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. Reference herein to any social initiative (including but not limited to Diversity, Equity, and Inclusion (DEI); Community Benefits Plans (CBP); Justice 40; etc.) is made by the Author independent of any current requirement by the United States Government and does not constitute or imply endorsement, recommendation, or support by the United States Government or any agency thereof.

Pacific Northwest
NATIONAL LABORATORY

PNNL-36720 Rev 0
DVZ-RPT-106 Rev 0

A Prototype Software to Demonstrate a Data Catalog for Hanford Environmental Datasets

September 2024

Rebecka L.B. Iveson
Aaron Moreno
Juliane Schneider
Christian D. Johnson

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <https://www.ntis.gov/about>
Online ordering: http://www.ntis.gov

PNNL-36720 Rev 0
DVZ-RPT-106 Rev 0

A Prototype Software to Demonstrate a Data Catalog for Hanford Environmental Datasets

September 2024

Rebecka L.B. Iveson
Aaron Moreno
Juliane Schneider
Christian D. Johnson

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Summary

Ensuring that data on long-term environmental remediation at the Hanford Site is high-quality, traceable, and easily accessible is an ongoing challenge, complicated by decades of data collection, multiple contractors maintaining data sources, and the wide range of data types. A centralized data catalog, known as the Hanford Environmental Information and Data Index (HEIDI), has been under development as part of the Hanford Environmental Data Management (HEDM) program to address these challenges. HEIDI fulfills a critical need to bring together a wide range of data types and sizes from multiple authoritative data sources, while documenting the data pedigree and quality information (i.e., traceable to the data source/originator).

This document describes additional development and maturation of the HEIDI prototype. Key accomplishments included deploying the catalog software, Esri Geoportal Server, on a server accessible to Hanford Local Area Network users, conducting cybersecurity evaluations, investigating integrated authentication solutions, and conducting functional testing of the catalog prototype. The server-based deployment enabled targeted feedback, leading to enhancements including improved accessibility features and an expanded metadata schema. Specifications for the server-based deployment of the prototype catalog and the HEIDI metadata schema are provided in this document to support subsequent HEIDI deployment by the U.S. Department of Energy Richland Operations Office.

Acknowledgments

Funding for this work was provided by the U.S. Department of Energy Richland Operations Office under the Deep Vadose Zone – Applied Field Research Initiative. The Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract DE-AC05-76RL01830.

Acronyms and Abbreviations

API	Application Programming Interface
ARIA	Accessible Rich Internet Applications
AR	Administrative Record
CPCCo	Central Plateau Cleanup Company
CSET	Cyber Security Exemption Tool
DOE	U.S. Department of Energy
EDA	Environmental Dashboard Application
FY	fiscal year
HEIDI	Hanford Environmental Information and Data Index
HLAN	Hanford Local Area Network
HMIS	Hanford Mission Integration Solutions
HTML	Hypertext Markup Language
ISO	International Organization for Standardization
NQAP	Nuclear Quality Assurance Program
PIR	Public Information Repository
PNNL	Pacific Northwest National Laboratory
RL	Richland Operations Office
SECB	Systems Engineering Control Board
URI	Uniform Resource Locator
URL	Uniform Resource Locator
XML	Extensible Markup Language

Contents

Figures

Tables

1.0 Introduction

Access to information that is of known quality and traceable to the underlying data and analysis is vital for supporting defensible decision-making at the Hanford Site, where environmental remediation is complex (DOE 1994). As documented in the Tri-Party Agreement (aka the Hanford Federal Facility Agreement and Consent Order; Ecology et al. 2024) and subsequent information management plans (DOE 1990, 1994, 2012), it has long been recognized that data/record management is important for near- and long-term operations at the Hanford Site. The Hanford Long-term Stewardship (LTS) program required the establishment of an Information Management (IM) program to preserve information generated during the Hanford cleanup mission for future use (DOE 2010). Yet ongoing challenges exist for identifying data quality, documenting data sources/lineage, maintaining information security, and addressing technology incompatibilities, with the ultimate goal of ensuring that data can be discovered, retrieved, interpreted, and reused over the expected century-long timeframe to achieve site closure.

The LTS IM mission specifies data cataloging as one component to overcome these challenges and to enable the success of Hanford's LTS program in eventually transiting information from existing Hanford repositories to the U.S. Department of Energy (DOE) Office of Legacy Management (DOE 2012). A centralized platform to catalog Hanford environmental data fulfills a critical need to bring together a wide range of data types and sizes from multiple authoritative data sources, while documenting the data pedigree and quality information (i.e., traceable to the data source/originator).

Since 2021, a catalog of environmental datasets known as the Hanford Environmental Information and Data Index (HEIDI) has been under development for the Hanford Site. The catalog is part of a renewed effort by the DOE Richland Operations Office (RL) to develop and implement a formal program for managing environmental data at the Hanford Site, known as the Hanford Environmental Data Management program. Prior work on HEIDI achieved the following milestones:

- Identified a suitable software tool to provide the necessary functionality of documenting Hanford data sources while integrating the metadata within established Hanford Local Area Network (HLAN) enterprise systems (Ham and Crockett 2021).
- Formed a collaborative HEIDI working group composed of staff from Hanford Mission Integration Solutions (HMIS), Central Plateau Cleanup Company (CPCCo), DOE RL, and Pacific Northwest National Laboratory (PNNL) for a site-managed approach to cataloging Hanford environmental data (Ham and Crockett 2021).
- Developed Hanford-specific requirements for metadata and data archiving and tested how those requirements could be implemented in the schema (the structure of how information is represented and stored) of the selected metadata standard (Ham 2022).
- Documented metadata specifications and data catalog requirements and illustrated selected components in the prototype Esri Geoportal Server data catalog software via a local installation (Bence et al. 2023).

This document describes the most recent phase of work on maturing the catalog prototype to a server-based application securely accessible to HLAN users. The catalog contains metadata describing publicly available Hanford datasets and authoritative Hanford environmental databases. Features necessary to demonstrate the required data catalog functionality in a Hanford-specific context were investigated and implemented, where possible. Focused feedback from HMIS and CPCCo staff through direct testing of the application was enabled by cybersecurity approval of the prototype for use by HLAN users.

2.0 Server Deployment Overview

The Esri Geoportal Server application was deployed in fiscal year (FY) 2024 on an externally accessible server to facilitate further testing and refinement of the HEIDI prototype. Up-to-date versions of Esri Geoportal Server and associated software applications were used during the deployment to avoid security vulnerabilities identified in previous versions. Appendix A describes these software version updates. Appendix B documents the specifications for the server-based deployment7.0Appendix B.

2.1 Security Requirements

As described in Bence et al. (2023), successful HEIDI implementation requires that the chosen software fulfill all cyber-security requirements unique to the HLAN system (e.g., Systems Engineering Control Board [SECB] review and approval). While a full implementation of the chosen prototype software (Esri Geoportal Server) has not been tested on the HLAN, the server-based deployment of the prototype enabled a preliminary evaluation of the HLAN cyber-security credentials. The externally accessible prototype catalog was successfully approved for an SECB pilot-testing period. The pilot testing phase enabled users to access HEIDI while connected to the HLAN network. This preliminary testing further supports the likelihood that the core components of Catalog and Harvester modules should be compatible with full HLAN security requirements.

2.2 Testing

Deploying a prototype data catalog on an external server accessible to HMIS and CPCCo staff enabled targeted feedback on the software functionality, user interface, and metadata schema. Functional testing was conducted by asking testers to complete a series of steps designed to evaluate the basic functions of the software. This included testers with different specified roles logging into the catalog with a username and password, filtering data, and testing user permissions (e.g., ability to create, edit, and delete catalog entries). Further feedback on the usability of the software and data organization (i.e. metadata schema) was captured in comments made by individual testers. These comments were combined into a table by grouping similar requests and clarifying comments through collaborative working meetings.

All suggested changes to the software were investigated according to the following levels: (a) issue resolved; (b) issue not resolved but resolution possible with more time/effort; (c) issue not resolved and resolution not possible with current software configuration. Testing results are provided in Appendix C. 4.0 describes implementation of the requested changes in the server-based deployment of the prototype catalog.

3.0 Hanford Data Catalog Requirements for Server Deployment

Previous studies tested and identified HEIDI requirements using a local installation environment for the prototype data catalog (Ham 2022; Bence et al. 2023). While requirements have not been tested directly within the HLAN environment, the chosen software (Esri Geoportal Server) was launched on a an externally accessible server for access by HLAN users. The previously identified requirements were re-verified using the new deployment configuration. The testing process verifies that chosen catalog software appropriately handles the unique components of the metadata and the metadata itself in the context of the HLAN environment.

Table 3.1 through Table 3.4 present the previously identified data catalog requirements and their evaluation status (pass/fail) for the testing in FY23 (local deployment) and FY24 (server deployment).. Ham (2022) provides detailed information on the origin of these requirements.

Table 3.1. Geoportal Server testing results: Hanford-specific theming implementation.

Implementation Area	Required Functionality	FY23 Pass/Fail	FY24 Update
Metadata Schema	Support custom lists for keywords	Pass	Pass
Metadata Editor	Enforce custom lists for keywords	Pass	Pass
Data Catalog	Make custom keywords accessible in search criteria	Pass	Pass
HLAN	Consistent use of theming nomenclature across environmental data information systems	Fail	Fail

Table 3.2. Geoportal Server testing results: Catalog entry access restriction implementation.

Implementation Area	Required Functionality	FY23 Pass/Fail	FY24 Update
Metadata Schema	Support differentiating classes of records by access limitation	Pass	Pass
Metadata Editor	Require access limitation entry	Pass	Pass
Data Catalog	Enforce limitation on access to catalog entry based on HLAN identity or role	Roles are implemented, but further testing is required.	Partial pass. Catalog restricts user access to entry based on role. HLAN identities remain untested.
HLAN	Maintain lists of identities with similar roles and access	Pass	Pass

Table 3.3. Geoportal Server testing results: Resource access restriction implementation.

Implementation Area	Required Functionality	FY23 Pass/Fail	FY24 Update
Metadata Schema	Support identifying access limitations for the resource	Pass	Pass
Metadata Editor	Support or require access limitation entry	Pass. Support for access limitation entry confirmed. Ability to enforce an access limitation statement metadata element confirmed.	Pass
Data Catalog	Enforce limitation on access to resource based on identity or role	Pass. Roles implemented. Pass with workaround.	Pass
HLAN	Maintain lists of identities with similar roles and access. File systems may also be involved in limiting access to resources.	Pass	Pass

Table 3.4. Geoportal Server testing results: Use limitation implementation.

Implementation Area	Required Functionality	FY23 Pass/Fail	FY24 Update
Metadata Schema	Support specifying constraints on use	Pass	Pass
Metadata Editor	Support or require use limitation entry	Pass. Support for use limitation entry confirmed. Ability to enforce a use limitation statement metadata element confirmed.	Pass
Data Catalog	Prominently display constraints on use	Pass	Pass
HLAN	*No requirement in this area*	*N/A*	*N/A*

Testing the requirements against the new server-based prototype catalog demonstrated that nearly all of the 15 requirements identified in Table 3.1 through Table 3.43 were met. This indicates the server-based deployment is as effective as the local installation. The data catalog requirement to enforce limitations on access to a catalog entry based on HLAN identity or role was not tested, but an overview of investigations into authentication for HLAN users is provided in Section 4.6.

4.0 Prototype Data Catalog: FY24 Updates

As described in Section 2.2, the server-based deployment of the prototype catalog enabled direct user feedback from HMIS, CPCCo, and RL staff. The process also included adhering to PNNL-specific cyber-security (described in Appendix B) and accessibility requirements. Changes made to the prototype catalog in response to the user feedback were demonstrated during collaborative working meetings and are described below.

4.1 Accessibility Enhancements

To make the application more accessible and user-friendly, several important changes were implemented within the user interface of the prototype catalog. These enhancements were necessary for compliance with web accessibility standards (e.g., 29 U.S.C. §794d) and to ensure an effective user experience, particularly for users relying on assistive technologies. Below is a detailed description of accessibility issues identified in the standard application deployment and the custom modifications made to address them.

4.1.1 General Changes to All Pages

The logo in the upper left corner of all pages was missing alternative text, making it inaccessible to screen readers. Descriptive alternative text was added to ensure users with visual impairments can understand the presence and purpose of the logo.

4.1.2 “Home Page” Changes

To clearly identify the “Home Page” of the application, the name of the catalog was added to the initial landing page. The added title, “Hanford Environmental Information and Data Index (HEIDI),” was formatted as an enlarged heading to aid in visual identification (Figure 4.1).

Figure 4.1. Catalog name in an enlarged format to clearly identify the application “Home Page.”

In addition to a clear application heading, the dropdown menu in the upper right of the application that allows a user to sort metadata was updated with an Accessible Rich Internet Applications (ARIA) label and a display label. ARIA labels are attributes that can be added to Hypertext Markup Language (HTML) elements to provide additional information about the element's purpose and functionality to assistive technologies, such as screen readers. Both the ARIA label and display label were updated to “Sort List” to clearly communicate the purpose of the HTML element (Figure 4.2). A user can now sort the data catalog entries according to clearly labeled options (e.g., “By Relevance,” “By Title,” “By Date”), providing a more intuitive experience.

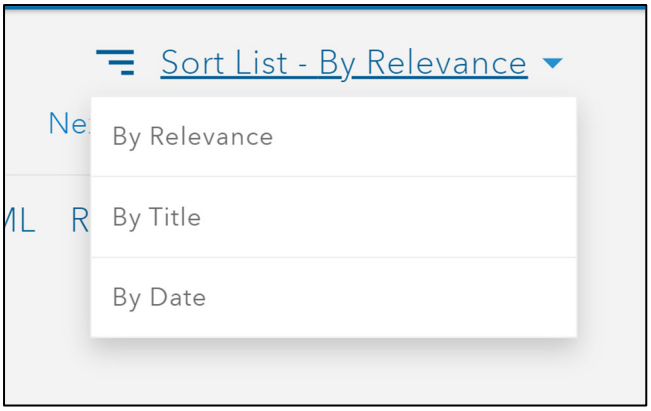

Figure 4.2. Improved labeling of catalog sort options.

Insufficient text color contrast in the metadata information for each record was also addressed across the "Home Page" by adopting a darker gray color (#757575) compared to the standard light gray application color (Figure 4.3). This change ensures a minimum 4.5:1 contrast ratio, meeting the AA rating for accessibility per the Web Content Accessibility Guidelines (W3 2023). Improved contrast makes text more readable for users with low vision and improves the overall reading experience for all users.

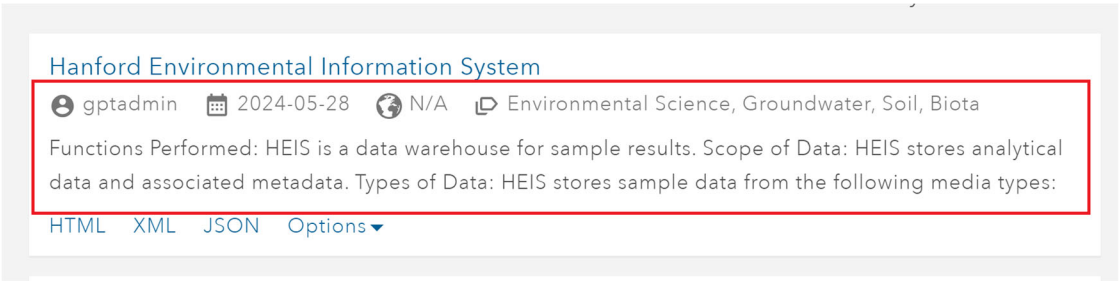

Figure 4.3. Enhanced color contrast applied within the application

4.1.3 "About Page" Changes

An "About Page" was added to the catalog to provide general information about HEIDI, its creators, and its purpose. Titles and subheads were organized with proper HTML hierarchy, with the page title formatted as an H1 and subheads as H2s (e.g., "Who We Are," "What We Do"). This structure improves readability and navigability, especially for screen reader users. The "About Page" feature was not part of the standard application configuration. The demonstration of this customization provides a pathway for HLAN-specific instructions, links, and user guides to be integrated into the application during full-scale deployment.

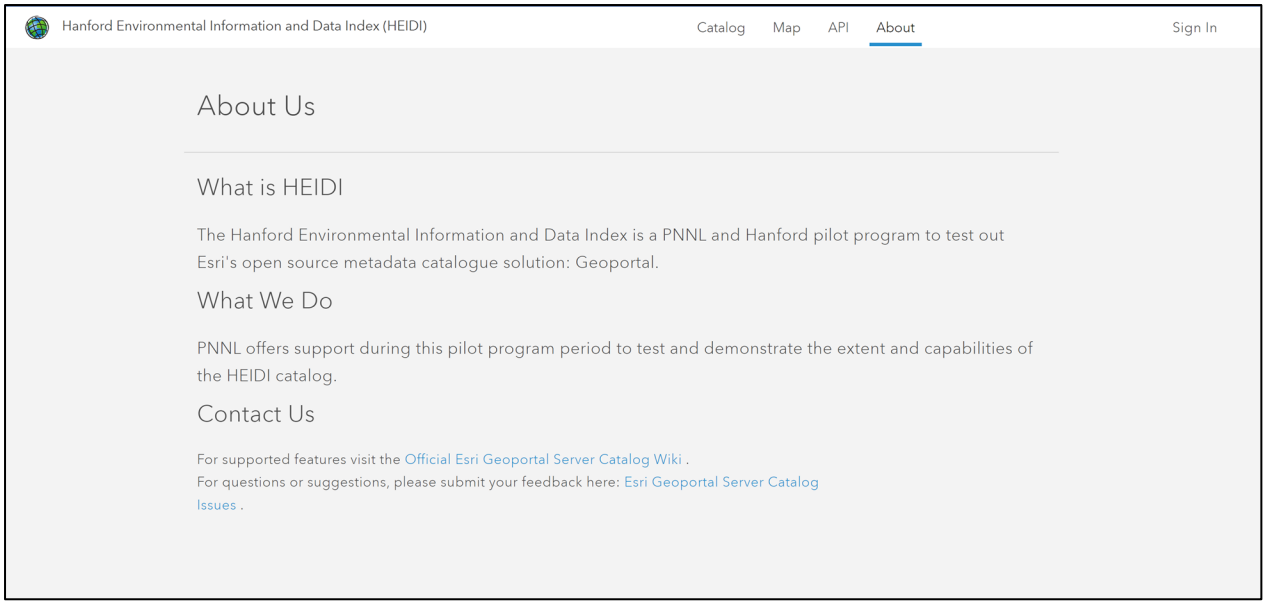

Figure 4.4. HEIDI "About Page" demonstration.

Additionally on the "About Page," hyperlinks under the "Contact Us" section were updated to a different color than the surrounding text, making them easily identifiable. This is important for users to distinguish links from regular text. The links in were also given identifying titles and Aria labels.

4.2 Filter and Search Panel Enhancements

A key functional requirement of a data catalog is the ability for a user to find data. While the standard deployment of Esri Geoportal Server is configured with a variety of filter and search functions, there is no labeling or instructions to help a user navigate the features effectively. This section describes customizations implemented to make the prototype catalog more user-friendly.

4.2.1 Adding Labels and Descriptions to Filters

Within the application, a user can filter catalog entries by a variety of criteria, including geospatial extent, date, and/or keywords. The filters are presented as dropdown menus on the left-hand side of the main catalog interface. However, individual testers identified difficulties in navigating these standard filter data features. To address this, a title was added to the left-hand filter menu to clearly identify the purpose of the dropdown menus, and instructional text was added at the bottom of the panel (Figure 4.5). This was accomplished by adding a section to the SearchPanel.html file with the associated styling for the guide in the main.css file.

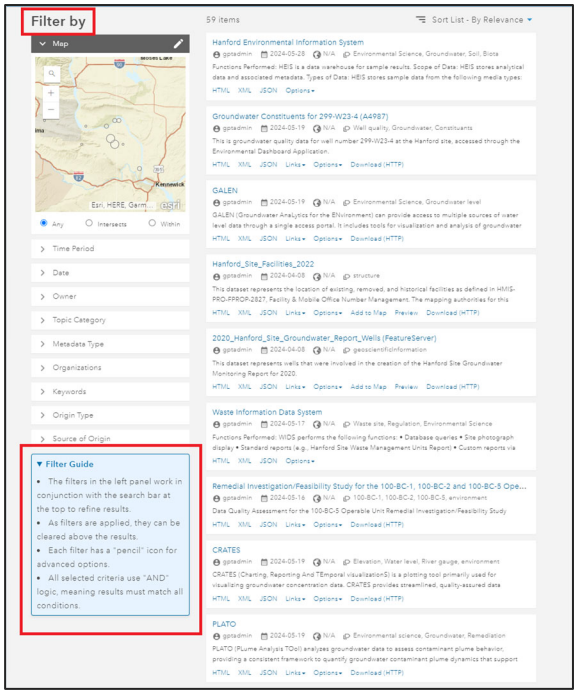

Figure 4.5. Filter panel updates.

4.2.2 Custom Filter Options

Testing of the catalog prototype identified the need to add custom data source filters, in addition to or in place of the standard filter options provided. Within the prototype catalog, users expressed the desire to filter catalog search results by the database origin of the metadata. For example, a user could limit all search results to only data from the Environmental Dashboard Application (EDA) by using this new filter option. A new data filter, "Source of Origin," was successfully added by identifying the appropriate metadata field to filter on and making the following changes to the application code:

- The GitHub documentation was followed to make modifications to the EvaluatorFor_ISO.js file (see Figure 4.6).
  - This script is critical because it handles the evaluation and extraction of metadata fields based on standards from the International Organization for Standardization (ISO).
  - Adding to this evaluator enables the custom filter to reliably parse and interpret the necessary metadata elements.

```
34      /* general */
35      G.evalProp(task,item,root,"fileid","gmd:fileIdentifier/*/text()");
36      G.evalProp(task,item,iden,"title","gmd:citation/gmd:CI_Citation/gmd:title/*/text()");
37      G.evalProp(task,item,iden,"description","gmd:abstract/gco:CharacterString");
38      G.evalProps(task,item,root,"keywords_s","//gmd:MD_TopicCategoryCode | //gmd:descriptiveKeywords/gmd:MD_Keywords/gmd:keyword/*/text() ");
39      G.evalProp(task,item,iden,"thumbnail_s","gmd:graphicOverview/gmd:MD_BrowseGraphic/gmd:fileName/gco:CharacterString");
40      G.evalProps(task,item,root,"contact_organizations_s","//gmd:CI_ResponsibleParty/gmd:organisationName/*/text()");
41      G.evalProps(task,item,root,"contact_people_s","//gmd:CI_ResponsibleParty/gmd:individualName/*/text()");
42 +    G.evalProps(task,item,root,"data_source_s","//gmd:CI_OnlineResource/gmd:name/*/text()");
```

Figure 4.6. The EvaluatorFor_ISO.js file was modified to create a custom data filter.

- The new field was integrated into the Elasticsearch indexing process to ensure that, when metadata is indexed, the "Source of Origin" information is appropriately captured and made searchable.

- New metadata tags were mapped to the corresponding fields in Elasticsearch, as shown in Figure 4.7.

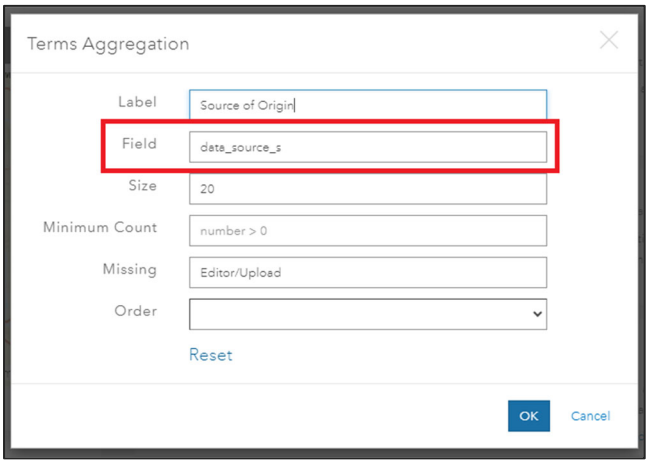

Figure 4.7. Metadata tags mapped to the filter field.

4.3 Implementing Additional Feature Descriptors

Additional feedback provided during testing of the catalog indicated the need for "user-friendly" navigation tips to make the application more intuitive for users. The ability to effectively use the catalog to find and retrieve data has been identified as a key requirement for the final implementation of HEIDI (Ham 2022). Based on this feedback, "hover over" descriptions of useful HTML elements were added by modifying each desired nav-link and nav-item tag within the App.html file and adding a title attribute. Figure 4.8 presents an example "hover over" description.

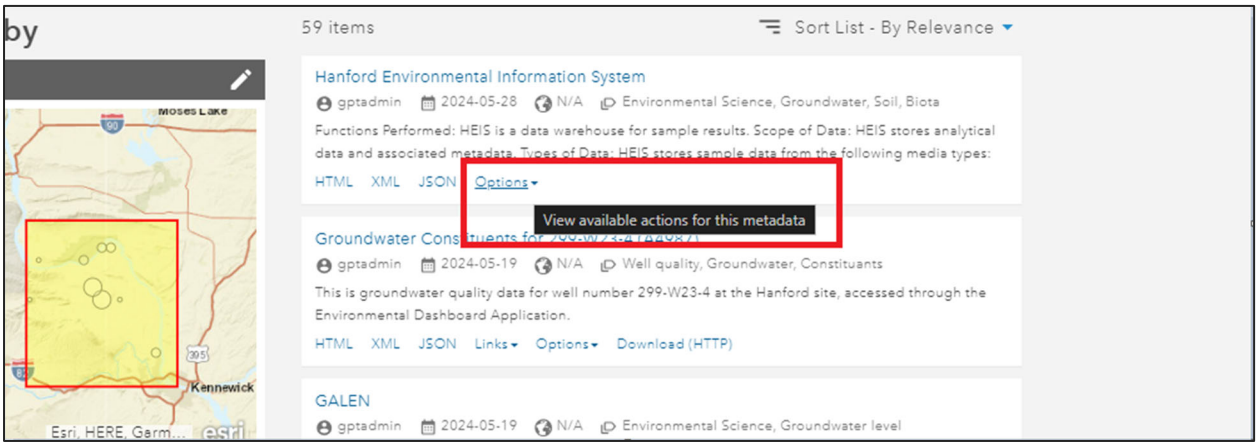

Figure 4.8. "Hover over" description example for the "Options" link on a metadata file.

4.4 Metadata Schema

Metadata standards offer a set of predefined guidelines that dictate the structure and format of metadata, ensuring consistency in describing and managing data. Ham (2022) determined that the ISO 19115:2003 North American Profile was a suitable metadata standard for HEIDI (NAP-Metadata Working Group 2007). Bence et al. (2023) described the process of "crosswalking" metadata typically recorded for Hanford data with ISO 19115. A crosswalk is "a specification for mapping one metadata standard or application profile to another" (NISO n.d.).

The mapping of the Hanford profile to ISO 19115 was expanded in this study, and a preliminary metadata schema for HEIDI was identified. The metadata schema is the part of the standard that outlines the overall structure of the metadata and addresses how to handle common components like dates, names, and places. Components of each metadata field are defined in the schema, including the location in the schema, label, data type, requirement, cardinality, and definition. A description of each of these components is provided below.

- Location in the schema: Where the information is stored/classified in the schema.
- Label: Common name for the information contained in the metadata field.
- Data type: Identifies the type of information that should be stored within the metadata field such as open-text, dropdown list, or a date.
- Requirement: Indicates whether the metadata field is required to be catalogued within HEIDI (i.e., minimum metadata requirements).
- Cardinality: Describes how many data types can be selected to fulfill the respective field.
  - "1-1" cardinality describes a "one-to-one" relationship. One response is required for exactly one data type.
  - "0-n" cardinality describes a "none-to-many" relationship. This is an optional field that may repeat any number (n) of times.
  - "1-n" cardinality describes a "one-to many" relationship. This is a required field that may repeat n times.
- Definition: Provides a standardized description of the metadata field.

A metadata schema can be adopted or adapted to fulfill the catalog requirements. To ensure maximum interoperability for the data stored within HEIDI, the identified schema adopted fields from ISO19115:2003. Table 4.1 provides an example of the metadata schema for proposed required fields, as informed by user feedback. Examples in Table 4.1 do not represent one resource, although most of the examples are based on a data report (RJ Lee Group 2022) within the Hanford Administrative Record (AR. This approach was taken to demonstrate all potential metadata fields in the schema. Appendix D presents the entire schema defined for HEIDI.

Table 4.1. Proposed required metadata schema for HEIDI. Examples are based on RJ Lee Group (2023) where applicable.

Location in Schema	Label	Data type	Required	Cardinality	Definition	Example
Metadata/Identifier	File Identifier	System supplied	y	1-1	Unique identifier for the metadata file	1686241220912r3853225710481458
Metadata/Identifier	Language	Text	y	1-1	Language of the metadata schema	en
Metadata/Contact	Organization Name	Text	y	1-n	The organization responsible for managing the metadata schema	RJ Lee Group Inc.
Metadata/Date	Metadata Date (now)	Date	y	1-1	The date the metadata was used to create a record	2023-06-08
Metadata/Standard	Metadata Standard Name	System supplied	y	1-1	The name of the metadata schema being used	ISO 19139/19115 Metadata for Datasets
Metadata/Standard	Metadata Standard Version	System supplied	y	1-1	The version of the metadata standard being used	2003
Identification/Citation	Resource Title	Text	y	1-1	The name of the resource	F21-047, 200W Pump & Treat - Extraction Well Water Sampling
Identification/Citation	Resource Date (creation date)	Date	y	1-n	A date related to the resource	2022-09-13
Identification/Citation	Date Type	Dropdown	y	1-1	Type of date	Publication
Identification/Description	Abstract	Text	y[(a)]	1-1	Description of the resource	F21-047, 200W Pump & Treat - Extraction Well Water Sampling
Identification/Description	Purpose	Text	y[(a)]	0-1	Purpose of the resource	This data was collected for the purpose of analyzing the composition of groundwater in an unspecified well in the 200-ZP-1 Operable Unit.
Identification/Contact	Point of contact	Text	y[(a)]	0-n	Contact for the resource	kgilman@rjleegroup.com
Identification/Keywords/ Keyword Collection	Keywords	Text	y	1-n	Keywords describing the resource, separated by commas	200-ZP-1
Identification/Keywords/ Keyword Collection	Keyword Type	Dropdown	y	1-1	Type of keyword	Place

Location in Schema	Label	Data type	Required	Cardinality	Definition	Example
Identification/Constraints/ Usage Constraints	Use Limitation	Text	y[(a)]	0-n	Limits of usage of the data	No special access is required to use the external Hanford AR, which is available to the public.
Resource/Language	Resource Language	Country code/text	y	1-n	Language of resource	en
Resource/Classification	Topic Category	Bullet list	y	1-n	Category that describes the resource	Environment
Distribution/Distribution/ Distribution Format	Format Name	Text	y[(a)]	0-n	The format the resource is distributed or downloaded in	Excel
Quality/Lineage	Lineage Statement	Text	y[(a)]	0-1	The Lineage field offers a free-text space to record a statement about the data's chain of custody. This will include the individual or organization that created or recorded the data, who has maintained the data, and how the data arrived at its current location. Any additions or modifications to the data should also be noted.	This data was created by the CCPO and recorded by Heather Medley. The data has been maintained in the EDA database and is currently collected monthly.

(a) Specific metadata requested by the Hanford working group, beyond typical minimum metadata requirements.

4.5 Metadata Additions to the Catalog

Efforts in FY24 included the expansion of the cataloged resources within the prototype to enable more effective testing. Metadata files (i.e., Extensible Markup Language [XML] files) were created manually and added to the data catalog for the main Hanford Environmental Databases (HMIS 2023) and sub-pages within those systems. A total of 59 metadata files are currently cataloged within the prototype (see Figure 4.9).

An additional 105 XML files were generated using the text-scraping tool described in Bence et al. (2023). The workflow and the associated text-scraping tool were reformatted to align to the metadata schema details described in Section 4.4. The tool was executed to ensure the generated files represented reports, studies, and/or data from the AR and Public Information Repository (PIR) from a variety of the 11 groundwater operable units. All generated XML files can be saved to a known file system directory, and then a task can be configured to harvest any new XML files added to that directory for automated ingestion into the prototype.

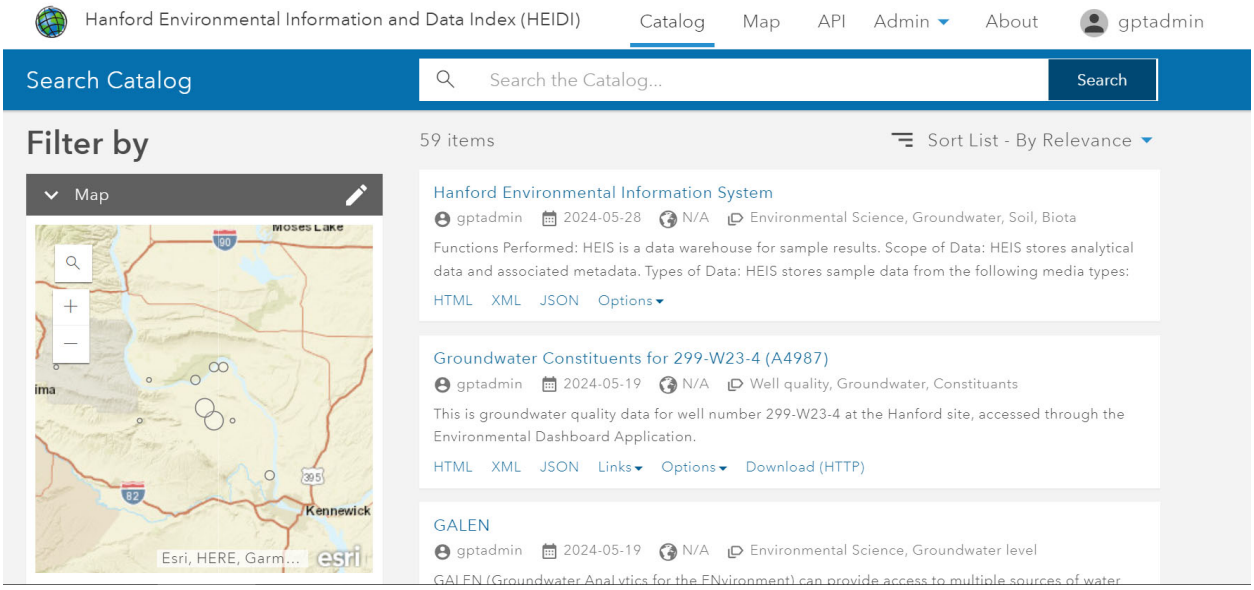

Figure 4.9. "Home Page" of HEIDI prototype displaying 59 metadata files

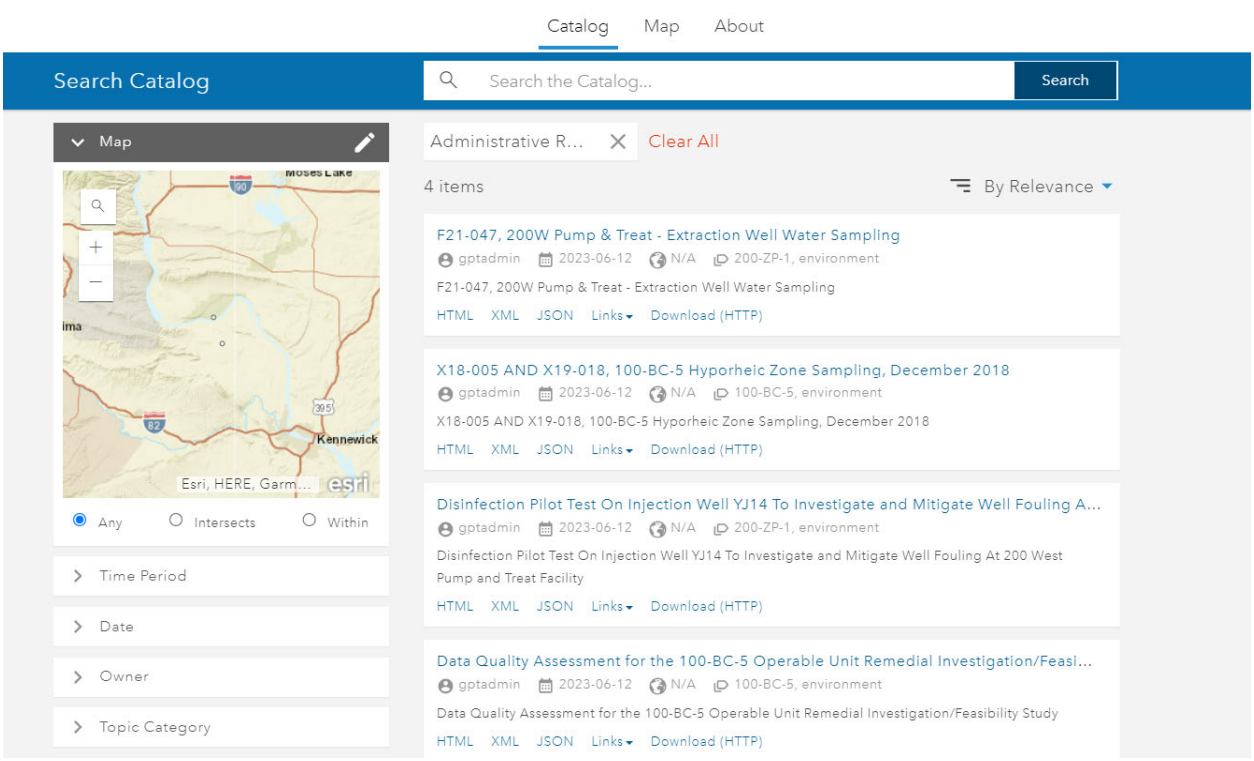

Figure 4.10. Display of four data resources ingested from the AR/PIR and uploaded to the catalog.

4.6 OneID Evaluation and Planning

Integration of authentication methods already used at Hanford was identified as a key component to a full-scale HEIDI deployment (Ham 2022). Bence et al. (2023) identified OneID as a potential solution for authentication of both Hanford and PNNL employees. Implementation of OneID was challenging in a prototype setting, so PNNL senior staff members who have implemented OneID in similar projects were consulted for practical recommendations. Based on prior experience the “out-of-the-box” authentication options provided within the geoportal’s GitHub documentation are not recommended. The preferred approach involves creating custom middleware within the Java Spring Application Programming Interface (API) to integrate OneID authentication using Spring Interceptors given that the application utilizes a Java Spring backend. The preliminary design to accomplish OneID integration would involve the tasks described below.

4.6.1 Saving the Authentication Token and Sending It with Each Request

Successful OneID implementation involves the secure capture of an HLAN authentication token and “capture” of the token with each request made by the application. The different methods of storing the token in a way that balances security with performance need to be considered. Options include storing tokens in HTTP cookies, or secure HTTP headers. Ultimately, the solution needs to ensure that tokens are stored securely to prevent unauthorized access while also being easily retrievable to maintain a seamless user experience.

4.6.2 Catch Token in Interceptor to Validate User Through OneID

Next, a mechanism to catch the authentication token using Spring Interceptors would need to be implemented. This middleware component would serve as a gatekeeper, intercepting each incoming request to the backend. This step ensures that only authorized users can access HEIDI's resources. The interceptor would need to be robust enough to handle edge cases, such as expired or malformed tokens, to prevent unauthorized access and potential security breaches. Further discussions would be needed to define the logic required for intercepting requests, extracting tokens, and verifying their validity through the OneID service.

4.6.3 Store Users by Distinguished Unique Identifier in a Configuration File

To streamline user management and enhance security, it is recommended that authenticated users be stored by their Distinguished Unique Identifier (DUID) in a ".env" or configuration file. This approach enables tracking and management of user access rights within the application. The use of a ".env" file provides a secure way to store sensitive information, ensuring that user data is not exposed or hardcoded into the application (as would be required with the "out-of-the-box" authentication options). This approach would also support scalability, since new users can be added, and respective Geoportal Server related permissions could be adjusted quickly by updating the configuration file.

5.0 Next Steps

Work conducted this delivered an externally accessible prototype data catalog to better document a full pathway to HEIDI implementation on the HLAN. Changes were implemented within the prototype to test customization of the application and metadata schema. Finalizing the metadata schema is crucial to finding and retrieving data and will be a key element of FY25 work. Groundwater operable units were incorporated into the metadata as keywords during this phase of work, and therefore could be searched for within the catalog. Further testing is needed to incorporate other authoritative Hanford names within the metadata. High-level goals regarding the incorporation of Hanford names include the following.

- Identify authoritative names for the most common environmental entities at Hanford, including waste sites, facilities, and wells.
- Determine the appropriate schema configuration to capture multiple entity name types within the metadata of a resource (e.g., Hanford-specific thesauruses for keywords).
- Enforce the use of authoritative names during the creation of new metadata.
- Link synonyms for authoritative names within the metadata to enable a HEIDI user to find data using non-authoritative names.
- Investigate the potential for automatic authoritative name updates within the catalog metadata editor to support dynamic features of Hanford, such as wells.

The final component of HEIDI prototype testing is to create a transition plan to document the prototype functionality and approaches for implementation by the site contractors. This transition plan will feed into the overarching timeline of activities in the Hanford Environmental Data Management Plan. Once implemented on the HLAN, HEIDI would improve awareness of and access to existing environmental data that may be used to support site cleanup activities and remediation decision-making. Capitalizing access to existing data would inform sampling plans and other environmental studies without incurring the additional costs or time required to reproduce data that might otherwise be overlooked. Overall, HEIDI facilitates effective, data-driven strategic decisions to meet long-term remediation goals.

6.0 Quality Assurance

This work was performed in accordance with the PNNL Nuclear Quality Assurance Program (NQAP). The NQAP complies with DOE Order 414.1D Chg 2 (LtdChg), *Quality Assurance*. The NQAP uses NQA-1-2012, *Quality Assurance Requirements for Nuclear Facility Application*, as its consensus standard and NQA-1-2012, Subpart 4.2.1, as the basis for its graded approach to quality.

Any data presented in this document are preliminary, for information only, and subject to revision. The information associated with this report should not be used as design input or operating parameters without additional qualification.

7.0 References

Bence, Rebecka L., Aaron Moreno, Kenneth D. Ham, and Christian D. Johnson. 2023. *Specification and Prototype Software To Demonstrate a Data Catalog for Hanford Datasets*. Pacific Northwest National Laboratory, PNNL-34960, Rev. 0. Richland, WA.

DOE (U.S. Department of Energy). 1990. *Environmental Information Management Plan*. Washington, D.C. Available at https://www.osti.gov/biblio/7162238 (accessed September 2024).

DOE (U.S. Department of Energy). 1994. *Strategic Plan for Hanford Site Information Management*. Washington, D.C. Available at https://inis.iaea.org/collection/NCLCollectionStore/Public/27/051/27051915.pdf (accessed September 2024).

DOE (U.S. Department of Energy). 2010, *Hanford Long-Term Stewardship Program Plan*. Richland Operations Office, DOE/RL-2010-35, Rev. 0. Richland, WA. Available at https://reviewboard.ca/upload/project_document/EA0809-001_Hanford_Long_Term_Stewardship_Plan.PDF (accessed September 2024).

DOE (U.S. Department of Energy). 2012. *Long-Term Stewardship Information Management Program Plan*. Washington, D.C. Available at https://www.hanford.gov/files.cfm/HNF-50340_Hanford_Long-Term_Stewardship_Information_Mgmt_Plan_Rev1.pdf (accessed September 2024).

DOE Order 414.1D Chg 2 (LtdChg), *Quality Assurance*. Washington, D.C.

Ecology (Washington State Department of Ecology) et al. 2024. *Hanford Federal Facility Agreement and Consent Order*, as Amended through June 30, 2024. Olympia, WA. Available at https://www.hanford.gov/page.cfm/TriParty/TheAgreement (accessed September 2024).

Esri. 2019. "Esri's Open Source GeoPortal Server: Building Discovery Solutions." Presented at Esri 2019 Developer Summit Tech Sessions (webinar). Redlands, CA. Available at https://www.youtube.com/watch?v=MW4W_0RPoF4 (accessed on July 24, 2023).

Esri. 2023a. *Esri Geoportal Server Harvester | Wiki Documentation*. Redlands, CA. Available at https://github.com/Esri/geoportal-server-harvester/wiki (accessed on July 24, 2023).

Esri. 2023b. *Esri Geoportal Server Catalog (GitHub repository)*. Redlands, CA. Available at https://github.com/Esri/geoportal-server-catalog (accessed on July 24, 2023).

Esri. 2023c. *geoportal-server-harvester (GitHub repository)*. Redlands, CA. Available at https://github.com/Esri/geoportal-server-harvester (accessed on July 24, 2023).

Ham KD and DL Crockett. 2021. *Evaluation of Data Catalog Software for Hanford Site Environmental Datasets*. Pacific Northwest National Laboratory, PNNL-31960. Richland, WA.

Ham KD. 2022. *Requirements for Cataloging Hanford Geophysical Datasets*. Pacific Northwest National Laboratory, PNNL-33396. Richland, WA.

HMIS (Hanford Mission Integration Solutions). 2023. *Tri-Party Agreement Databases, Access Mechanism, and Procedures*. DOE/RL-93-69. Richland, WA. Available at https://www.hanford.gov/files.cfm/rl-93-69.pdf (accessed 24 September 2024).

NAP-Metadata Working Group. 2007. *North American Profile of ISO19115:2003 - Geographic information – Metadata (NAP – Metadata, version 1.1)*. Federal Geographic Data Committee, Reston, VA. Available at https://www.fgdc.gov/standards/projects/incits-l1-standards-projects/NAP-Metadata/napMetadataProfileV11_7-26-07.pdf (accessed on July 24, 2023).

NISO (National Information Standards Organization). n.d. *Building a Metadata Schema – Where to Start*. ISO/TC 46/SC11/N800R1. Available at https://committee.iso.org/files/live/sites/tc46sc11/files/documents/ (accessed September 2024).

NQA-1-2012, *Quality Assurance Requirements for Nuclear Facility Applications*. American Society of Mechanical Engineers, New York, NY.

RJ Lee Group. 2022. *F21-047, 200W Pump & Treat - Extraction Well Water Sampling*. AR-21427. Pasco, WA. Available at https://pdw.hanford.gov/document/AR-21427 (accessed 25 September 2024).

W3 (Worldwide Web Consortium). 2023. *Web Content Accessibility Guidelines (WCAG) 2.1, W3C Recommendation 21, September 2023*. Available at https://www.w3.org/TR/WCAG21/ (accessed 24 September 2024).

Appendix A – Updated Software

Preliminary software specifications were derived using a functional prototype environmental data catalog in fiscal year 2023 (Bence et. al., 2023). The configuration and programming of the catalog, metadata editor, and harvester components of Esri Geoportal Server were reviewed and updated based on (1) available software updates and (2) the server-based deployment of the prototype. Newly identified software versions are described in this appendix. More information about Geoportal Server (Esri 2019) is provided in Bence et al. (2023).

A.1 Esri's Geoportal Server Required Software

Esri's Geoportal Server Catalog v2 is based on Elasticsearch (https://www.elastic.co/) (version 8.8.2 or higher) and Apache Tomcat (https://tomcat.apache.org/) (version 9.x):

- Elasticsearch version 8.11.3 was used for this specification.
  - The required download is available at https://www.elastic.co/downloads/elasticsearch.
- . Apache Tomcat version 9.0.84 was used for this specification.
  - The required download is available at https://tomcat.apache.org/download-90.cgi.
- Java JDK 21 (AdoptOpenJDK v. 21 preferred) is a prerequisite for both the Catalog and the Harvester modules.
  - The required download is available at https://adoptopenjdk.net/.

A.2 Esri's Geoportal Server Harvester Required Software

The Harvester requires the same software as the Catalog (Elasticsearch, Apache Tomcat, and AdoptOpenJDK). Harvester v. 2.7.1 was used for this specification.

Appendix B – Server Specifications

B.1 Specifications

The internal Research Computing department at Pacific Northwest National Laboratory (PNNL) handled the provisioning and setup of the prerequisites of the application in a dedicated server. The server instance for the prototype is an OpenStack cluster instance with the following specifications:

- OS: CentOS/Linux
- Memory: 16GB
- CPU: 8 cores
- Disk space: 16GB
- Installations:
  - Elasticsearch 8.11.3
  - JDK-21
  - Apache Tomcat 9.0.84

To comply with security guidelines and make the prototype externally accessible to users outside of the PNNL network, it was necessary to apply for an exemption via the Cyber Security Exemption Tool (CSET). Along with the CSET application, creation of a PNNL TechDirectory entry for the project was required (Figure B.1).

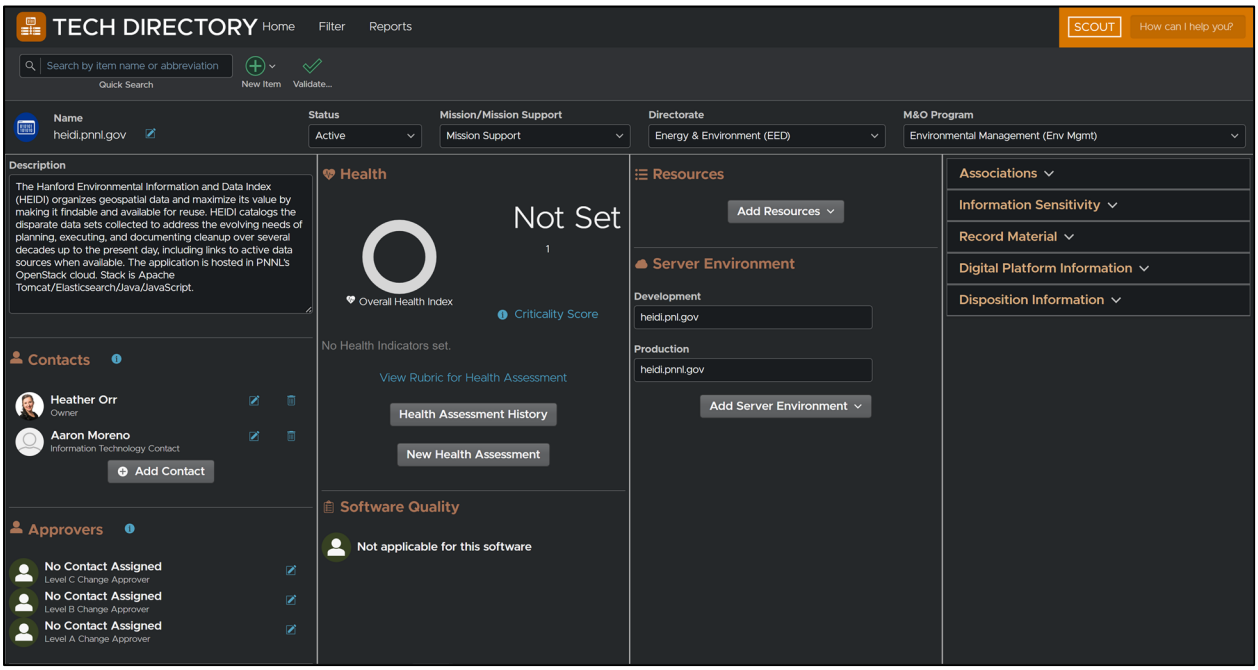

Figure B.1. TechDirectory entry for the Hanford Environmental Information and Data Index (HEIDI) application.

The internal Research Computing team at PNNL then provisioned the domain name of heidi.pnnl.gov for the externally accessible prototype. The application instance also needed to be proxied to not be dependent on a Uniform Resource Identifier (URI), so the domain was set up with an F5 virtual IP to act as a firewall.

To access the server instance and modify any of these specifications, one would need to have administrator privileges. All data within the prototype is publicly available (cleared for public use) to comply with restrictions on data access outside of the Hanford Local Area Network environment.

B.1.1 Catalog Installation

1. Deploy Geoportal Server
    a. Download the "geoportal-server-catalog-2.7.1.war" file from the GitHub repo: https://github.com/Esri/geoportal-server-catalog/releases (Esri 2023b).
        i. Geoportal v. 2.7.1 was used for this local installation guide.
    b. Extract the files, then move the "geoportal-server-catalog-2.7.1.war" file to the Apache Tomcat webapps directory.

B.1.2 Harvester Installation

1. Deploy Geoportal Harvester
    a. Download the "geoportal-harvester-war-2.7.1.war" file from the GitHub repo: https://github.com/Esri/geoportal-server-harvester/releases/tag/v2.7.1 (Esri 2023c).
    b. Extract the files, then move the "geoportal-harvester-war-2.7.1.war" file to the Apache Tomcat webapps directory.

B.2 Modifications to Allow Addition of F5 Firewall

To deploy the application within a secure and controlled network environment, it was necessary to address compatibility issues between the application's routing system and the laboratory's network security infrastructure, specifically F5 firewalls and reverse-proxies. F5 firewalls and reverse-proxies play critical roles in managing and securing network traffic, but this also introduced challenges when dealing with the application's Uniform Resources Location (URL) structures, particularly the use of hash-based routing.

Hash-based routing, which uses the "#" symbol in URLs (e.g., https://heidi.pnnl.gov/#search) as seen in Figure B.2, is commonly employed in single-page applications to manage client-side navigation. This routing method maintains the frontend navigation state while ensuring that the browser doesn't trigger a full page reload. However, F5 firewalls and reverse-proxies can mishandle these hash symbols, leading to issues like improper URL redirection or blocked access due to perceived security risks. This clash can disrupt the user experience and hinder the application's functionality.

Figure B.2. URL with hash-based routing, per the "out of the box" software configuration.

To ensure seamless integration with the F5 firewall and reverse-proxy setup, it was imperative to modify the Geoportal's routing and navigating system to eliminate reliance on hash symbols, as seen in Figure B.3. Key files in the application that were affected included App.js, OpenSearchLinksPane.js, Paging.js, and SearchComponent.js.

Figure B.3. URL without hash-based routing, which allowed the prototype to be deployed on an externally accessible server.

B.2.1 Modifications to App.js

This core component of our application, responsible for initializing routing and loading various modules, was updated to support non-hash-based routing. Changes were made to the startup logic, event bindings for the application tabs, topic subscription, and the _onHome() method. All these areas relied heavily on hash-based routing, so the updated code refactors those areas to directly display the relevant panel tab with logic to manually set the page displayed: $("a[href='#searchPanel']").tab("show"). This changes the routing handled in this core component to manually be setting which page to display rather than relying on the hash-fragments in the URL.

B.2.2 Modifications to OpenSearchLinksPane.js and SearchComponent.js

The OpenSearchLinksPane component manages the display and interaction of OpenSearch links. Modifications were made to replace hash-based navigation with a more manual path-based routing, ensuring links could be processed correctly by the reverse-proxy and firewall. The change was as simple as removing all instances of "#searchPanel" in any variable defining the URL. Modifications to SearchComponent.js were similar, with the removal of the history.replaceState call within the search method, which appended hash fragments to the URLs.

B.2.3 Modifications to Paging.js

The paging functionality, which often manipulates URL fragments to keep track of the current page, was refactored to modify the URL directly without hashes. This change was accomplished by removing the hash symbol from the URL during button click actions in the search component. By eliminating the history.replaceState calls that appended hash fragments to the URLs, the application improves its compatibility with network security configurations and results in cleaner URLs containing no hash symbols.

Appendix C – HEIDI Prototype Testing

The server-based deployment of the Hanford Environmental Information and Data Index (HEIDI) prototype enabled direct user feedback from staff at Hanford Mission Integration Solutions (HMIS) and Central Plateau Cleanup Company (CPCCo). Table C.1 presents the suggested changes to the software and actions Pacific Northwest National Laboratory (PNNL) to address the changes.

Table C.1. Issues identified by HMIS and CPCCo staff during testing of the HEIDI prototype.

No.	Issue Identified	Desired Resolution	Issued Resolved	If Not Resolved, Reason	Description of Actions Taken by PNNL
1	Well 199-N14 assigned incorrect bounding box.	Update bounding box to reflect 100-NR-2.	Yes	N/A	Bounding boxes fixed for all well data.
2	Date displayed on met data entry box is date of metadata creation.	Change display date to resource creation date.	No	Possible but more time needed.	Investigated the current metadata field read by Geoportal server. Determined filters and display date can be changed in the code of the software. More time needed by developers to adjust this.
3	The metadata upload option is there but it doesn't say what file types or formats are required. I tested it on a Word file and got an error. I then tested it on a txt file and also got an error.	Either need instructions, a form/screen in the admin portal with fields to enter data into and the system will format the information, or a user guide. I was not able to test this feature because there is no information on the file types or formats accepted.	No	Possible but more time needed.	A guide to uploading metadata was included in DVZ-RPT-097. A mini-user guide can be made available via the “About” tab in the portal interface.
4	I cannot edit the metadata for any items that are already in the catalog. The metadata editor only provides the option to edit files that I can open from a file location on my computer.	The admin portal should contain an option to select from existing items in the system and edit them, example to update a URL or a project contact.	Yes	N/A	Only users with administrator permissions can edit metadata added by a different user.
5	I cannot delete any metadata for items in the catalog.	The admin portal should contain an option to select from existing items in the system and delete them.	Yes	N/A	Only users with administrator permissions can delete metadata added by a different user.
6	One of the results “ARIUS” makes a red vertical line show up on the map	N/A	Yes	N/A	Bounding box corrected to reflect Hanford site.

No.	Issue Identified	Desired Resolution	Issued Resolved	If Not Resolved, Reason	Description of Actions Taken by PNNL
	instead of a bounding box. It looks like the left border of the Hanford Site bounding box.				
7	When moving between views (Catalog/Map/API/About) and the back arrow is clicked it logs out and returns to the “Notice to Users”	I would expect the back button to return me to the previous screen.	No	Not possible with current configuration	Investigated and discovered solution but requires the use of hashing ('#' in URL) which conflicts with the proxy we have set up to share the site with HLAN.
8	Searching the map is not intuitive and sets the map extent to something other than Hanford	Map should stay on relevant map extent.	No	Possible but more time needed	Discovered where map extent is supposedly configured but would not change the map extent when updated.
9	Legend on left is not intuitive as a set of filters by metadata. It took a bit to realize that the filters could be cleared above the selection set.	It needs to be obvious how filters work with relation to search and the selection set. Also, all filters are This and That there does not appear to be an option for an Or filter.	Yes	N/A	A label “Filter by” was added to the top of the left-hand panel filter options. A guide to using the filters was also added at the bottom of the left-hand menu.
10	Better organization of the metadata filters by category would make them more useful.	Filter categories need to be chosen based on Hanford Operations. In many cases these could be derived from existing Hanford Systems. For example, Wells should come from the HWIS Well Table (Hanford Environmental Information System: WELL_ADM.Wells)	No	Possible but more time needed	Custom filters are tied to the metadata and the unique information tagged to each uploaded metadata file. Work is ongoing to create Hanford specific filters.
11	Add a resource type filter (dataset, repository, ???)	N/A	No	Possible but more time needed	Determined the appropriate metadata field to record this information. More development work is needed to update the dropdown options available in the metadata editor. This will allow for a custom filter on this field.

No.	Issue Identified	Desired Resolution	Issued Resolved	If Not Resolved, Reason	Description of Actions Taken by PNNL
12	Add documentation for how to do Boolean searches	N/A	Yes	N/A	"Filter Guide" documenting how to do Boolean searches added to left-hand search panel.
13	Filter on the data source (EDA, ARIUS, etc.)	N/A	No	Possible but more time needed	Identified the filter field to change to be able to search on the parent repository. Still updating feature.
14	Link datasets to the repository they're in.	N/A	Yes	N/A	Added repository information to identification information in metadata (identification -> contact -> name)
15	Add headers so people know left panels are filters.	N/A	Yes	N/A	Added a label at the top of the filter panel to say "Filter by"
16	Add hover over descriptions so people know the features.	N/A	Yes	N/A	Added hover over descriptions.
17	See if you can add a filter that will filter based off of an official spatial zone.	N/A	No	Possible but more time needed	Similar to how the software can filter based on "intersects" or "within" map view pane, it should be possible to filter on spatial zone. More time needed to investigate.
18	See if you can add different layers to the map.	Test the ability to have multiple base maps. At the least, test the ability to change the basemap.	No	Possible but more time needed	Basemap is able to be changed, but still testing the ability to have a dynamic map interface with multiple base maps that a user can choose.

Appendix D – Proposed HEIDI Metadata Schema

Table D.1. Metadata Schema Proposed for HEIDI. Examples based on RJ Lee Group (2022) where applicable.

Location in Schema	Label	Data type	Required	Cardinality	Definition	Example
Metadata/Identifier	File Identifier	System supplied	y	1-1	Unique identifier for the metadata file	1686241220912r3853225710481458
Metadata/Identifier	Language	text	y	1-1	Language of the metadata schema	en
Metadata/Identifier	Hierarchy Level	dropdown	n	0-n	Specifies how the current record relates to a parent record in the type of data the current record offers	Dataset
Metadata/Identifier	Hierarchy Level Name	text	n	0-n	Specifies how the current record relates to a parent record in the type of data the current record offers	Software
Metadata/Contact	Organization Name	text	y	1-n	The organization responsible for managing the metadata schema	RJ Lee Group Inc.
Metadata/Contact	Email address	text	n	0-n	The contact email for the organization responsible for the metadata schema	RJLeeGroup@outlook.com
Metadata/Contact	Role	dropdown	n	1-1	The role of the organization regarding the metadata.	Author
Metadata/Date	Metadata Date (now)	date	y	1-1	The date the metadata was used to create a record	2023-06-08
Metadata/Standard	Metadata Standard Name	System supplied	y	1-1	The name of the metadata schema being used	ISO 19139/19115 Metadata for Datasets
Metadata/Standard	Metadata Standard Version	System supplied	y	1-1	The version of the metadata standard being used	2003
Metadata/Reference	Reference System	text	n	0-n	The system used to determine positions on the earth's surface	WGS84

Location in Schema	Label	Data type	Required	Cardinality	Definition	Example
Identification/Citation	Resource Title	text	y	1-1	The name of the resource	F21-047, 200W Pump & Treat - Extraction Well Water Sampling
Identification/Citation	Alternate Title	text	n	0-n	An alternative name for the resource	AR-21427
Identification/Citation	Resource Date (creation date)	date	y	1-n	A date related to the resource	2022-09-13
Identification/Citation	DateType	dropdown	y	1-1	Type of date	Publication
Identification/Citation	URI	text	n	0-n	Unique identifier for the resource	RJLG22C0529, Rev 0
Identification/Description	Abstract	text	y[(a)]	1-1	Description of the resource	F21-047, 200W Pump & Treat - Extraction Well Water Sampling
Identification/Description	Purpose	text	y[(a)]	0-1	Purpose of the resource	This data was collected for the purpose of analyzing the composition of groundwater in an unspecified well in the 200-ZP-1 Operable Unit.
Identification/Description	Credits	text	n	0-n	Credit statement for the resource	Heather Medley conducted the groundwater sampling for the RJ Lee Group.
Identification/Contact	Point of contact		y[(a)]	0-n	Contact for the resource	kgilman@rjleegroup.com
Identification/Contact/Point of Contact	Individual Name	text	n	0-1	A person related to the resource	Krystin Gilman
Contact/Point of Contact	Organization Name	text	n	0-1	An organization related to the resource	RJ Lee Group Inc.
Contact/Point of Contact	Position Name	text	n	0-1	The job title of a person related to the resource	Manager
Contact/Point of Contact	Contact Info		n	0-1	Contact information for the contact	509-792-1955
Contact/Point of Contact/Contact Info	Phone | Address | Online Resource | Hours of Service | Contact Instructions	text	n	0-n	Varieties of way to contact (there are subfields for these not recorded here)	N/A

Location in Schema	Label	Data type	Required	Cardinality	Definition	Example
Contact/Point of Contact/Contact Info	Role	dropdown	n	0-1	Role of contact vis-à-vis the resource	Point of Contact
Identification/Keywords/ Keyword Collection	Keywords	text	y	1-n	Keywords describing the resource, separated by commas	200-ZP-1
Identification/Keywords/ Keyword Collection	Keyword Type	dropdown	y	1-1	Type of keyword	Place
Identification/Keywords/ Keyword Collection	Associated Thesaurus	text	n	0-1	Thesaurus used to generate keywords	Link to Hanford Groundwater Operable Units Thesaurus
Identification/Keywords/ Keyword Collection	Specification Title	text	n	0-1	Name of a thesaurus	Hanford Groundwater Operable Units Thesaurus (or, if not Hanford place theme, could use generic thesaurus like ESIP)
Identification/Keywords/ Keyword Collection	Specification Date	date	n	0-n	Date related to a thesaurus	2014
Identification/Keywords/ Keyword Collection	Date Type	dropdown	n	0-1	Type of date for a thesaurus	Revision date
Identification/Constraints/ Usage Constraints	Use Limitation	text	y[(a)]	0-n	Limits of usage of the data	No special access is required to use the external Hanford AR, which is available to the public.
Identification/Constraints/ Legal Constraints	Access Constraints	text	n	0-n	Legal limits on accessing the resource	This data is legally restricted for access by DOE employees and contractors
Identification/Constraints/ Legal Constraints	Use Constraints	text	n	0-n	Legal limits on using the resource	This data is legally restricted for use by DOE employees and contractors
Identification/Constraints/ Legal Constraints	Other Constraints	text	n	0-n	Other legal restrains on the resource	This data is legally restricted for use by DOE employees and contractors, unless vetted through required DOE channels
Identification/Constraints/ Security Constraints	Use Limitation	text	n	0-n	Limitations on use of the resource	This data is restricted for use to people with a Q clearance.
Identification/Constraints/ Security Constraints	Classification	dropdown	n	0-1	Restricted	Restricted
Identification/Constraints/ Security Constraints	User Note	text	n	0-n	Note advising the user of a security constraint	This data is restricted for use to people with a Q clearance.

Location in Schema	Label	Data type	Required	Cardinality	Definition	Example
Identification/Constraints/ Security Constraints	Classification System	text	n	0-n	The classification system used for restricting the resource	Governmental Security Classification
Identification/Constraints/ Security Constraints	Handling Description	text	n	0-n	Note on how the resource must be handled	This data must be shared through a closed, secure system
Identification/Resource/R epresentation	Spatial Representation Type	dropdown	n	0-n	The type of spatial representation used	Vector
Resource/Language	Resource Language	country code/text	y	1-n	Language of resource	en
Resource/Classification	Topic Category	bullet list	y	1-n	Category that describes the resource	Environment
Resource/Extent/Spatial Extent	West Bounding Longitude	text	n	0-1	West Bounding Longitude	-119.6678578
Resource/Extent/Spatial Extent	East Bounding Longitude	text	n	0-1	East Bounding Longitude	-119.2618553
Resource/Extent/Spatial Extent	South Bounding Longitude	text	n	0-1	South Bounding Longitude	46.31274653
Resource/Extent/Spatial Extent	North Bounding Longitude	text	n	0-1	North Bounding Longitude	46.72263515
Resource/Extent/Tempora l Extent	Begin Date	date	n	0-1	The date the resource was begun	1998-01-01
Resource/Extent/Tempora l Extent	End Date	date	n	0-1	The date the resource ended	2020-08-15
Distribution/Distribution/ Distribution Format	Format Name	text	y[(a)]	0-n	The format the resource is distributed or downloaded in	Excel
Distribution/Distribution/ Distribution Format	Format Version	text	n	0-n	The version of the format the resource is distributed in	2023
Distribution/Distribution/ Online	URL	text	n	0-n	The link to the site where the resource can be obtained	https://pdw.hanford.gov/download/a635f2dc-133f-4a28-8921-93cf8a37a903
Distribution/Distribution/ Online	Function	dropdown	n	0-1	The function the online link performs	Information | download
Quality	Scope	dropdown	n	0-1	The quality the information applies to	Dataset

Location in Schema	Label	Data type	Required	Cardinality	Definition	Example
Quality/Conformance/Report/Conformance Result	Specification Title	text	n	0-n	The title of the quality control performed	NQAP
Quality/Conformance/Report/Conformance Result	Specification Date	date	n	0-n	Date related to the quality control specification	2024-01-01
Quality/Conformance/Report/Conformance Result	Date Type	dropdown	n	0-1	Type of date	Publication Date
Quality/Conformance/Report/Conformance Result	Explanation	text	n	0-1	Description of the quality control process	This work was performed in accordance with the Nuclear Quality Assurance Program (NQAP). The NQAP complies with DOE Order 414.1D Chg 2 (LtdChg), Quality Assurance.
Quality/Conformance/Report/Conformance Result/Degree	Validation Performed	dropdown	n	0-1	Indicates if the validation was conformant or non-conformant	Conformant
Quality/Lineage	Lineage Statement	text	y[(a)]	0-1	The Lineage field offers a free-text space to record a statement about the data's chain of custody. This will include the individual or organization that created or recorded the data, who has maintained the data, and how the data arrived at its current location. Any additions or modifications to the data should also be noted.	This data was created by the Central Plateau Cleanup Company and recorded by Heather Medley. The data has been maintained in the EDS database and is currently collected on a monthly basis.

(a) Specific metadata requested by the Hanford working group, beyond typical minimum metadata requirement.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov