&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&

NATIONAL GEOSCIENCE

DATA REPOSITORY SYSTEM

PHASE III: IMPLEMENTATION AND OPERATION OF THE REPOSITORY

PROGRESS REPORT
DE-FG22-97BC-14847

3rd Quarter
July - September, 1997

Submitted by the

AMERICAN GEOLOGICAL INSTITUTE

to the

Office of Fossil Energy, U.S. Department of Energy

October 1997

American Geological Institute, 4220 King Street, Alexandria, VA 22302-1502
(703) 379-2480   Fax: (703) 379-7563

&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&

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.

NGDRS Phase III — Implementation
Third Quarter Progress Report

October 1997

Overview

Most of our activities during the third quarter dealt with the deployment and the starting of the utilization of the NGDRS technological infrastructure. We dealt primarily with the installation of the Sun Microsystems computer servers. The installation has been very smooth with no particular problems to highlight. Installation of the Oracle database management system on the Sun servers is currently underway. On the software side, work continued with the development of the front-end pages for the web site, and with modifications and additions to the GeoTrek™ software to fulfill the specific needs of the repository system. Activity on Project 1.0, the conversion and installation of electronic databases continued. We have located and obtained from the BLM the software needed to translate the Oklahoma Geological Survey data set Township/Range/Section coordinates into Lat/Long. We have received and are currently loading the new updated data sets (cores and well log headers) from the BEG. We attended a meeting at the Kansas Geological Survey (KGS) on September 24 to define the participation of the KGS in the NGDRS project. No project review was held in August. Details follow.

Progress by Specific Project

1.0 ESTABLISHING THE CLEARINGHOUSE

Task 1.1 -- Project Plan and Management

The following projects had their plan sheets updated:

PROJECT NAME	PROJECT TYPE	PROJECT STATUS
1. Houston Core Facility Project	Transfer of Core and Cuttings	Underway
2. Well Log Data Tape Conversion Project	Transfer of Digital Data	*Not defined*
3. Seismic Data Tape Conversion Project	Transfer of Digital Data	Underway
4. BEG Project	Electronic Database Transfer	Underway
5. PGS Project	Electronic Database Transfer	Underway
6. Kansas Geological Survey Project	Electronic Database Transfer	Planned
7. Mobil Seismic Sales Project	Electronic Database Transfer	*Not defined*
8. Metadata Repository Infrastructure	Creation of the Metadata Repository	Underway
9. Metadata Repository Utilization	Utilization and Operations of the Metadata Repository	Underway
10. Mineral Management Services (MMS)	Electronic Database Transfer	Underway
11. Oklahoma Geological Survey (OGS)	Electronic Database Transfer	Underway

A **project review meeting** was held on August 27 at the offices of The Information Store in Houston. The meeting was attended by M. Milling, Jan F. van Sant, R. Graebner, Christopher Keane, G. Breed, O. Teoh and A. Eloy. The agenda included the following topics:

- An overview of the Phase III progress and different projects. Refer below for details on each project.
- The difficulties faced by the NGDRS project, resulting from the current industry level of activity, led to the suggestion to reconvene the NGDRS Steering Committee during October 97 to reassess the priorities and opportunities for <u>data preservation</u>. The possibility was brought up of transferring most (or all) of the project efforts and resources to the <u>data access</u> activity. Switching to a data access strategy will require installing and incorporating a much larger number of electronic databases than those currently contemplated.
- Possible new targets for data access were proposed by C. Keane. They could consist of well established mature repositories (such as EROS) that manage data from satellites, gravity surveys, geochemical surveys, bathymetry, and other data types. The inclusion of these additional repositories, containing geosciences or geospatial data, was described as being useful not only to the geoscientist community (in the petroleum industry) but also potentially to a much larger audience in academia and government. This expansion of the NGDRS objectives could lead to much broader and intensive use of the NGDRS services by all communities.
- An agreement was reached regarding the contents and tone of the Internet site that will house the NGDRS metadata supercatalog. The following general principles were adopted:
  - Site should have the look-and-feel of a **"yellow pages"** advertising or announcing data from public as well as private repositories.
  - Site will emphasize **breadth** (versus depth) of information. As a "yellow pages" the objective is to display abbreviated contents for **all** relevant repositories (i.e., to be a complete catalog).
  - Site will strive for **simplicity**. This should be reflected in the user interface, user access, and uniform (possibly standard) metadata for the different data types.
  - The site will be clearly identifiable as an **AGI site**. References to The Information Store will be kept to a minimum to avoid any mis-perceptions. The Information Store will be treated just as other private data repositories.
  - **GeoTrek** (version 37) was identified as the software browser for the NGDRS supercatalog.
- We have agreed on targeting mid-October as the date for the official opening of the NGDRS to the public.

Task 1.2 - **Houston Core/Cuttings Facility (HCF)**

Meetings at Amoco have yet to produce the desired results. M. Milling and Jan van Sant reported on their meetings with Amoco management and described the difficulties obtaining commitment from all parties involved in the donation of a facility for the HCF.

Task 1.3 - **Digital Data Media Conversion**

Discussions were held with Mobil E&P confirming the possibility of having a donation of 65,000 seismic tapes promised sometime ago. Unfortunately, because Mobil E&P now believes these tapes have a commercial value, they are no longer available for donation. New targets are being considered.

Task 1.4.1 - **BEG - Databases Conversion**

BEG contracted with Oracle for installation of the Oracle DBMS and Oracle Web Server at their facilities in Austin. This task in now complete. We have requested a new version of their databases in 4mm tape TAR format. We expect to install GeoTrek in Austin during September. Meanwhile, a template contract for the use of the GeoTrek software at the BEG facilities was sent to the AGI. The AGI is responsible for all contractual aspects relating to GeoTrek.

Task 1.4.2 - PGS (and Fairfield) - Databases Conversion

Loading of the seismic surveys database continued in August. Several additional meetings were held to review data quality and other technical details. NGDRS activities related to the Fairfield account will begin to be reported under this project. Fairfield data is also being prepared for input into the NGDRS supercatalog. Both clients (PGS, Fairfield) have requested that their data sets (that we are loading for the NGDRS) not be displayed until after the SEG trade show takes place (first week in November).

Task 1.4.3 -- *Kansas Geological Survey (KGS)* - Databases Conversion

Susan Bolton, from the Geoscience Data Resource Management group, organized a meeting on September 25 at KGS headquarters in Lawrence. Besides Susan Bolton, the meeting was attended by: William (Bill) Harrison (Deputy Director), Tim Carr (Section Chief for Petroleum Research Section), and Ken Stalder (Web Content Assistant for the Petroleum Research Section). The meeting agenda included several items: an overview of the NGDRS project, a demo of the GeoTrek browser, a discussion of the current status of KGS projects, and scoping a preliminary plan of action for KGS's participation in the NGDRS. Some general comments follow:

- The KGS has just completed some internal reorganization. Bill Harrison, the Deputy Director for the Survey, now has most of the day-of-day operations responsibilities for KGS. Lee Gerhard's (Director) activities will deal more with the research aspects. Bill attended part of our session in the morning and seemed positively impressed with the NGDRS goals.
- As part of the changes, the director for Information Services (Joe Ritt) has left the KGS and Bill has assumed his responsibilities. Tim Carr and Susan Bolton appear to share the responsibilities for technical information systems. They both seem to be in agreement that the NGDRS and GeoTrek are of value to them.
- The following is a preliminary, prioritized list of databases that Tim and Susan identified as the potential targets for installation in the NGDRS supercatalog: electrical logs, well cuttings, plugged wells, cores, scout tickets, production data, list of operators, cultural data, and other geophysical (magnetic and gravity).
- Most of the KGS data is identified using the Township/Range/Section (TRS) coordinate system which for our purposes will need to be converted to Lat/Long. Susan remarked that their customer data requests are made in the TRS system. KGS client requests normally read as "send me all well logs within TRS."
- Susan and Tim mentioned most of their clients are majors or independents typically not located in the state of Kansas, more commonly in Houston or Calgary. There is only a small group of Kansas-based operators in the eastern part of the state. Remote access to Kansas data for all operators was acknowledged to be very important.
- Tim Carr is also an officer of the Kansas Geological Society and he believes their log database (about 2,000 wells) is a good candidate for the NGDRS. The Society is a for-profit organization. He promised to involve the Society at the right time.
- Most of KGS databases are Oracle-based or are being converted to Oracle. Susan remarked that the databases were recently loaded into Oracle in their original form and schema, with no data model used as a standard. In the future they would like to take advantage of a standard E&P data model and they are leaning towards the PPDM data model. They would appreciate any help we could give them with their data modeling efforts.
- Susan mentioned that she is *definitely* interested in installing GeoTrek at KGS, probably when the conversion of their databases is completed.

The following action items originated from the meeting:
- Susan requested permission to use the GeoTrek Internet-based demo. We have sent her instructions and given her access privileges to the GeoTrek demo.
- Susan agreed to put together a preliminary action plan for their participation in the NGDRS leading to an implementation during the November-December (97) time frame.

Task 1.4.4 -- **MOBIL Seismic Sales Organization** - **Databases Conversion**

There is no progress to report for this project.

Task 1.4.5 -- **OGS** - **Databases Conversion**

We have located a software program to help us convert the OGS core database locations from Township/Range/Section (TRS) into Lat/Long coordinates needed by GeoTrek. The program, named "TRS2LL" was identified and provided to us by the BLM. The program is in executable code only and we are defining procedures to ensure it generates reasonable approximations for the Lat/Long location. If the program proves reliable it will be very useful with several other databases, such as the those described above for the Kansas Geological Survey.

Task 1.4.6 -- **MMS** - **Databases Conversion**
We have identified all the information elements present on the "Borehole" data set that we will need to create the NGDRS well information database. We are beginning to define the procedural aspects of how to periodically update the NGDRS supercatalog from the MMS databases.

Task 1.5 -- **Hardcopy Data Media Conversion**
This project has no funds allocated in the FY96 budget. No activity is planned at this time.

2.0 CREATING THE METADATA REPOSITORY

Sun Microsystems

The Sun Microsystems servers - Ultra II and 3000 - arrived respectively on 9/5 and 9/9. The remaining hardware components (DLT tape unit for backups and storage array for Ultra IIs) arrived on 9/30. Since the arrival of the equipment the following was achieved:
- Compilers (C, C++, Perl) installed and tested.
- Round-robin DNS capabilities implemented and tested.
- FrontPage web page authoring extensions installed and tested.
- Oracle 7.3 database management system was installed on Sun Server 3000 and is being tested.
Loading of the software components, including GeoTrek, is currently underway.

Internet Software Components Development and Installation

The development of the front-end pages of NGDRS continued. It was determined that a new page be created to facilitate user access to data from different parts of the country. This page, to be accessed just before entering GeoTrek, will present a map with hyperlinks into the geographic areas for which supercatalog data is available. This map will initially point to our target areas — GOM, Texas, Kansas and Oklahoma.

Below is the status of different sections of the NGDRS web site:

About the NGDRS project	completed
What's New (1997)	completed
NGDRS Installed Data Sets	not started
What's Planned (1998 and beyond)	started
Site Purpose and Intended Audience	completed
Tutorials	started
Frequently Asked Questions (FAQ)	completed
Ask A Question	started
System Requirements	started
Contact Us	started

The "streamlined" version of the GeoTrek software required by the NGDRS for Internet access continues to be developed by the software team headed by I. Morison. Early versions are being installed and tested on newly installed Sun Ultra II servers.

3.0 UTILIZING THE METADATA REPOSITORY

Marketing NGDRS

The following is an updated list of events or tasks that have been identified as part of the on-going marketing campaign:

EVENT	WHERE	WHEN	STATUS	RESPONSIBLE
1. Repositories Meeting	Calgary	June 16-17	Completed	Tom Woolley
2. DOE Contractors Workshop	Houston	June 20	Completed	MM, GB and AE
3. PTTC Board Meeting	DC	July 13-15	Completed	GB
4. AGI/NGDRS Steering Committee Meeting	Houston	Oct. 14	Invitations issued	MM, JVS
5. GSA Meeting	Salt Lake City	Oct. 20-23	*Not defined*	MM, D. Archer
6. SEG Annual Meeting	Dallas	Nov. 2-7	*Not defined*	---
7. AAPG - Annual Meeting	Salt Lake City	May 17-20	*Not defined*	---
8. POSC/AGI Standards Workshop	DC	Not scheduled	Discussions	MM, D. Archer
9. NGDRS Marketing Plan	Houston	Oct./Nov.	Not Started	A. Eloy

An NGDRS Steering Committee meeting was organized by M. Milling, J. van Sant and J. Deery (Amoco) for October 15 at Amoco facilities in Houston. The tentative agenda includes a demonstration of NGDRS and GeoTrek, the status of the proposed donations to NGDRS, and a status report from each company on their data management activities.

Utilization of the Metadata Repository

Two tasks were started in this project during September. One deals with the several options for registration of NGDRS users. The intent is to keep access to NGDRS as simple and as open as possible. However, there is a need to install a registration/password process that will permit gathering of statistics on systems utilization and provide some basic auditing capabilities. A team lead by O. Teoh is investigating the different alternatives for registration and make a final decision during October.

Meanwhile O. Teoh is also testing systems performance and identifying potential issues related with access of the metadata supercatalog via Internet. This will require, for example, to test the access behind corporate "firewall" systems. This will be done initially from several sites around Houston, specifically from BPX, Amoco, Fairfield and PGS. From this effort we expect to get a realistic and relevant understanding of system performance and other issues that may confront NGDRS users.

<u>Alternate NGDRS Site at the AGI</u>

During our conversations with Christopher Keane in August, Chris suggested the possibility of establishing an NGDRS "backup server" at the AGI facilities. This would provide fall back capabilities for those times when the network services provided by our Internet provider (PSINet) are not available. We know from C. Keane's experience at the University of Maryland, this eventuality is not an uncommon occurrence and hence deserves our attention. We have, therefore, begun to make plans with C. Keane for an AGI site.