**NATIONAL GEOSCIENCE
DATA REPOSITORY SYSTEM**

**PHASE III: IMPLEMENTATION AND OPERATION
OF THE REPOSITORY**

**PROGRESS REPORT
DE-FG26-99BC-15115**

**1st Half FY00
October 1999-March 2000**

Submitted by the

AMERICAN GEOLOGICAL INSTITUTE

to the

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

April 2000

*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 Part 2— Implementation

First Half FY00 Progress Report

April 2000

Overview

In the past six months the NGDRS program has continued to engaged new contacts, identify additional data transfer targets, and improve the metadata catalog for both easier use and long-term maintainability. With industry conditions continuing to rapidly change and evolve, the primary core and cuttings preservation strategy has evolved as well. With the severe lack of available public data repository space and the establishment of a major national geoscience data repository facility unlikely in the near future, the focus is on increasing public awareness and access to non-proprietary company data holdings that remain in the public and private sector. Efforts still continue to identify and facilitate the entry of new repository space into the public sector. Additionally, AGI has been working with the National Academy of Sciences Board on Earth Sciences and Resources staff to initiate a study and workshop to develop a policy recommendation on geoscience data preservation and prioritization of efforts.

Additional data transfer efforts were undertaken during the first half of FY00. AGI is working with the Texas Bureau of Economic Geology to assist in the transfer of Altura's midland core holdings to the University of Texas. Phillips has made selected seismic data from the Santa Barbara Channel available for transfer. A pilot test has been initiated to determine the cost and potential success rate at transcription of the original tapes.

Additionally, redesign of the GeoTrek metadata catalog was initiated, including both a redesign of the user interface as well as making GeoTrek fully a broker, accessing multiple databases at remote locations in real time.

Program Activities Summary

1.1 Pre-planning of Geoscience Data Preservation Workshop – National Academy of Science

Discussions began between the American Geological Institute and the Board on Earth Sciences and Resources staff of the National Academy of Science concerning holding a workshop on the issue of geoscience data preservation. The discussions have focused on defining the scope, goals, outcomes, and timing of the workshop. The American Geological Institute encouraged the National Academy of Science to conduct the workshop in the later part of 1999, with a primary focus on determining the priorities, optimal methods, and development of a national geoscience data preservation strategy. The expected outcome will be a report of the National Academy of Science based on the results of the workshop detailing a recommended national geoscience data preservation strategy, including an assessment of the need and priorities for preservation. It now appears the study will be initiated in 2000.

The Board on Earth Sciences and Resources is raising support for this study from various agencies and private organizations, including the US Geological Survey, US Department of Energy, POSC, and AGI.

1.2 NGDRS Steering Committee Meeting

On November 19, 1999, the NGDRS Steering Committee, chaired by Robert Merrill, Chief Geologist Spirit 76, met at the offices of Unocal in Sugar Land, Texas. The committee discussed current issues regarding the NGDRS project, including the future of GeoTrek, the NRC study, and the issue of repository space. Most of the remaining items in this report were discussed and guidance provided during the meeting. Minutes and action items from the meeting are attached as appendix A.

1.3 Discussions with DOSECC

Discussions with DOSECC began in the second quarter of 1999. DOSECC is a consortia of 48 universities and research laboratories who are engaged in research on onshore crustal studies and drilling techniques. Given DOSECC's interest in onshore cores, AGI made contact with their Executive Director, Dennis Neilson.

DOSECC currently has two major operations underway, drilling 5000 meters of core from the flank of Mauna Kea and deploying a mobil floating drill rig for coring of lake bottoms, such as the Great Salt Lake. DOSECC recognizes the long term core preservation issues and recognizes that all projects face similar circumstances in being unable to find data repositories willing to accept the core for curation. This situation represents a potential point of collaboration.

With their focus on core and equipment, DOSECC has found itself with an immediate need for storage space. DOSECC, in communication with AGI, contacted the agent for the former Toole Army Depot west of Salt Lake City to inquire about potential storage space. At this point the property prices are too high to be viable for acquisition by either DOSECC or the NGDRS. However, both organizations will continue to monitor the situation and if property prices decline, then this option once again may be revisited.

Establishing the Clearinghouse

2.1 Project status

The ongoing component projects for the NGDRS and their status:

Project Name	Project Type	Status
Metadata Repository Infrastructure	Creation of the Metadata Repository	Completed
Metadata Repository Utilization	Utilization and Operations of the Metadata	Completed
Mineral Management Services (MMS)	Electronic Database Transfer	Completed
Oklahoma Geological Survey (OGS)	Electronic Database Transfer	Completed
PGS Project	Electronic Database Transfer	Completed
Eastern Gulf Region PTTC	Electronic Database Transfer	Completed

BEG Cores and Cuttings	Electronic Database Transfer	Completed
Seismic Data Tape Conversion Project	Transfer of Digital Data	Underway
Altura Midland Core Facility	Transfer of Cores and Cuttings	Completed
Mobil Seismic Surveys	Electronic Database Transfer	Underway
Unocal Onshore Cores & Cuttings Metadata	Electronic Database Transfer	Completed
Unocal/Spirit Energy Cores and Cuttings	Transfer of Core and Cuttings	Underway
Oregon Department of Geology & Mineral Industries	Electronic Database Transfer	Underway
Phillips Seismic Surveys	Transfer of Digital Data	Pilot Underway
Chevron Cores and Cuttings	Cataloging and Opening of Data	Underway
State of Arizona Core and Well Log Records	Electronic Database Transfer	Underway
Utah Geological Survey Cores and Cuttings	Electronic Database Transfer	Underway

2.2 Data Targeting and Transfer

2.2.1 Unocal/Spirit Energy Cores and Cuttings

Unocal has completed cataloging and indexing their non-proprietary onshore domestic cores and cuttings. The metadata for the conterminous United States was transferred to the NGDRS and is being hosted by the American Geological Institute in its database servers. This transfer covered 2082 cores from across the nation. Initial quality control by AGI allowed the inclusion of 1198 core records into the metadata catalog. Additional work on the data will be performed during the second Quarter of 2000 to complete the inclusion of all of the Unocal core records. Unocal has indicated that they will be providing the NGDRS metadata for their Alaska, California, and Offshore data during 2000. Unocal continues to express interest in the transfer of ownership of their cores and cuttings to the NGDRS. However, their core and cutting holdings remain in storage at C&M Storage in Schulemberg, Texas. Users of the GeoTrek metadata catalog can arrange for access to listed cores by contacting C&M Storage directly.

2.2.2 Unocal/Spirit Energy Utah Core Transfer

The NGDRS is assisting in the transfer of Unocal's Utah cores and cuttings from Shulemberg, Texas to the new core repository at the Utah Geological Survey. Unocal has made as a condition of this transfer, that all of Utah's data holdings, including the Unocal data, needs to be listed in GeoTrek. Discussions have been suspended until a new Director at the Utah Geological Survey is named.

2.2.3 Chevron Cores and Cuttings

Chevron has agreed to index their core and cutting holdings at the C&M Storage facility. The non-proprietary Chevron cores and cuttings records will be released to the NGDRS for inclusion in the metadata catalog. Similar to Unocal, once an agreement between AGI, Chevron, and C&M Storage is reached for access to the non-proprietary data, the Chevron holdings will be made publicly accessible. Chevron has provided AGI with a pilot test sample of their metadata for quality control analysis. The Chevron metadata is in extremely good form and should result in the rapid

integration of their data holdings into the metadata catalog once they have completed indexing and transfer.

2.2.4 Altura Midland Core Facility

Altura has transferred ownership of some 80,000 boxes of core and cuttings to the Bureau of Economic Geology at the University of Texas at Austin. The materials are from Altura's Permian Basin properties. The BEG is expanding its existing repository facility in Midland, Texas to accommodate the new Altura acquisition. The actual transfer is expected to take place this Fall on completion of the repository expansion.

2.2.4 Phillips Seismic Tapes

Phillips Petroleum has transferred selected seismic holdings for the Santa Barbara Channel in California to the NGDRS in the Fall of 1999. AGI initiated a pilot project in December 1999 to evaluate the feasibility and costs for transcribing the data to current media and format. The data is currently stored on 1-inch analog tapes, for which there are few known working readers. It is believed that if feasible, only selected lines can be economically retrieved. The pilot project should conclude in 2Q 2000 with potential data transfers occurring shortly thereafter.

2.2.5 Kansas Geological Survey

During March 2000 AGI staff met with the Kansas Geological Survey's Oil & Gas Division. An agreement was made to link the current metadata catalog of the Kansas Geological Survey's cores and cuttings holdings to the NGDRS metadata catalog. This linkage planned for 2Q2000 will add over 350,000 new wells to the metadata catalog, and complete the integration of the main oil and gas states in the conterminous United States into GeoTrek. The Kansas Geological Survey integration will represent the first State Geological Survey to have their data integrated via a remote real-time linkage between the GeoTrek broker and onsite databases.

2.2.6 Miram County, Wisconsin Mining Core

In February 2000 two faculty members from the University of Wisconsin inquired about the ability for the NGDRS to assist in the transfer of 62,000 boxes of mining core from sulfide exploration in Miram County, Wisconsin. The county council had decided to dispose of the core immediately to save storage costs since the Wisconsin Department of Natural Resources had functionally terminated any potential for economic use of the sulfide prospects. Inquires to surrounding states yielded no available space for storage of the core, and given the short time horizon to disposal (less than 1 week), transfer of the core into permanent storage was infeasible.

2.2.7 Texas Railroad Commission

The Texas Railroad Commission is working with the Texas Bureau of Economic Geology to make more of their data available electronically to the public. Through GeoTrek, the RRC and BEG have been running access experiments. The production process for digitization of logs for the RRC is 24 min per log (16 min/log to scan and 8 min/log for Qc and editing).

The BEG and RRC have entered Phase II of the effort to evaluate:

- Quality
- Value of the data. (What needs to be preserved, e.g. hearing files and submitted logs)

Viability of having tiff images of logs as a legal substitute for hard-copy logs (this will begin on a voluntary basis for the near term).

The pilot test data is available via the GeoTrek metadata catalog, and as additional data is completed, they will be integrated into GeoTrek.

Utilizing the metadata repository

3.1 Operating the Metadata Catalog

The operation of the metadata catalog continued during the second quarter of 1999. The following databases are currently available on the metadata catalog:

- Fairfield Seismic
- A2D Well Logs
- MMS Well Logs
- Alabama Eastern Gulf PTTC Well Logs
- BEG Well Logs
- BEG Cores
- Oklahoma Geological Survey Cores
- MMS Block and Lease Boundaries
- Texas Railroad Commission Well Logs

PGS has opted to delay listing their holdings on the metadata catalog until they can provide sufficient quantity and quality of sample seismic sections for online viewing. A number of additional vendors are considering adding their catalogs to the NGDRS system as well.

The current access statistics are provided, as well as Project to Date (PTD).

	1998	1Q-3Q99	4Q99	1Q00	PTD
NGDRS Website Hits	30,911	37,947	23,205	25,184	117,247
Unique Visitors to NGDRS	1331	3171	1165	1523	7190
Unique GeoTrek Users	167	160	143	134	604
Number of GeoTrek Searches	2000	1301	137	1288	4726
Total GeoTrek databases transactions	17,930	1588	809	3069	23,396

Appendix A

Minutes and Action Items from November 11, 1999 Steering Committee Meeting

NGDRS issues:

- 1. Storage facilities**
- 2. Funding for continued storage is under pressure.**

Bob Merrill, Steering Committee Chairman reviewed the central issues facing the NGDRS initiative. These are:

Appropriate storage facilities

- The NGDRS should link Federal, State, Academic and non-proprietary geoscience data repositories: Geotrek has the ability to do this.
- Core and sample storage space for nonproprietary, industry core is in short supply: Such storage needs to be identified or built
- Funding is ongoing for Federal and State storage, but is always under reduction pressure.
- Funding for non-proprietary, industry material is ongoing, but pressures to eliminate funding are strong, so material is being destroyed.

Geotrek is the link and is operational.

Geotrek has the ability to link all facilities' databases and is operational. Geotrek is being streamline to meet the IT industry standard "3-click rule". In other words, a user can create a query in 15 seconds or less. An enhanced help section has been added. It is now easy to separate data queries by type and only bring up that data-type within a defined area of interest. Further, there is also an expert user option with fewer screens (command driven). The software is also moving toward dynamically accessing repository metadata, so updates do not have to be sent to a specific server periodically.

There has been limited success adding vendors to Geotrek. A few seismic companies are participating and well log vendors are also joining the effort.

The Board on Earth Sciences and Resources has been commissioned to study the issue of geoscience data repositories. This study will bring additional visibility to geoscience data storage, particularly with respect to funding sources.

The Board on Earth Sciences and Resources has been asked by its parent, the National Academy of Sciences/National Research Council to produce a report on the preservation of geoscience data and collections that is sponsored by the DOE, NSF and USGS. The charge of this study is to:

1. Develop a strategy for prioritization of geoscience data and collections;
2. Examine options for long-term storage curation and archiving of geoscience data and collections;
3. Use case studies accession and repositories as examples of successes and failures; and
4. Discuss the roles of the public and private sectors in geoscience data preservation.

The overall goal of the study is to develop a comprehensive national strategy and policy for managing geoscience data in the United States.

A committee of about 12 is expected to be assembled by early January, 2000 and expected to complete the report in about 18 months. In mid-November, about \$195,000 of the required \$250,000 was promised. If \$400,000 were available, the study would be completed in less than one year.

Bob Merrill volunteered to help in whatever capacity needed, including committee member.

DOSECC will probably have significant storage requirements.

DOSECC, the Drilling, Observation and Sampling of Earth's Continental Crust program is a 501(C)(3) corporation supported by 44 universities, 3 laboratories and 1 state geological survey. The charter is to support continental scientific drilling. It is a focal point for data generated by NSF funds.

With the closing of the Grand Junction Core Repository, material was disbursed to Denver and the University of Utah. New material is coming in and expected to accelerate with the development of their hybrid coring system. As an example, the Hawaii drilling program is expected to generate at least 10,000 feet of core. The current plan is to build or contract additional storage space at the former Toole Army Depot west of Salt Lake City.

The Texas RRC is working with the Texas BEG to make more data available to public electronically.

The Texas Railroad Commission is working with the Texas Bureau of Economic Geology to make more of their data available electronically to the public. They have been experimenting with Geotrek to do this. Their scanning efforts take 24 min per log (16 min/log to scan and 8 min/log for Qc and editing).

Phase II of the work will address the

- Quality
 - Value of the data (What needs to be preserved, e.g. hearing files and submitted logs)
- Viability of having tiff images of logs as a legal substitute for hard-copy logs (this will begin on a voluntary basis for the near term).

Reports

Texas – Marcus Milling: Texas is adding a contribution from Altura to the Midland Core Repository.

Kansas – Bill Harrison: Kansas received 4500 boxes of core from Amoco with an additional 10,000 boxes to follow for archival. Kansas currently has 8 – 10 years of space available for additional core storage.

Oklahoma – Charlie Mankin: Oklahoma has received 109,000 boxes of core from Amoco accompanied by a cash contribution to defray ongoing storage fees. New storage space is added by purchasing shipping containers, each of which holds about 4000 boxes of core. Currently Oklahoma has 12 such units in Norman, OK.

Wisconsin – Chris Keane: Miram County, Wisconsin has inquired about placing 62,000 boxes of core remaining from sulfide exploration in that county. Proper storage facilities are not available on an ongoing basis.

Phillips Petroleum – Tom Moore: Phillips is looking for a place for all their Santa Barbara Channel seismic data. As far as core storage is concerned, Phillips has sufficient space for 8 – 10 years of additional material.

C&M Storage – Robert Shafer: C&M Storage, Schulenberg, Texas presented a proposal to maintain public storage of industry core. Their current facility contains about 205,000 ft² of core storage with associated layout and processing facilities. For future growth C&M has options on adjacent acreage. Volume requirements for C&M are roughly: 879 boxes can be contained in a volume of 250ft² and 8 ft high. Their costs to move core from Houston to Schulenberg is about \$3.19 per box. This includes associated shelving.

Marathon – Jean Hamilton : Marathon has core scattered in several facilities with available space in Littleton, Colorado.

U.S. Geol. Survey – Barney Lewis: Ice core storage is full. Conventional core storage is full. The USGS is looking at ways to repackage material to more efficiently use their space.

AAPG – Chuck Noll: The new Executive Director, Rick Fritz, has a keen interest in the NGDRS initiative.

MMS – George Dellagrino: Seismic data will take on more prominence in the future as older 2-D data becomes obsolete.

Westport Labs – Mike Dix: There is also a need to maintain a system to archive and curate thin sections and paleo slides.

Chevron – Ed Donovan: Chevron has been trying to find places for core that has historically been stored near their district offices such as New Orleans and in Richmond, California.

Other: Nuevo is looking to find a home for the California core left by Unocal following Nuevo's purchase of Unocal's California properties. ARCO has also been looking to find a solution to preserve their California material.

Action Items:

1. The NGDRS will support the Board on Earth Sciences and Resources study to bring additional visibility to geoscience data preservation. – Bob Merrill, Marcus Milling
2. The NGDRS will pursue discussions with C&M to find a solution for storing industry non-proprietary core that is in Texas or might be shipped to Texas. –Marcus Milling, Bob Merrill
3. Determine the need and storage needs for preserving thin sections and paleo slides. – Mike Dix
4. Connect additional repositories to the NGDRS via Geotrek. – Glen Breed
5. Continue to upgrade Geotrek. – Glen Breed

Attendance:

Jean Hamilton (for Dale Beach)	Marathon Oil	jlhamilton@marathonoil.com
Barney Lewis	USGS	Bdlewis@usgs.gov
C.R. "Chuck" Noll	AAPG	Crnexp@aolcom
George Dellagriano	MMS	George.dellagriano@mms.gov
Bill Harrison	Kansas Geological Survey	Harrison@kgs.ukans.edu
Mike Partridge	Texas Railroad Commission	Mike.partridge@rrc.state.tx.us
Tammy Dickinson	National Research Council	Tdickins@nas.edu
Tony de Souza	National Research Council	Adesouza@nas.edu
Dennis L. Nielson	DOSECC	Denielson@egi.utah.edu
Ed Donovan	Chevron	Ewdo@chevron.com
Donna Shafer	C&M Storage	Cmstore@cvtv.net
Robert Shafer	C&M Storage	Cmstore@cvtv.net
Tom R. Moore	Phillips Petroleum	Trmoore@ppco.com
Bob Graebner	AGI	Bob.graebner@beg.utexas.edu
Christopher Keane	AGI	Keane@agiweb.org
Marcus Milling	AGI	Mmilling@agiweb.org
Robert Lemmon	DOE	Blemmon@ntpo.doe.gov
Mike Dix	Westport Technology	Mike.dix@westport1.com
Susan Cisco	Texas Railroad Commission	Susan.cisco@rrc.state.tx.us
Sigrid Clift	Texas Bureau of Economic Geol.	Sigrid.clift@beg.utexas.edu
George Bush	Texas Bureau of Economic Geol.	George.bush@beg.utexas.edu
Bob Brown	Chevron	Brrp@chevron.com
Robert Merrill	Unocal	Rmerrill@unocal.com