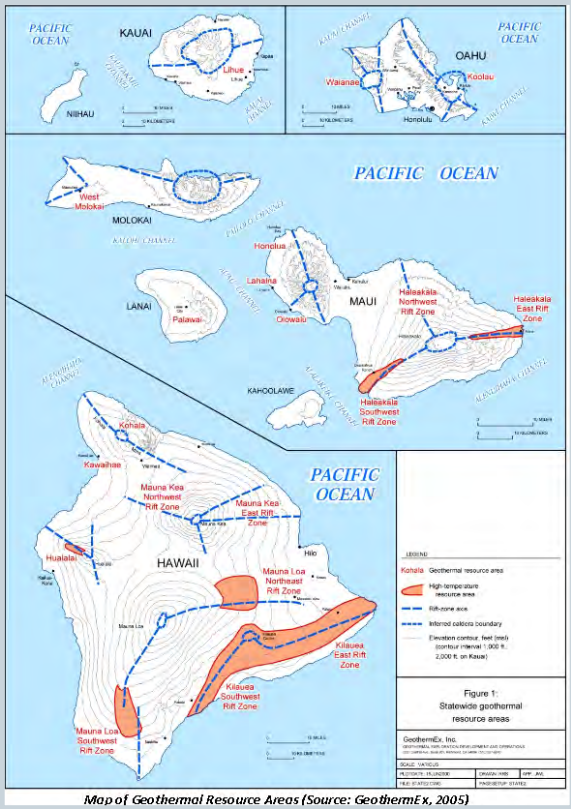
This glowing vent developed at the base of the east wall of Halemaumau Crater at the summit of Kilauea Volcano, Hawaii, in March 2008. (CREDIT D. Dzurisin (USGS).

Document sources include the Hawaiian Electric Company, the DLNR, the State Archives, Hawaii Natural Energy Institute, and UH researchers. The acquired documents serve as a reference resource for the DLNR-Engineering Division, which is responsible for regulating geothermal drilling in the state. In 2012, a total of nearly 10,000 files were downloaded from this site by ~6000 viewers.

Applied Research

Geothermal resources represent Hawaii's only economically viable source of base-load power. The NGDS project has stimulated cooperation between key players in geothermal development, provided enhanced accessibility to a broad suite of data relevant to geothermal resources, and galvanized new projects on geothermal exploration. The NGDS project is therefore making a critical contribution to Hawaii's transition to a more sustainable energy supply.

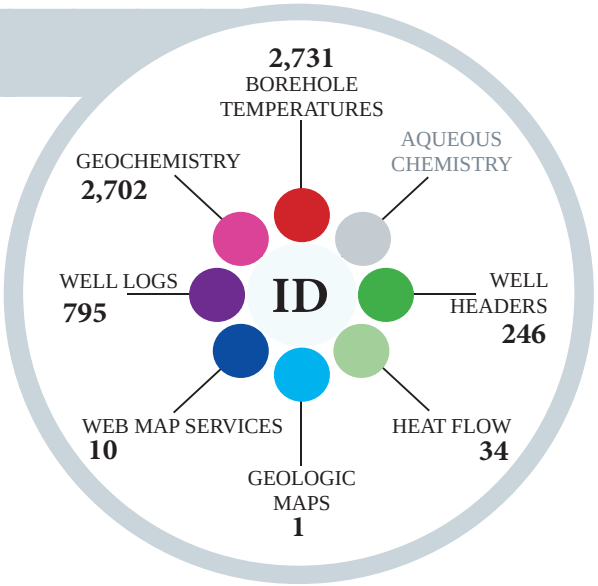
One of the most significant projects that will benefit from the NGDS is a new state-wide assessment of geothermal resources. This work involves magnetotelluric (MT) surveys, covering all the major geologic structures in the state with geothermal resource potential.



Geothermal Resource map shows potential geothermal energy production sites on the Hawaiian Islands.

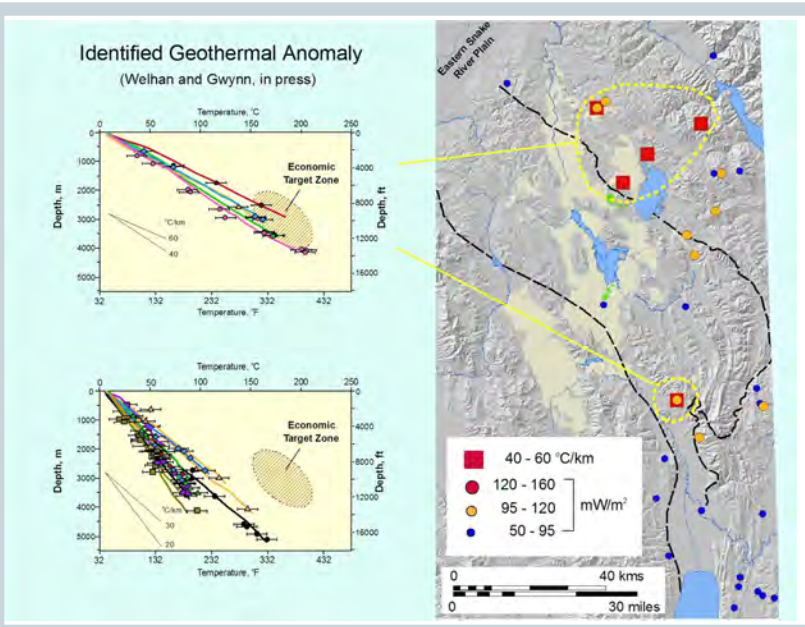
Easy access to the results of prior geothermal work, as well as to the site-specific hydrologic and geologic data contained with the Water Commission files, will tremendously enhance the state's ability to interpret and analyze the MT datasets. NGDS project activities are also enabling the State's Commission on Water Resources Management to transition entirely to an e-document system of water resource oversight and permitting. ■

Idaho



Geothermal Exploration & New Discoveries

Data compiled by the Idaho Geological Survey (IGS) for the NGDS project led to the detection of a previously unrecognized, drillable high-temperature geothermal resource. Idaho's NGDS supplemental heat flow drilling project focused on understanding the high-temperature potential of the Blackfoot volcanic field, one of the largest regions of young, bimodal (basalt-rhyolite) volcanism in the interior U.S. Thought to be a potential geothermal resource in the 1970s and 80s, temperatures from preliminary drilling near the China Hat rhyolite dome were disappointing, and due to advective thermal overprinting by shallow groundwater flow, shallow heat flow drilling was believed to be ineffective. It was only when the agency began to compile NGDS data on deep oil and gas exploration wells that a pattern began to emerge, revealing bottom-hole temperatures consistently higher than expected. Using Looking at the multiple data sets from the NGDS project, IGS re-analyzed data and found four wells with acceptable flow rates and temperatures of 150–200°C at just 2.5 to 4 km deep as The Blackfoot volcanic field may prove to be an economically valuable geothermal resource. IGS credits the NGDS project for facilitating the discovery of these exciting new resources and is compiling additional information for the NGDS project.



Evidence of a hidden, high-temperature magmatic-hosted geothermal resource—a potential location for geothermal development.

IGS recently collaborated with the Utah Geological Survey and Utah's Energy and Geoscience Institute on a Department of Energy proposal to examine the reservoir characteristics of the Blackfoot Volcanic Field and the Black Rock Desert prospect in central Utah (another resource revealed thanks to the NGDS project). The proposal was unsuccessful, but IGS obtained seed funding from NSF's "SedHeat" Research Coordination Network and convened a proposal development workshop to focus research interest on the Idaho prospect. Additional papers and projects are planned, including research and exploratory drilling to gather ▸▸

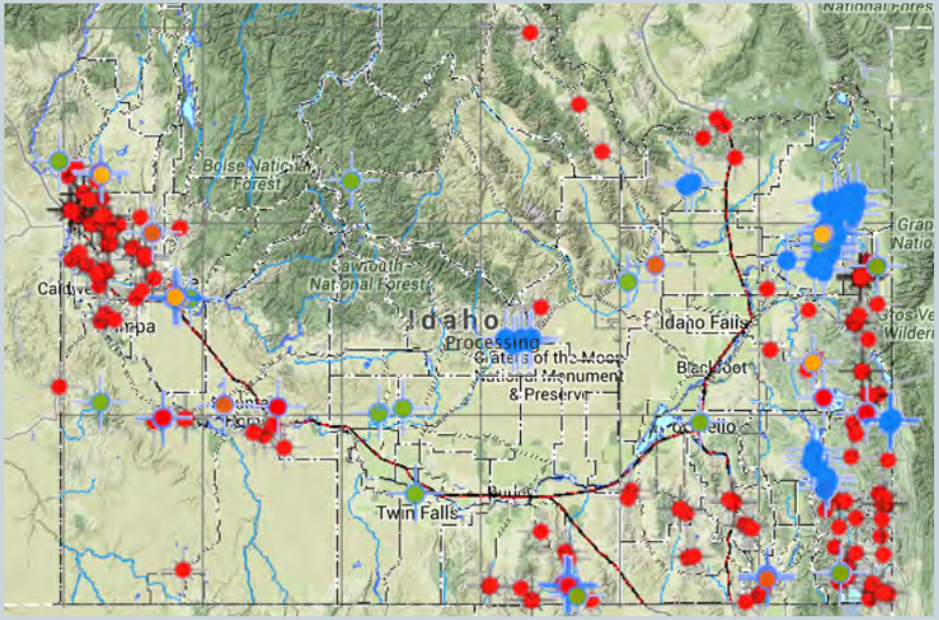
The China Hat rhyolite dome (ca. 2 km across at its base) is the largest of three domes in the central Blackfoot Volcanic Field. Having erupted ca. 58,000 years ago, it represents the most recent phase of volcanism from a magma chamber at 12–14 km depth that is believed to supply the heat encountered in shallow sedimentary reservoir rocks of the Idaho Thrust Belt.

data regarding the suggested presence of large quantities of lithium in some deep wells – possibly another previously unrecognized potential economic resource!

Data Management & Best Practices

In addition to data compiled and exposed via the NGDS, several new features were created for the IGS web site to increase exposure of geothermally relevant data:

- a searchable oil and gas well database and interactive map interface (*http://www.idahogeology.org/services/Oilandgas/*)
- an Idaho interactive map interface to geothermally relevant data, including geology, active faults (Miocene and younger), thermal springs and permitted geothermal wells (*http://www.idahogeology.org/webmap/*)
- fully digital 1:750,000-scale geologic coverage of the state (*http://cloud.insideidaho.org/arcgis/rest/services/IGS/GeologicMapIdaho/MapServer*)
- an Idaho geothermal web page that is currently under development ■



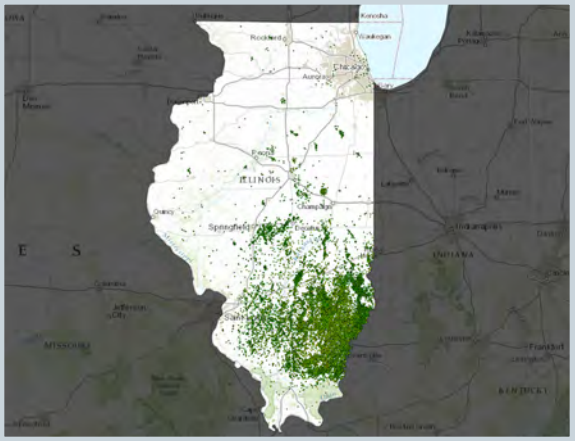
Borehole temperature measurements in the region of Boise National Forest; borehole temperatures can be indicators of geothermal resources.

Illinois

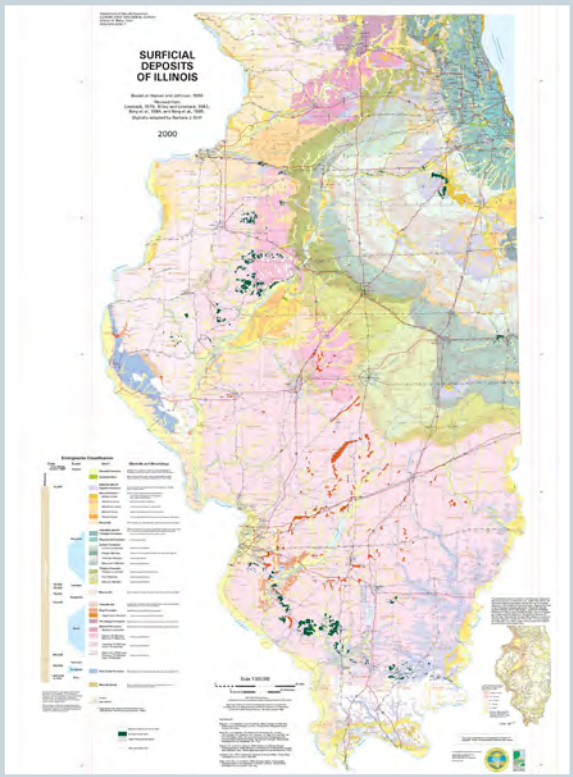
New Technology & Skills

As an NGDS hub hosting data from multiple states, the Illinois State Geological Survey (ISGS) not only compiled data but developed new software. In 2010, ISGS created programing to extract data from their existing state oil and gas database to extract borehole temperature data and drill stem test data to map to the NGDS content models.

In addition, ISGS developed software to data mine information and also manually add DST data to the database in early 2011. These proprietary tools were uploaded to GitHub to create and maintain services for use by other hubs or states that are interested in self-hosting their data.



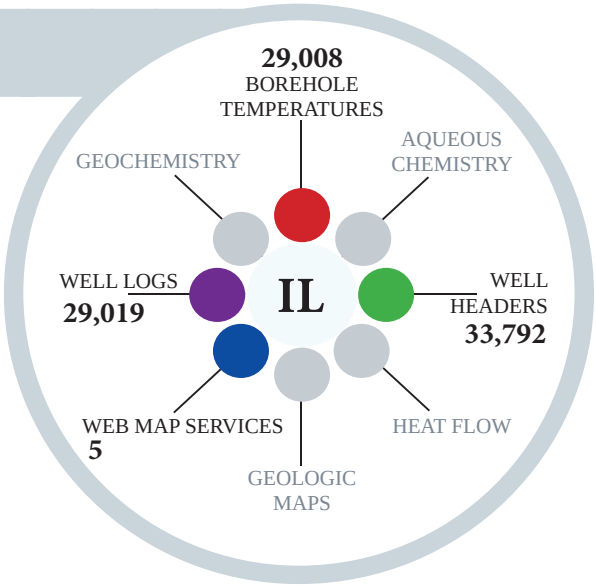
A web map service conveying borehole temperature measurements provided by Illinois is part of their massive data collection efforts for the NGDS project. The ISGS contributed over 29,000 borehole temperature observations.



A 2000 geologic map of surficial deposits in Illinois produced by the Illinois State Geological Survey describes sediments useful for a variety of geospatial analyses.

Data Collection & Standardization

Illinois provided tens of thousands of bottom-hole temperature values and associated data, as well as more than 16,000 drill stem test results. The ISGS contributed digital copies of all temperature logs in the ISGS collection, compiled a bibliography of ISGS geothermal-related publications, scanned oil well folder contents, and researched collections of old geophysical logs donated by industry to find logs lacking in the Survey's collection.



A team of approximately a dozen students identified electronic record information and adding numerous new data fields to files in preparation for their contribution to the NGDS and ensuring Tier 3 interoperability with NGDS schemas and standards. Over 145,000 logs were queried for temperature data to transfer to the borehole temperature Tier 3 template for interoperable web services.

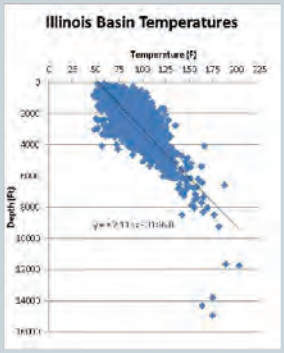
Derivative Products

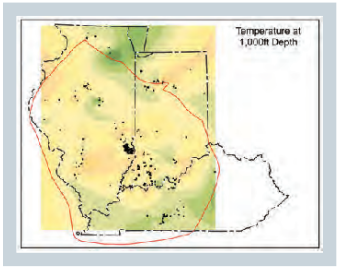
One derivative product of the effort was a joint report by the ISGS, Indiana Geological Survey, and Kentucky Geological Survey titled "Development and Application of a New Geothermal Database for the Illinois Basin," by Tiffany A. Proffitt, Kevin M. Ellett, Charles W. Zuppann (Indiana Geological Survey, Bloomington), Melony Barrett, Bryan Huff, Chris Korose, Alison Lecouris ▸▸

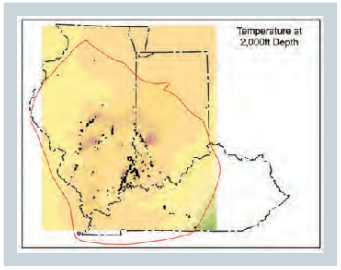
(ISGS), and T. Chase Noakes (Kentucky Geological Survey, Lexington), which was presented at a meeting of the American Association of Petroleum Geologists. Such a comprehensive effort at subsurface temperature data compilation and analysis had not been attempted since the 1970s. ■

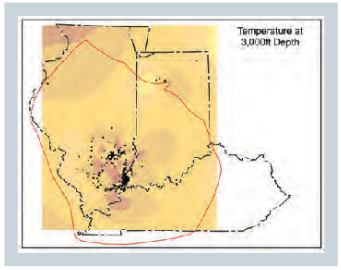
Depth-Based Temperature Slices

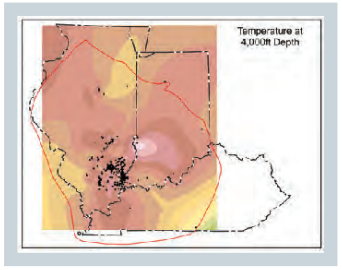
Temperatures and gradients for wells drilled to depths of 1,000 ft, 2,000 ft, 3,000 ft, and 4,000 ft were mapped. In order to ensure a suitable data size, a buffer of up to 75 feet was used for each depth range. Maps confirmed the general trends of increasing temperature, and decreasing gradients with depth. Previous researchers such as Harrison et al. (1982) have found similar results in other regions and further analysis is underway to determine how low thermal conductivity sediments may be enhancing temperature gradients due to a thermal blanket effect. This depth-based analysis also helps to identify regions where advective heat transport may be significant from shallow groundwater flow systems.

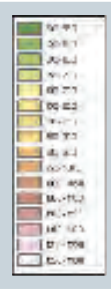


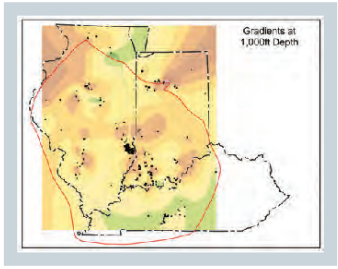


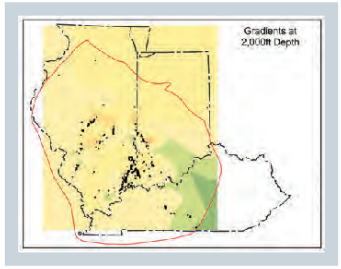


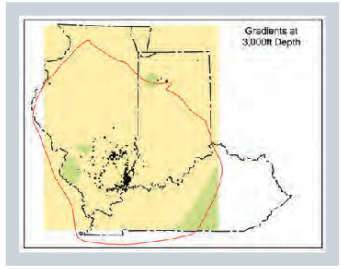


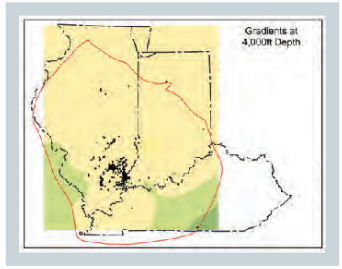


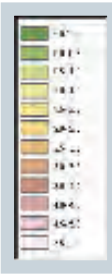








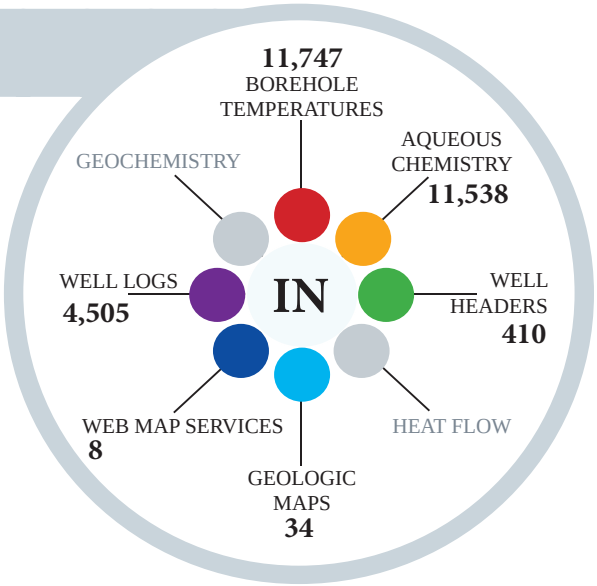




Indiana

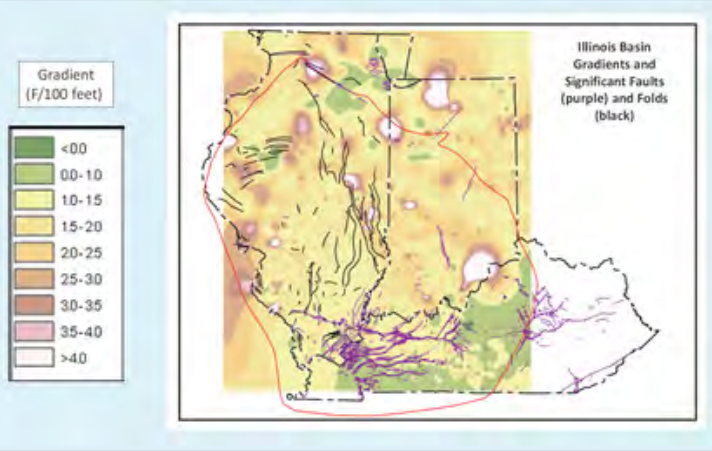
Ground Source Geothermal

Geothermal interest in Indiana pertains largely to ground source heat exchange systems. The NGDS project helped bolster this industry with new maps, standardized data, and better information tracking. In addition, the project excavated new pits and installed new monitoring data for shallow near-surface temperatures and thermal properties.



Data Management & Best Practices

Indiana Geological Survey (IGS) is now a self-serving node of the NGDS, serving data using GeoPortal. The survey reports improved capabilities for serving real-time data from sites throughout Indiana. The project also implemented a standard method for attributing geologic maps in tabular/vector GIS formats for public distribution. Several geologic maps previously not in digital format are now georeferenced or are raster images for viewing in GIS.



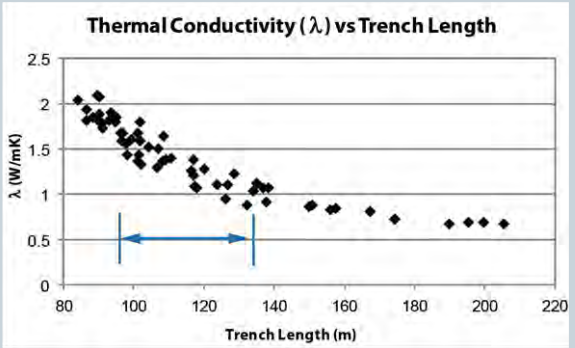
A map of geothermal gradients and significant faults in the Illinois Basin; can be used to facilitate geothermal exploration

Geothermal Discoveries & Exploration

Geothermal gradient analysis for the Illinois Basin region from over 26,000 wells was published in an American Association of Petroleum Geologists (AAPG) search and discovery article. These results are being used to parameterize reservoir temperatures in new modeling studies of carbon capture utilization and storage funded by the DOE.

Industry Benefits

Data from the NGDS project has contributed to two peer reviewed publications: "Controls on ground-coupled heat exchange in glacial sediments" by Shawn Naylor, Kevin M. Ellett, and Andrew R. Gustin (submitted to the journal Renewable Energy); and "Evaluation of regional-scale estimates of shallow thermal conductivity for application in horizontal geothermal heat pump design" by Kevin M. Ellett and Shawn Naylor, an invited paper that will be published in a Geological Society of America Special Paper titled "Geothermal Energy: An Emerging Resource." ■



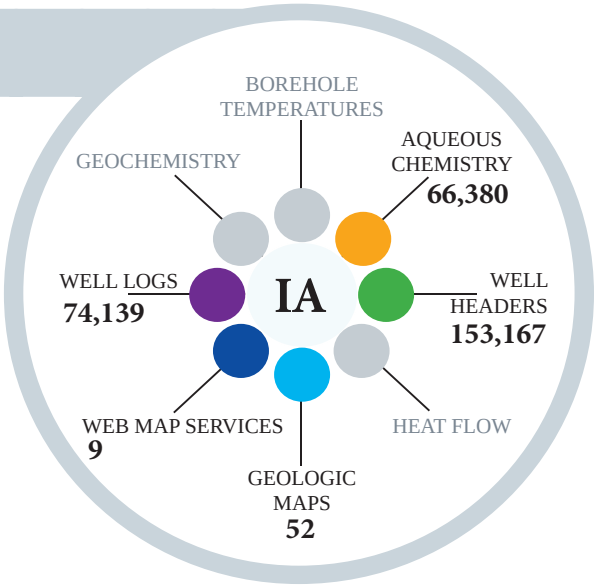
Thermal conductivity of unconsolidated sediments plotted vs. calculated horizontal trench length; this graph demonstrates optimal trench lengths for ground-source heat pumps.

Iowa

Interagency Cooperation

Data collected from the Iowa Department of Natural Resources (IDNR) are deployed in Iowa Suitability Maps through Web Map Services. The suitability map resource is a metadata compilation for 20 map data layers, including vector and raster data, developed to assess site suitability for heat pump development. Information includes karst potential, known wells, contaminant sources, and hydrogeologic properties in the Dubuque County area. The data layers are hosted by the Iowa Geological and Water Survey. Data are stored in Web Mercator projection ESPG:102113 and are available as ESRI ArcGIS Server layers at *http://programs.iowadnr.gov/arcgis/rest/services/Projects/Geothermal/MapServer*. This dataset is delivered as an ESRI file geodatabase. Iowa Hydrolgeologic Map Server provided for 45 various layers including formation thickness, water yield, surface elevation, water levels, sinkholes, underground mines, and aquifers.

The Illinois hub, serving all data the IDNR contributed to the NGDS project, also serves the IA OneGeologyMap with 3 NGDS layers: *http://repository.stategeothermaldata.org/repository/resource/50ec3aefb656b70647f32e38bcf8521a/* ■



Iowa OneGeology map has three layers. Three (3) layers included in the Web Map Service (WMS) at 1:500,000 map scale and are as follows: geologic age, lithology and lithostratigraphy.

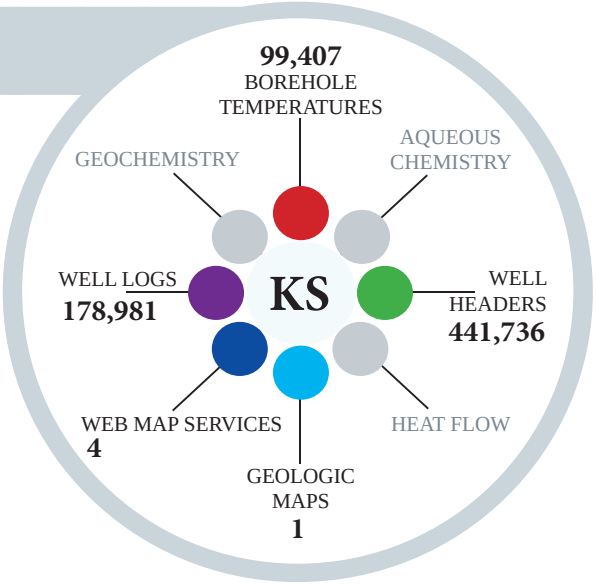
Kansas

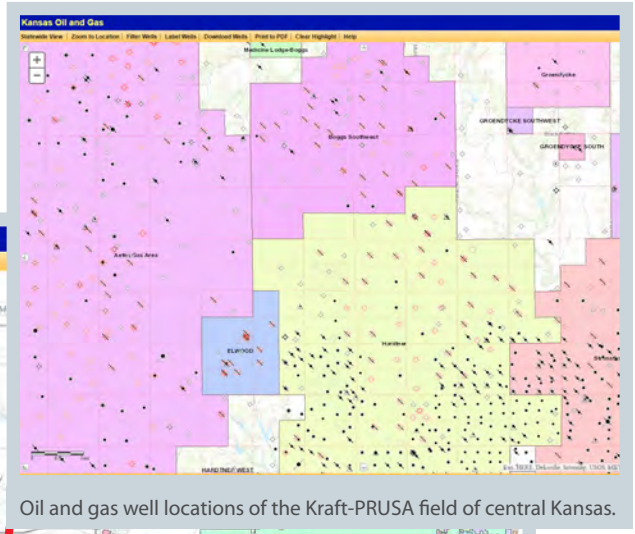
New & Legacy Data

Over the course of the project, the Kansas Geological Survey (KGS) provided NGDS with data for well headers, drill stem tests, borehole temperatures, and well logs. Map Services now available through NGDS include all of these features. Data discovery and processing was performed by KGS to compile Drill Stem Tests from multiple sources, including the Kansas Online Automated Report (KOLAR) system. This data was previously unavailable to the general public.

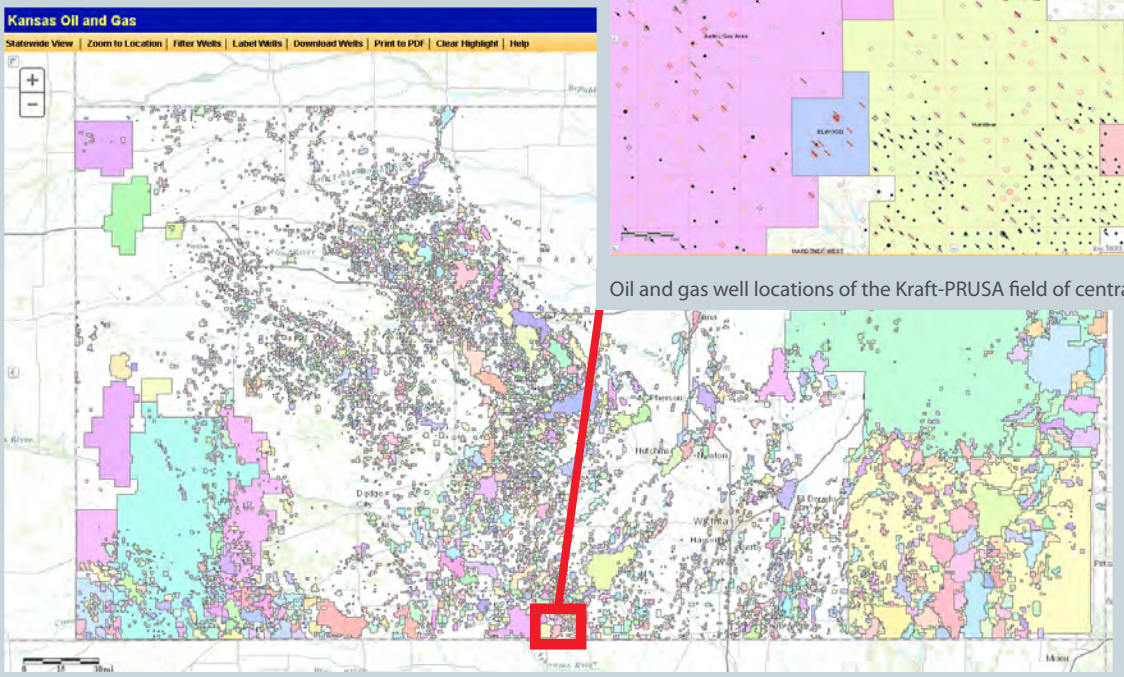
Geologic Maps

The 1:24K Surficial Geology map service was enhanced with 15 additional counties in the first year and 24 counties in the second year. ■



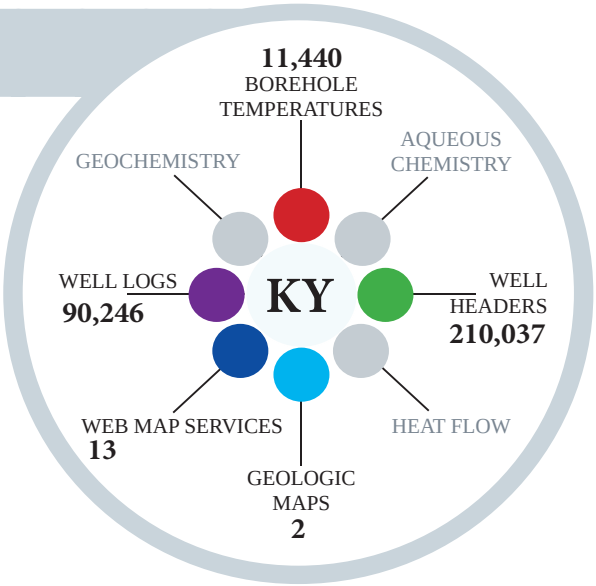


Oil and gas well locations of the Kraft-PRUSA field of central Kansas.



Interactive Kansas oil and gas fields with legend, located online at *http://maps.kgs.ku.edu/oilgas/index.cfm*.

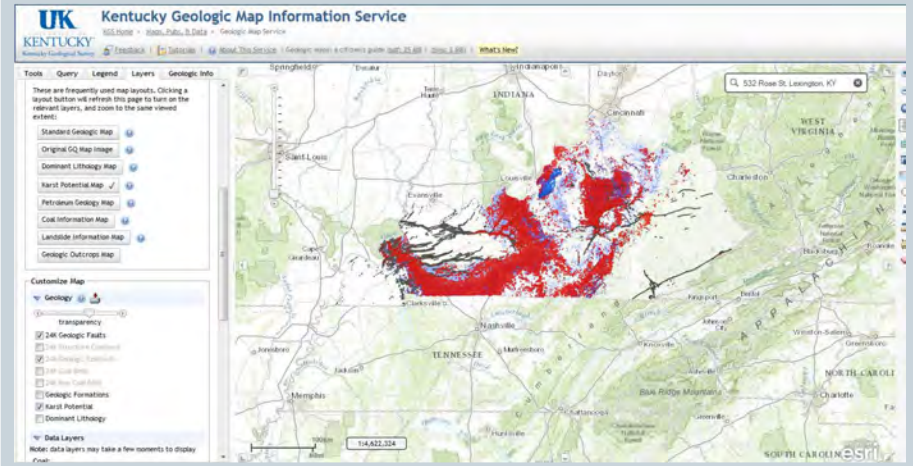
Kentucky



New Data Viewer

The Kentucky Geological Survey (KGS) acts as the Southeastern hub for the NGDS, serving data for ten additional states. The KGS submitted the state's 707 published 1:24,000 geologic quadrangle (GQ) maps in digital format (1996-2006). The 1:500K map data are also made available through an interactive web map application, the "Kentucky Geologic Map Information Service," (*http://kgs.uky.edu/kgsmap/kgsgeoserver/viewer.asp*) to provide interactive and customized geologic maps with links to the related data. These data provide users with detailed earth science information for Kentucky concerning geologic rock units and mapped surface faults. ■

Standard Geologic Map layer of the Kentucky Geologic Map Information Service.

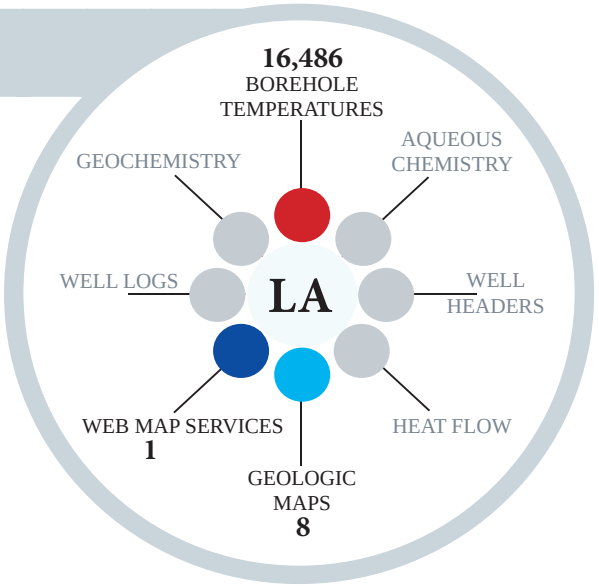


Kentucky Worse – Karst Potential layer of the Kentucky Geologic Map Information Service.

Louisiana

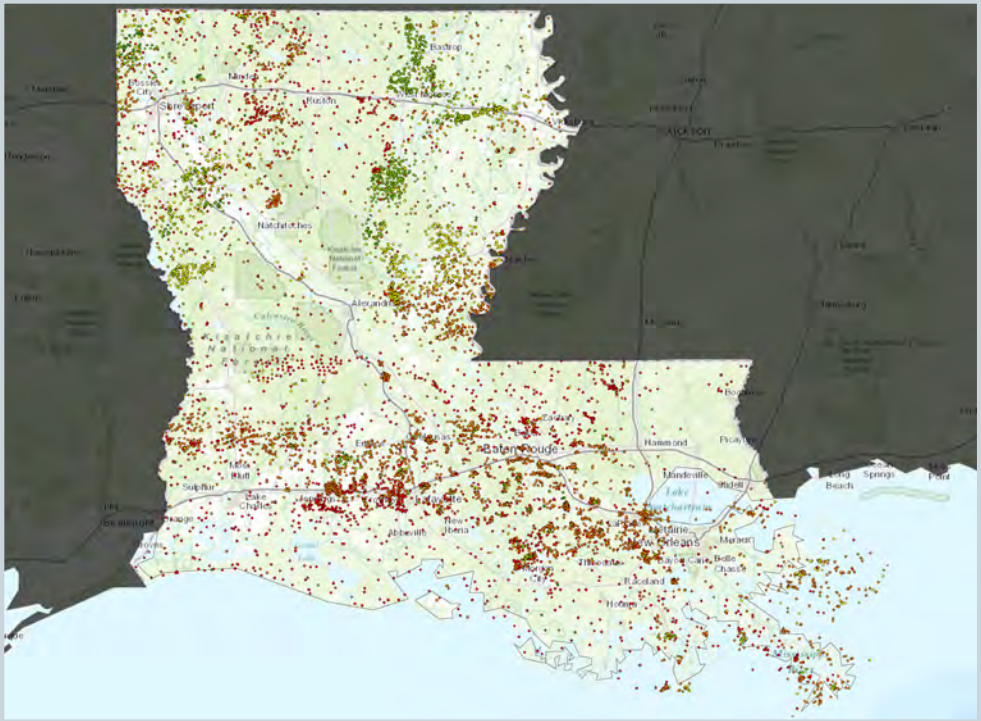
Inter-industry Legacy Data

The Basin Research Energy Section (BRES) is the oil, gas, and coal research section of the Louisiana Geological Survey (LGS), and it performs research into the nature and occurrence of oil, gas, and coal both in Louisiana and around the world. In addition to supplying much of Louisiana's data contribution to the NGDS, BRES disseminates its research and technical information widely, especially to the independent oil and gas producers of Louisiana, thereby enabling more economically efficient exploration and production operations. Services include geology maps, updated aero-mag maps and gravity maps.



Interest in Geothermal Energy

The January 2012 edition of the *AAPG* (American Association of Petroleum Geologists) *Explorer*, which has worldwide circulation, features Chacko John, LGS director, state geologist and research professor, discussing the DOE's research program on geopressured-geothermal resources. The article recognizes LGS's previous work in geothermal energy along the Gulf Coast, including a DOE-funded program from 1975 to 1992 that gathered geological, engineering, environmental and economic information to determine the viability of development of the resource. John said the current interest in alternative energy sources indicates that now may be the appropriate time to pursue geopressured-geothermal resource development. ■

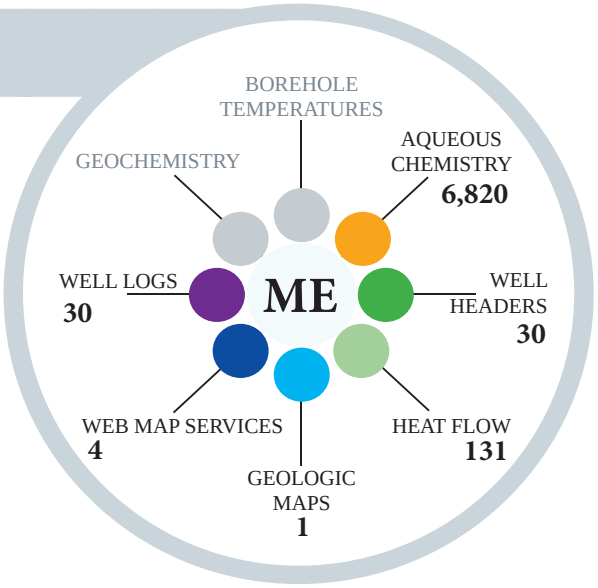


Borehole temperatures collected from BRES available through NGDS displayed on a map of Louisiana though arcMap.

Maine

New & Legacy data

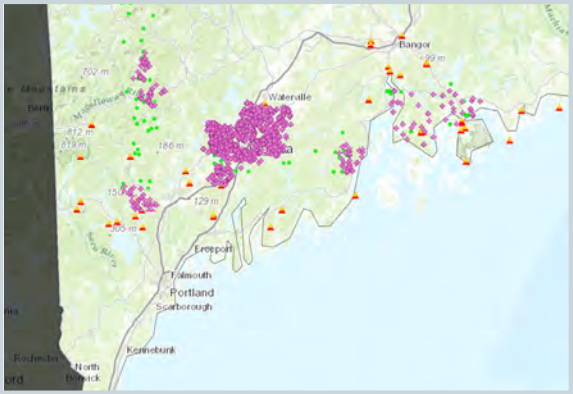
In the first phase of the project, existing unpublished temperature, thermal conductivity, and heatflow data were compiled from various sources, reformatted, and submitted for inclusion in the National Geothermal Data System (NGDS). Existing unpublished groundwater quality data, including near-surface groundwater temperatures, collected from private water wells through various Maine Geological Survey (MGS) projects were also compiled, reformatted, and submitted. A total of nearly 1,500 new records were added to the NGDS in this phase. In the second phase, new temperature profiles for 30 deep bedrock boreholes were collected and formatted for NGDS.



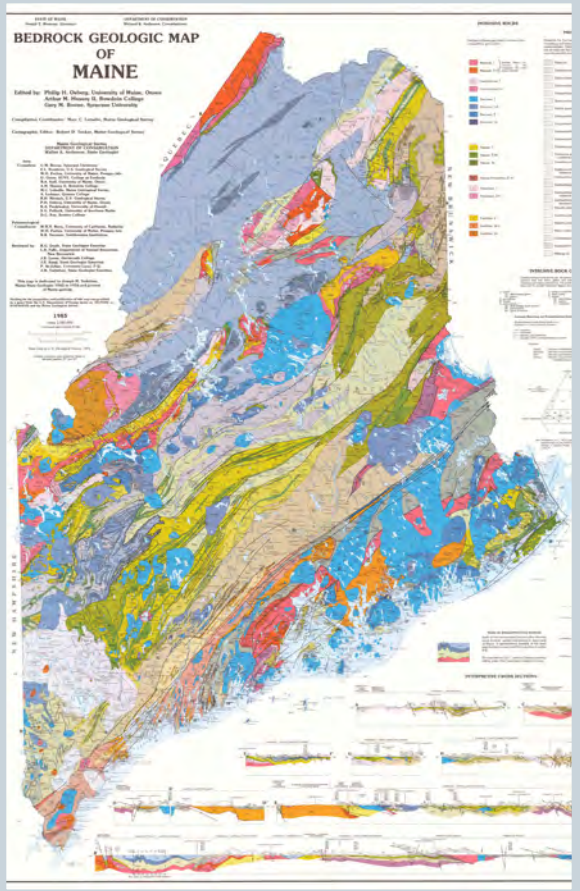
Industry Benefits & Applied Research

The Geothermal Map of North America (Blackwell and Richards, 2004) identifies New England as having among the highest heat-flow values in the eastern United States (60-80 mWm-2). Funded through the NGDS Supplemental proposal process, MGS acquired new geothermal data. New derivative products include the Maine Heat Flow database, and the Maine Aqueous Well Chemistry data map, available in WMS and WFS.

One data set, the Columbia Groundwater Quality database, contains 340 records collected in 2007 from private homeowner water wells. The database is part of a larger study on the occurrence of naturally occurring arsenic in private water wells. ■



Composite map of Maine aqueous chemistry and heat flow data points as they appear in the NGDS Web Map Service.



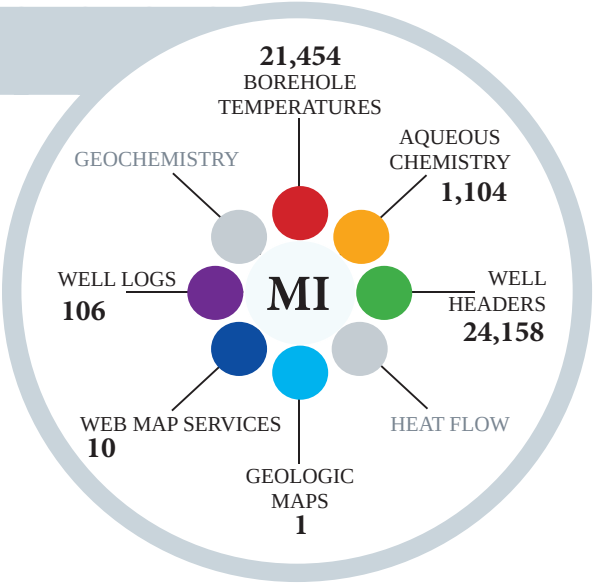
Bedrock geologic map of Maine.

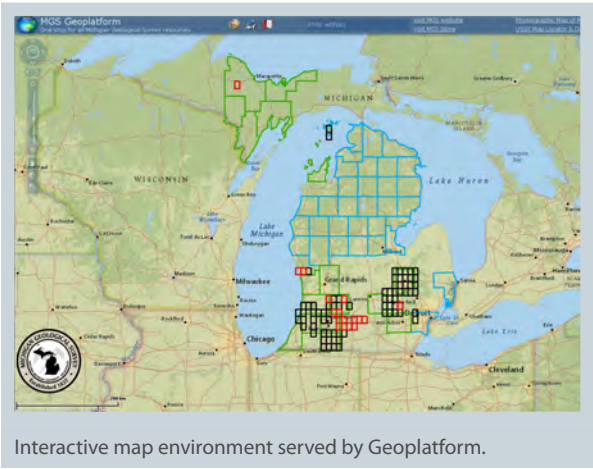
Michigan

New & Legacy Data

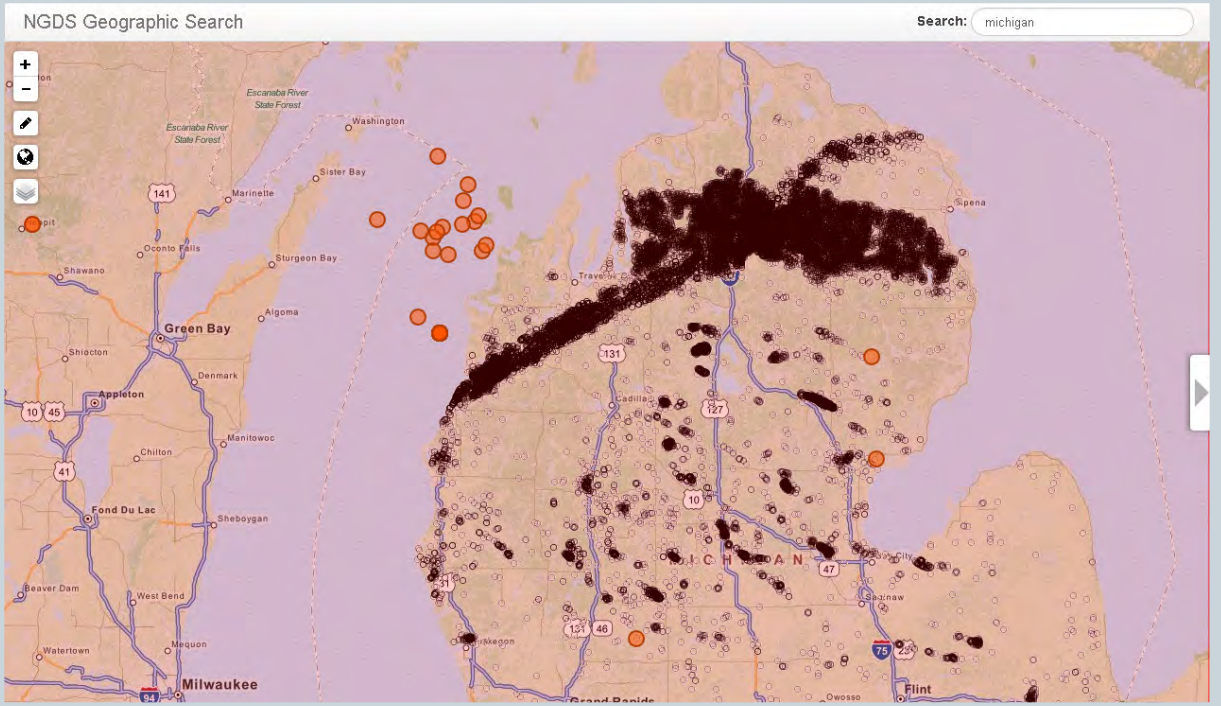
Data collected over two years for the NGDS project in Michigan includes bottom-hole temperature measurements, drill stem tests with temperature and pressure data, and water chemistry measurements with a total of 36 elements or compounds analyzed. In the third year of the project, focus shifted to shallow temperature data, including the characteristics of soils and shallow glacial deposits, thickness of glacial deposits, and depth to water table. More detailed information was gathered on the composition of geological materials, monthly and annual temperatures, and deeper subsurface flow to the surface. Locations of current ground source heat pump installations were also collected.

It is hoped that these measurements will bolster interest in geothermal heat exchange systems in Michigan. ■





Interactive map environment served by Geoplatform.

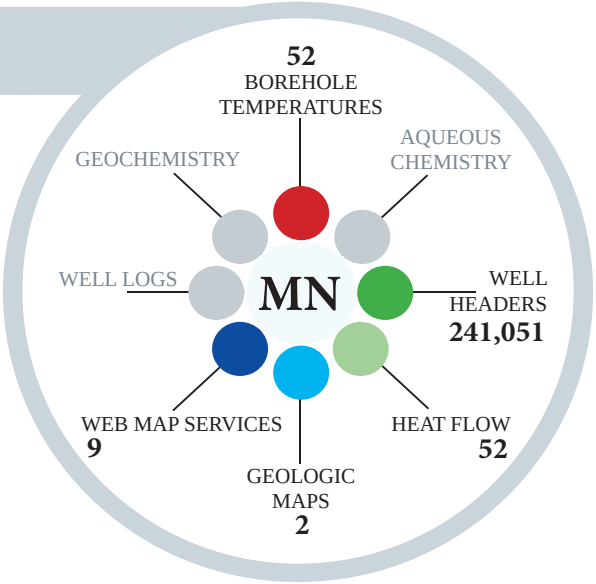


An NGDS well header search of Michigan shows numerous results.

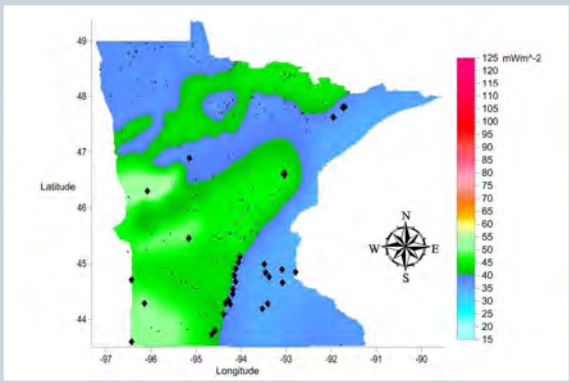
Minnesota

Applied Research

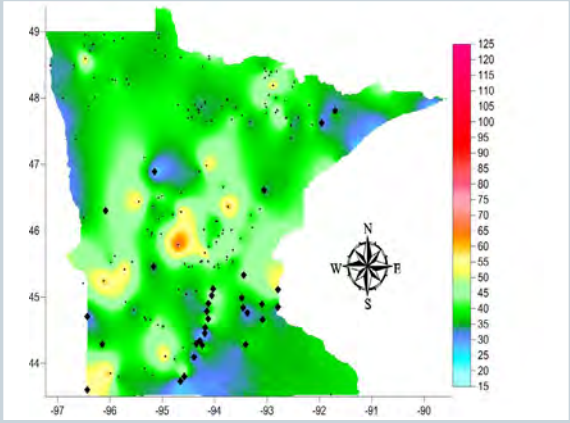
Compiled by the University of North Dakota team in conjunction with data on the Dakotas and Nebraska, Minnesota data has been used in two theses and in the Natural Resources Research Institute Technical Report, "New Heat Flow Map of Minnesota Corrected for the Effects of Climate Change and an Assessment of Enhanced Geothermal System Resources."



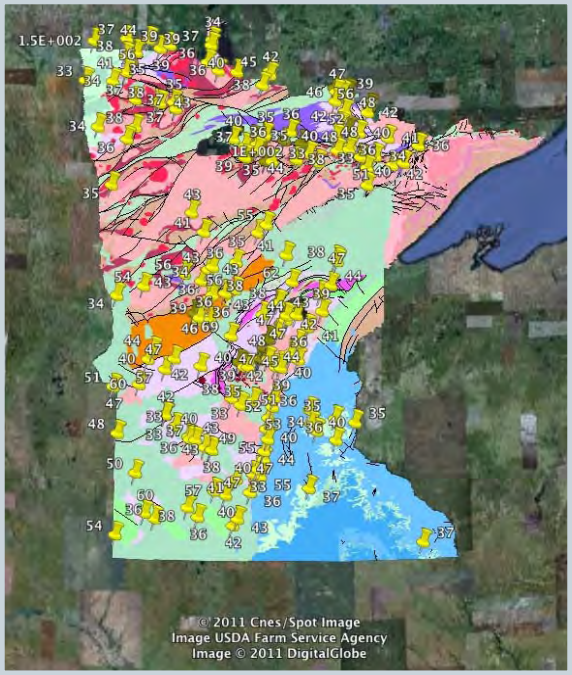
Researchers were able to pull data from the NGDS to show possible climate change/global warming signals. Particularly relevant information is the curvature of temperature against depth plot for 30 boreholes in Minnesota. Some data (from 43.5 N to 49 N latitude, logged in 2010 and 2011) show the combined effects of warming of the ground surface during the past 120 years and since retreat of the Pleistocene ice sheet ~11,000 years ago. ■



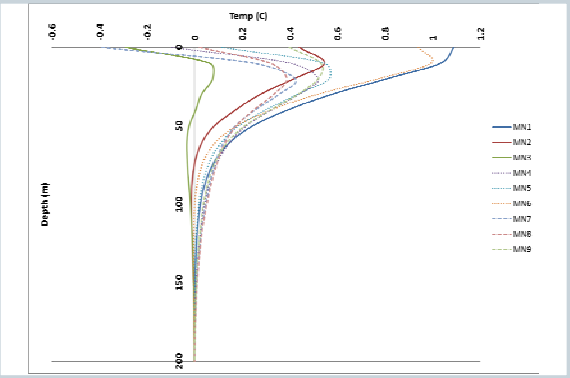
New heat flow map of Minnesota using both traditional heat flow measurements (diamonds) and calculated heat flow measurements (dots) based on the Q-A relationship. The colored vertical scale is heat flow in mW m-2. Horizonal and vertical exes on the map are in degrees latitude and longitude.



Heat flow measurements (diamonds) are systematically lower than the heat flow calculated from radioactivity (dots). This difference gives further evidence that a correction for post-glacial warming needs to be applied. The colored vertical scale is heat flow in an mW m-2 . Horizontal and vertical exes on the map are in degrees latitude and longitude.



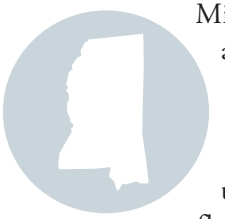
Heat flow locations on the geologic map of Minnesota. These data were used to create the new heat flow map of Minnesota by averaging the heat flow data over the geologic terrain it represented.



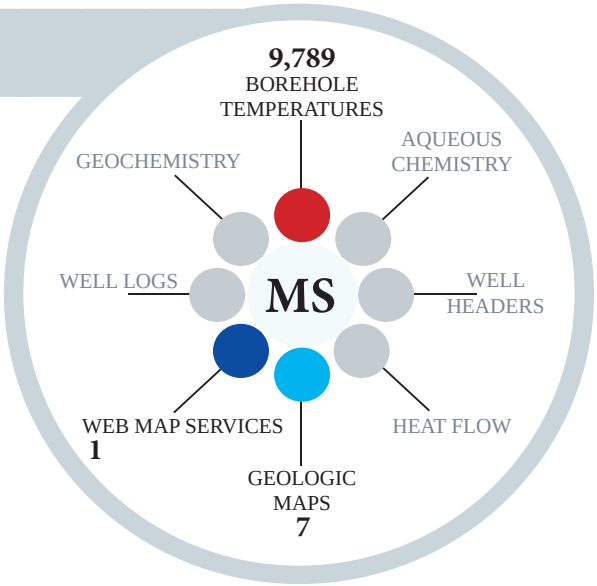
Recent warming perturbations affecting the nine different climate regions in Minnesota. Information captured from "Implications of Post-Glacial Warming for Northern Hemisphere Heat Flow", Gosnold, W. et al., 2011, GRC Transactions, Vol. 35.

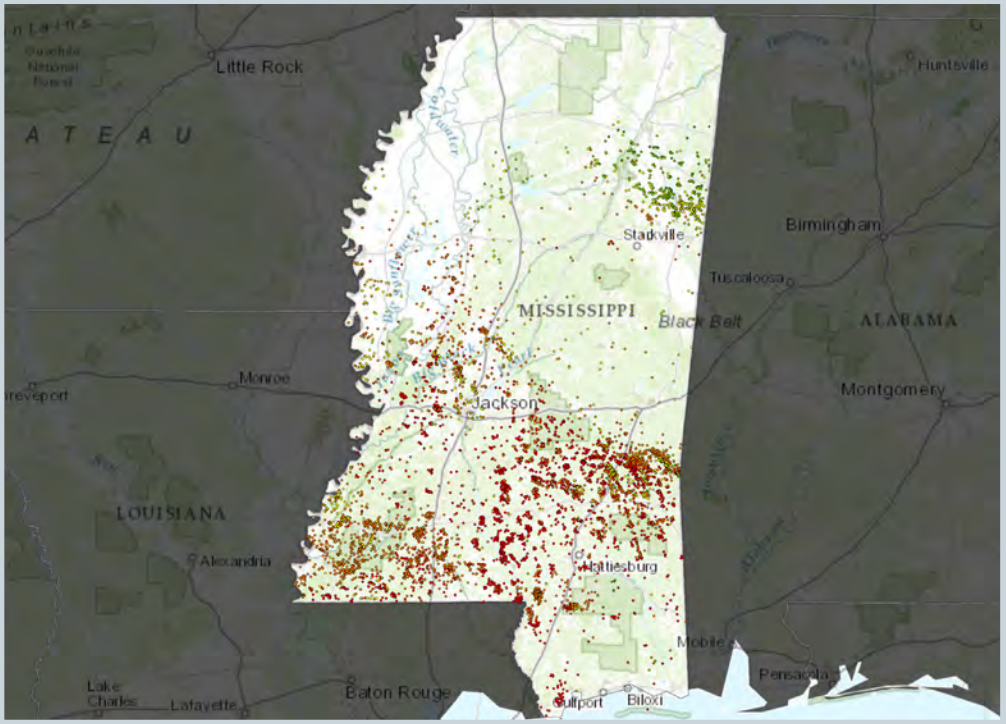
Mississippi

Legacy Data



Mississippi may lack geysers and active volcanoes, but has large volumes of warm salt water coproduced from oil and gas wells. Technology exists to extract useful energy from such warm fluids. Therefore the Mississippi Department of Environmental Quality (MDEQ) is concentrating efforts on collection of data about bottom-hole temperatures from the headers of geophysical logs in old files. The MDEQ delivered metadata for 18 Isotherm maps and documents and submitted production data for oil & gas field pools in machine-readable format. Borehole temperatures can appear as a web feature service in the NGDS catalog. ■





A web map service conveying borehole temperature measurements provided by Mississippi as part of the NGDS.

Missouri

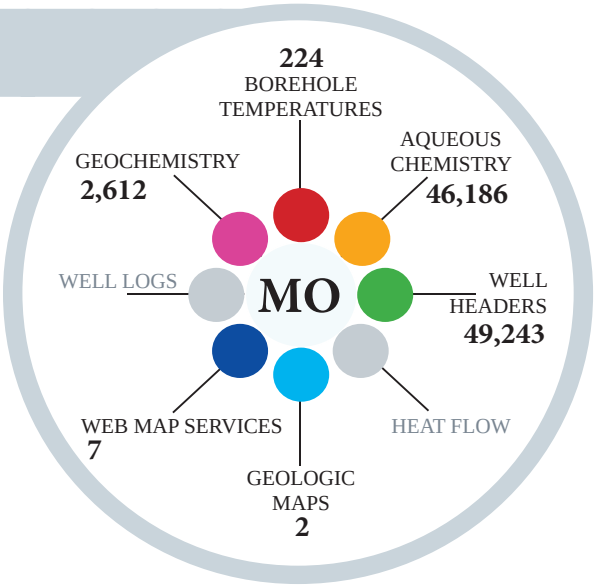
New & Legacy Data

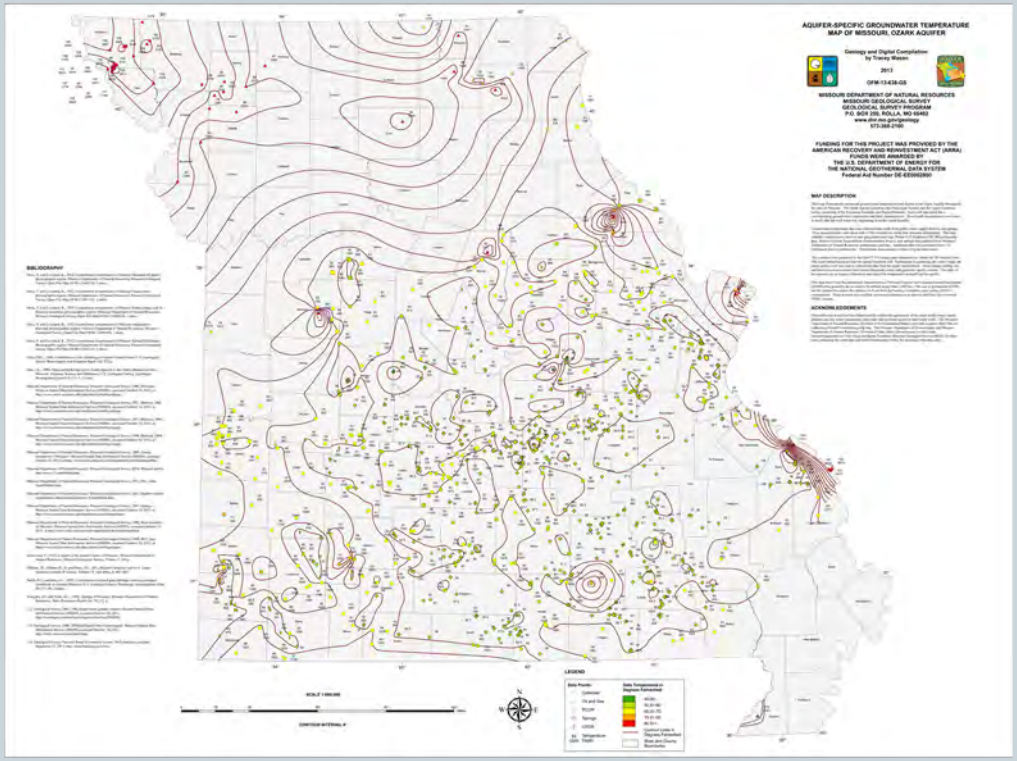
Though the state of Missouri does not have hot springs, fumaroles or any of the high to moderate temperature geothermal features found in some western states, Missourians use heat pumps to take advantage of the moderating influence of the earth to heat their homes in winter and cool them in summer. Data and information found on the NGDS and the Missouri Geological Survey Program website provide valuable assets to aid in development of geothermal resources.

Between May of 2011 to November, 2013, the Missouri Geological Survey produced several datasets and maps, including information on heat pump facilities, borehole temperatures, lithology logs with well headers, aqueous chemistry, rock chemistry and a geodatabase for a 1:500,000 scale map service. The maps include a depth to bedrock map as well as a set of aquifer-specific groundwater temperature maps. Much of the data were collected prior to 2011, but were not readily available.

Data Management

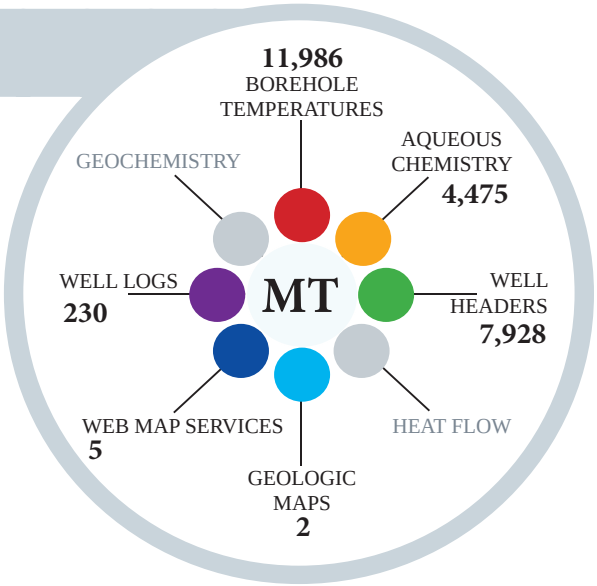
The Missouri Geological Survey developed a web page (*http://dnr.mo.gov/geology/geosrv/geores/geothermal.htm*) devoted to geothermal resources in the state with links to NGDS. This web page also includes information about heat pumps and geothermal usage in Missouri and an interactive geologic map. The Aquifer-Specific Groundwater Temperature Map of Missouri, Ozark Aquifer (right) is one of a series of aquifer-specific temperature maps presenting raw temperature and depth data for Missouri aquifers. ■





This map illustrates groundwater temperatures and depths in the Ozark Aquifer throughout the state of Missouri. Data used to make this map are available through NGDS.

Montana

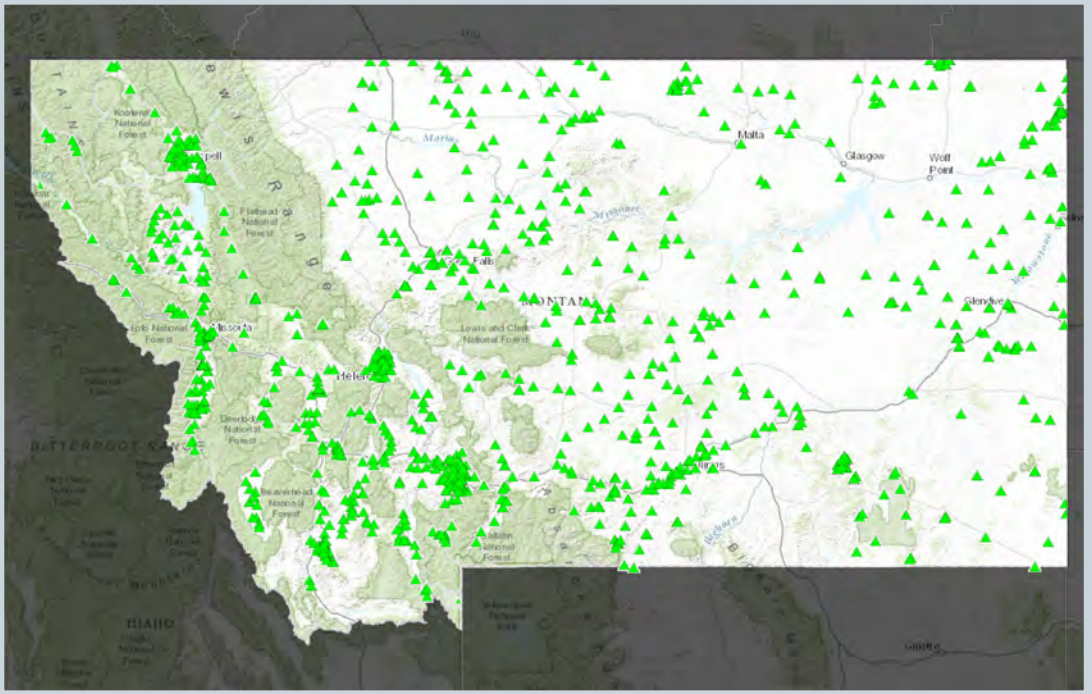


New Data Development Platform

The Montana Bureau of Mines and Geology (MBMG) is migrating their GIS data development processes from an Arc/Info environment to an ArcGIS platform. The MBMG's digital data availability has expanded to include web mapping services, made available through their ArcGIS Server, making spatial data available online. As of November 2013, services include our 1:500,000 geology maps, the Ground Water Iinformation Center database, and the earthquake catalog. Users can access MBMG's ArcGIS web services at *http://www.mbmg.mtech.edu/gis/gis-ArcGISservices.asp*.

New Data Format Standard

MBMG is part of the national effort to standardize the data format for geologic maps and data among state geologic survey agencies. The National Cooperative Geologic Mapping Project (NCGMP) is a USGS-sponsored program to develop standardized data layers for geologic data. The Montana Bureau of Mines & Geology participates in both the STATEMAP and EDMAP portions of the NCGMP. For more information, visit the NCGMP website at: *http://ncgmp.usgs.gov/*. ■

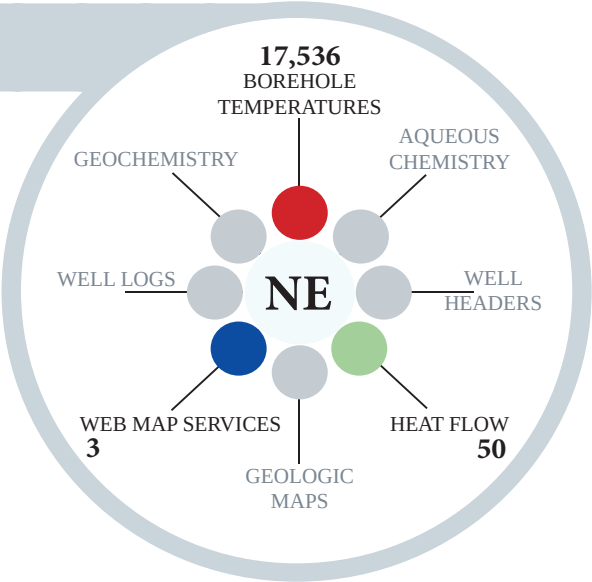


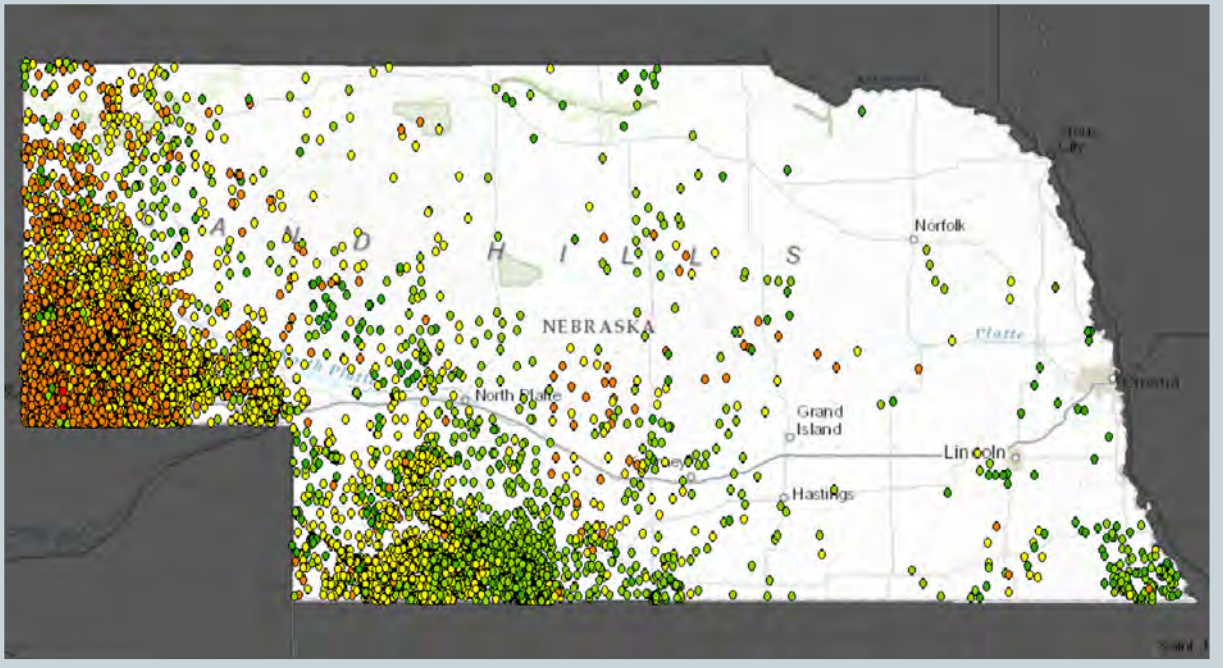
A Web Map Service conveying well construction fields for statewide long-term monitoring network wells provided by Montana's Ground Water Information Center. These wells are used to monitor groundwater, but the data they collect can be used to locate geothermal resources.

Nebraska

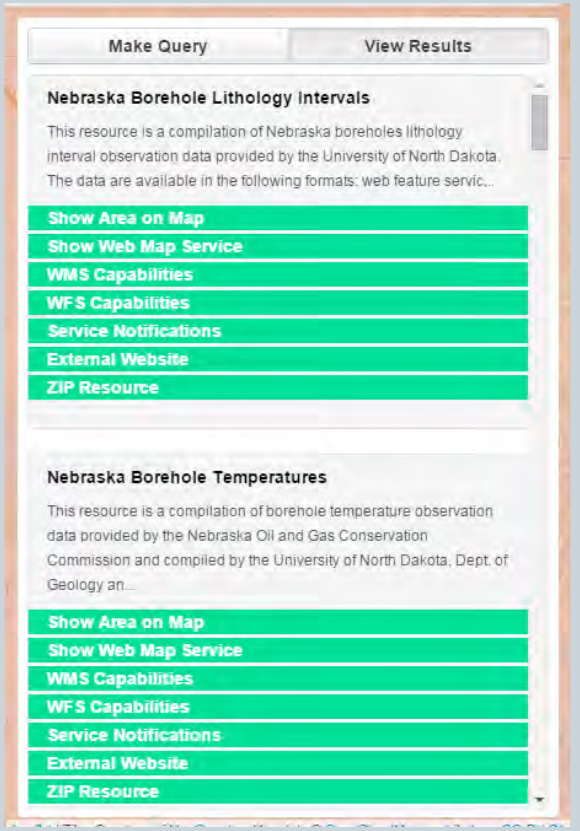
New & Legacy Data

Data for Nebraska was compiled largely by the University of North Dakota. Derivative data from Nebraska fueled one academic thesis, a North Dakota Academy of Science paper on the Chadron Arch heat flow anomaly, and Geothermal Resources Council papers on borehole temperature corrections in 2012 and 2013. ■





Borehole temperature data in the State of Nebraska.

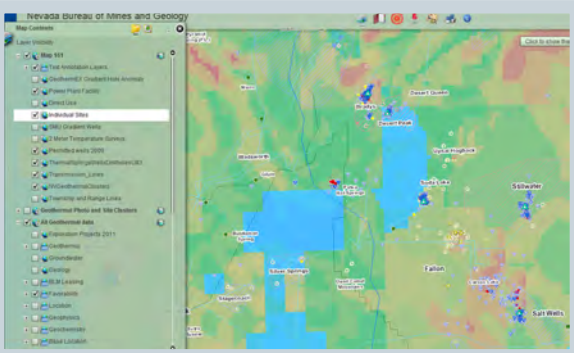


Data on borehole lithography, borehole temperatures, and heat flow are available through Web Map Services on the NGDS interface.

Nevada

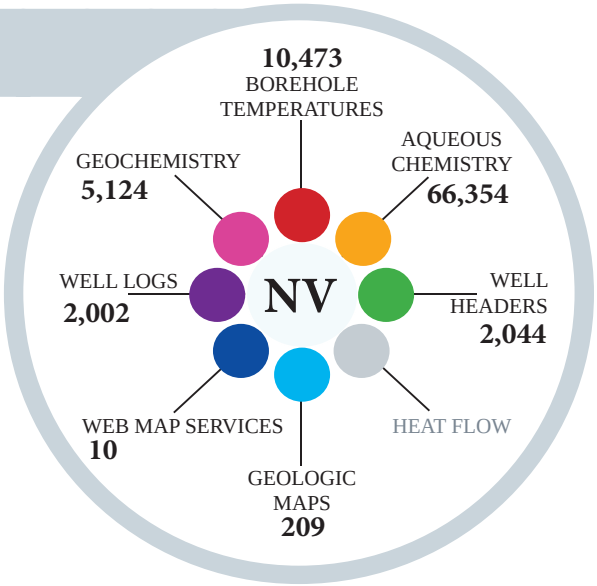
New Map Services

The Nevada team has created an interactive 80 layer mapping tool to help visualize NGDS data, available at *http://gisweb.unr.edu/geothermal/*. Creation of map services has been an important component of the Nevada Bureau of Mines and Geology (NBMG) modernization and adding of new capabilities to geothermal maps. Located here is one single map service where all the geothermal data layers can be accessed, and can be downloaded as a .cvs file.



80 layer mapping tool includes query, search, and extract functions.

The map application will allow the ArcGIS user to directly use any geothermal data layers at NBMG in their own projects after downloading each data set as needed. As databases, are updated, the map application will always reflect the most up-to-date data at NBMG.



Industry Benefits

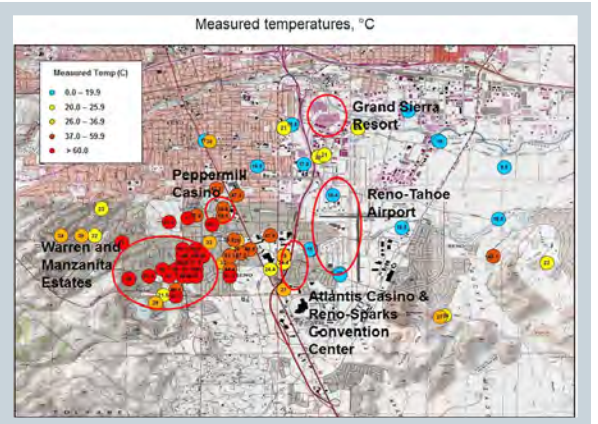
Nevada's NGDS customers include companies conducting exploration:

- ATLAS Geosciences
- Geothermal Development Associates
- Lumos Engineering
- Navy GPO
- Nevada Geothermal
- Ormat
- Zonge Engineering

New Geothermal Exploration

No less than 26 publications are directly related to project activities or use of compiled data from the project.

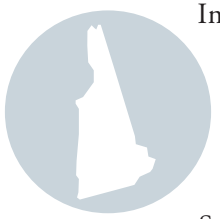
Geothermal site assessments were conducted at twelve sites, including: Baltazor HS, Bradys, Fallon NAS, Fernley area, Grand Sierra Property, Lahontan Valley, McGinnis Hills, Pumpernickel Valley, Black Warrior, Reno VA, Trego HS, and Washoe Tribal Lands. ■



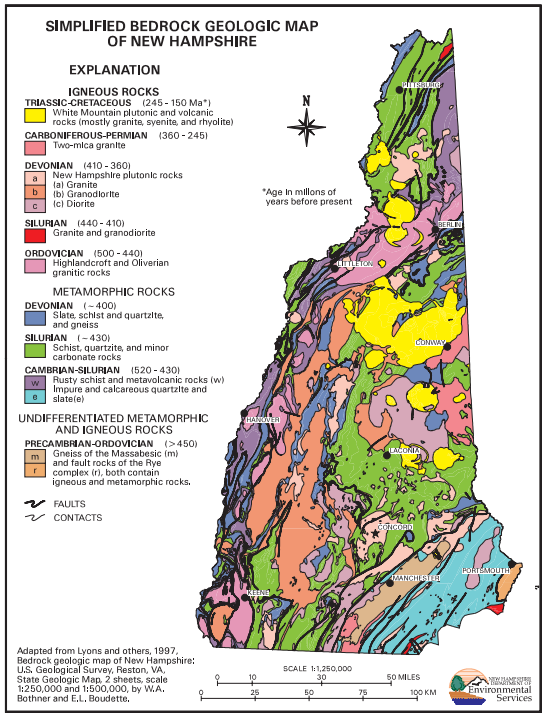
New site assessment of Reno area reveals numerous wells above 60 degrees C.

New Hampshire

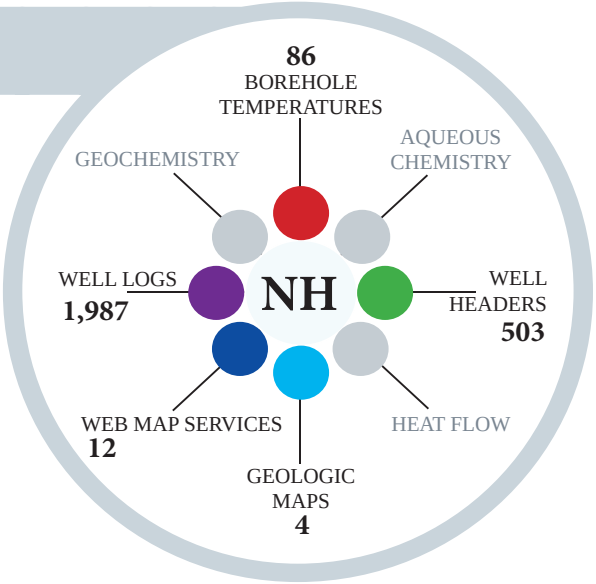
New & Legacy Data



In addition to placing well logs, bedrock sample analyses that include 103 newly collected samples, and temperature and heat flow measurements online, the New Hampshire Geological Survey (NHGS) exposed nine map products, all accessible as web services with fourteen data services.



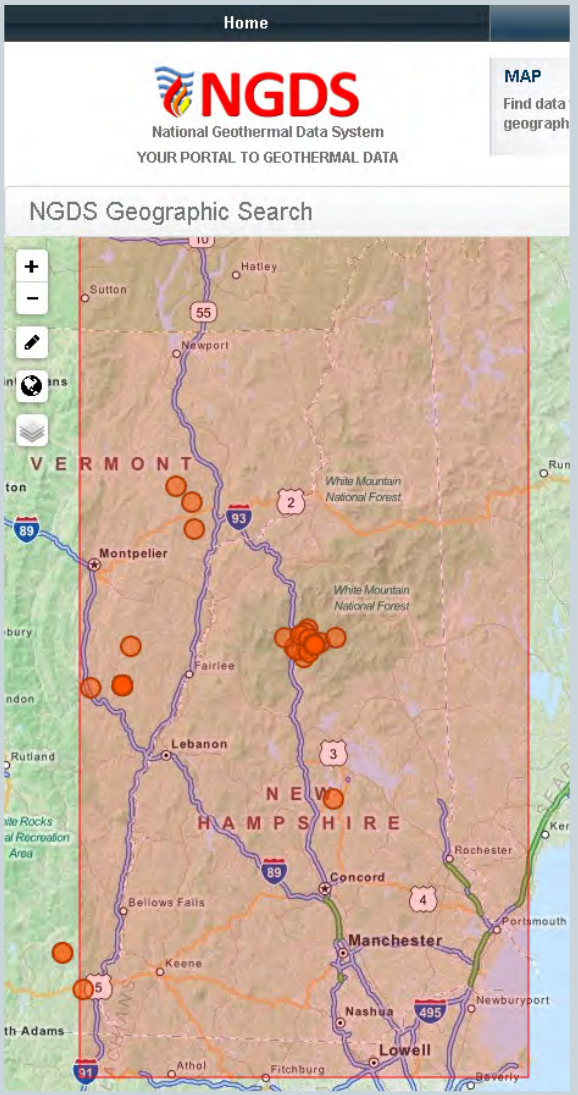
The ESRI shapefiles for this and other geologic maps of NH are at NGDS.



Outcomes & Collaborations

The NGDS project improved understanding of available mapping technologies, especially open source technologies and their applications. New Hampshire is now an independent NGDS node, serving their own data using GeoPortal software.

The work has improved cooperation across state and federal agencies. Two joint publications with the U.S. Geological Survey were released. In addition, the NHGS enjoys a collaborative relationship with the University of New Hampshire Earth Science Department (UNH). The UNH plans to perform an analysis of heat flow for New Hampshire using many of the data sets and is using the bedrock sample collections data service as a teaching tool. ■

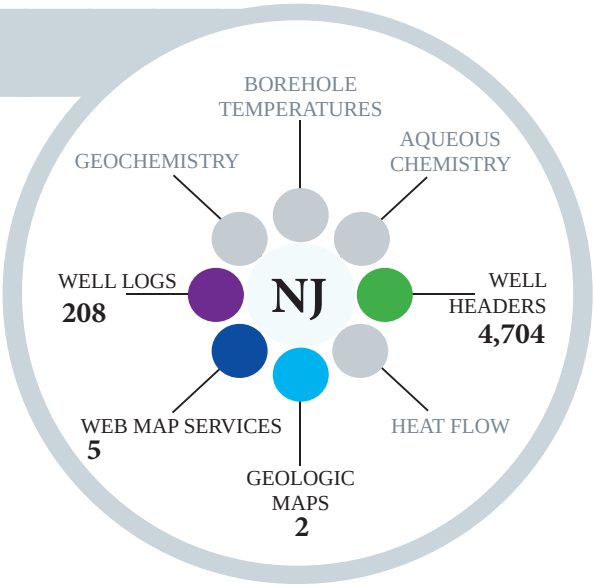


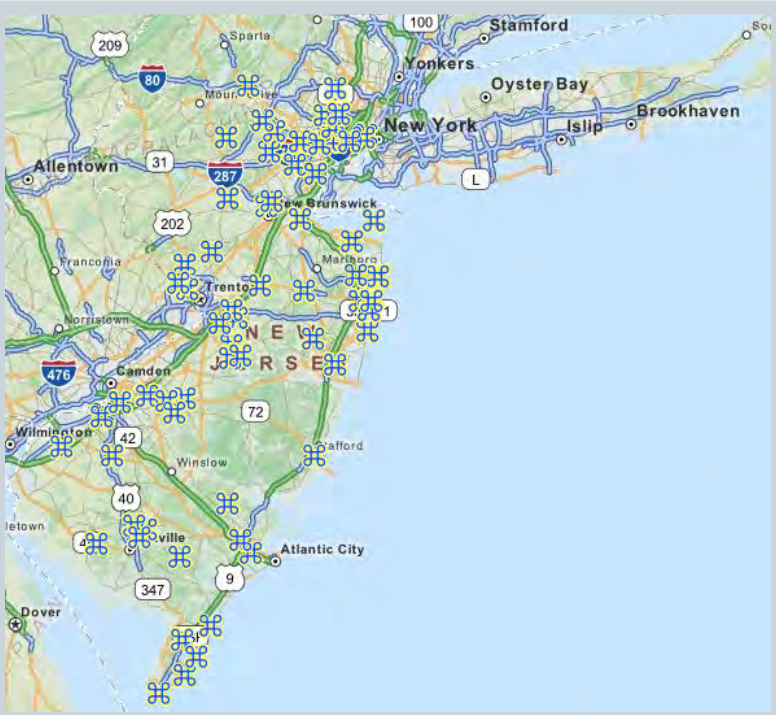
The individual orange circles represent more than 50 data sets loaded into NGDS, including: thermal springs, gravity stations, well logs, well headers, borehole temperatures, thermal conductivity, rock chemistry, depth to bedrock, and more.

New Jersey

Cost Comparison

A state goal of the NGDS project set by the New Jersey Geological and Water Survey (NJGWS) is to provide a cost-benefit comparison of geothermal energy with other energy options for New Jersey homeowners. The NJGWS gathered data on 122 test locations provided by two companies: Alderson Engineering, Inc. and Geothermal Services, Inc., to illustrate how different subsurface material affects the efficiency of geothermal heat exchange systems. Thermal conductivity and diffusivity values were plotted on graphs and physiographic provinces maps. ■



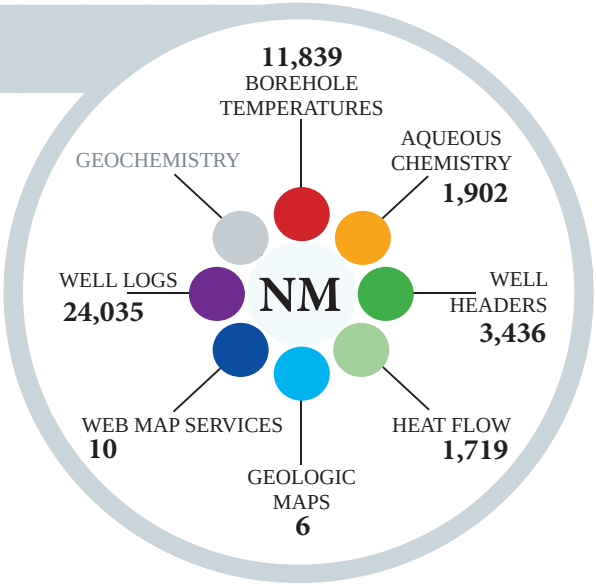


Thermal conductivity map of New Jersey as shown through the NGDS web map service.

New Mexico

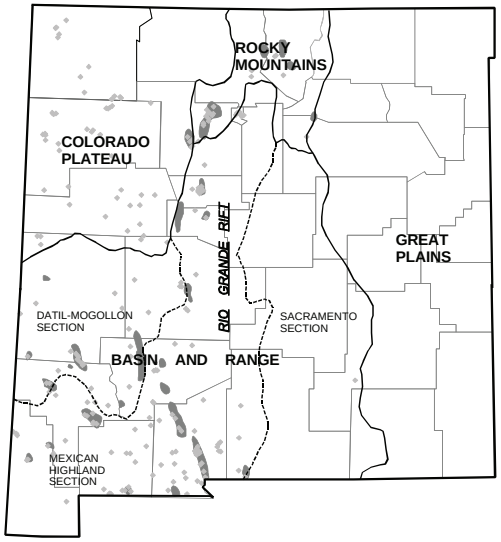
New & Legacy Data

The New Mexico Bureau of Geology and Mineral Resources organized and digitized many of their legacy records as part their NGDS-related work. They report major accomplishments as: 1) the transformation and additions to a petroleum well database, which now includes geothermal data; and, 2) merging project-specific water databases into one master database that can be mined for unexplored resources. The project also made a push toward building and deploying map services.

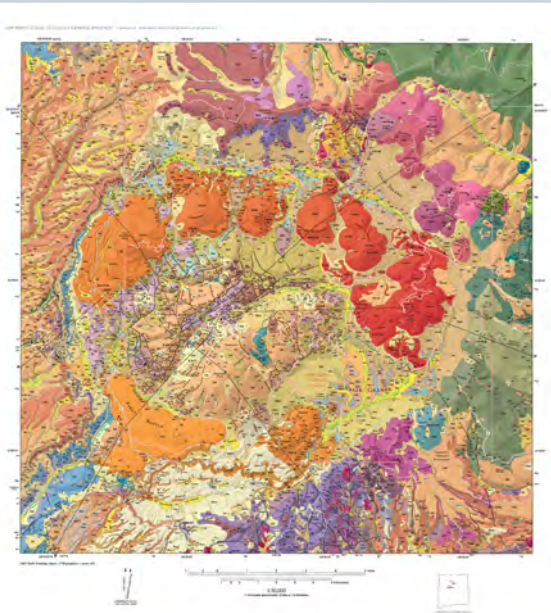


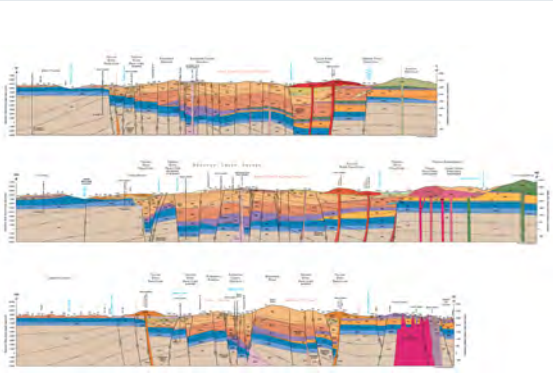
Geothermal Exploration

Rescued data helped reassess information on a geothermal prospect located at the intersection of a Laramide fault system, a caldera margin, and Rio Grande rift faults. Investigations using thermal profiles, water geochemistry, major and trace elements, isotopes and temperature and flow measurements revealed changes from baseline data collected in 1941. ■



Geologic map of the Valles Calera by Fraser Goff and others.



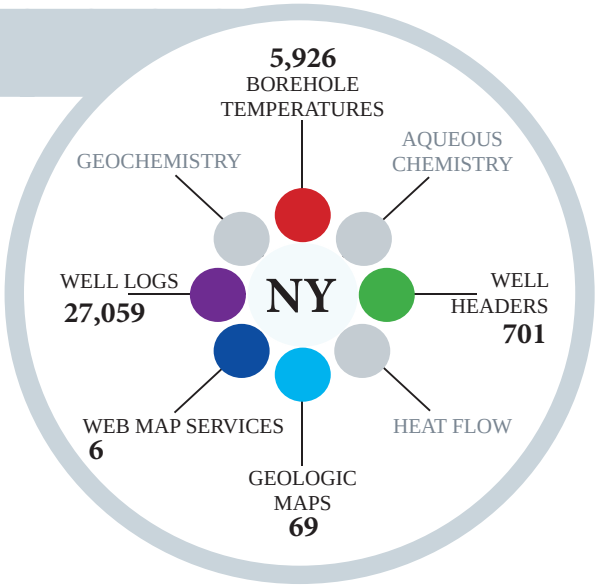


General geothermal resource map by Jim Witcher—available through NGDS.

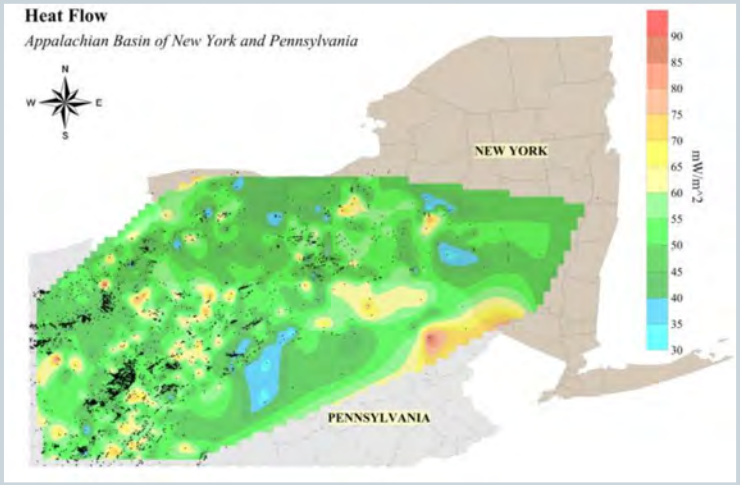
New York

New & Legacy Data

The New York State Geological Survey (NYSGS) chiefly contributed data inventories and collection statistics to the NGDS program. This includes several excel spreadsheets with information regard- ing well logs, drill cuttings, earthquake database, and products such as isopach and structure-contour maps for select sedimentary units. While some of this information was readily available, there were three tasks in particular that required significant time and effort. These three products are anticipated to be especially useful for future geothermal studies.



1) The NYSGS maintains a database known as ESOGIS (Empire State Oil & Gas Information System) which contains records for all of the oil and gas operations in the state. The NYSGS staff digitized over 700 temperature curves from scanned well logs and submitted them to NGDS (also offered through ESOGIS).



Surface heat flow in the Appalachian Basin of New York and Pennsylvania, calculated as the product of thermal gradient and average thermal conductivity for a specified location. Heat flow can be used to facilitate geothermal exploration.

2) Over 5,900 bottom-hole temperatures pulled from 35,000+ well logs were digitized and submitted to the NGDS. Researchers at Cornell University are mining these data as part of their research on geothermal resources in New York and Pennsylvania (Shope et al. 2012, Proceedings from the 37th Workshop on Geothermal Reservoir Engineering).

3) The NYSGS maintains a collection of rock cores collected from 150 wells across the state. These cores are stored in a climate controlled warehouse where they are available for study. Due to limited resources, the inventory of this collection was outdated and lacked crucial pieces of information. In just 6 months, the entire core collection was re-cataloged with a focus on capturing all available data from each core box/container. A copy of the completed core collection catalog was submitted to the NDGS and is expected to help future geothermal research by providing information on the characteristics of reservoir rock, a valuable tool for any scientist attempting to describe a potential geothermal reservoir, not just in New York, but across the entire Appalachian Basin. ■

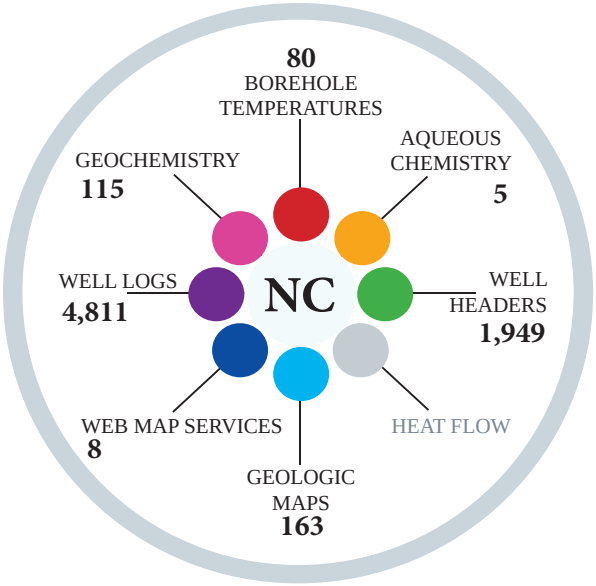
Core of the NYSGS collection; a copy of the core collection catalog was submitted to NGDS to facilitate geothermal research.

North Carolina, Georgia, and South Carolina

North Carolina

Legacy Data

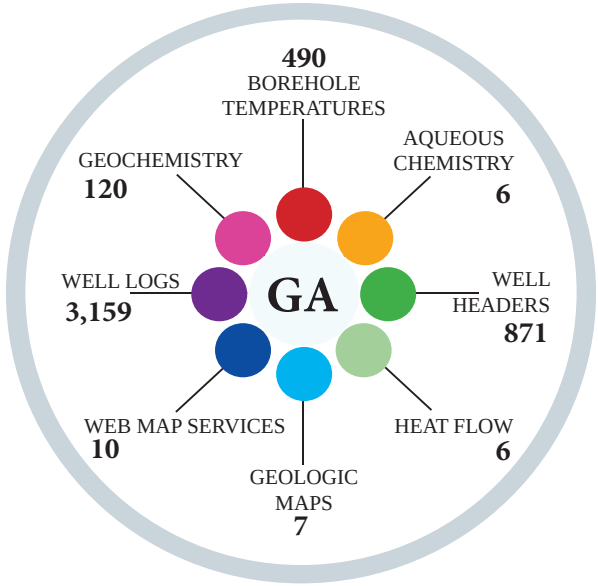
When the North Carolina Geological Survey (NCGS) requested funding under the NGDS, a hiring freeze proved to be an impediment. NCGS funding award was redirected to the Virginia Geological Survey, where their staff made two trips to the NCGS Coastal Plain Office and Core Repository to scan original paper reports, geophysical logs and lithology logs. This project encompassed the largest conversion of paper records to digital files in the NCGS history. While the original paper records have been returned to the Core Repository, digital copies are now available through the NGDS. ■



Georgia

Legacy Data

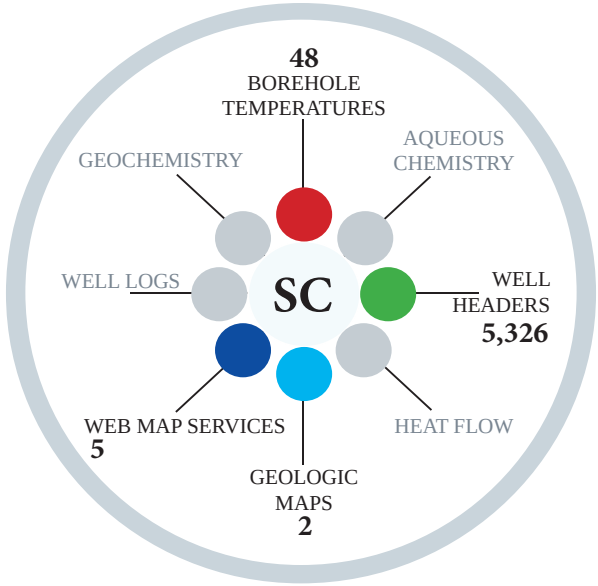
Data from the Department of Natural Resources (GDNR) was collected by Virginia and is served at the Kentucky Hub. Metadata is available for geologic maps, (~250 documents), Georgia geothermal documents and references, and gravity maps. Available services for the state of Georgia include aqueous chemistry (version 1.10), borehole temperatures, gravity stations, heat flow 1.23, rock chemistry, thermal conductivity, thermal springs, well headers, and well logs. ■



South Carolina

Legacy Data

The South Carolina Geological Survey provided a limited amount of data to the Virginia Geological Survey, including metadata for geologic maps and geothermal articles. Map services are now in place for aeromagnetic and gravity maps (as part of USGS data), limited borehole temperatures, well logs, and 733 records for seismic event hypocenters. ■

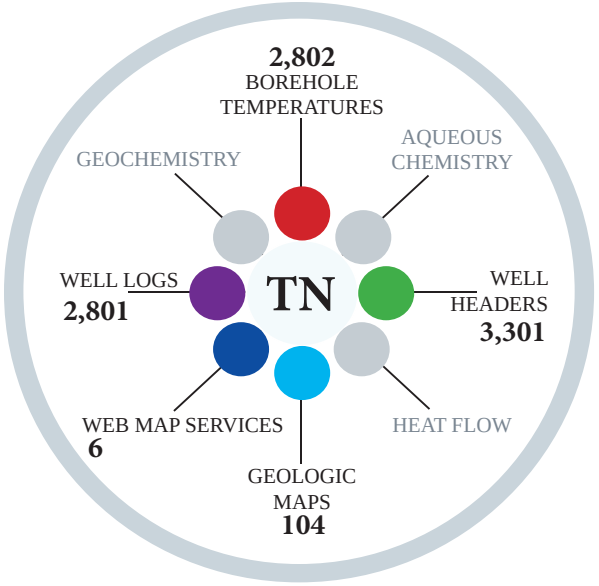


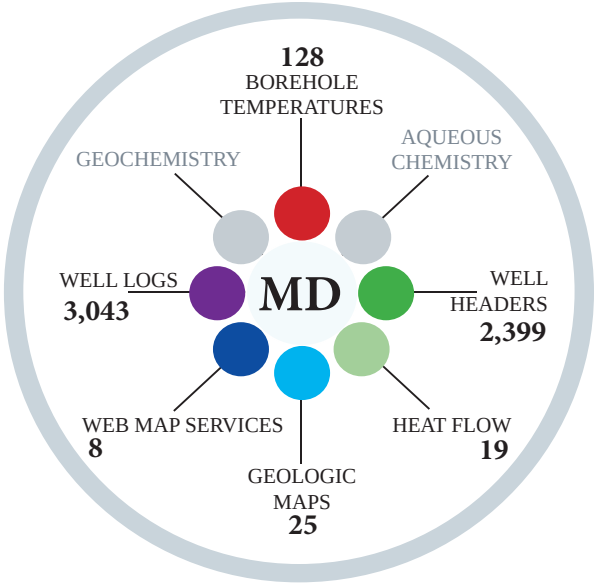
Tennessee and Maryland

Tennessee & Maryland

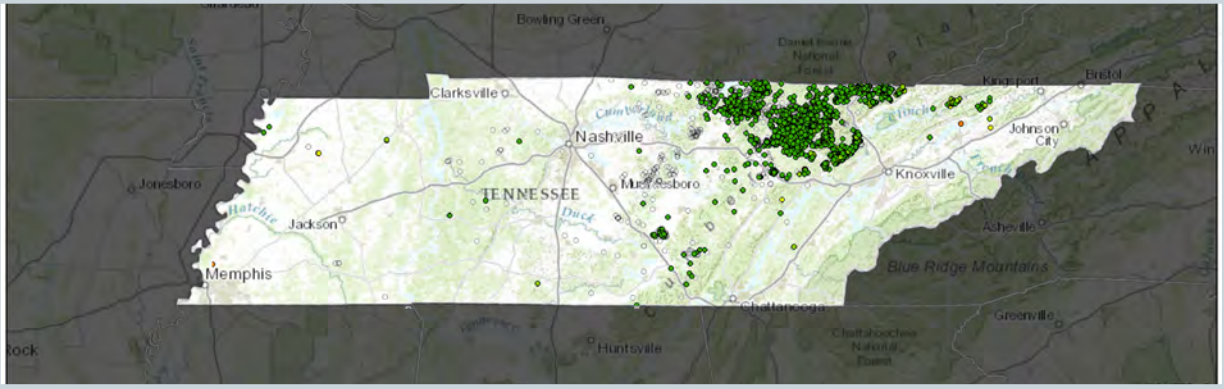
Legacy Data

All data from the Tennessee Geological Survey (TGS) and the Maryland Geological Survey was collected by the Virginia Geological Survey and is served through the Kentucky Hub. Both agencies contributed metadata for maps and documents. The TGS contributed 21 metadata records for Tennessee heat pump installation and geothermal test reports, 31 records for direct use sites, and 66 conductivity test reports by the Tennessee Valley Authority. ■



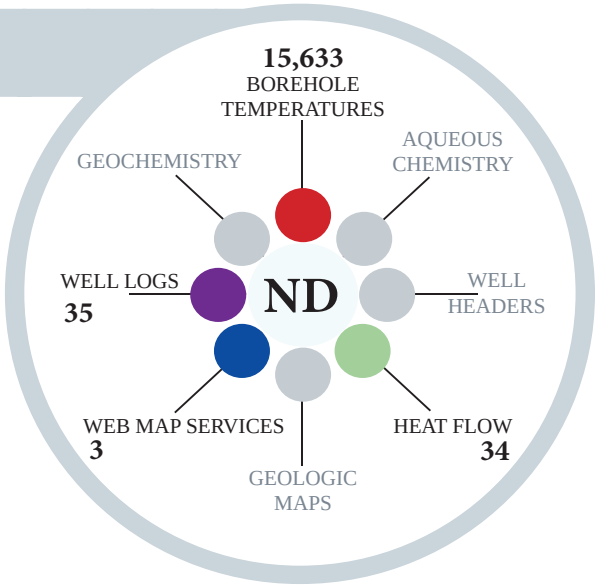


Borehole temperature measurements provided by Tennessee as part of the NGDS. The TGS contributed 2800 borehole temperatures and 3300 well header measurements that appear as a Web Map Service.



Maryland's heat flow and well header services are pictured on this composite map.

North Dakota



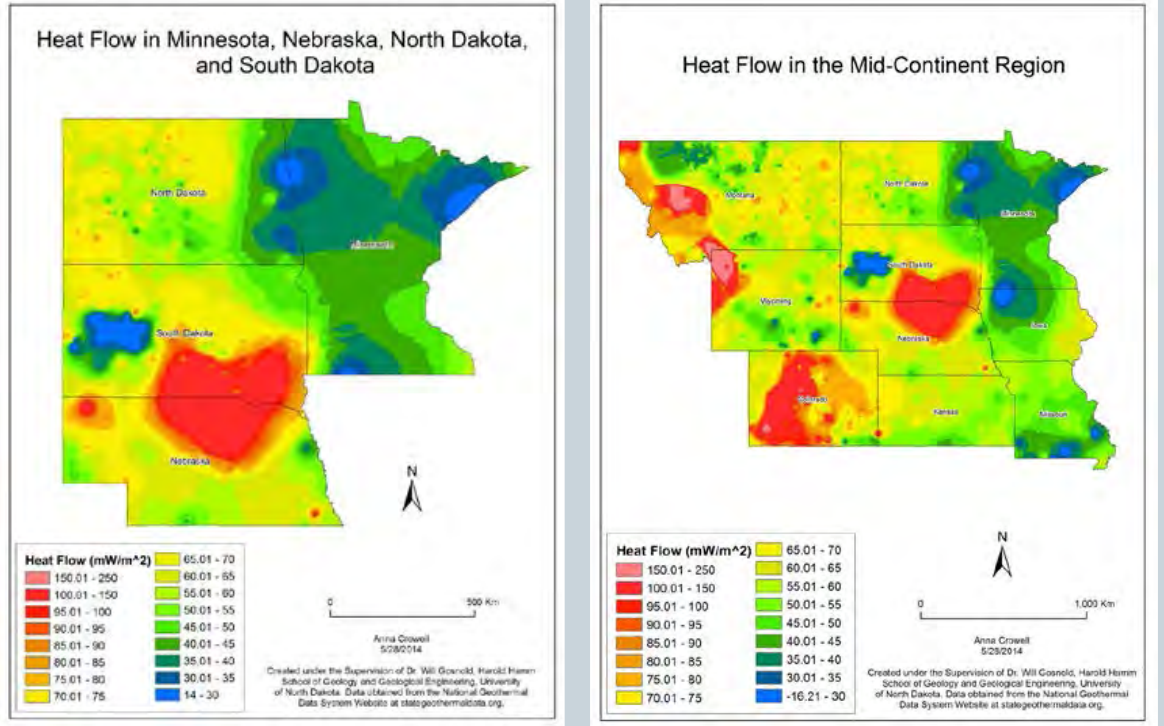
Multiple-State Effort

The University of North Dakota (UND), compiled geothermal-related data for Minnesota, South Dakota, and Nebraska, and forged a successful collaboration with Southern Methodist University (SMU) to aggregate data from both SMU and International Heat Flow Commission (IHFC) databases for the NGDS project. UND used available templates to produce maps on heat flow, thermal conductivity, and radioactive heat production featuring uranium, thorium, and potassium, and to plot temperature against depth measurements. The team at UND also produced a georeferenced NOAA geothermal resource map. This map covers Minnesota, Nebraska, and South Dakota as well.

UND is now independently serving data using the free-and-open-source GeoServer software.

Geothermal Exploration

Data compiled for the project was used in a geothermal analysis of Williston Basin and on at least six thesis projects ranging from geophysical properties of reservoirs to sedimentary enhanced geothermal systems (EGS). Associated research papers explore geothermal resource potential in the northern Midwest. ■



Heat flow in Minnesota, Nebraska, North Dakota, and South Dakota.

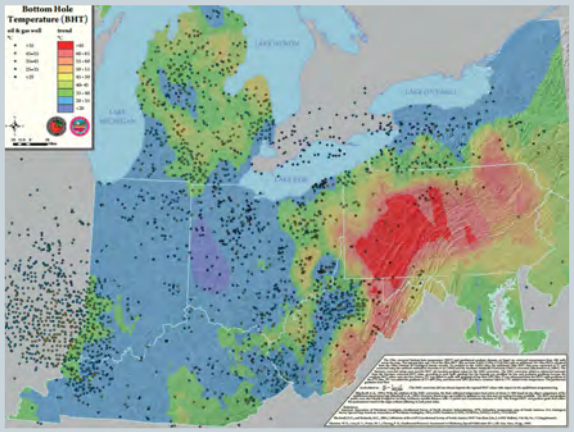
Ohio

Applied Energy Mapping

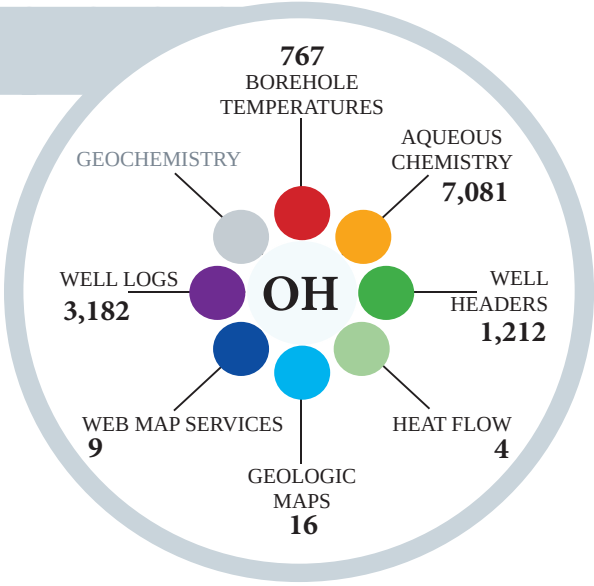
The Ohio Geological Survey (OGS) is currently engaged in a number of projects to appraise Ohio's geologic resources as they apply to developing shale oil and gas, storing CO2, and geothermal potential. This research is being conducted with funding, in part, provided to address specific project objectives for the Ohio Coal Development Office (OCDO), the Midwest Regional Carbon Sequestration Partnership (MRCSP) funded by the DOE, and the NGDS.

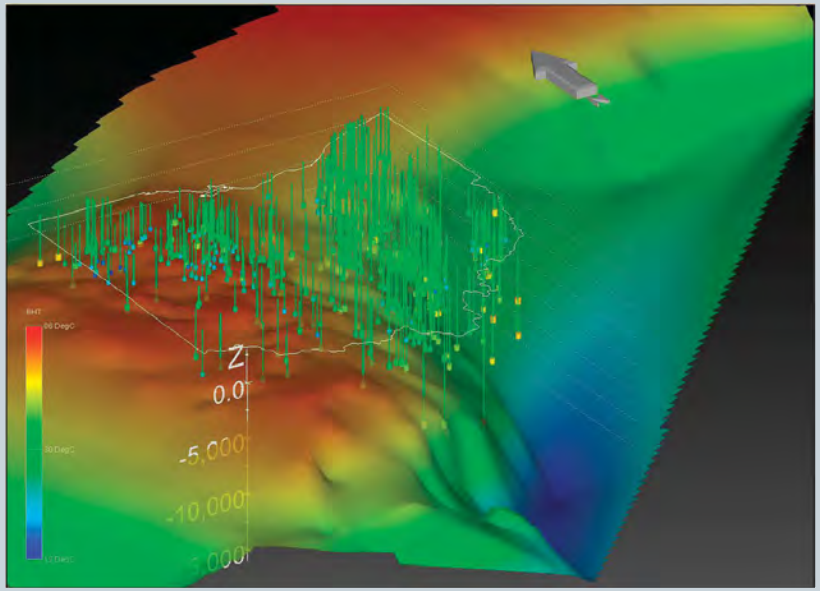
The OGS began research into the state's geothermal resources in July 2010 and is evaluating its very large dataset of bottom-hole temperatures and the American Association of Petroleum Geologists corrected bottom-hole temperature dataset. Specifically, selected bottom-hole temperature data was corrected and used with the AAPG dataset to construct bottom-hole temperature and gradient 3-D plots and maps in order to help evaluate the regional, subsurface geothermal environment.

Maps and information showcased here are from the poster presentation "Applied Energy Mapping " and the paper "Applied Energy Mapping at the Ohio Geological Survey" by the Ohio Department of Natural Resources, Division of Geological Survey. ■



Ohio and surrounding environs with BHT data values and the generalized Krieged BHT surface. The 2,643 data points are from AAPG (1994) and 334 new BHT values for Ohio. The BHT data generally reflect the depth of the well with hotter values found in deeper wells.





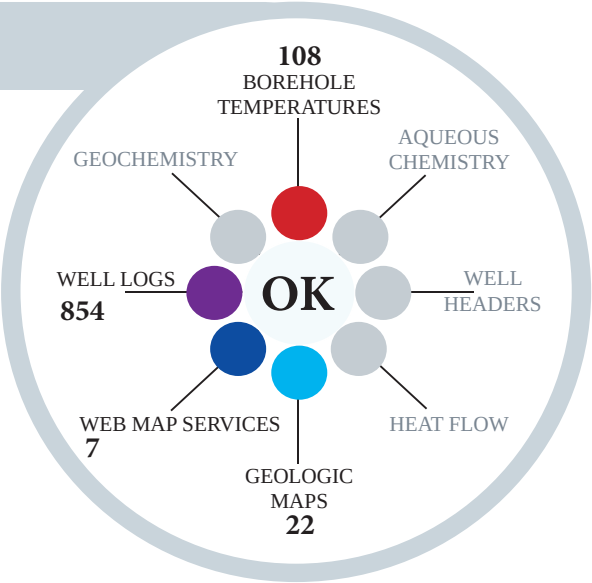
3D perspective of Ohio and surrounding environs Precambrian surface with color-coded barrels that indicate corrected borehole temperature observations. The straws show well locations and observation depths. These data will help to evaluate where geothermal energy, CO2 injection, enhanced recovery, and electricity coproduction may be viable.

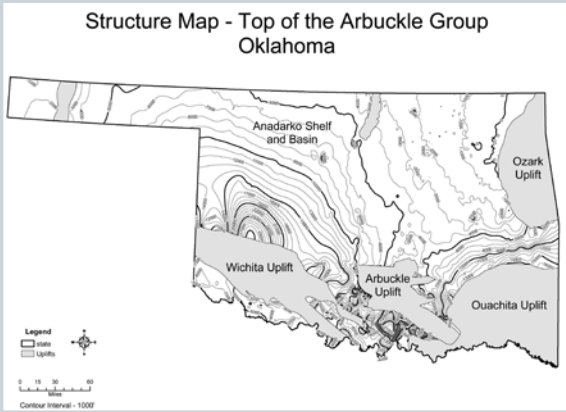
Oklahoma

New & Legacy Data

The Oklahoma Geological Survey (OGS) revamped and expanded existing resources through collaboration fostered by the NGDS project. OGS collaborated with agencies and industry to acquire Depth to Bedrock measurements from HIS Energy Group, Natural Resources Information System (NRIS) OGS publications, personal correspondence, and log interpretations.

One accomplishment is correction of errors in the location of active fault lineaments. Recently collected GPS waypoints reveal that the USGS database was inaccurate. OGS integrated new data with the fault trace identified on recently-released, high-resolution aerial photography to produce an updated shape file defining the position and extent of the Meer's Fault in southern OK.



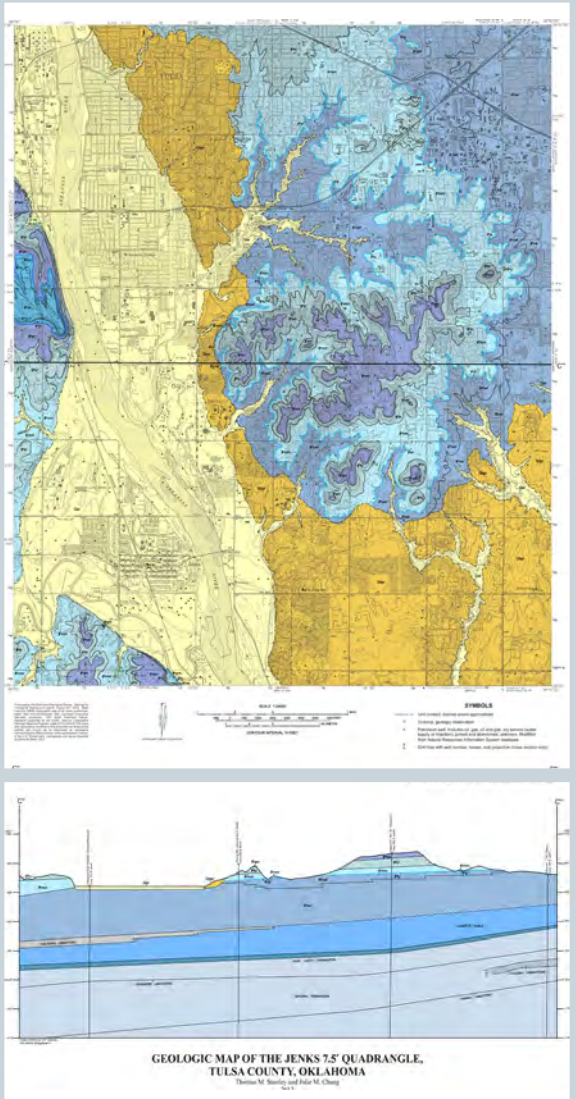


Top of the Arbuckle Group Oklahoma—captured from geographic search of OK geospatial data at NGDS. This resource is a metadata compilation of map images and associated NetCDF data for Oklahoma gravity and magnetic measurements provided by the OGS.

New Map Products

During the course of NGDS work, OGS also produced gravity maps of Oklahoma, including grids and images based on the Total Magnetic Intensity, various residual and upward-continued anomalies, and other effects. OGS uploaded files for a complete Bouguer anomaly map to the NGDS document repository (*www.stategeothermaldata.org*). Additionally, gravity data extracted from a recently located source are in the process of being opened for public access.

Additional data were compiled for Abruckle Group well penetrations, where the OGS merged data from multiple sources to create the most complete data set on the Arbuckle Formation to date. ■



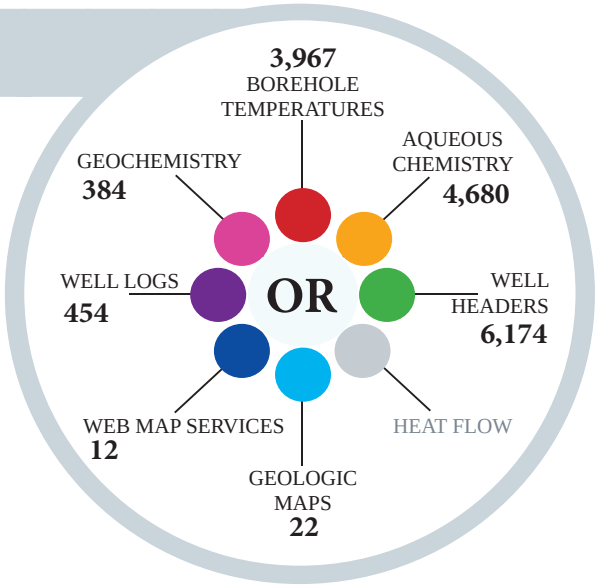
Geologic map of Jenks 7.5 minute quadrangle, Oklahoma.

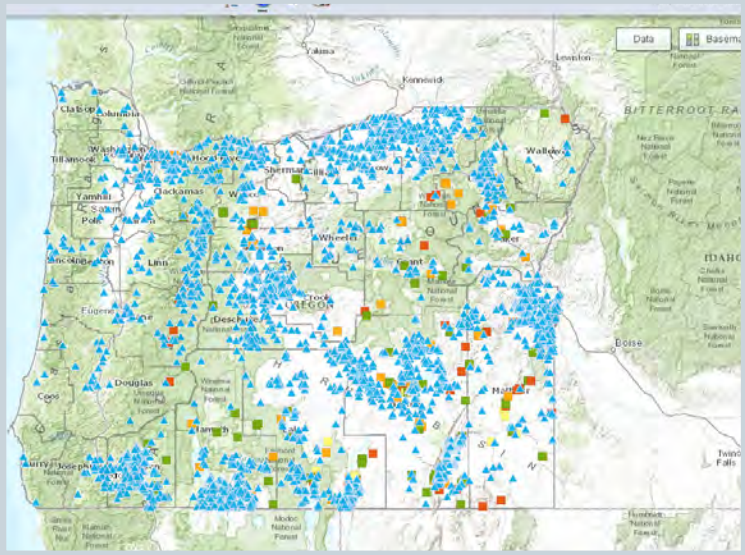
Oregon

Data Management & Best Practices

The Geothermal Information Layer for Oregon (GTILO) is a geospatial database system that stores and manages Oregon's geothermal resource information. The map includes 690 sites representing Oregon's thermal springs, and 4,344 sites representing various types of low-temperature wells in Oregon—domestic/irrigation/other water supply wells.

The Oregon Department of Geology and Mineral Industries (DOGAMI) hosts GTILO. The resource map spotlights areas investigated for geothermal resource potential or which show geothermal potential. (Additional details are available at *http://www.oregongeology.org*).





Interactive geothermal map constructed from the GITLO database.

For each year of the award, DOGAMI digitized its holdings of at-risk, legacy geothermal-relevant data and published this existing digital data by exposing databases to the NGDS. DOGAMI added over 39,700 online records to the NGDS and is currently acting as a distribution node to the system, serving data online and hosting in-house.

New Exploration Data

Supplemental NGDS funds targeted collection of the following:

- A pilot study employing airborne thermal infrared (TIR) and co-acquired lidar
- The drilling of a thermal gradient well at three different locations in Oregon
- Geologic characterization of the quadrangle where each well is located

The experiment using high-resolution, high-accuracy TIR and lidar imagery for geothermal exploration demonstrated that it is possible to collect large swaths of very high quality data and to achieve high spatial accuracy with the thermal imagery by using the lidar data to rectify it.

The AltaRock Co. boom truck and crane during the inspection of the Newberry NWG 55-29 geothermal injection well in 2013.

The combined imagery is extremely effective at locating geologic structures and cold and warm springs, both on land and in water bodies. The data also allow accurate estimation of spring temperature, within the limited testing accomplished.

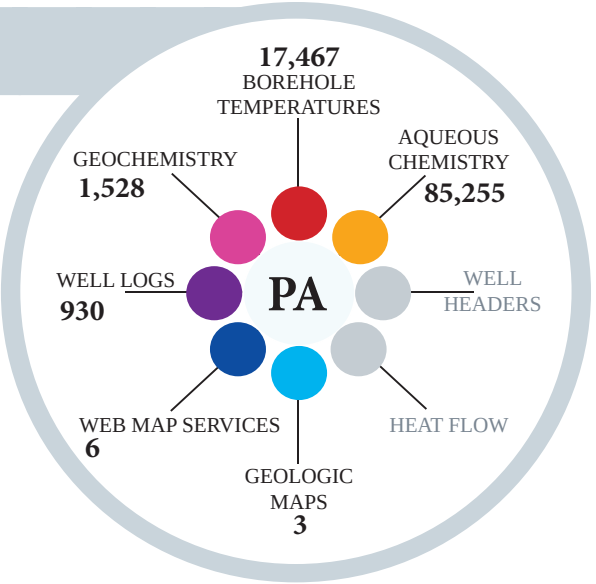
The EGS company AltaRock Co. performed the drilling and testing of new wells. ■

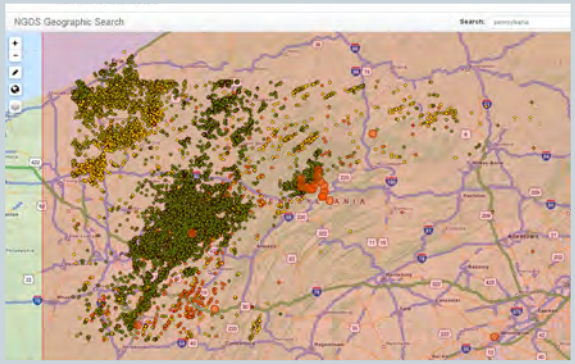
Pennsylvania

Geothermal Exploration

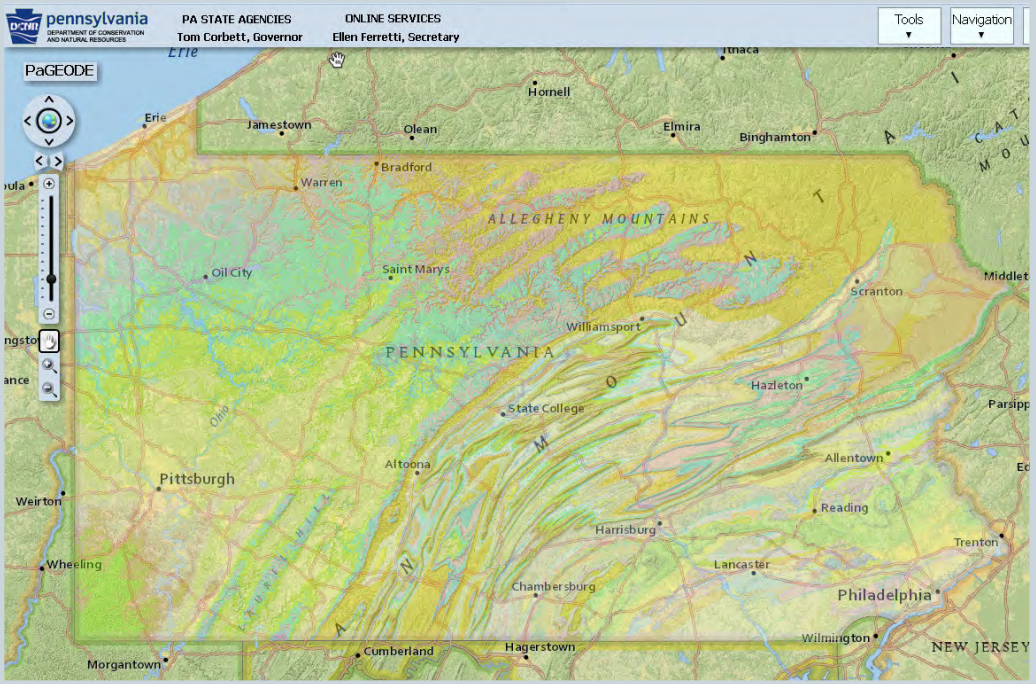
Geologic and geothermal-related data collected, digitized and compiled as part of the NGDS project is now slated for use in identifying areas for more thorough characterization of potential geothermal resources. Bottom-hole temperatures from over 17,000 oil and gas wells, as well as supporting geological, geochemical and geophysical information, are now readily available in one location through NGDS.

Leveraging information gained from work on NGDS, the Pennsylvania Geological Survey is now a supporting partner of Cornell University's "Low Temperature Geothermal Play Fairway Analysis for the Appalachian Basin" project, recently funded by the DOE. This preliminary one-year project will assess subsurface data for potential geothermal reserviors in the central Appalachian basin. The Survey will support the project by providing geologic expertise and access to oil and gas well data contained in the PA*IRIS/WIS System (Wisconsin's self-directed supports program). ■





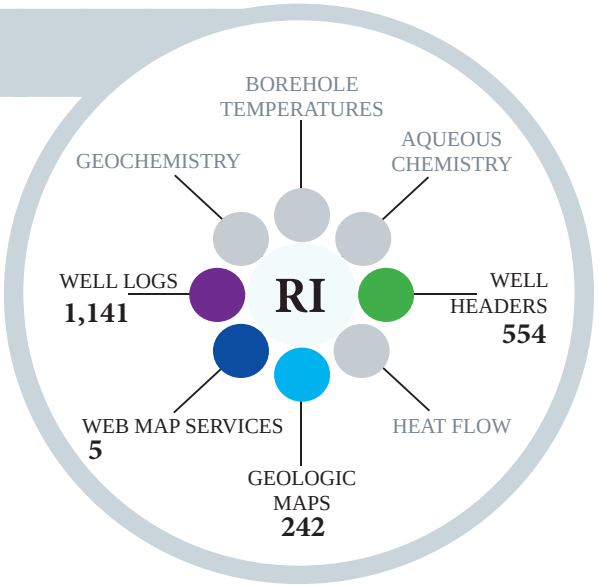
Borehole temperature data for Pennsylvania curated at the NGDS. These data are being used by staff at the Pennsylvania Geological Survey to further assess geothermal potential in that state.



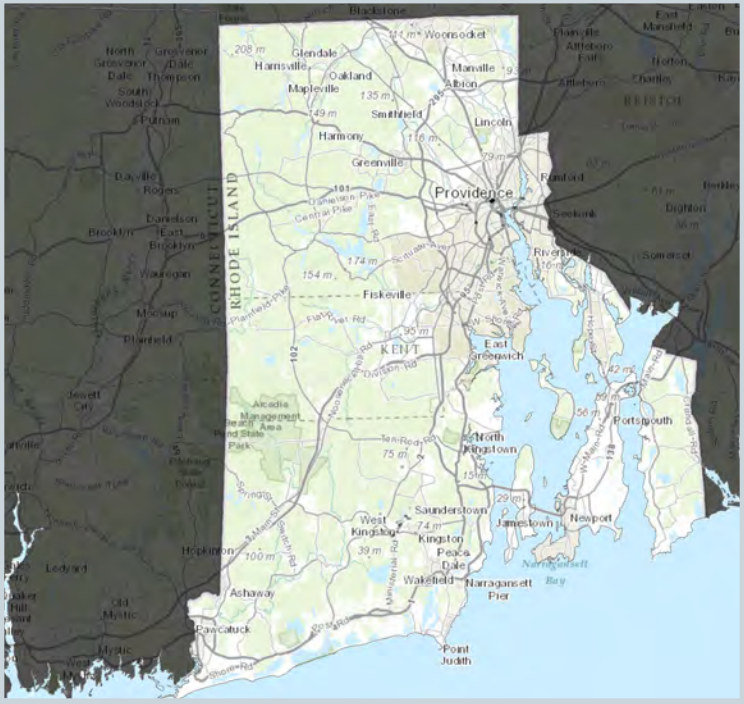
Interactive geologic map of Pennsylvania produced by the Pennsylvania Geological Survey, *http://www.gis.dcnr.state.pa.us/geology/index.html.*

Rhode Island

Legacy Data



The Rhode Island Geological Survey began work on contributions to the NGDS by supervising an undergraduate Energy Fellow funded by URI Energy Center to digitize 400 water well logs, and making digital copies of Rhode Island bedrock and Quaternary (surficial) maps with a drum plotter-scanner. During the project, students were an essential resource in assisting with digitizing and uploading data. In August 2012, RIGS delivered Vector data (polygons) of digitized Quaternary maps that were digitized and re-projected on raster topo images by Co-PI Oakley, along with a total of 8, 1:24,000 scale quadrangles and digital boring logs created using StraterTM software. In 2014, work on rock chemistry was part of undergraduate projects at the University of Rhode Island in partnership with CCRI-Physical Sciences. ■



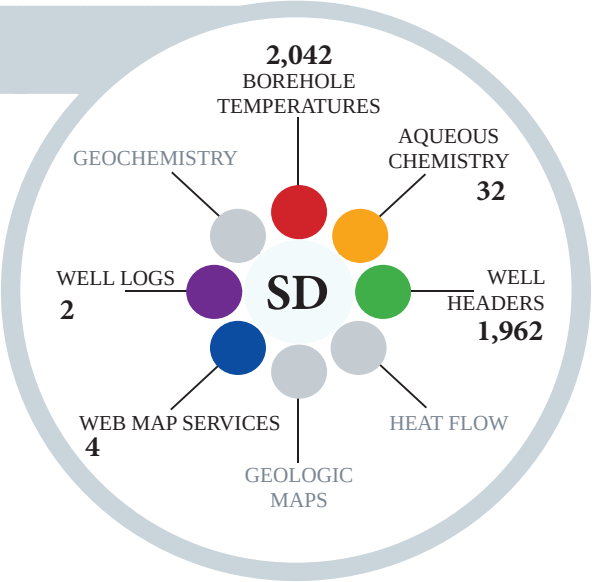
A Web Map Service conveying well headers provided by Rhode Island as part of the National Geothermal Data System.

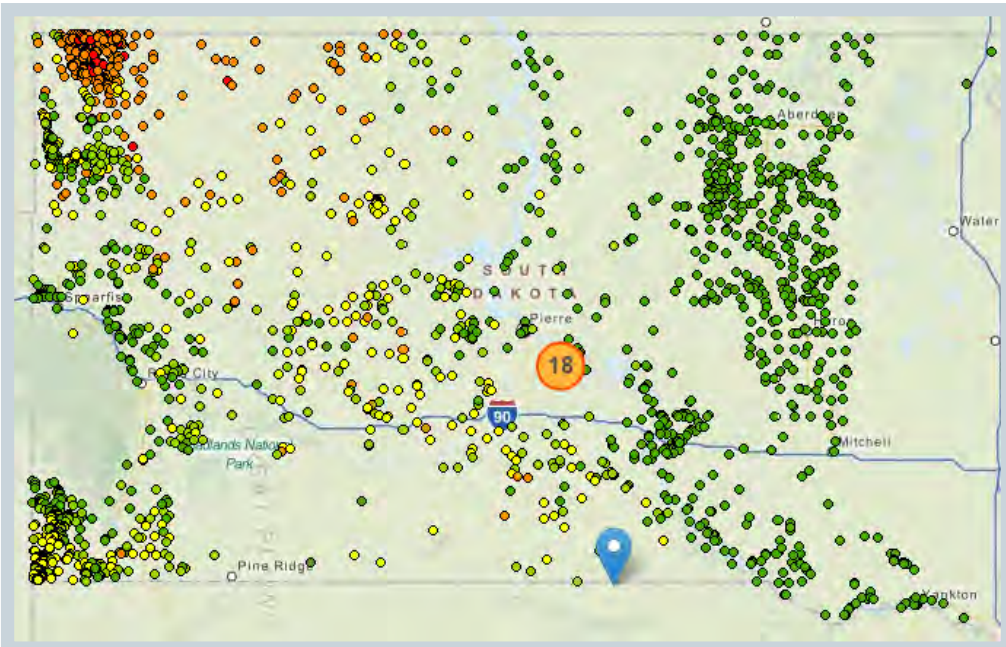
South Dakota

Legacy Data

Geothermal data for South Dakota was compiled by a team at Sinte Gleska University. Major contributions include temperature, flow volumes, and chemical analysis of water for 12 privately owned artesian wells in this area.

Other NGDS-related products include lithological data, logs, and detailed water analysis for two new geothermal water wells near White River, South Dakota, as well as several geological/lithological maps for South Dakota, with special emphasis on the Rosebud Indian Reservation. ■

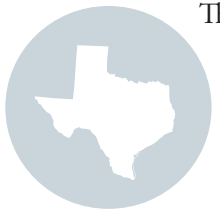




A compilation of borehole temperature observation data for wells in SD, shown as a web map service on the NGDS interface.

Texas

New & Legacy Data



Though Texas already had an abundance of data in the form of oil and gas logs, the Texas Bureau of Environmental Geology (BEG) team worked to acquire raw data in the form of new well logs, bottom-hole temperatures, reservoir water quality, production information and reservoir characteristics related to the geometry and characteristics of potential geothermal reservoirs found throughout Texas, with special attention to the area west of Interstate 35.

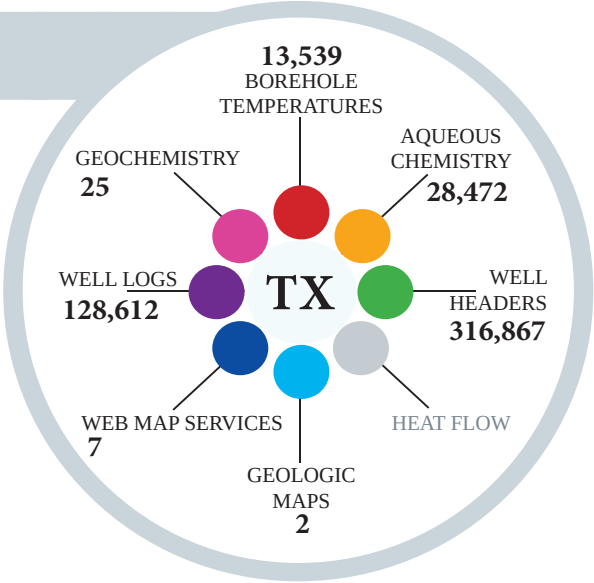
Geothermal Potential

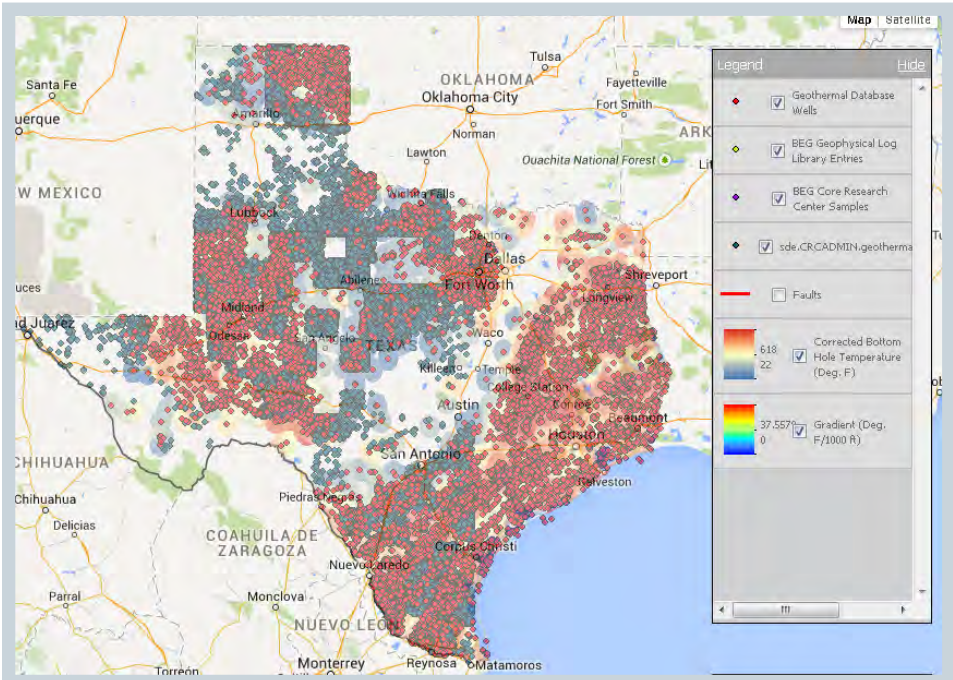
These raw data are being analyzed to provide derivative data products that measure potential productivity of deep reservoirs. A database of these analyzed data is being generated through analysis of railroad commission well production documents and well logs. The team initially completed a detailed analysis for the areas in Crockett and Val Verde Counties as a test example, and then focused on the southwestern Gulf Coast and West Texas regions.

As time and resources permit, the new exploration efforts ongoing in the Anadarko Basin may produce information useful to geothermal energy assessment. If possible, the BEG team will assemble information for the Anadarko Basin area that may be relevant to geothermal production.

Data Management & Best Practices

The BEG developed visualization tools to display the raw data contributed to the NGDS and the analyses via their website; (*http://igor.beg.utexas.edu/geothermalmap/*). They are in the process of building an interactive system to provide insight and motivation for further study, and for development of commercial scale geothermal energy production. ■





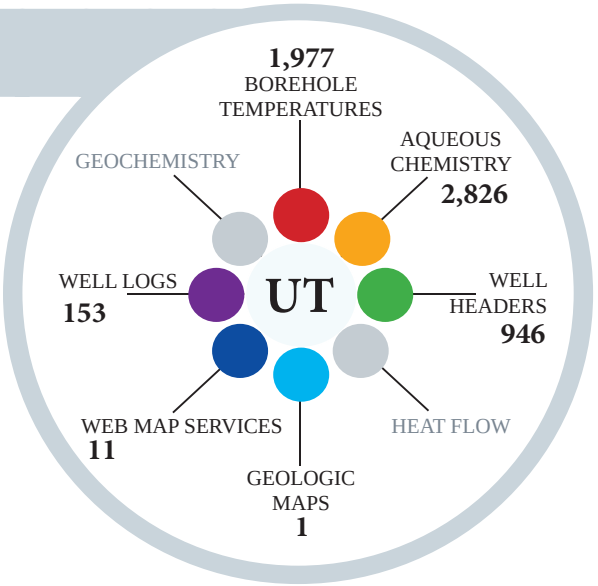
Texas geothermal data available at the NGDS. Includes well, geophysical log library entries, core research samples, faults, geothermal gradient, well log data and more.

Utah

New & Legacy Data

The Utah Geological Survey (UGS), in cooperation with the University of Utah, produced and compiled a large quantity of Utah geothermal data between 2010 and 2013 for inclusion in the NGDS. Published and unpublished legacy data include Quaternary fault and volcanic data, data on individual thermal springs and wells, corrected bottom-hole temperatures and temperature and pressure information from selected oil and gas wells, digital geologic maps of Utah, and geothermal-related bibliographic information.

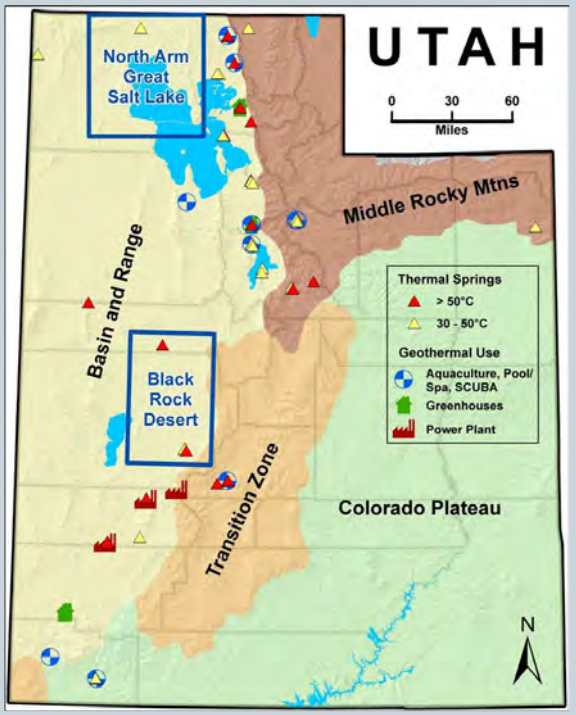
New or supplemental data were collected from studies, mainly in west-central and northwestern Utah, that focused on the Sevier and Black Rock Deserts of Millard and Juab County. These data consist primarily of temperature-depth plots, cutting and core samples, geophysical and lithologic logs, and thermal-conductivity measurements collected from 12 new thermal-gradient boreholes and five deep oil exploration wells. Supplemental temperature-depth data were also collected from 14 previously drilled "wells of opportunity" and new water analyses were obtained from 26 additional well and spring sources. Nine technical papers have been published, and several unpublished reports were written, detailing various aspects of the supplemental data collection work.



A map of thermal springs, wells, Quaternary features, and sediments in northern Utah that can be used to facilitate geothermal exploration and development.

Geothermal Exploration

Though Utah currently has multiple high temperature geothermal power plants, the NGDS effort contributed to the discovery of a major geothermal resource in the Black Rock Desert of west-central Utah where temperatures exceeding 150°C exist over an area of 285 km² at a depth of about 3 km. Temperatures of 200–to230°C at similar depths are present in a centrally-located well and are expected over an area of about 40 km² (Figure 1). Heat flow data were derived from 10 new and 25 previously drilled thermal-gradient boreholes (some not shown in Figure 1) and 12 oil exploration wells. The discovery of a viable geothermal resource has garnered local and national attention since the fall of 2012. ■



A map of thermal springs and direct use geothermal sites in Utah; can be used to facilitate geothermal development.

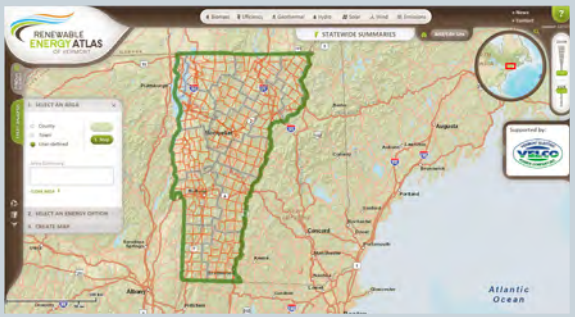
Vermont

Legacy Data Preservation

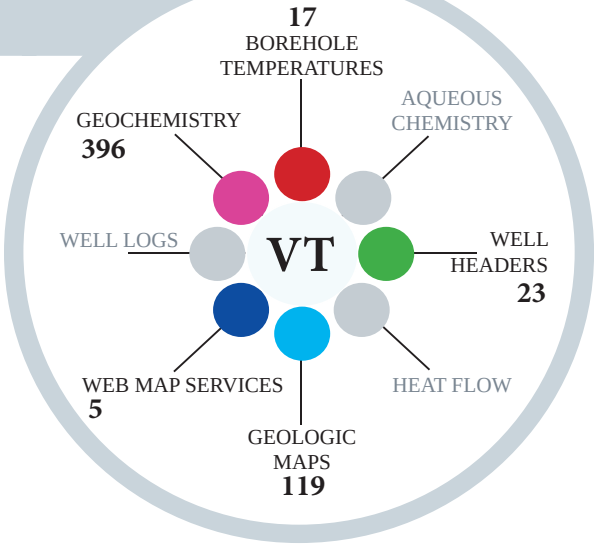
Participation in the NGDS project enabled the Vermont Geological Survey (VGS) to scan, document, and expose its entire library, providing online access to previously unavailable documents including published, draft and open file maps, bound books, field trip guidebooks, volumes with multiple articles, State Geologist Reports and publications of the VGS. Legacy "basement" files including mineral resource documents and maps (i.e. Brandon lignite, mineral springs, peat, phosphate etc), town and county reports, and sand and gravel reports also were brought online.

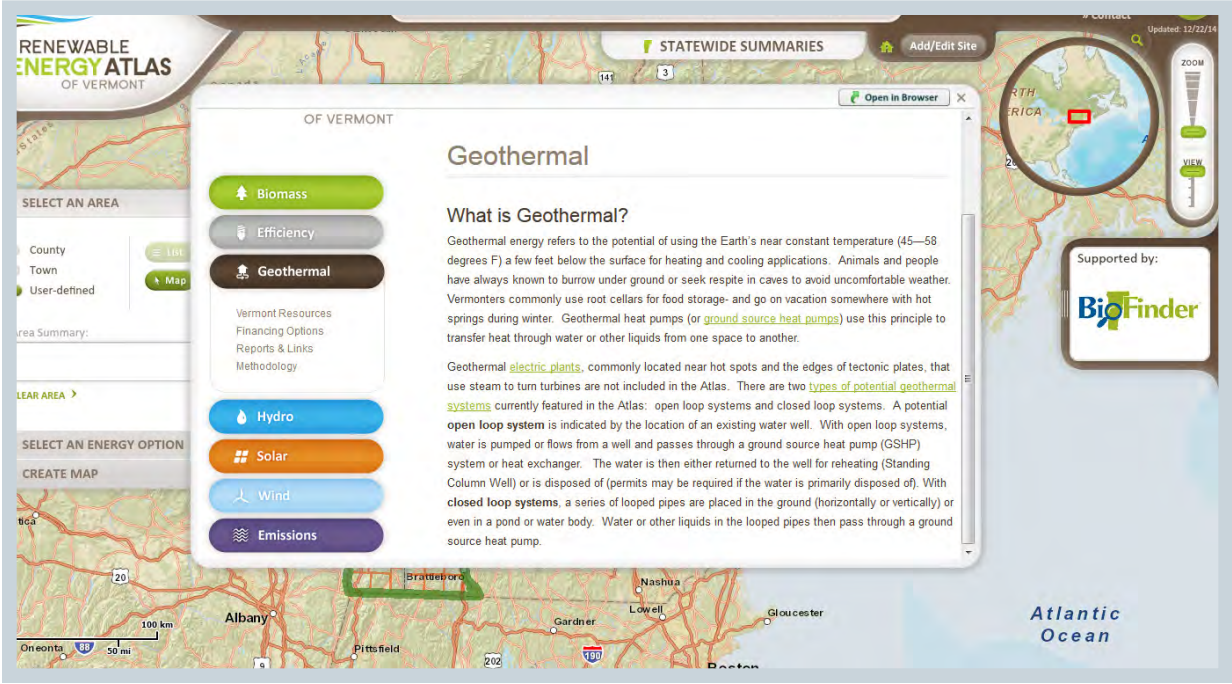
Hard copies of all these documents had been physically relocated to the office of the scanning contractor in August, 2011 when Tropical Storm Irene hit Vermont. The VGS basement was completely flooded and the bookcases previously containing the archives were substantially affected by flood waters. It is likely that files currently on the affected shelves would have been damaged beyond repair. Instead, the digitization and OCR project is complete and the data are available in computer formats with the paper versions returned and on library shelves. This event clearly illustrates the benefits and necessity of digitizing legacy information.

Now the VGS hosts the Renewable Energy Atlas, which provides information on energy resources in the state at *http://www.vtenergyatlas.com/#*. ■

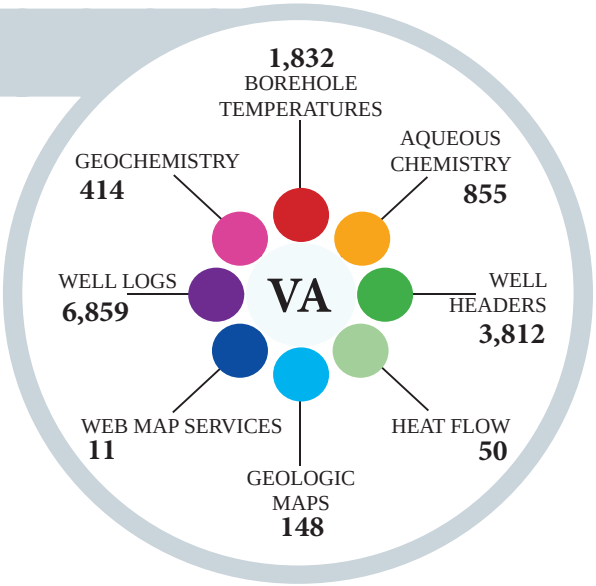


The Vermont Renewable Energy Resource Map - above and below—helps people locate available renewable energy installations.





Virginia



Interagency Cooperation

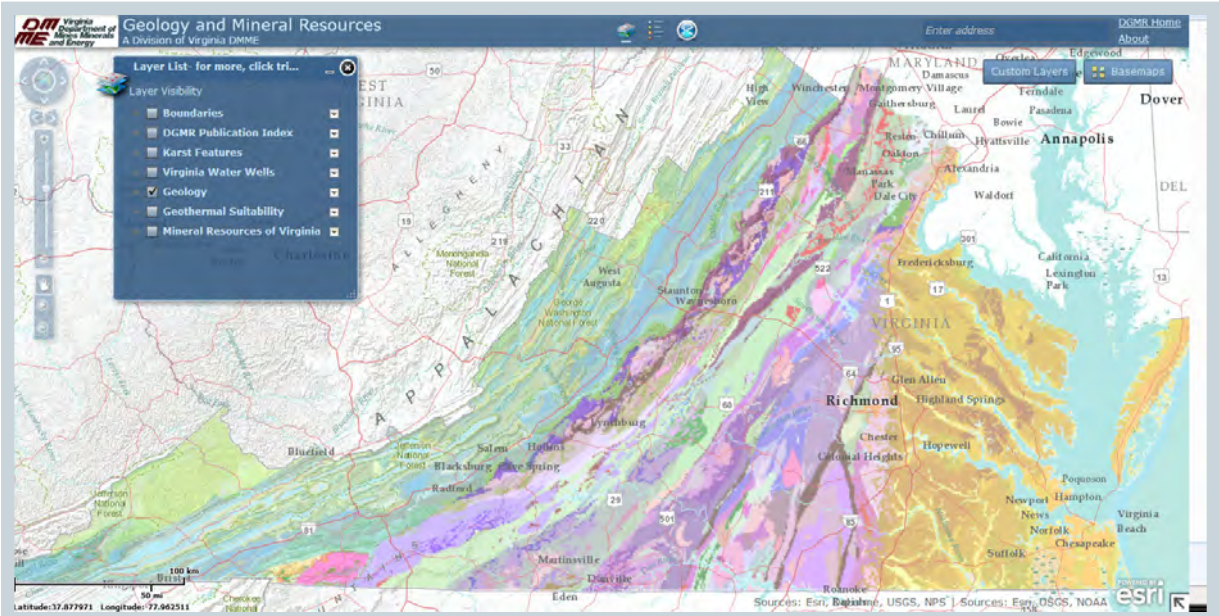
Virginia collected data from Georgia, Maryland, North Carolina, South Carolina, and Tennessee, contributing over 50,000 geologic data records from private industry, utility, and state and public agencies to the NGDS project. Numerous collaborators on the NGDS project increased inter-agency awareness and knowledge.

The Virginia Geological Survey anticipates further collaborative efforts, leading to improved data preservation and efficiency of data collection.

Industry Benefits

The NGDS project resulted in the production of derivative tools that provide a foundation for advancing the eastern geothermal industry. One important product is a comprehensive interactive map located at: *http://dmme.virginia.gov/webmaps/DGMR/*.

This map brings together state data from multiple agencies and geologic maps, allowing users to create custom data proiles. Map features include temperature gradient and heat flow maps using VA Division of Gas and Oil database and VTech Geothermal Program data, heat pump installations, depth to bedrock maps, and more. ■

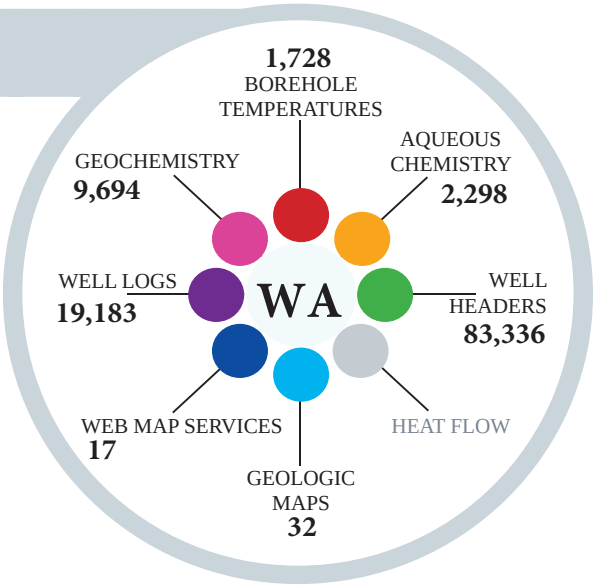


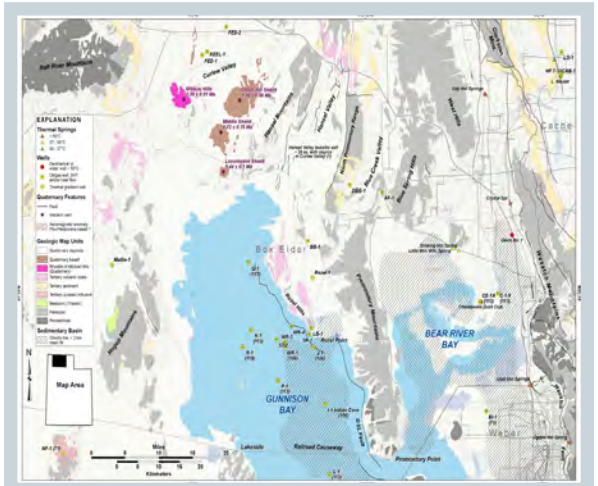
Interactive geologic map of Virginia.

Washington

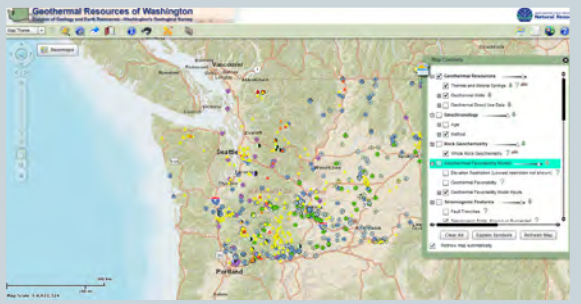
NGDS Stimulates Growth

The Washington Department of Natural Resources (WDNR) reports that the NGDS project was a boon for digital data compilation in Washington State. The datasets created during the project have direct application in geothermal and other geologic disciplines and are highly valued by the geologic community. As a result of this work, WDNR will produce several publications, including digital datasets and open-file reports. Much of the data appears on an interactive map application available at: *https://fortress.wa.gov/dnr/geology/?Theme=geothermal*





Geothermal data available through the NGDS for Washington: thermal springs and more.



Graphic expression of geothermal favorability model for Washington. Applying weighted raster calculations and fuzzy transformations, including weights permeability, heat, elevation, slope, and distance to transmission lines.

Industry Benefits

Private geothermal exploration companies are mining Washington's NGDS data to pinpoint potential resources in their geothermal lease holdings in the state. Maps made from the data yield insight in areas with little previous exploration. Enhanced and rapid data-gathering, coupled with reduced cost and reduced risk, is kicking off new geothermal exploration.

This project has prompted cooperation between the WDNR and the U.S. Army Corps of Engineers in exploring for geothermal resources on the Yakima Firing Center. It has also enabled collaboration with county utilities, as they look to develop renewable resources in their jurisdictions.

Geothermal Exploration

The NGDS is the cornerstone of a government-industry-academia Geothermal Play Fairway Analysis program to evaluate geothermal potential in Washington.

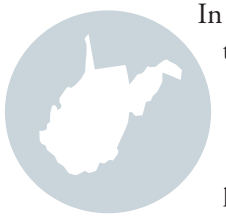
Additionally, thermal and mineral spring chemistry data collected for the NGDS project has been used recently by scientists with the USGS studying noble gas isotopes in the Cascadia Forearc (McCrory and others, 2014). Other USGS scientists used heat flow data and geothermal gradients from NGDS to study heat flow loss in the Columbia Basin. New geologic mapping in the Wind River Valley near the Columbia River recently facilitated the siting of wells that have elevated warm water temperatures and could lead to new geothermal possibilities.

New Data Collection

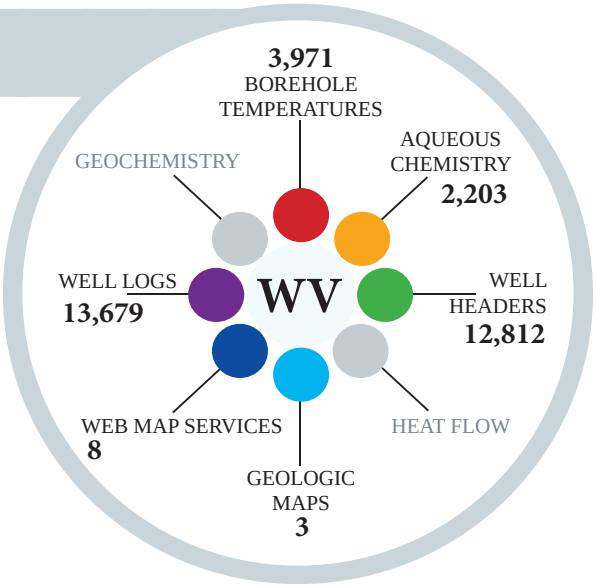
New Data Collection (October 2011 - June 2013) enabled maping of the geology of 61 square miles in the Wind River Valley. New information includes: thermal and mineral spring sampling, temperature-gradient logging, and drill temperature gradient boreholes. ■

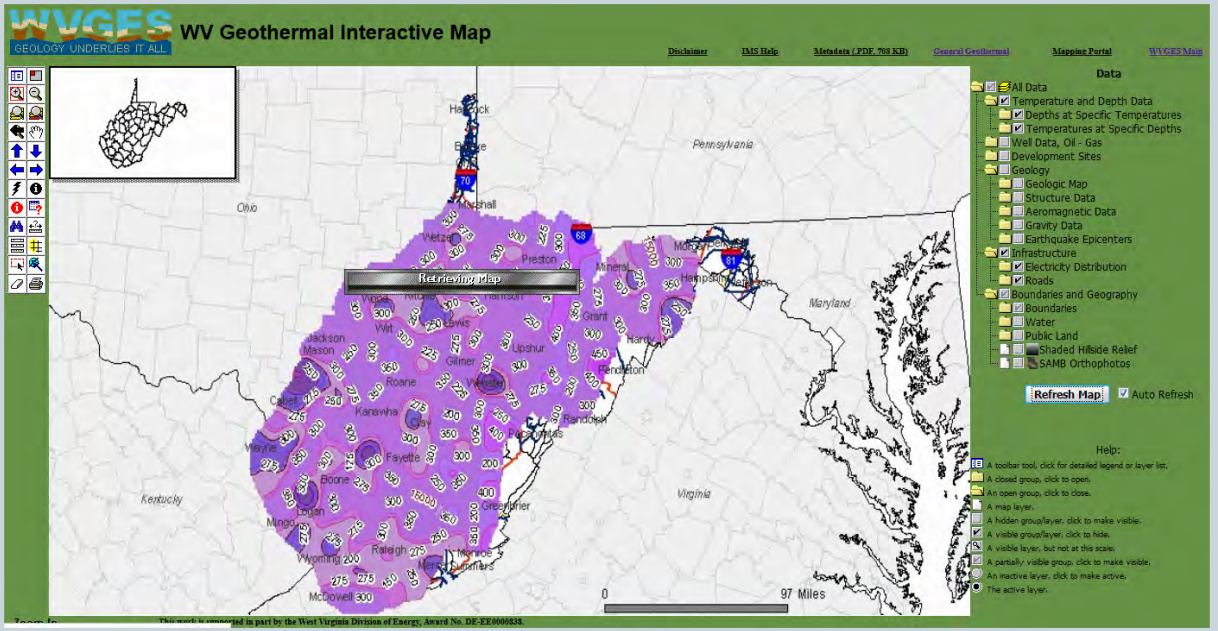
West Virginia

New & Legacy Data



In 2010, it was discovered that the temperature of the earth under West Virginia was much warmer than previously believed, fueling extensive research into a new potential "hot spot" for geothermal energy. Over the course of the NGDS project, West Virginia published seven maps as both Web Map Services and Web Feature Services. Map services now available include a georeferenced basement structure map displaying over 115 individual features (e.g., faults), gravity and magnetic maps, a georeferenced fault map for the state for the Devonian Onondaga Limestone, an earthquake epicenter map, and a geologic map. ■





Temperature data (pictured here) can be viewed in relation to other geological features, infrastructure, development sites, and more.

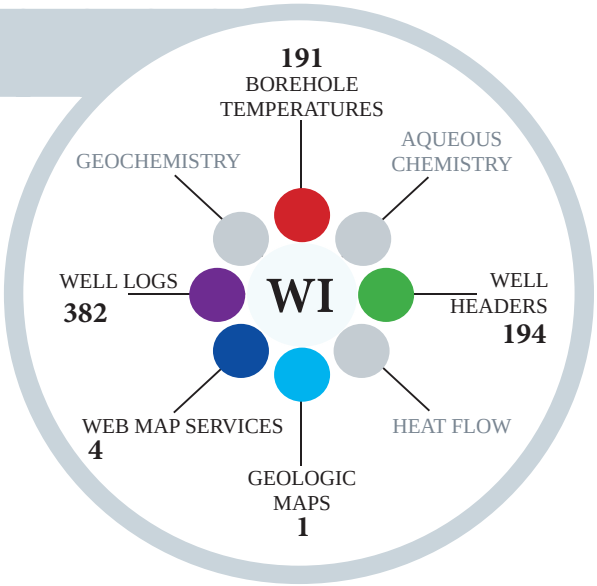
Wisconsin

Applied Research

Wisconsin's participation in the NGDS project culminated in the state's first geothermal energy program, with positive results for funding, installation of energy efficient systems, and additional research. The project digitized existing data, drilled six new boreholes and established a viable state-wide geothermal gradient, and characterized thermal properties of Wisconsin's bedrock.

Staff from the Wisconsin Geological and Natural History Survey (WGNHS) teamed up with the Geological Engineering program at the University of Wisconsin-Madison to assess Wisconsin's resources. Working with local engineering firms and geothermal heat pump installers, the group studied geothermal heat pump systems with a focus on applications in Wisconsin.

To date, this state-private industry consortium has received additional grant funds from the National Science Foundation (NSF) and the University of Wisconsin (UW). The total of the grants is more than $700,000, with just over $300,000 in initial funding from the NGDS. UW is now home to an NSF Research Experience for Undergraduates that focuses on alternative energies; each year an average of twelve students engage in research on wind, solar, geothermal and other alternate energy systems.



New Discoveries

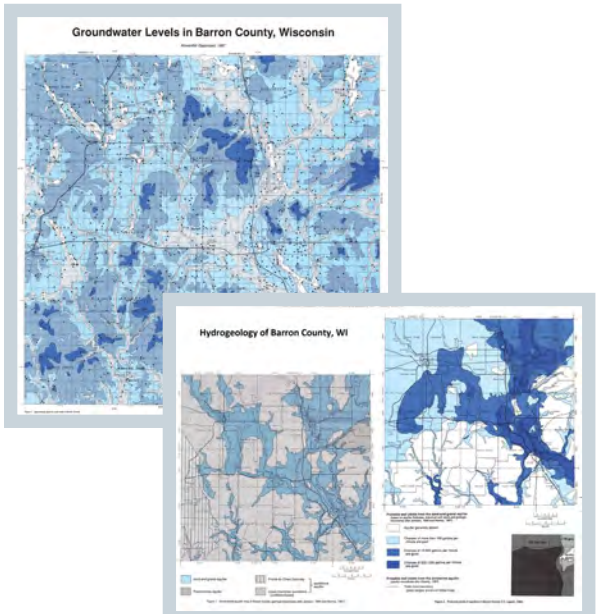
The project enabled seven 1,000-ft deep borings at selected locations in Wisconsin where geothermal temperature profiles were previously unavailable to be drilled and characterized. The project also enabled collection of complete data sets on over 180 existing wells. One of the 1,000-ft borings (Geothermal 7) is being converted to a long term groundwater monitoring well. Another boring (Geothermal 4) will likely be used as part of a ground source geothermal study. A short report on the hydrogeology of Geothermal 5 was used to help locate additional water supply for Fort McCoy, Wisconsin.

The research provided new insights into Wisconsin's hydrogeology, especially in northern Wisconsin where little data was previously available. This data is being used to help regulatory and policy professionals understand flows in the fractured rock, especially with regard to mining issues in Wisconsin and Minnesota.

Data Management & Best Practices

Wisconsin is now an NGDS node, independently serving their data through free-and-open-source GeoPortal software.

The NGDS project improved data management by creating NGNHS procedure for handling geophysical data from collection to serving via the internet. This procedure enabled identification of areas with little or poor geophysical and geologic data where additional effort should be focused and identified specific questions that will direct future research on the state's geothermal properties. ■



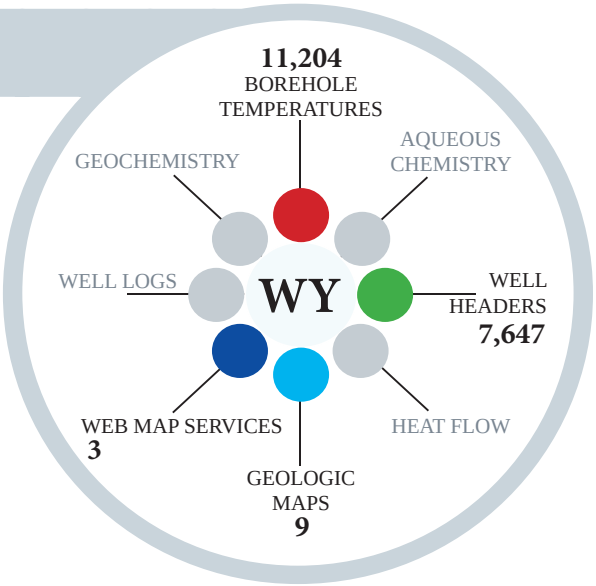
Examples of Wisconsin's maps now available online.

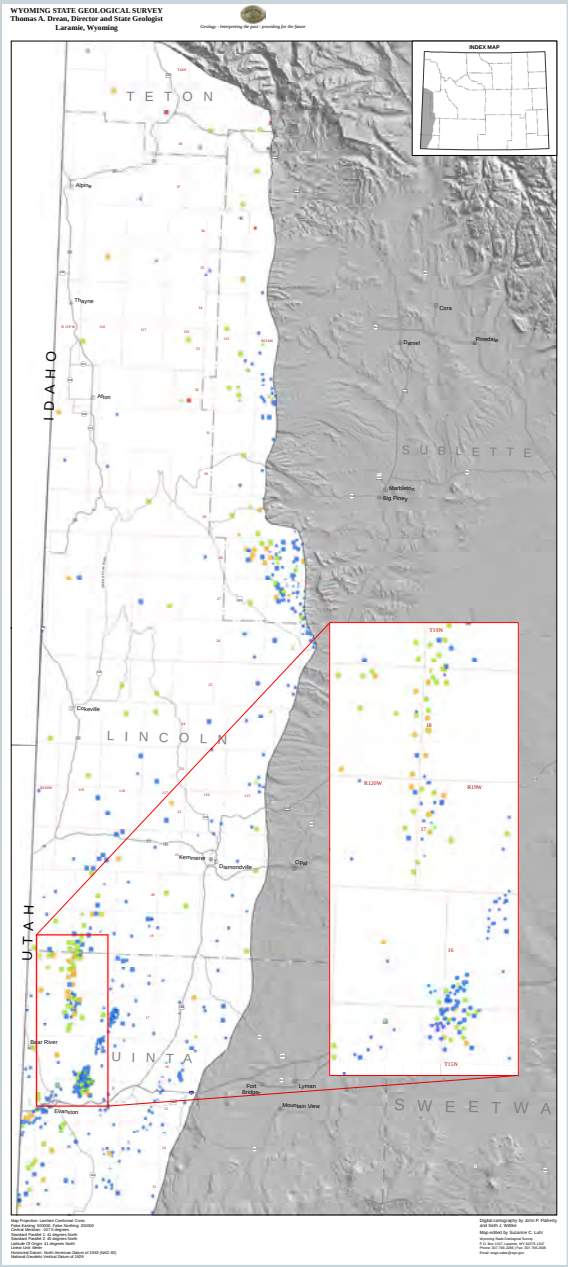
Wyoming

Geothermal Potential

Wyoming is a state with high geothermal potential. Currently, tourism is the primary use of geothermal energy in Wyoming. Geothermal features draw hundreds of thousands of tourists to Wyoming each year, allowing them to enjoy thermal springs in the state, primarily in Yellowstone National Park and Hot Springs State Park (located in Thermopolis). Direct use of geothermal energy is also harnessed as a means to heat buildings, water, and roadways in some areas of the state.

The Wyoming State Geological Survey (WSGS) has a number of publications related to thermal springs and geothermal data in Wyoming. A number of digital reports and maps are available on their website at *http://www.wsgs.wyo.gov/Research/Energy/Geothermal.aspx*. Other Wyoming related geothermal data, including bottomhole temperatures, are available via the NGDS. ■





Map of Reported bottomhole temperatures in the Overthrust Belt, Wyoming.

NGDS Featured in May 2013 *Scientific American*

"Heated to the Core"

"The ultimate goal [of the NGDS] is to capture the full geologic and geophysical profile of geothermal energy reserves across the country."

— ***Scientific American***, "New Geothermal Data System Could Open Up Clean Energy Reserves," by William Ferguson, February 25, 2013

Access Videos to Promote and Explain NGDS Data

13 "Scenario Card" NGDS Access Tutorial Videos – These videos, available on the Geothermal Data YouTube channel and on the main NGDS website, walk the user through use scenarios such as viewing data in GeoServer, and importing to a WFS or database program.

Playlist: https://www.youtube.com/playlist?list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-

- NGDS Data Access Scenario 01: Viewing a Well Log Service with the NGDS Data Explorer https://www.youtube.com/watch?v=YCvivdGmgPM&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=1

- NGDS Data Access Scenario 02: Wildcard Search https://www.youtube.com/watch?v=RBNO4TVtwyc&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=2

- NGDS Data Access Scenario 03: Merging Datasets from Different Sources https://www.youtube.com/watch?v=mE3r8u5x6-I&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=3

- NGDS Data Access Scenario 04: Basic Measurements with the NREL Geothermal Processor https://www.youtube.com/watch?v=ZkrESgY0Lxk&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=4

- NGDS Data Access Scenario 05: Creating an OpenLayers Web Application https://www.youtube.com/watch?v=VPsY3YxDfls&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=5

- NGDS Data Access Scenario 06: Creating a Leaflet Web Application https://www.youtube.com/watch?v=bj4_5kF11Ss&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=6

- NGDS Data Access Scenario 07: Filtering NGDS Catalog Results https://www.youtube.com/watch?v=YOvB5nlz98s&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=7

- NGDS Data Access Scenario 08: Performing Analyses with NGDS Data
- https://www.youtube.com/watch?v=YOvB5nlz98s&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=8

- NGDS Data Access Scenario 09: Performing Geospatial Analyses with the USGS National Map Viewer https://www.youtube.com/watch?v=eiF_3gCzplI&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=9

- NGDS Data Access Scenario 10: Performing Geospatial Analyses with ArcMap
  https://www.youtube.com/watch?v=JQiSI8eqpEg&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=10

- NGDS Data Access Scenario 11: Discovering Related Datasets in uDIG
  https://www.youtube.com/watch?v=ZU9sNFD_r10&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=11

- NGDS Data Access Scenario 12: OneGeology Maps and URIs
  https://www.youtube.com/watch?v=utLBacxNniM&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=12

- NGDS Data Access Scenario 13: WFS Symbology in QGIS
  https://www.youtube.com/watch?v=QIEYFEPoGrU&list=PLxRc6-3iV2qctZII8QTy6IUhd5G_n0_L-&index=13

Appendix H: Sustainability Plan

1. Sustainability plan draft (original-previously submitted to DOE)……2
2. USGIN Foundation development
    a. USGIN articles of Incorporation……23
    b. USGIN bylaws……28
3. USGIN website (description)……39

NATIONAL GEOTHERMAL DATA SYSTEM SUSTAINABILITY PLAN April 30, 2014

V1.0

M. Lee Allison, PI

Arizona Geological Survey Geoinformatics Section
Stephen M. Richard, Chief
Christy Caudill Daugherty, Deputy Chief

Kim Patten, editor

Acknowledgements and Disclaimer

This material is based upon work supported by the Department of Energy, Office of Energy Efficiency and Renewable Energy, Geothermal Technologies Office, under Award Numbers DE-EE0001120 and DE-EE0002850.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Contents

Background 2
Summary Conclusions 3
The National Geothermal Data System (NGDS) Vision 3
Goal 3
User Communities 3
System Architecture 4
Data Provider Software Package 4
End User Software 4
US Geoscience Information Network 5
NGDS Data Provider Software Package and End User Software 5
NGDS Software Requirements & Achievements 7
System Operations 8
Code Repository 8
NIAB Code 8
Licensing 9
Website and User Portal 9
Service Notifications 10
Maintenance Staffing 10
Maintaining Content: Data, Services, Applications 11
Regional Server Hubs 11
Regional Server Hardware 12
Replication and Disaster Recovery 12
Nodes on the Network 13
Possible Extensions & Improvements on the NIAB 14
Content Model (Information Exchange) Maintenance & Sustainability 15
Maintaining the Existing Data 16
Organization & Management 16
Non-Profit Organization 16
Technology Transfer (Education, Outreach, Training) 17
Foundation Website 17
Products & Services 17
Business Models and Client Development 19
Conclusions 20

Background

Since the 2009 American Recovery and Reinvestment Act (ARRA) the U.S. Department of Energy's Geothermal Technologies Office has funded $33.7 million for multiple data digitization and aggregation projects focused on making vast amounts of geothermal-relevant data available to industry for advancing geothermal exploration and development. These projects are collectively part of the National Geothermal Data System (NGDS), a distributed, online networked system for maintaining, sharing, and accessing data in an effort to lower the levelized cost of electricity (LCOE). Most of the projects involved in the NGDS conclude on or before April 30, 2014. The exception is the "State Geological Survey Contributions to the NGDS" project which received a no-cost extension through December 31, 2014 to deploy the Node-in-a-Box (NIAB) developed under the Architecture, Design, and Maintenance award to Boise State University. The investment in building and populating the NGDS has been substantial, both in terms of dollars and time; it is critical that this investment be protected by ensuring sustainability of the data, the software and systems, and the accessibility of the data. Only then, will the benefits be fully realized.

The NGDS vision is that the functions, accessibility, and availability of data will encourage continued participation within the NGDS, from all system users – data consumers, data providers, and application developers. As each data node adds to its data repositories, the system-wide NGDS functions become increasingly valuable to it. Each data provider will have created a value-added service that is transportable and scalable to cover all data in its possession, an example of such is the recent adoption by the State of Arizona to deploy a network built on the NGDS underlying framework for natural resource, environmental, and transportation data. Thus, there are benefits to each participant to continue to add data to the system and maintain it. The long term goal is that the data network reach a 'tipping point' at which it becomes like a data equivalent to the World Wide Web – where everyone will maintain the function because it is expected by its clientele and it fills critical needs. Applying this vision to the NGDS, it also opens the door for additional data providers external to geothermal development, thus increasing the value of NGDS and its underlying data integration platform, USGIN.

This plan is submitted to the NGDS Architecture, Design, and Testing project as a final deliverable for the April 30, 2014 project end-date. It should be noted that the plan is a living document and will be continuously updated throughout the course of 2014 to reflect the most up-to-date developments, information, and resources available to the NGDS. The plan builds off of the May 2013 vision paper "Sustainability Plan for National Geothermal Data System (NGDS) Operations," available via the NGDS GitHub, and describes the four pillars of sustainability: System Operations, Maintaining Content from Providers, Organization and Management, and Business Models.

In May of 2013, it was anticipated that NGDS would not be in full independent sustainable mode upon the completion of the Architecture, Design, and Testing project in April of 2014; this conclusion remains. This is due to a number of factors including the original ambiguity on the ownership and functionality of NGDS prior to the redirection of the Architecture, Design, and Testing project, which completed development of one of the core pieces of the NGDS in April of 2014, an easy mechanism for nodes to provide and share data.

Ownership of the NGDS has been assigned to the USGIN Foundation, Inc. the non-profit organization that serves as the support group to the USGIN platform. USGIN Foundation, Inc. is set up initially under the auspices of the Arizona Geological Survey. Additional details are provided in the text of this document. As anticipated in May 2013, there is a required transition period post April 30, 2014 to

continue maturation of the product, including deployment of the NIAB, requiring transition and maintenance support from the primary funding agency, US DOE. Transition and maintenance support is currently guaranteed through April 30, 2015 through a contract from Sandia National Laboratories (option to extend support up to 4 years in one-year increments).

Summary Conclusions

Based on the recommended actions in the May 2013 Sustainability Document, and the current state-of-development for the NGDS, the following can be concluded:

- USGIN Foundation, Inc. will be the long term owner of the USGIN data integration platform and the NGDS software, including the CKAN node package.
- USGIN Foundation, Inc. will include a Geothermal Advisory Committee to ensure that NGDS interests are pursued
- USGIN Foundation, Inc. will market the USGIN system as a solution to the Federal Open Data Access Executive Order and OMB Implementation Memo
- Upon conclusion of the primary data digitization and aggregation projects a snapshot repository of the data available via the NGDS will be backed-up via cloud storage and utilized should one of the contributors nodes fail
- Upon conclusion of the primary projects at least 8 data nodes and 3 regional hubs will be in operation with the hubs providing replication and disaster recovery
- During the transition period, , the NGDS CKAN extension will be cleaned of unused and redundant code to ensure ease of use for external developers in the free and open source CKAN network, enhancing opportunities for reuse (and thus updates to the code)
- During maintenance mode the NGDS CKAN extension will receive added improvements based on user testing, including web site revisions to improve the user experience

The National Geothermal Data System (NGDS) Vision

Goal

The ultimate goal of the National Geothermal Data System (NGDS) is to support the discovery and generation of geothermal sources of energy. The NGDS will provide online access to important geothermal-related data from a network of data providers in order to:

- Increase the efficiency of exploration, development and usage of geothermal energy by providing a basis for financial risk analysis of potential sites
- Assist state and federal agencies in making land and resource management assessments
- Foster the discovery of new geothermal resources by supporting ongoing and future geothermal-related research
- Increase industry and public awareness of geothermal energy development potential

User Communities

The National Geothermal Data System is a network of three linked communities:

1. **Data providers** who will expose information to the system through standardized, internet-accessible interfaces and interchange formats
2. **Software developers** who will build applications that utilize the data in the system, and make it easier for end-users to interact with the system.
3. **End-users** who will utilize the software and information provided by the system in order to understand and develop geothermal resources.

System Architecture

The NGDS includes data covering a wide range of topics, from well logs and drilling data to temperature, geochemical, and geophysical measurements. Standardized data access to important datasets will facilitate utilization of these information resources.

A key component of the system is the catalog service through which data providers register the availability of resources, and through which users discover, evaluate and access resources. A resource is considered part of the system when it can be located by searching the catalog service, which returns a metadata record describing how the resource can be accessed. Data providers maintain nodes in the network, connected through the use of standardized metadata for describing resources, content models for geothermal data, and common web-service protocols for exchanging information. These standards were developed in conjunction with the US Geoscience Information Network (USGIN), thereby providing interoperability with a wider range of geoscientific information.

Data Provider Software Package

A redistributable, open-source software package has been created to give data providers a simple way to register data sources, load data and expose those data as a node in the NGDS. The software supports batch import and upload of shared datasets in supported formats adhering to standard content models. The use of this software is not required in order to participate as a node in the network; data providers may use whatever tools they wish to expose their data, as long as they utilize interchange formats and web-service protocols conforming to NGDS specifications.

End User Software

End-users may interact with the system through a variety of entry points, but the project implemented two primary access points. As much as possible, these will be integrated in order to appear to the end-user as a single web-based experience. The two primary access points are as follows:

NGDS Website

The website is designed to provide information about the NGDS. It serves as an entry point to the system, allowing users to discover data and applications that utilize NGDS resources. The site includes information on the project's progress, NGDS specifications, access to the Map-Centric Search Application described below as well as other software applications utilizing NGDS data, presentations, documentation and tutorials, a catalog of NGDS nodes, and any other results as they become available.

Map-Centric Search Application

A user-friendly, web-based application has been created to support finding, visualizing, mapping, and acquisition of data by end-users. This application allows users to discover and access resources made available across all NGDS nodes, and to search for data across the system based on topic, location, time, provider, or key words. Standardized metadata describing each dataset provide the user with the information necessary to determine the utility of that dataset for their purposes. Geographic datasets visualized through a map interface also allow users to inspect the details of individual data points (e.g. wells, temperature measurements, etc.) from properly formatted datasets. In addition to visualization within the application, the interface provides the information necessary for users to access the data from other, third party software applications.

US Geoscience Information Network

The NGDS is designed based on a service-oriented approach using open standards to support data access by a wide variety of software applications, promote novel approaches to data analysis, and foster the development of tools by third parties. NGDS is based on the U.S. Geoscience Information Network (USGIN) which enables users to efficiently find, access, and share geoscience data, reducing the time and effort spent locating and integrating useful information and document new data by providing for information registration (by providers) and discovery (by users) based on standardized catalog services and metadata. The system accommodates resources in various forms, from unstructured text and images to documented, community Web services and interchange formats. To simplify, USGIN is a collection of Web-accessible resources that are registered in online catalogs and conform to data-sharing practices. Developer specifications for USGIN are available via the USGIN GitHub site at https://github.com/usgin.

The core components of the network are information exchange specifications and the catalog function. Information exchange specifications are community agreements on the conventions necessary for the interoperable exchange of some particular information. By using an interchange exchange it allows data in a variety of formats, organization, and structure to be integrated without having to manually transform it. The catalog function comprises a collection of metadata records that describe resources accessible through the network, and a special information exchange that defines metadata content, how the metadata collection is searched, and how metadata are encoded in search responses.

As a system, USGIN is a loosely coupled system of independent data providers, client applications, and infrastructure. The infrastructure includes 1) tools for registration of new resources, searching metadata catalogs, authentication, and resource validation; 2) registries for vocabularies, agents, specifications, and interchange schema; and 3) documentation and educational resources. Because network operation is based on information exchange specifications that are independent of any particular hardware or software implementation, all of the operational components can evolve as technology evolves. Use of standard protocols enables data access using off-the-shelf software, both commercial and open-source.

The distributed nature of the system means that stewardship of resources is determined by the resource owner. Participation in the network requires that a resource provider create metadata that conforms to the profile, and make the metadata and the described resource available. The network is open; anyone can deploy new nodes and components that implement one or more USGIN specifications, without requiring approval. New specifications can be introduced for service protocols, interchange formats, or vocabularies. Similarly, the system back end is decoupled from the front end, meaning anyone conforming to system configuration can build a portal into the system, consuming the data feeds and adding their own applications, visualizations, or other value-added functions. The International Renewable Energy Agency is already taking advantage of this, consuming key data sets for inclusion in the Global Atlas of Renewable Energy, accessible via the IRENA website www.irena.org. Keeping resources under the stewardship of the parties responsible for the information promotes system sustainability because the stewards have a direct connection with the quality of the product rather than submitting the data to a centralized database and related data manager.

NGDS Data Provider Software Package and End User Software

While USGIN permits application of a variety of service protocols, to simplify participation by NGDS data providers a simplified software stack, Node-in-a-Box (NIAB) was developed under the NGDS

Architecture, Design, and Testing Project (NGDS Design Project). The NIAB software targets organizations or individuals who wish to host at least one of the following:

- An online repository containing resources for the NGDS;
- An online site for creating metadata to register resources with the NGDS
- NGDS-conformant Web APIs (application programming interface) that enable access to NGDS data (WMS, WFS, WCS)
- NGDS-Conformant Web APIs that support the discovery of NGDS resources via catalog service (CSW)
- A web site that supports discovery and understanding of NGDS resources.

While a number of different open-source frameworks were identified for the development of this software stack, including existing systems in use by the NGDS, the NDGS Design Project determined to use the Open Knowledge Foundation's CKAN (http://ckan.org) implementation as it provided the closest match between out-of-the-box functionality and the NGDS NIAB requirements (see https://github.com/ngds/documents).

A major consideration in the evaluation of the software system was the long-term viability (sustainability) of the eventual application. By building on an existing, active and widely used open-source project, including by the US Office of Management and Budget for use on http://data.gov, it ensures that external developers can assist in the ongoing maintenance of the system. Other factors considered in the development framework included; 1) adaptability of the user interface to be compatible with an independently developed user experience concept for NGDS users; 2) ease of extensibility, with a plug-in architecture that allows addition of functionality without having to modify the core codebase; 3) support for geographic data and map-based search and data browsing; 4) support for administrative activities like user management, access control, and activity logging.

CKAN is written in Python and makes use of a variety of open source frameworks including Pylons, which itself is a combination of various open source frameworks integrated to form the basis for Web-based Enterprise-level applications. The primary CKAN user scenario is data storage and management and includes file storage, metadata management, and management of structured data. In addition, it offers a plug-in mechanism enabling developers to rapidly extend CKAN's core functionality; most importantly to the NGDS, the ability to support geographic features as well as exposing metadata according to the OGC standard catalog (CSW). CKAN implements crucial housekeeping features such as user management and logging, which can be tedious to implement but crucial for the site's usability.

While there are many attributes to the out-of-the-box CKAN functionality, there are a number of items that required NGDS developer modification including:

1. Modification of the User Interface (UI) to reflect the user experience testing as completed through the NGDS Design Project UI testing
2. Creation of the NGDS/USGIN metadata, including spatial extent
3. Consistency check for well-known structured data files (described in the next section)
4. Providing OGC services for uploaded structure files with geospatial information
5. Full-text indexing of documents
6. Role-based right for uploading and publishing data
7. User feedback and rating of uploaded data

The NGDS CKAN code is currently available for deployment at https://github.com/ngds/ckanext-ngds/ as well as registered with the CKAN developer code repository (https://github.com/okfn/ckan). During the

sustainability maintenance phase, USGIN developers will interact directly with the Open Knowledge Foundation to provide the following:

1. A more robust installation script
2. A mechanism for script updates to installed instances
3. Clean-up unused/redundant code
4. Ongoing organization and cleanup of the install instructions and repository

Using CKAN as the NIAB solution and primary UI meets many of the needs and software specifications outlined in the NGDS Software Requirements Summary, version 2.7, November 5, 2012 developed by the NGDS Design Project.

NGDS Software Requirements & Achievements

Based on the Requirements Summary, the following have been achieved and are presented as they have implications on the maintenance and sustainability of the network.

2.1 Maintenance:

- All project source code has comments, at least on a per-class level
- The system's architecture and configuration parameters are well documented
- The system's source code is covered by unit tests to at least 50% coverage

2.2 Usability and Accessibility:

- The system has a reasonably simple to use installation tool, with detailed instructions that guide the user through the process of installation and joining the network.
- The system is cloud ready, archived, and available via a Virtual Machine using Amazon Web Services
- The system uses project developed graphical user interfaces for a uniform look-and-feel for web applications
- The project-developed applications provide online help explaining how to perform user-related functions
- Key data import operations are transactional, i.e. the user can abort operations before completion without any negative consequences
- The project has a status indicator for data uploads
- The system underwent review for compliance with section 508 amendment to the Rehabilitation Act of 1973 and ISO/TS 16071 "Ergonomics of human-system interaction – Guidance on accessibility for human-computer interfaces."

Performance and Scalability:

- The system was designed to support a maximum of 1,000 concurrent users
- Each data provider node is capable of maintaining at a list of at least 100 other NGDS nodes for harvest or distributed search
- Each data provider node is capable of supporting up to 50 simultaneous authenticated, logged-in users, at least 50 HTTP requests every minute, and takes no more than 10 seconds to respond to each request
- They system can handle the import of data files up to 2BG in size, import up to 1,000 data files in any one import operation, store up to 100,000 data files in the import director of each data provider, and store up to 500GBs of data files (this is completed by the use of Amazon Web Services and thus is expandable to meet the needs of the system)

Security:

- Valid login authentication is required for all data submitters, stewards, and administrator functions
- Permits only valid users to write data they have permissions to write and delete data they have permissions to delete
- The data provider node maintains the integrity and availability of all data stored in its local data store and repository

Supportability:

- NGDS components are written using standard coding styles for the programming languages used (Python, Java)
- NGDS software is designed utilizing the concept of encapsulation, that is, components are encapsulated by related functionality

System Operations

Code Repository

As part of the systems commitment to open source software development two GitHub repositories for system design, specifications, documents, and code were created and are maintained by the Geoinformatics staff at the Arizona Geological Survey. The repositories are open for participation and available for public viewing.

Main NGDS Repository: https://github.com/ngds
CKAN Development Repository: https://github.com/ngds/ckanext-ngds/
NIAB Installation Repository: https://github.com/ngds/install-and-run

Programming languages used for the NGDS CKAN extension build are Python and JavaScript, with OS (open-source) components Apache SOLR and Tomcat, PostGIS, GeoServer, and GDAL. These were selected due to their support and use in the open source community and have strong communities of practice surrounding their use. Likelihood that the languages and components will be obsolete in the near future is slim.

NIAB Code

This section is specifically on the maintenance of the NIAB code and does not include details on the data maintenance and installed nodes. For information on data maintenance and installed nodes, please view the next section "Maintaining Content."

AZGS is currently in the process of working directly with the Open Knowledge Foundation to increase the viability and usability of the NGDS NIAB as well as to better align the NGDS extensions to the CKAN project code. This work will occur from May 2014 to December of 2014 to help improve the installation of at least six NIAB at State Geological Survey's across the country (additional information provided in the Maintaining Content section). This work will also make additional developers aware of the NGDS extensions, adding developers to the code team, and thus the maintenance of the extension code.

The current version of the code is tagged at version 1.0.1 and available at https://github.com/ngds/ckanext-ngds/. As improvements are made and new code is pushed to the

GitHub master branch, newer versions will always be tagged. Current users will be alerted by RSS feed, email, or other notification system when a new version is available, along with a list of upgrades and changes indicated for that version.

Licensing

Per the requirements of the NGDS Design Project, the software produced would have an open source license variant allowing users to copy, distribute, and transmit the software; adapt software for other applications; and make commercial use of the software under the condition that the attribution of the software is included in any copy or derived work. Thus, NGDS selected the GNU Afferno General Public License v.3 and the Creative Commons Attribution 3.0 Unported License. All NGDS software and development materials are licensed in this capacity, including the website and site materials.

- GNU Afferno General Public License, v.3 – specifically for software designed to ensure cooperation with the community, ensuring that users can share and change all versions of a program while maintaining its free use in the community
- Creative Commons Attribution 3.0 Unported License – requires that users provide appropriate credit, provide a link to the license and indicate if changes were made when they share (copy and redistribute the material) or adapt (remix, transform, and build upon the material)

Currently NGDS does not track or monitor usage or popularity of search terms due to the intellectual property associated with exploration activity. The NGDS website at http://geothermaldata.org does monitor page visits, DNS locations, and response time of the site to assist with improvements of the site.

In addition to the licenses for NGDS materials, the website is dedicated to the public domain through the Creative Commons 0 1.0 Universal (CC0 1.0) Public Domain Dedication. Thus the http://geothermaldata.org site includes the following statement:

> *The content of this web site is dedicated to the public domain. All rights to the work are waived worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.*
>
> *The NGDS requests that it be acknowledged as the source in any subsequent use of its information. Some materials on this site have been contributed by private individuals, companies, or organizations and include a copyright notice. It is the user's responsibility to contact copyright owners and obtain the written permission required under U.S. copyright law before using these materials.*
>
> *Links may be made to the NGDS Web site from personal and organization Web pages. NGDS requests that you link to its site rather than downloading portions of the site to another Web server so viewers will see the most up-to-date information.*
>
> *NGDS materials may not be used to state or imply the endorsement of DOE EERE, AZGS or any of their employees of a commercial product, service, or activity, or be used in any other manner that might mislead the public.*

Website and User Portal

The site http://geothermaldata.org is the primary site that provides information about the NGDS to all participants as well as serves as the gateway to the system to discover data and applications that utilize

NGDS resources. In addition the site provides informational resources on the project's progress, access to NGDS specifications (on the GitHub), access to a map-centric and library style search interface, access to other software applications utilizing NGDS services (Resources), NGDS tutorials (via YouTube and USGIN site), and a link to the catalog of NGDS nodes.

While the primary site is http://geothermaldata.org, we have also procured the following domain names and provided redirects to the primary site:

- Geothermaldatasystem.org
- Usgeothermaldatasystem.org

In addition, we have the http://stategeothermaldata.org and http://usgin.org domain names (as well as all associated subdomains of the five total URLs). All domain names are registered through the year 2020 to provide consistency in URLs. Currently all domains are hosted via Site5 on a shared server with unlimited disk space, bandwidth, and websites. However, due to recent issues in support with Site5 it was decided that the NGDS and related sites would move to KnownHost which guarantees an independent Virtual Private Server, thus providing additional power and speed to the sites. The system at KnownHost provides 1536MB of guaranteed RAM, 60GB of disk space, and 4000GB of Bandwidth for approximately $40 per month. This should accommodate the NGDS and related sites well into the future. However, should the sites require additional RAM, disk space, or bandwidth, there are five additional options for hosting at varying costs. Service includes a 99.9% SLA, free backups, high performance VPS hosting, and 24/7 support.

Additional advantages to managed hosting includes no hardware requirements, all downtime is covered by the hosting company, and the information/sites are easily transferrable should a new entity require management of the sites. Finally, the new hosting service provides a full dashboard for site statistics and monitoring.

In addition to general website hosting and maintenance, the NGDS requires cloud services for the NIAB GeoServer and multitude of Web Services. Throughout the duration of this project we have used Amazon Web Services (AWS) to meet these needs. AWS costs are established via the traffic and requested redundancy. Currently NGDS services cost approximately $150 per month. While we expect traffic to increase due to the v1 release of NIAB, we anticipate costs of approximately $200 per month for AWS for the near future (12-18 months). During the maintenance phase through April 30, 2015 we will evaluate traffic and adjust this plan accordingly.

Service Notifications

NGDS nodes and hubs are encouraged to use the USGIN RSS feed for service notifications when deploying web services. The tool is available at: https://github.com/usgin/service-notifications. This permits data providers the ability to quickly distribute notifications regarding the services that they host, such as planned downtime, unforeseen server issues, new service availability, etc. Subscription to the service notification is available via the USGIN notifications RSS feed on the GitHub.

Maintenance Staffing

Ongoing maintenance of each of these components will be the responsibility of the Arizona Geological Survey and in time, the USGIN Foundation Inc. It is expected that pure maintenance of such a system will require approximately ¼ developer/sys admin time. This does not include significant updates to the software code or user-experience, nor does it include help-desk or technology transfer assistance.

Maintaining Content: Data, Services, Applications

Regional Server Hubs

Four NGDS Regional Server Hubs (hubs, IL, KY, AZ, and NV) have been thus far tasked with hosting web services for their surrounding data providers (other state surveys or institutions) who lacked the technical capability to host their own data in web service format. Deployment of the Server Hubs was estimated at $100,000 per year. As the hubs move into a maintenance mode, we estimate maintenance at approximately $30,000 to $50,000 per year depending on the institution. This includes limited troubleshooting time, monthly data updates, backup and disaster recovery costs, and institutional facilities and administration costs. Currently 2 of the 3 hubs are using shared ArcGIS servers (1 has a dedicated ArcGIS server) and independent institutional infrastructure, as the concept of USGIN enables participants to use existing infrastructure systems with specified protocols and standards.

NGDS tier 3 data services are discrete data types where aspects are controlled by static schemas. One aspect that is specified in the schema is the layer name of a given data type; this assists with search and access of the data. As the hubs host multiple Tier 3 services from multiple states, they therefore serve multiple services of the same data type. This becomes an issue with the current CKAN NIAB application.

The current construct of the CKAN NIAB application is problematic for hubs or other entities who might serve multiple Tier 3 web services of the same type. This is due to the GeoServer web server platform used in NIAB where only one layer name can exist in any given GeoServer instance. Thus, even though NGDS hubs could use NIAB for service deployment and management, there are undesirable consequences. First, the web services may become too large for the current processing capabilities of most desktop/laptop machines, and second, the data management scheme (each state having its own dedicated service per data type) and proven successful workflow of the hubs would need to change. Over the life of the project, the hubs have been functioning through use of individualized workflows, which usually include ArcGIS Server and PostGIS for web services management coupled with metadata management at the state geothermal data repository (http://repository.stategeothermaldata.org/repository/).

Considering the possible NIAB limitations for those entities serving multiple datasets of the same type, two options are recommended for hubs moving forward; 1) continued use of the state data repository for metadata management and ArcGIS Server for web service deployment; 2) continued use of the current ArcGIS Server workflow for web service deployment coupled with NIAB metadata management.

Scenario 1: Hubs continue to use ArcGIS Server; Hubs continue to use the UI at repository.stategeothermaldata.org

Currently, NGDS hubs use ArcGIS software to deploy web services and manage the associated data formats (geodatabases, feature classes, data tables). When project deliverables (datasets) are deployed as web services, a metadata entry is made at repository.stategeothermaldata.org for that dataset with distributions that include the links to the live web services (WMS, WFS, ESRI rest page).

The current workflow of AASG deliverables (data) processing includes the creation of metadata for the services, where a 'New Resource' is created at repository.stategeothermaldata.org. Moving forward with this scenario, hubs will be responsible for creating metadata at repository.stategoethermaldata.org, which will require minimal additional training for hubs as they are

already familiar with the interface. The biggest hurdle will be to assist the hubs in properly integrating the metadata entry point into their current workflow.

Scenario 2: Hubs continue to use ArcGIS Server; Hubs use NIAB for creation and management of metadata for services, but not service deployment

As creating Tier 3 web services using NIAB becomes problematic for hubs, this scenario proposes that services are deployed as usual with ArcGIS Server. Instead of entering the metadata for the services at repository.stategeothermaldata.org, the hubs would use NIAB simply for the creation and management of that metadata.

When a "Link to a data service" is added as a resource in NIAB (creating a metadata record), it is possible to add multiple service distributions. These distributions are the WMS/WFS links. Once a metadata record is created for a service in NIAB, locate the Edit button on the page of the resource. Here, go through the "Link to a data service" workflow once more to add an additional link. This will be added to the same metadata record. The hurdle here is much the same as Scenario 1; hubs will decide which methods of metadata management best fit their workflow.

Until the NIAB can handle multiple layer names or GeoServer instances, a potential development target over the next year, we recommend that the hubs maintain Scenario 1 as installing the NIAB simply for metadata management would cause unnecessary costs for the regional hubs (for additional cloud services and server maintenance on services that they will not be implementing, i.e. GeoServer).

Regional Server Hardware

While the USGIN system is designed to accommodate a variety of data sharing "nodes" and platforms, due to the importance of the multi-state hosting of the Regional Server Hubs, we are recommending a more standardized platform for the hubs while moving forward.

In order to increase the Regional Network Hubs reliability, maximize uptime, and ongoing sustainability we recommend that the Hubs virtual machine host servers be dedicated services rather than using a shared server solution (as is currently the case). We estimate that a Dell tower server in the $2,000 range would be powerful enough to support this type of solution. We also recommend that each of the Hubs use a USB hard drive ($150) attached to the server as a backup drive. The free Windows Backup software included with Windows Server 2012 has the ability to perform scheduled backups to a USB drive that includes a complete image of the entire Hyper-V host server including any virtual machines on the server.

We also recommend setting up Hyper-V Replication (next section) between the hubs Hyper-V host servers for redundancy. If a hub server or virtual machine goes offline, Hyper-V Replication would provide a copy of the hub that could be quickly booted up on the backup Hub server until the main Hub server problem is resolved.

Replication and Disaster Recovery

AZGS is assisting with the replication and disaster recovery for the other hubs. Delays in virtualizing, due to institutional requirements, have postponed full implementation, although as of March of 2014 each hub was running a virtual machine. Hyper-V is a virtualization solution from Microsoft designed to create and operate virtual machines. Hyper-V replication is a feature of Hyper-V that provides the

ability to make a copy of a virtual machine at a different geographic location for disaster recovery purposes. In 2013, AZGS installed Hyper-V 2012 on an Arizona Geological Survey server to support the hubs.

In September of 2013, the NGDS virtual server at the Nevada Bureau of Mines and Geology (NBMG) failed catastrophically with a loss of 200GB of data. Fortunately, Hyper-V replication was engaged and the server mirrored, insuring that all NBMG data survived. Using the Hyper-V mirror to restore 100% of the data avoided a catastrophic loss of data that would have taken hundreds of hours to reconstitute; in this case it took approximately 3 days. AZGS has published a guide for installing and creating a Hyper-V replication environment. It is available via the AZGS repository at http://repository.azgs.az.gov/uri_gin/azgs/dlio/1567.

Nodes on the Network

An NGDS node is a web-accessible server that hosts at least one of the functional capabilities enumerated below to play a role in making geothermal-relevant data assets accessible.

To be considered a node in the NGDS, the following criteria must be met:

1. The capabilities offered must play a role in making geothermal-relevant data assets accessible.
2. The capabilities offered must be publicly Web-accessible.
3. A node must offer at least one of the following capabilities:
    a. Host a Web-accessible repository of geothermal-relevant data assets with metadata conforming to the USGIN ISO profile (USGIN Standards and Protocols Drafting Team, 2010-11) published through an NGDS catalog node.
    b. Host a web-accessible folder that 1) contains NGDS-conformant metadata files and 2) is registered for harvesting by an NGDS catalog node.
    c. Host NGDS-conformant web services (WMS, WFS, WCS, etc.) that are registered in an NGDS catalog.
    d. Host an NGDS catalog node, which is a server operating a CSW 2.0.2 service that offers metadata conforming to the USGIN ISO profile.

NGDS nodes should self-identify by providing an **NGDS node self-description document**. Additional information on Node Specifications are available at: https://github.com/ngds/install-and-run/tree/master/Becoming%20an%20NGDS%20Node

There are currently two types of NGDS nodes; the NIAB installs and organizations that use their in house systems to meet the requirements above. Currently, the following nodes are hosting their own data via NGDS Tier 3 web services:

- Southern Methodist University's Geothermal Laboratory (independent project with independent sustainability plan, SMU has guaranteed to house the web services and data for at least 10 years; how the SMU node will respond to issues with metadata and general server maintenance are unknown)
- Illinois Geological Survey (also a hub)
- Kansas Geological Survey
- Kentucky Geological Survey (also a hub)
- Minnesota Geological Survey
- New Hampshire Geological Survey
- Nevada Bureau of Mines and Geology (also a hub)

- Oregon Department of Geology and Mineral Industries
- Washington Department of Natural Resources, Geology & Earth Sciences
- Wisconsin Geological Survey
- Arizona Geological Survey (also a hub)
- Minnesota Geological Survey

As part of the NGDS Design project additional installations of the NIAB were to be deployed, including the metadata aggregator node. Currently the metadata aggregator node is housed at the Arizona Geological Survey (AZGS). Two instances of the NIAB are successfully installed at the Arkansas Geological Survey and the New Mexico Bureau of Geology and Mineral Resources. An additional node at the University of Utah's Earth and Geosciences Institute was to be installed by the time of project completion however, ongoing server issues at EGI prevented the completion of this node by April 30. We will continue to attempt NIAB installation at EGI, if at all possible.

During the May to December 2014 time period, AZGS will work to install a minimum of six NIAB installations across the country at State Geological Surveys. During that time we will also work to maintain the NIAB installs. The following Surveys have been identified as targets:

First Tier Targets:

- Alaska Division of Geological & Geophysical Surveys
- University of Hawaii, Manoa, Geological Department
- Idaho Geological Survey (Confirmed)
- Oregon Department of Geology and Mineral Industries (Confirmed)
- Utah Geological Survey
- Virginia Division of Geology and Mineral Resources (Confirmed)
- Wyoming Geological Survey

Second Tier Targets:

- Geological Survey of Alabama
- Indiana Geological Survey
- Missouri Geological Survey

For the one time setup and maintenance of the NIAB installations, we estimate approximately $12,500 to $25,000 a year in virtual server costs, developer/sys admin time, and facilities & administration costs. As with all costs, this will vary dependent on salary and F&A rates at the institutions; however, institutions should estimate approximately an 80%/20% split between personnel and equipment. Participation in the setup and ongoing testing (May to December 2014) may run as high as $50,000 to $75,000 a year as this requires significant developer/sys admin time.

Possible Extensions & Improvements on the NIAB

The software specifications and requirements document outline a series of requirements that were not met for the Version 1 release. For example, the bulk upload of data sets was a target that was not met. In addition, testing of the v1 release over the next 6-12 months will generate a series of additional user-inspired requirements that will require implementation. One such example is for a metadata repository extension of the NIAB, that is, a NIAB that does not require installation of the GeoServer platform and can be used by entities or organizations that simply want a mechanism for contributing metadata to the network while hosting their own services. As these requirements are developed, issues will be logged

and prioritized on the NGDS GitHub. The AZGS and USGIN Foundation Inc. team will accommodate as many of these priorities as possible.

Content Model (Information Exchange) Maintenance & Sustainability

Content models, or information exchanges, in use for the NGDS were established using a community developed mechanism and approved by the Department of Energy's Geothermal Data System Development and Population Working Group (GDSDPWG), which was comprised of the technical leads from each project and a series of technical monitors. When available, existing exchanges were used. Information in deprecated exchanges were generally synthesized with another exchange and thus, are still available. All available exchanges are listed at http://schemas.usgin.org/home/.

Should the likely event occur that additional exchanges are required to expand the usage and applicability of the system, the following methods are proposed. Full details are available at https://github.com/usgin-models.

The decision to define a new information exchange should be based on the likelihood that others will want to publish similar datasets in the future. Members of the USGIN community propose specifications for data sharing exchanges. Exchange documents are developed and reviewed using a publicly accessible repository on GitHub (https://github.com/usgin-models). Each exchange has a separate repository associated with the USGIN-models pseudo organization. A proposed model must have an identified steward, and a working group of at least three participants with relevant domain knowledge and understanding of the interchange technology. There is no formal process for defining workgroup membership; normally the challenge is finding a sufficient number of qualified individuals to provide meaningful reviews and comment.

The exchange steward is responsible for assembling the workgroup and assuring sufficient expertise in the group to generate a sound content model and implementation. The exchange steward requests creation of a new model repository at the USGIN-models GitHub from the organization members (currently primarily AZGS Geoinformatics staff), and identifies workgroup members who will have commit privileges on the repository. Any community member can create a repository branch to propose changes using standard GitHub procedures, and request consideration for merging back into the developing model.

After review and approval by the workgroup, a call goes out to a USGIN technical review e-mail list or by RSS feed (http://notifications.usgin.org/) for comments from the community. An open review period of 4 weeks is normal, after which any comments from the community must be resolved to the satisfaction of the commenter. When issues are resolved to the satisfaction of the stakeholders (workgroup and engaged community), the exchange specification is adopted.

When a specification is adopted, all associated documents are copied to a 'tag' branch in the GitHub repository and are not changed after they are 'tagged'. The Specification documents are also copied to the exchange repository at http://schemas.usgin.org, which is a web site set up to provide public access to exchange specifications and any related xml schema documents or other artifacts required for the deployment and operation of the information exchange.

This mechanism ensures that the development of specifications continues to be community led as well as providing a source-trail for the development of such exchanges.

Maintaining the Existing Data

Experience has dictated that server outages can and will occur. In addition, the potential loss of data from a distributed network, while not as great as "closing the doors" on a traditional database, may occur. In order to preserve the vast amount of data digitization and preservation, the Arizona Geological Survey plans to conduct a one-time data harvest to capture the data available within the NGDS and place it as a back-up and disaster recovery mechanism on both cloud storage and external hard drive. This process will be complete by May 30, 2014. The external hard-drive will be submitted to DOE as a deliverable under the State Geological Survey Contributions project (end date 12/2014).

Maintaining the Client Developed Applications

NGDS and USGIN support client side developed applications similar to applications for the iPhone or Android phones. As such, externally developed applications are subject to the maintenance schedules of the developers and not the NGDS or USGIN.

Organization & Management

Non-Profit Organization

Per the May 2013 sustainability paper, we suggested that a stand along organization offered the greatest independence and likely adherence to the organizations vision, goals, and missions and thus the longevity of a system such as the NGDS. In the fall of 2013, the Arizona Geological Survey, with the guidance of the Arizona Attorney General appointed counsel of Rusing, Lopez, and Lizzardi, PLLC, have successfully incorporated the USGIN Foundation, Inc. with the State of Arizona's Corporation Commission as well as filed initial requests with the Internal Revenue Service. Not only is this step in alignment with the recommendations of the May 2013 paper, but also aligns with the Hutchison and Richard 2011 "Recommendations for the Future of the U.S. Geoscience Information Network," joint report to the Association of American State Geologists and U.S. Geological Survey.

Thus far, we have confirmed the following for the USGIN Foundation Inc.

- Articles of Incorporation with the Arizona Corporation Commission, 12/05/13, file number 1890284-5 (full details are available at: http://starpas.azcc.gov/scripts/cgiip.exe/WService=wsbroker1/names-detail.p?name-id=18902845&type=CORPORATION)
    - Articles of Incorporation
    - Publication of the Articles of Incorporation
- Completed SS4 and received EIN from the IRS: 46-4569299
- Corporate by-laws signed
- Board members being identified – solicitations to the AASG
    - Advisory board members being identified – potential overlap with GRC/GEA

We are still awaiting IRS designation as a 501c3 scientific organization. In the interim, we have been working with our marketing consultants to establish proposed service rates, similar to the Socrata structure or Open Knowledge Foundation service package (additional information on the service packages is described below in the Technology Transfer section).

At this time, the USGIN Foundation Inc. is co-located with the Arizona Geological Survey at 416 W. Congress Street, Suite 100, Tucson, AZ 85719. The State of Arizona has precedence for such activity as

the Arizona Game & Fish Department and the Arizona Department of Tourism have related non-profit organizations. While the USGIN Foundation Inc. is co-located in the physical property of the AZGS, the eventual goal is to move the Foundation off state property to ensure that the Foundation is not subject to political activities.

The Foundation provides the following expertise:

- Earth Science Data
- Project Management Excellence & Streamlining of Workflows
- Managing Geospatial Data
- Knowledge of Unstructured AND Structured data formats
- Interoperable Systems
- Data Provider Community Building
- Visions for Long-Term Sustainability

Duties of the USGIN Foundation Inc. will include:

- Management of the Foundation
    - Strategic Planning
    - Fiduciary
- Client Systems Development & Maintenance
    - Programming/Development
    - Data Curation
    - Metadata Maintenance and Curation
    - Systems Architecture
    - Systems Administration
- Business Development, including Outreach & Training
    - Contract Development and Negotiation
    - Government and Foundation Grant Development and Negotiation

Technology Transfer (Education, Outreach, Training)

Foundation Website

As part of our improved outreach and future business development we have created a more user friendly website to introduce potential clients to the system. The revised page is at the existing http://usgin.org page with the technical specifications and tutorials moving to a subdomain at http://tech.usgin.org. All product descriptions, once finalized, will be available and accessible via the USGIN homepage.

In addition to the traditional website at http://usgin.org, specifications are available via http://github.org/usgin, and scenario and user video tutorials at www.youtube.com/geothermaldata. The videos available on the YouTube site include 13 data access scenarios guiding users through the search interface as well as a variety of software clients to show the diversity of the system.

Products & Services

In cooperation with our marketing and product development firm, Blue Canoe Marketing, we have four deliverables, in varying stages of completion, outlining next steps for business development, outreach, and training. These include: 1) the USGIN Suite of Solutions for Open Data Access, 2) a Compliance Guide for Federal Open Data Compliance, 3) the NIAB Marketing Plan, and 4) a 12 month marketing and communications plan for business development.

USGIN Suite of Solutions

There three types of services that the Foundation will offer: Marketing/Pre-Consultation, Consulting Services, and Support Services. Brief information on each of the types of services is offered below. Note that the services are designed to build upon each other.

Marketing/Pre-Consultation

- Workshops
- Training & Education

Consulting Services

- Assessments – provides a plan and recommendations for Open Data Compliance
- Open Data Strategy and Action Plan – provides a scope of work, specifications, and a proposal to build a basic data sharing system
- Interoperable Data Strategy and Action Plan - provides a scope of work, specifications, and a proposal to build an advanced interoperable data sharing system
- Deploying Your Own Node – provides individualized data node with documentation (installation guide, technical support, help desk requirements) and independent data control for the host
- Hosting – provides a migration plan for sustainability of the independent node

Support Services

- Open Data Development – provides hands-on modification of data, catalog creation, and customized workflows (based on the results of the Assessment)
- Interoperable Data Network Development - provides hands-on modification of data, catalog creation, customized workflows (based on the results of the Assessment), community building efforts (data sharing affiliates), and information exchange specifications
- Annual Subscription & Support – provides hosting services, data maintenance, replication services

Fees associated with each service will depend on the scope and size of the proposed project. These services, while intended to be stand-along services offered to potential clients, can also be revised to respond to funding opportunities in government and foundations. For example, the State of Arizona's Natural Resources Review Council recently approved a plan to develop a Natural Resources Decision Support System (NRDSS) to link data, documents, and GIS layers among nine state natural resources, environmental, and transportation agencies, as well as with corresponding agencies, using a USGIN style system. In response, we will be able to provide a best estimate for total cost of such a system using the clearly delineated services defined above. In this instance, we would include services for assessment of available data resources, and then provide development of the interoperable data network, concluding with a set of proposed annual subscription fees for maintenance and support.

Compliance Guide for Federal Open Data Compliance

To aid and assist with marketing of a USGIN style network, and to communicate the benefits of a USGIN style network to non-data or technical stakeholders, we have developed an Open Data Compliance Guide in response to the Executive Order 13642 and OMB Memorandum M-13-13. The Guide provides descriptions on the open source tools and resources available for potential client applications that go beyond the OMB requirements to create an interoperable, federated data network, similar to the NGDS. Details described within the guide discuss the benefits of USGIN and NGDS, information on how to

create and maintain an Enterprise Data Inventory, how to achieve interoperability with that inventory, and how to create and maintain a public data listing (or USGIN style catalog). In addition, the Guide describes mechanisms for engaging with stakeholders and how to document data that is currently under moratorium. Once completed, this guide will also provide a brief overview of how USGIN can help provide these services.

NIAB Marketing Plan

As described in the Data Maintenance section of this plan, a series of State Geological Survey's will be selected to deploy the NIAB software stack and migrate their services from the Regional Server Hubs to their own node. In order to assist them in this effort, we will provide an easy to follow web page with concept graphics, a series of overview slides, and then work with the Survey's to conduct continual user-feedback which we can incorporate both into the software and the long-term marketing of the NIAB software stack.

12 Month Marketing & Communications Plan

Target markets for the NGDS and USGIN Foundation Inc. are other geoscience data providers in the Federal, State, and Private Sectors. We plan to maintain and build our connections within the Geothermal Energy Sector and strategically expand to other Natural Resource sectors, then the Natural Sciences, and broader Environmental communities. To do this, we are currently crafting a 12 month marketing strategy for NGDS and USGIN that will include the following components:

- Prioritized customer segments and targets
- Development of core messaging and unique selling positions for the top tier segments
- Development of an integrated marketing and communication strategy aimed at the target segments that addresses: email marketing, national and international open data/digital data events and conferences; trade association participation and events; public relations and press releases; social media strategy; digital and traditional advertising; and lead generation materials
- A set of recommended marketing activities by month
- Estimated budget to execute the marketing plan

We have already begun implementing the revised marketing strategy with a test-exhibit at the American Association of Petroleum Geologists Annual Meeting & Expo with a revised exhibit pop-up and literature.

Business Models and Client Development

The primary services offered by USGIN Foundation Inc. in support of new interoperable data systems are primarily described in the "USGIN Suite of Solutions" section. However, the section does not address maintenance of existing systems, beyond a general overhead expenditure of the Foundation or through sustained annual subscriptions. As such, we have proposed the following mechanisms for sustained funding:

- Sponsorship – from interested companies such as Esri, Microsoft Research (which offered Azure cloud computing storage), Google
- Donations – from individuals similar to Wikimedia Foundation or Foundation support such as through the Gordon & Betty Moore Foundation or Sloan Foundation
- Advertisement – primarily web based
- Premium User Subscriptions – through the development of improved applications interfaces

These funding mechanisms will be tested through user-feedback during the maintenance phase of the NGDS, and refined based on the success of the 12 month marketing plan.

Conclusions

The National Geothermal Data System and USGIN Foundation Inc. are still largely in development mode; however, with the release of the Node-in-a-Box v1 and NGDS user interface v1 in April, as well as the closure of the primary data collection efforts, we can now focus on the next steps to ensure NGDS and USGIN are sustained and maintained well into the future. These steps include significant user experience testing, marketing of the existing system, and promotion of new interoperable data networks.

ARTICLES OF INCORPORATION
OF
USGIN FOUNDATION, INC.
an Arizona Nonprofit Corporation

The undersigned individuals, acting as incorporators under Arizona Nonprofit Corporation Law, adopts the following Articles of Incorporation.

ARTICLE 1.

The name of the Corporation shall be: USGIN Foundation, Inc.

ARTICLE 2.

The principal office and place of transacting the business of the Corporation shall be 416 W. Congress Street, Suite 100, Tucson, AZ 85701; however, its Board of Directors may meet for the transaction of business at such other places as the Directors may from time to time designate.

ARTICLE 3.

This Corporation is organized exclusively for charitable and scientific purposes, including, for such purposes, the making of distributions to organizations that qualify as exempt organizations under section 501(c)(3) of the Internal Revenue Code, or the corresponding section of any future federal tax code.

ARTICLE 4.

This Corporation will not have members.

ARTICLE 5.

No part of the net earnings of the corporation shall inure to the benefit of, or be distributable to its members, trustees, officers, or other private persons, except that the corporation shall be authorized and empowered to pay reasonable compensation for services rendered and to make payments and distributions in furtherance of the purposes set forth in Article 3 hereof. No substantial part of the activities of the corporation shall be the carrying on of propaganda, or otherwise attempting to influence legislation, and the corporation shall not participate in, or intervene in (including the publishing or distribution of statements) any political campaign on behalf of or in opposition to any candidate for public office. Notwithstanding any other provision of these articles, the corporation shall not carry on any other activities not permitted to be carried on (a) by a corporation exempt from federal income tax under section 501(c)(3) of the Internal Revenue Code, or the corresponding section of any future federal tax code, (b) by a corporation, contributions to which are deductible under section 170(c)(2) of the Internal Revenue Code, or the corresponding section of any future federal tax code or (c) act in violation of A.R.S. §10-3860, or the corresponding provisions of any future Arizona nonprofit corporation law.

ARTICLE 6.

Upon the dissolution of the corporation, assets shall be distributed for one or more exempt purposes within the meaning of section 501(c)(3) of the Internal Revenue Code, or the corresponding section of any future federal tax code, or shall be distributed to the federal government, or to a state or local government, for a public purpose. Any such assets not so disposed of shall be disposed of by a Court of Competent Jurisdiction of the county in which the principal office of the corporation is then located, exclusively for such purposes or to such organization or organizations, as said Court shall determine, which are organized and operated exclusively for such purposes.

ARTICLE 7.

The corporation will distribute its income for each tax year at a time and in a manner as not to become subject to the tax on undistributed income imposed by section 4942 of the Internal Revenue Code, or the corresponding section of any future federal tax code.

ARTICLE 8.

The corporation will not engage in any act of self-dealing as defined in section 4941(d) of the Internal Revenue Code, or the corresponding section of any future federal tax code.

ARTICLE 9.

The corporation will not retain any excess business holdings as defined in section 4943(c) of the Internal Revenue Code, or the corresponding section of any future federal tax code.

ARTICLE 10.

The corporation will not make any investments in a manner as to subject it to tax under section 4944 of the Internal Revenue Code, or the corresponding section of any future federal tax code.

ARTICLE 11.

The corporation will not make any taxable expenditure as defined in section 4945 of the Internal Revenue Code, or the corresponding section of any future federal tax code.

ARTICLE 12.

The address of the registered office of the corporation is:

416 W. Congress Street, Suite 100
Tucson, AZ 85701

ARTICLE 13.

The name and business address of the agent for service of process:

Rusing Lopez & Lizardi, P.L.L.C.
6363 N. Swan Rd., Ste. 151
Tucson, AZ 85718

ARTICLE 14.

The names and addresses of the Incorporators are:

Dr. Stephen M. Richard	416 W. Congress Street, Suite 100, Tucson, AZ 85701
Dr. M. Lee Allison	416 W. Congress Street, Suite 100, Tucson, AZ 85701

ARTICLE 15.

The names and addresses of the persons who shall serve as the board of directors of the Corporation, until their successors are elected and qualified, are:

Dr. Stephen M. Richard	416 W. Congress Street, Suite 100, Tucson, AZ 85701
Dr. M. Lee Allison	416 W. Congress Street, Suite 100, Tucson, AZ 85701

IN WITNESS WHEREOF, these Articles of Incorporation have been executed on this ____ day of November, 2013.

Dr. Stephen M. Richard
Incorporator

Dr. M. Lee Allison
Incorporator

Rusing Lopez & Lizardi, P.L.L.C., having been designated to act as Statutory Agent, hereby consents to act in that capacity until removal or resignation is submitted in accordance with Arizona Revised Statutes.

By: _____
Oscar S. Lizardi, Member
Rusing Lopez & Lizardi, P.L.L.C.

DO NOT WRITE ABOVE THIS LINE; RESERVED FOR ACC USE ONLY.

CERTIFICATE OF DISCLOSURE

Read the Instructions C003i

1. **ENTITY NAME** – give the exact name of the corporation in Arizona:

   USGIN Foundation, Inc.

2. **A.C.C. FILE NUMBER** (if already incorporated or registered in AZ):_____
   Find the A.C.C. file number on the upper corner of filed documents OR on our website at: http://www.azcc.gov/Divisions/Corporations

3. **Check only one of the following to indicate the type of Certificate:**
   - ☒ Initial (accompanies formation or registration documents)
   - ☐ Annual (credit unions and loan companies only)
   - ☐ Supplemental to COD filed _____ (supplements a previously-filed Certificate of Disclosure)

4. FELONY/JUDGMENT QUESTIONS : Has any person (a) who is currently an officer, director, trustee, or incorporator, or (b) who controls or holds over ten per cent of the issued and outstanding common shares or ten per cent of any other proprietary, beneficial or membership interest in the corporation been:		
4.1 Convicted of a felony involving a transaction in securities, consumer fraud or antitrust in any state or federal jurisdiction within the seven year period immediately preceding the signing of this certificate?	☐ Yes	☒ No
4.2 Convicted of a felony, the essential elements of which consisted of fraud, misrepresentation, theft by false pretenses or restraint of trade or monopoly in any state or federal jurisdiction within the seven-year period immediately preceding the signing of this certificate?	☐ Yes	☒ No
4.3 Subject to an injunction, judgment, decree or permanent order of any state or federal court entered within the seven-year period immediately preceding the signing of this certificate, involving any of the following: a. The violation of fraud or registration provisions of the securities laws of that jurisdiction; b. The violation of the consumer fraud laws of that jurisdiction; c. The violation of the antitrust or restraint of trade laws of that jurisdiction?	☐ Yes	☒ No
4.4 If any of the answers to numbers 4.1, 4.2, or 4.3 are **YES**, you **MUST** complete and attach a form C004.		

5. BANKRUPTCY QUESTION:		
5.1 Has any person (a) who is currently an officer, director, trustee, incorporator, or (b) who controls or holds over twenty per cent of the issued and outstanding common shares or twenty per cent of any other proprietary, beneficial or membership interest in the corporation, served in any such capacity or held a twenty per cent interest in **any other corporation** (not the one filing this Certificate) on the bankruptcy or receivership **of the other corporation**?	☐ Yes	☒ No
5.2 If the answer to number 5.1 is **YES**, you **MUST** complete and attach a Certificate of form C005.		

IMPORTANT: If within 60 days of the delivery of this Certificate to the A.C.C. any person not included in this Certificate becomes an officer, director, trustee or person controlling or holding over ten per cent of the issued and outstanding shares or ten per cent of any other proprietary, beneficial or membership interest in the corporation, the corporation must submit a SUPPLEMENTAL Certificate providing information about that person, signed by all incorporators or by a duly elected and authorized officer.

SIGNATURE REQUIREMENTS:	
Initial Certificate of Disclosure:	This Certificate must be signed by all incorporators. If more space is needed, complete and attach an form C084.
Foreign corporations:	This Certificate may be signed by a duly authorized officer or by the Chairman of the Board of Directors.
Credit Unions and Loan Companies:	This Certificate must be signed by any 2 officers or directors.

Dr. Stephen M. Richard
Name

416 W. Congress Street, Suite 100
Address 1

Address 2

Tucson | AZ | 85701
City | State | Zip

UNITED STATES
Country

SIGNATURE – *see Instructions C003i:*

By typing or entering my name and checking the box marked "I accept" below, I intend to affix my electronic signature and (or through my physical signature appearing below) I acknowledge *under penalty of perjury* that this document together with any attachments is submitted in compliance with Arizona law.

☒ I ACCEPT

Signature

Dr. Stephen M. Richard 12-3-2013
Printed Name Date

REQUIRED – check only one:

- ☒ **Incorporator -** I am an incorporator of the corporation submitting this Certificate.
- ☐ **Officer -** I am an officer of the corporation submitting this Certificate
- ☐ **Chairman of the Board of Directors -** I am the Chairman of the Board of Directors of the corporation submitting this Certificate.
- ☐ **Director** – I am a Director of the credit union or loan company submitting this Certificate.

Dr. M. Lee Allison
Name

416 W. Congress Street, Suite 100
Address 1

Address 2

Tucson | AZ | 85701
City | State | Zip

UNITED STATES
Country

SIGNATURE – *see Instructions C003i:*

By typing or entering my name and checking the box marked "I accept" below, I intend to affix my electronic signature and (or through my physical signature appearing below) I acknowledge *under penalty of perjury* that this document together with any attachments is submitted in compliance with Arizona law.

☒ I ACCEPT

Signature

Dr. M. Lee Allison 12-3-13
Printed Name Date

REQUIRED – check only one:

- ☒ **Incorporator -** I am an incorporator of the corporation submitting this Certificate.
- ☐ **Officer -** I am an officer of the corporation submitting this Certificate
- ☐ **Chairman of the Board of Directors -** I am the Chairman of the Board of Directors of the corporation submitting this Certificate.
- ☐ **Director** – I am a Director of the credit union or loan company submitting this Certificate.

Filing Fee: None All fees are nonrefundable - see Instructions.	Mail: Arizona Corporation Commission - Corporate Filings Section 1300 W. Washington St., Phoenix, Arizona 85007 Fax: 602-542-4100

Please be advised that A.C.C. forms reflect only the **minimum** provisions required by statute. You should seek private legal counsel for those matters that may pertain to the individual needs of your business.
All documents filed with the Arizona Corporation Commission are **public record** and are open for public inspection.
If you have questions after reading the Instructions, please call 602-542-3026 or (within Arizona only) 800-345-5819.

BYLAWS
OF
USGIN FOUNDATION, INC.
an Arizona Nonprofit Corporation

ARTICLE I
OFFICES AND CORPORATE SEAL

SECTION 1. Name of Corporation. The name of the Corporation shall be USGIN Foundation, Inc., hereinafter referred to as the "Corporation."

SECTION 2. Principal Office. The principal office of the Corporation shall be located at 416 W. Congress Street, Suite 100, Tucson, AZ 85701, but other offices may be established and maintained within or outside of the State of Arizona at such places as the board of directors may designate.

SECTION 3. Registered Office and Registered Agent. The Corporation shall maintain a registered office and registered agent in the State of Arizona. The registered office may, but need not be, the same as any of its places of business. The Corporation shall have and continuously maintain within the State of Arizona a statutory agent as required by law and the statutory agent may be changed from time to time by the Board of Directors.

SECTION 4. Corporate Seal. A corporate seal shall not be requisite to the validity of any instrument executed by or on behalf of the Corporation, but nevertheless if in any instance a corporate seal be used, the same shall be a circle having on the circumference thereof the name of the Corporation and in the center the words "corporate seal," the year incorporated, and the state where incorporated.

ARTICLE II
PURPOSE AND POWERS

SECTION 1. General Purpose. The Corporation is organized and shall be operated exclusively for educational, scientific, literary and charitable purposes within the meaning of Section 501(c)(3) of the Internal Revenue Code of 1986, as amended, and the corporation shall receive and administer, expend and distribute funds exclusively for such purposes.

SECTION 2. Powers. The Corporation shall have all powers permitted by law and may do and perform all things, in addition to those expressly enumerated herein or in the Articles of Incorporation, which are necessary or appropriate to the purposes and activities of the Corporation except as otherwise set forth in these Bylaws or the Corporation's Articles of Incorporation.

SECTION 3. Public Charity Status. The Corporation is expected to operate at all times as an organization that is not a private foundation. If at any time the Corporation is a private foundation, as described in Code § 509(a), then during such time the Corporation shall distribute its income for each tax year at such time and in such manner as not to become subject to the tax on undistributed income imposed by Code § 4942, and further, the Corporation shall not engage

in any act of self-dealing as defined in Code § 4941(d), nor retain any excess business holdings as defined in Code § 4943(c), nor make any investments in such manner as to incur tax liability under Code § 4944, nor may the Corporation make any taxable expenditures as defined in Code § 4945(d) or as such laws are amended from time to time.

ARTICLE III
MEMBERSHIP

SECTION 1. Members. The Corporation shall have no Members.

ARTICLE IV
BOARD OF DIRECTORS

SECTION 1. Function. The chief function of the board of directors is to manage the affairs of the Corporation and to amend, when necessary, the Bylaws, as provided by **Article XV.** Subject to the foregoing limitations and to further limitations contained in these Bylaws or in any statute as to action to be authorized or approved by the directors, all corporate powers shall be exercised by or under authority of, and the business and affairs of the association shall be controlled by, the board of directors.

SECTION 2. Number. There shall be at least one (1) and not more than fifteen (15) directors on the board of directors, provided the number shall at all times be odd. The exact number of directors shall be determined by the then existing directors of the Corporation, in a manner consistent with these Bylaws.

SECTION 3. Term. The term of office of each director shall be two (2) years and until the election and qualification of a successor. The directors may, by the resolution or resolutions electing each director, provide that the terms of office shall be shorter or longer than as provided above, so that the terms of office may be staggered so that at least one-third of the directors shall be elected each year.

SECTION 4. Voting Rights - Proxies Prohibited. Each director shall be entitled to one vote, and voting rights may not be exercised by proxies. Voting may happen in person, via telephone, or email.

SECTION 5. Qualifications. No person under the age of twenty-one years shall be eligible for election as a director. There may be no "disqualified person" appointed or elected to the board of directors, as defined in Internal Revenue Code § 4946.

SECTION 6. Elections. Directors of this Corporation shall be selected by the majority vote of the then existing directors of this Corporation. Directors to fill vacancies arising from the expiration of terms shall be elected at the annual meeting of directors, but if any such meeting is not held or such directors are not elected thereat, they may be elected at any meeting of directors. Directors elected to fill vacancies shall hold office until the expiration of their respective terms and until their respective successors are elected or until their office becomes vacant as hereinafter provided.

Directors to fill vacancies arising for any reason other than the expiration of terms shall be elected at the annual meeting of directors or at a special meeting of the directors held for that purpose. Each director elected to fill such vacancy shall hold office for the balance of the unexpired term of his/her predecessor and until his/her successor is elected, or until his/her office becomes vacant as hereinafter provided. If no predecessor shall have previously occupied the vacancy to which such director is elected, his/her term shall be specified to be a full or partial term in the resolution electing him/her, and he/she shall hold office for the balance of such term and until his/her successor is elected, or until his/her office becomes vacant as hereinafter provided.

SECTION 7. Resignation, Removal and Vacancies. Any director may resign at any time by giving written notice to the chairman, to the secretary of the Corporation, or to the board of directors. Such resignation shall take effect at the time specified therein, and unless otherwise specified therein, the acceptance of such resignation shall not be necessary to make it effective. If any director tenders his/her resignation, to take effect at a future time, the directors shall have power to elect a successor to take office at such time as the resignation becomes effective.

Any director may be removed with or without cause by a majority vote of all directors at a meeting of directors called for that purpose.

A vacancy or vacancies shall be deemed to exist in the case of death, resignation, or removal of any director, or the failure of any director to attend three successive meetings without reason acceptable to the majority of the other directors, or if the directors shall increase the authorized number of directors but shall fail to elect the additional director so provided for, or in the event the directors fail at any time to elect the full number of authorized directors.

No reduction of the number of directors shall have the effect of removing any director prior to the expiration of his/her term of office.

ARTICLE V
MEETINGS OF THE BOARD OF DIRECTORS

SECTION 1. Place of Meeting. Meetings of the directors shall be held at the principal office of the Corporation or at any place designated from time to time by resolution of the Board of Directors or by written consent of all directors.

SECTION 2. Annual Meeting. The annual meeting of directors of the Corporation shall be held on or before the 1st day of November of each year commencing in 2014, if not a legal holiday, and if a legal holiday, then on the succeeding business day.

Other regular meetings of the board of directors shall be held at such times and places as the Board shall from time to time designate.

Special meetings of the board of directors may be called at any time by the chairman of the board or the president. If the chairman of the board or the president is absent or unable to act,

such special meetings may be called by any director in place of the chairman of the board and by any officer in place of the president.

SECTION 3. Notices of Meetings - Waiver. Notices of the annual meetings of directors shall be given at least ten days prior to the date thereof and notices of regular meetings or special meetings called by the chairman of the board or president shall be given at least five days prior to the date thereof. Each notice shall specify the place, the day, and the hour of the meeting, and in the case of special meetings, the general nature of the business to be transacted. Notices shall be given as provided in **Article VIII, Section 1** of these Bylaws and may be waived either before or after the meeting as provided in **Article VIII, Section 2** of these Bylaws. Notices of adjourned meetings need not be given except when the adjournment is thirty days or more. Notices shall be given by the secretary or, if he/she is absent, or is unable or refuses to act, by any other officer or director of the Corporation.

SECTION 4. Quorum - Adjournment. The presence of a majority of the directors shall constitute a quorum for the transaction of business. Any business, not otherwise requiring the vote of a majority of the whole board of directors under the Corporation's Articles of Incorporation or Bylaws, may be conducted by vote of a majority of the quorum then present.

In the absence of a quorum, any meeting may be adjourned from time to time by the vote of a majority of the directors present thereat, but no other business may be transacted. At any such adjourned meeting at which a quorum is present, any business may be transacted which might have been transacted at the meeting as originally noticed.

The directors present at a duly called or held meeting at which a quorum is present may continue to do business until adjournment notwithstanding the withdrawal of enough directors to leave less than a quorum.

SECTION 5. Consent to Meetings. The transactions of any meeting however called and noticed and wherever held, shall be as valid as though it had been a meeting duly held after regular call and notice, if a quorum be present, and if, before the meeting, each of the directors, not present in person, gives a waiver of notice as provided in **Article V, Section 3** of these Bylaws. All such waivers shall be filed with the corporate records or made a part of the minutes of the meeting.

SECTION 6. Meeting Format. At every meeting of the board of directors, the chairman of the board or, in his/her absence, the president, shall act as chairman. The secretary of the Corporation may act as secretary of all meetings. In the absence at any such meeting of the secretary, the chairman of the meeting may appoint another person to act as secretary of the meeting.

SECTION 7. Action without Meeting. Any action required to be taken at a meeting of the board of directors of the Corporation, or any other action which may be taken at a meeting of the board of directors, may be taken without a meeting if a consent in writing, setting forth the action so taken, shall be signed by all the directors entitled to vote with respect to the subject matter thereof. Such consent may be signed in one or more counterparts and shall be included in

the minutes filed with the corporate records. Action taken pursuant to this **Section 7** shall become effective when the last director signs the consent, unless such consent specifies a different effective date following such signatures. A consent signed pursuant to this **Section 7** has the effect of a meeting vote and may be described as such in any document.

ARTICLE VI
COMMITTEES AND BOARD OF ADVISORS

SECTION 1. Committees. There may be such committees as the board of directors may from time to time determine. Membership of such committees shall be by appointment by the chairman of the board, with the consent of a majority of the board of directors.

SECTION 2. Board of Advisors. The board of directors may appoint one or more board(s) of advisors, the number and qualifications of which shall be in the discretion of the directors, for the purpose of giving advice to the board of directors that they must attempt to follow with regard to goals, policy and philosophy. In the event there is an even number of directors and a deadlock vote occurs in the board of directors, the vote of the Chairman of the Board will serve to break this deadlock.

ARTICLE VII
OFFICERS

SECTION 1. Positions. The officers of the Corporation shall be determined by the directors, which may include a chairman of the board, a president, a vice-president, a secretary, a treasurer, and/or such other officers as may be appointed by the president of the Corporation with the advice and consent of a majority of the board of directors.

SECTION 2. Qualifications. No person under the age of twenty-one years shall be eligible to be an officer of this Corporation. Only the chairman of the board and the president need also be members of the board of directors of this Corporation.

SECTION 3. Election and Term. Each officer, except those allowed to be appointed by the board of directors in accordance with **Article VII, Section 1,** shall be chosen every two years by a majority vote of the board of directors and shall hold office until his/her successor shall have been duly chosen and qualified, or until his/her death, or until he shall resign, or shall have been removed in the manner herein provided. Such election shall take place at either a regularly scheduled meeting or a special meeting called for the purpose of electing such officer or officers.

The same person may occupy, at the same time, any two or more offices as the board of directors may deem expedient, except that the same person shall not at the same time be both president and secretary. No officer shall execute, acknowledge, or verify any instrument in more than one capacity.

SECTION 4. Removal. Any officer may be removed either with or without cause, by the vote of a majority of the whole board of directors at any meeting of the board, or, except in

case of any officer elected by the board of directors, by any committee or superior officers upon whom such power of removal may be conferred by the board of directors.

SECTION 5. Resignation. Any officer may resign at any time by giving written notice to the board of directors or to the president or to the secretary of the Corporation. Any such resignation shall take effect at the time specified therein; and unless otherwise specified therein, the acceptance of such resignation shall not be necessary to make it effective.

SECTION 6. Vacancies. A vacancy in any office because of death, resignation, removal, disqualification, or any other cause, may be filled for the unexpired portion of the term by the board of directors at a regular meeting, or at a special meeting called for that purpose.

SECTION 7. Chairman of the Board. The chairman of the board shall, in addition to presiding at board meetings, exercise and perform such other powers and duties as may from time to time be assigned to him/her by the board of directors or prescribed by the Bylaws. The chairman of the board shall not vote on any matter before the board unless his/her vote is required to break a deadlock.

SECTION 8. President. The president shall be the principal executive officer of the Corporation and shall have general supervision over the business of the Corporation and over its several officers, subject, however, to the control of the board of directors. The president may sign and execute in the name of the Corporation, deeds, mortgages, bonds, contracts, or other instruments authorized by the board of directors, except in cases where the signing and execution thereof shall be expressly delegated by the board of directors or by these Bylaws to some other officer or agent of the Corporation; and in general shall perform all duties incident to the office of the president of the Corporation, and such other duties as from time to time may be assigned by the board of directors.

SECTION 9. Vice-President. At the request of the president or in the president's absence or disability, the vice-president, if any, shall perform all the duties of the president and when so acting, shall have all the powers of, and be subject to all the restrictions upon, the president. The vice-president may also sign and execute, in the name of the Corporation, deeds, mortgages, bonds, contracts or other instruments authorized by the board of directors except in cases where the signing and execution thereof shall be expressly delegated by the board of directors or by these Bylaws to some other officer or agent of the association, and shall perform such other duties as from time to time may be assigned by the board of directors or the president.

SECTION 10. Secretary. The secretary shall:

a. Certify and keep at the principal office of the Corporation the original or a copy of its Bylaws as amended or otherwise altered to date.

b. Keep at the principal office of the Corporation or such other place as the board of directors may order, a book of minutes of all board of directors' meetings with the time and place of holding, whether regular or special and if special how authorized, the notices thereof given, and the names of those present at such meetings.

c. Ensure that all notices are duly given in accordance with the provisions of these Bylaws or as required by law.

d. Ensure maintenance of the records and of the seal of the association and ensure that it is engraved, lithographed, printed, stamped, impressed upon or affixed to all documents the execution of which are on behalf of the association under its seal is duly authorized in accordance with the provisions of these Bylaws.

e. Ensure that the books, reports, statements and all other records and documents, required by law are properly kept and filed.

f. Provide at all reasonable times to any director upon request the Bylaws and minutes of proceedings of the directors of the Corporation.

g. In general, perform all duties incident to the office of the secretary and such other duties as from time to time may be assigned by the board of directors.

SECTION 11. Treasurer. The treasurer, if required so to do by the board of directors, shall give a bond for the faithful discharge of his/her duties in such sum, and with such sureties, as the board of directors shall require. The treasurer shall:

a. Have charge and custody of, and be responsible for all funds and securities of the Corporation and deposit all such funds in the name of the Corporation in such banks, trust companies or other depositories as shall be selected by the board of directors.

b. Keep and maintain adequate and correct accounts of the Corporation's properties and business transactions, including accounts of its assets, liabilities, receipts, disbursements, gains, losses, capital and surplus.

c. Render a statement of the condition of the finances of the Corporation at any meeting when required by the board of directors, and a full financial report at the annual meeting of the Board.

d. Exhibit at all reasonable times the books of account and records to any director or officer during business hours at the principal office of the Corporation or where such books and records are kept.

e. Receive and give receipt for, monies due and payable to the Corporation from any source whatsoever.

f. In general perform all the duties incidental to the office of treasurer and such other duties as from time to time may be assigned by the board of directors.

SECTION 12. Salaries. In limited circumstances, the board of directors may authorize and set a reasonable and fair market value amount of compensation for professional services that

would otherwise be contracted out. In such an event, compensation shall be based on services rendered and time committed to the Corporation, in accordance with these Bylaws, applicable law, and those standards and practices of any accrediting organization of the Corporation. Nothing contained herein shall be construed to require any salary in the absence of a determination by the board of directors. In the event the board of directors votes to compensate an officer or director for professional services they are to provide, that officer and/or director must abstain from said vote.

ARTICLE VIII
NOTICES

SECTION 1. Notices. Whenever any notice is required to be given to any director by statute or by these Bylaws, whether of a meeting or for some other purpose, it may be given personally or sent to such director by mail, e-mail, facsimile transmission, or any other form of written communication, charges prepaid, addressed to him/her at his/her address as it is shown on the records of the Corporation or if it is not so shown on such records or is not readily ascertainable, at the place of the principal office of the Corporation. In case such notice is mailed, it shall be deemed given at the time when the same shall be deposited in the United States mail. Such mailing as herein provided shall be due, legal and personal notice to such director. In case such notice is sent by e-mail or facsimile transmission, it shall be deemed given at the time transmitted by the sender. Transmission by any of the methods provided for herein shall be due, legal and personal notice to such director. If given personally, the person effecting such notice shall sign a statement to that effect stating where and when such notice was effected. All such notices shall be filed with the records of the Corporation.

SECTION 2. Waiver of Notice. Whenever any notice is required to be given to any director by statute or by these Bylaws, whether of a meeting or for some other purpose, a director may waive such notice; and a waiver or waivers in writing, signed by the person or persons entitled to said notice, or by, e-mail or facsimile transmission, when given before the meeting or at the time at which such notice is required to be given, shall be deemed equivalent to such notice. All such waivers shall be filed with the records of the Corporation.

ARTICLE IX
INDEMNIFICATION OF DIRECTORS AND OFFICERS

SECTION 1. Indemnification. Any person (and the heirs, executors, and administrator of such person) made or threatened to be made a party to any action, suit, or proceeding by reason of the fact that he/she is or was a director or elected officer hereof shall be indemnified by the Corporation against any and all liability and the reasonable expenses, including attorneys' fees and costs, incurred by him/her in connection with the defense or settlement of such action, suit or proceeding, or in connection with any appearance therein, to the fullest extent permitted by law. Such right of indemnification shall not be deemed exclusive of any other right to which such director or officer (or such director or officer's heirs, executor, or administrator) may be entitled apart from this Article.

SECTION 2. Insurance and Other Indemnification. The board shall have the power, without limiting the other powers of the board, to take the following actions:

a. Purchase and maintain, at the Corporation's expense, insurance on behalf of the Corporation and on behalf of others to the extent that power to do so has been or may granted by statute.

b. Give other indemnification to the extent permitted by law.

ARTICLE X
CONFLICT OF INTEREST

SECTION 1. Transactions Presumed Valid. No contract or other transaction between the Corporation and one or more of its directors, officers or any other corporation, firm, association or entity in which one or more of its directors or officers are members, directors or officers, or are financially interested, is either void or voidable because of such relationship or interest or because any director or officer is present at the meeting of the board or a committee of the board which authorizes, approves or ratifies such contract or transaction or because their votes are counted for such purpose, if either of the following apply:

a. The fact of such relationship or interest is disclosed or known to the board or committee which authorizes, approves or ratifies the contract or transaction by a vote or consent sufficient for the purpose without counting the votes or consents of those interested directors.

b. The transaction or contract is fair and reasonable to the Corporation at the time the transaction is consummated or, if made pursuant to contract, when the Corporation becomes contractually obligated such that its unilateral withdrawal from the contract would entail significant loss, liability or other added damage. The fairness of the transaction or contract shall be judged according to the circumstances at the time of consummation or at the time when the Corporation becomes contractually obligated.

SECTION 2. Clear and Convincing Evidence Needed to Rebut Presumption. Any person seeking to establish that a contract or transaction described in this section is void or voidable must prove by clear and convincing evidence that the requirements of **Article X, Section 1** have not been satisfied.

SECTION 3. Directors Counted for Quorum. Common or interested directors or officers may be counted in determining the presence of a quorum at a meeting of the board or a committee which authorizes, approves or ratifies a contract or transaction.

SECTION 4. Abstention of Interested Directors. The board may require directors to abstain from, or to not be present during, any discussion with regard to or vote upon a contract or transaction between the Corporation and such director or any relative or affiliate of such director. For purposes of this section "affiliate" shall mean any partnership, corporation or other entity in

which the director or a director's relative has a ten percent (10%) or more interest or by which such director is employed.

ARTICLE XI
DISSOLUTION

SECTION 1. Dissolution. In the event of dissolution of the Corporation, all funds and assets of the Corporation remaining after payment of creditors shall be distributed for one or more exempt purposes within the meaning of Section 501(c)(3) of the Code, or the corresponding section of any future federal tax code, such as The Arizona Geological Survey ("AZGS") an independent state agency established on July 1, 1988, which shall have first priority after payment of creditors, or otherwise shall be distributed to the federal government, or to a state or local government, for a public purpose. Any such assets not so disposed of shall be disposed of by a court of competent jurisdiction of the county in which the principal office of the corporation is then located, exclusively for such purposes or to such organization or organizations, as said court shall determine are organized and operated exclusively for such purposes. No part of said funds shall inure, or be distributed to any director or officer.

ARTICLE XII
TERM OF EXISTENCE

The period of duration of this Corporation shall be perpetual unless the directors elect to dissolve the Corporation in accordance with the Articles and Bylaws of the Corporation and Arizona law.

ARTICLE XIII
FISCAL YEAR

The fiscal year of the Corporation shall begin the first day of January and end the last day of December, unless otherwise modified by resolution of the board of directors.

ARTICLE XIV
SIGNATURES

Any action requiring a signature from a director or officer shall be deemed effective upon receipt by the secretary or the designated officer if other than the secretary, and may be delivered personally, by mail, by email, or by any other form of written transmission (including facsimile transmission).

ARTICLE XV
AMENDMENT OF BYLAWS

The authority to adopt, repeal and amend these Bylaws is hereby expressly delegated to the board of directors and any of these Bylaws may be altered, amended or repealed or any new Bylaws may be adopted solely by the affirmative vote of a majority of the whole board of directors at any meeting.

ARTICLE XVI
VALIDITY OF BYLAWS

The invalidation of any section of these Bylaws, for whatever reason, shall have no effect on the balance of these Bylaws.

ARTICLE XVII
NON-DISCRIMINATION

The Corporation shall not discriminate on the basis of race, color, religion, creed, national origin, age, gender, marital status, or sexual preference. The Corporation encourages the full input of all relevant minority and cultural groups to its deliberations and in its policy making. The actions of the organization shall in all ways reflect sensitivity to the diverse groups within our pluralistic, democratic public.

The undersigned hereby certifies that the foregoing Bylaws have been duly adopted by the Corporation by unanimous written consent.

4/25/2014
Date of Execution

Dr. Stephen M. Richard

5-6-14
Date of Execution

Dr. M. Lee Allison

US Geoscience Information Network (USGIN) Website Description - *Appendix H3*

URL: http://usgin.org/

Description/Purpose: The USGIN website was launched in February 2014 at the recommendation of the National Geothermal Data System Sustainability Plan with the intent of enabling agencies and businesses to build highly functional, distributed data sharing systems by using the USGIN cyberinfrastructure framework. As part of phase 3 of the State Contributions to the National Geothermal Data System project, the site is designed to provide an information source to both governmental and non-governmental entities who wish to create or build on a distributed data sharing system for geoscience information. We anticipate that expanding the use of USGIN and encouraging new nodes to be deployed under this framework will help NGDS achieve sustainability by increasing developer familiarity with the standards and protocols upon which USGIN is built, increasing the profile of the USGIN platform to software and application developers, expanding the user base and the quantity of data available. USGIN's expansion allows NGDS to grow as new data is added into the system.

Examples of usgin.org pages: Home Page

Advantages to interoperability: Beyond Open Data

39

Home » Beyond Open Data

Beyond Open Data

In addition to facilitating interoperability, USGIN specifications ensure that USGIN data remains open. Any data shared according to these specifications can be hosted and accessed by free-and-open-source software; likewise, USGIN specifications conform to international standards set by the Open Geospatial Consortium.

What is Open Data?

Data is considered open when "anyone is free to use, reuse, and redistribute it - subject only to the requirement to attribute and/or share-alike," according to opendefinition.org.

Interoperability Enriches Data

Naturally, you wish to share your data with the largest possible audience. How do you make sure your data can be easily used and integrated with other data in new contexts? The answer is interoperability.

USGIN interoperable data is based on simple, well documented interchange formats and standardized data exchange protocols so users can view and analyze interoperable data through the program of their choice. USGIN accepts three tiers of data quality. Tier 1 type data is unstructured data can be viewed and often must be downloaded. Tier 2 can be analyzed in some programs but is not standardized, and conversion may take time and effort. Tier 3 data, or interoperable data, is standardized to open in multiple formats and to effortlessly merge with additional data sources in a content model for ease of analysis.

Tier 1
Good ★
unstructured data
- PDF
- Powerpoint
- image
- *etc.*

Tier 2
Better ★★
structured data
- Excel
- CSV
- XML
- *etc.*

Tier 3
Best ★★★
structured + standardized data
- NGDS content model
- Standardized Excel, CSV, XML, RDF, JSON, *etc.*

USGIN in Action Page: NGDS

Home » USGIN in Action: NGDS

USGIN in Action: NGDS

"One of the most successful [cyberinfrastructure] programs to date is the National Geothermal Data System (NGDS)..."
—Sara E. Pratt, Digitizing Earth: Developing a cyberinfrastructure for the geosciences, "Earth Magazine" August/September 2013

Data Sets from Many Agencies

USGIN's suite of standards, protocols, and practices is already in use by all fifty state geological surveys and more than 55 organizations through the National Geothermal Data System (NGDS). The distributed, interoperable NGDS network has received extensive data catalogs from Southern Methodist University's Geothermal Data Repository, Oregon Institute of Technology, Energy and Geoscience Institute, and the Nevada Bureau of Mines and Geology. In addition, OneGeology, an international digital mapping initiative, has agreed to conform to USGIN architecture as the potential platform for collaboration on a future virtual geological survey.

NGDS provides access to over 30 types of data representing more than 9 million data points and over 17,000 reports and documents, and is growing every day.

National Geothermal Data Platform

Originally funded by the U.S. Department of Energy's Geothermal Technology Office, NGDS is intended to provide access to information about geothermal resources to enable geothermal analysis and widespread public use in an effort to minimize the business-related risks associated with geothermal development. The system serves as a platform for sharing reliable geothermal information with users of all types, supplying tools relevant for their work.

Successful Design and Implementation

Data added to NGDS can be accessed over the Internet by free-and-open-source software around the world. The International Renewable Energy Agency (IRENA) is starting to harvest six types of data from NGDS to populate their own system.

NGDS has already attracted attention from the United States Energy Association and the Geothermal Energy Association, and is garnering international interest. Recently, NGDS managers have been contacted by the U.S.-East Africa Geothermal Partnership to assess data management needs in Kenya, and NGDS staff have been invited to speak to the China Geological Survey.

Explore the NGDS homepage at geothermaldata.org.

40