

**NATIONAL GEOSCIENCE  
DATA REPOSITORY SYSTEM**

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

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

**2<sup>nd</sup> Half FY02  
April 2002-October 2002**

**Submitted by the**

**AMERICAN GEOLOGICAL INSTITUTE**

**to the**

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

**October 2002**

*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.

## **ABSTRACT**

The NGDRS has facilitated 85% of cores, cuttings, and other data identified available for transfer to the public sector. Over 12 million linear feet of cores and cuttings, in addition to large numbers of paleontological samples and are now available for public use. To date, with industry contributions for program operations and data transfers, the NGDRS project has realized a 6.5 to 1 return on investment to Department of Energy funds.

Large-scale transfers of seismic data have been evaluated, but based on the recommendation of the NGDRS steering committee, cores have been given priority because of the vast scale of the seismic data problem relative to the available funding. The rapidly changing industry conditions have required that the primary core and cuttings preservation strategy evolve as well. Additionally, the NGDRS clearinghouse is evaluating the viability of transferring seismic data covering the western shelf of the Florida Gulf Coast.

AGI remains actively involved in working to realize the vision of the National Research Council's report of geoscience data preservation. GeoTrek has been ported to Linux and MySQL, ensuring a purely open-source version of the software. This effort is key in ensuring long-term viability of the software so that it can continue basic operation regardless of specific funding levels. Work has commenced on a major revision of GeoTrek, using the open-source MapServer project and its related MapScript language. This effort will address a number of key technology issues that appear to be rising for 2002, including the discontinuation of the use of Java in future Microsoft operating systems. Discussions have been held regarding establishing potential new public data repositories, with hope for final determination in 2002.

## Table of Contents

Table of Contents.....	4
List of Graphical Material.....	4
Introduction .....	5
Executive Summary.....	6
Experimental Approaches .....	7
Results and Discussion .....	8
National Academy of Science Geoscience Data Preservation Panel .....	8
NGDRS Steering Committee .....	9
Discussions with DOSECC.....	10
Data Transfer Status .....	11
Total Priority Data Transfers Relative to Phase I Targets.....	11
Assessing the Return on Investment of the NGDRS.....	12
Legacy Log Library Well Logs.....	13
Unocal/Spirit Energy Cores and Cuttings.....	13
Unocal/Spirit Energy Utah Core Transfer .....	14
Chevron Cores and Cuttings .....	14
Altura Midland Core Facility.....	14
Phillips Seismic Tapes .....	14
Marathon Oil Cores and Cuttings.....	15
Texaco/Chevron Midland Proposal.....	15
Houston Geoscience Research Center Contribution .....	15
Operating the Metadata Catalog.....	15
Redesign of GeoTrek.....	16
International Geoscience Data Repository Meeting.....	16
Conclusions .....	18
References .....	19

## List of Graphical Material

Figure 1. Financial Assessment of the NGDRS	11
---	----

## Introduction

All scientific and technical investigations require access to basic fundamental data. The capture and long-term preservation of data are required to address a wide range of scientific issues. The National Research Council released a report entitled *Preserving Scientific Data on Our Physical Universe (1995)* that took a broad look at the challenges of scientific data preservation and management in Federal Agencies. The report concluded "a general problem prevalent among all scientific disciplines is the low priority attached to data management and preservation by most agencies. Experience indicates that new research projects tend to get much more attention than the handling of data from old ones, even though the payoff from optimal utilization of existing data may be greater." No discipline is in greater need of an increased focus on data preservation than the geosciences, where private-sector downsizing and public-sector budgetary constraints have combined to jeopardize vast quantities of valuable geoscientific data critical to our understanding of the Earth's environment and natural resources.

The American Geological Institute's (AGI) National Geoscience Data Repository System (NGDRS) project was initiated in the face of the fact that billions of dollars worth of domestic geoscience data is in jeopardy of being irrevocably lost or destroyed as a consequence of the ongoing downsizing of the U.S. energy and minerals industry. Preservation and access to domestic geological and geophysical data are critical to the energy security and economic prosperity of our nation. The goal of the project is to act before valuable data are permanently displaced.

The NGDRS will serve as an important and valuable source of information for the entire geoscience community and the nation at large for a variety of applications, including environmental protection, water resource management, global change studies, and basic and applied research. It will also contain critical data that enable domestic energy and minerals companies to enhance their exploration and production programs in the United States for improved recovery of domestic oil, gas, and mineral resources.

A model for transferring data from the private to public sector is provided by the 1995 transfer of Shell Oil's core facility in Midland, Texas to the University of Texas at Austin. Shell deeded its collection of 2.2 million linear feet of core and cuttings from some 39 states to the university along with its warehouse and a \$1.3 million endowment to cover annual operating expenses. All of these data entered the public domain for the first time.

The 2002 report from the National Research Council entitled *Geoscience Data and Collections – National Resource at Peril* built upon the work done by AGI in the NGDRS program and reaffirms the critical nature of the problem and the distinct benefits to the science and society from active preservation of geoscience data.

## **Executive Summary**

With increased oil prices in 2000 and early 2001, the NGDRS has seen a distinctive increase in activity and interest. This increase in activity is the result of increased activities in the petroleum sector, including new funding to examine infrastructure issues facing many of the companies over the long-term. Over the past several years, the petroleum industry has been focused on short-term issues and cost-savings. However, with increased activities and continued industry consolidation, longer time horizons have reemerged.

Despite a lack of available public repository space, the NGDRS has pressed ahead in coordinating transfers both to existing facilities and to virtualize some transfers, whereby previously private data is made public. This has resulted in the NGDRS attaining 85% of the targeted cores and cuttings transfers, with over 14 million linear feet of cores and cuttings being transferred to in the public sector. Additionally, large-scale transfers of seismic data have been evaluated, but based on the recommendation of the NGDRS steering committee, cores have been given priority because of the vast scale of the seismic data problem relative to the available funding. The rapidly changing industry conditions have required that the primary core and cuttings preservation strategy evolve as well.

AGI remained actively involved in assisting the National Research Council with background materials and presentations for their panel convened to study the data preservation issue. The GeoTrek metadata catalog system continues to expand. Additionally, the metadata catalog is now served directly from AGI headquarters, and the software has been ported to Linux and all datasets transferred from Oracle to MySQL databases. Current revision efforts are being retargeted given the identification of new open source technologies which will provide for a flexible, user-friendly GeoTrek, Version 3. This includes adopting MapServer and MapScript for future versions, and examining user-interface issues given the downward pressure on the use of Java in web browsers with the introduction of Internet Explorer 6 and Windows XP.

## **Experimental Approaches**

The National Geoscience Data Repository System, Phase III is an operational project focused on coordinating and facilitating transfers of at-risk geoscience data from the private sector to the public domain. As such, the project does not have a consistent “experimental approach.” During the second half of FY02, no efforts undertaken required experimental approaches to arrive at specific conclusions.

## Results and Discussion

### *National Research Council Geoscience Data Preservation Panel*

The Board on Earth Sciences and Resources staff of the National Research Council (NRC) secured sufficient funding in the Fall 2000 to commence a study on the issue of geoscience data preservation. The study formally began in March 2001, with an initial meeting of the panelists on April 6, 2001. AGI worked closely with the NRC in developing its project scope:

*With budget cuts and the downsizing of the U.S. oil industry and some federal agencies, combined with the lack of space in private and public museums, the preservation of geoscience data (e.g., cores, cuttings, maps, paper reports, digital data) is becoming a critical issue for federal agencies, academic researchers, museums, institutes and industry. This study will (1) develop a strategy for determining what geoscience, paleontological, petrophysical and engineering data to preserve; (2) examine options for long-term archival of these data; (3) examine 3-5 accession and repository case studies as examples of successes and failures; and (4) distinguish the roles of the public and private sectors in data preservation. The overall goal of the study is to develop a comprehensive strategy for managing geoscience data in the United States.*

The NRC panel consisted of the following members:

Christopher Maples, University of Indiana (**Panel Chair**)

Beth Driver, National Imagery and Mapping Agency

Robert Schafer, Kinross Gold

Kevin Biddle, ExxonMobil

Robert Sneider, Sneider Exploration

Sally Zinke, Society for Exploration Geophysicists

Thomas R. Janecek, Florida State University

John Steinmetz, Indiana Geological Survey

Linda R. Musser, Penn State University

Warren Allmon, Paleontological Research Institute

Donald D. Clark, City of Long Beach

The NRC panel report was released in April 2002, entitled *Geoscience Data and Collections – National Resource at Peril*. The report builds on AGI's previous analyses of the issues facing the geosciences and society regarding data preservation. In particular, the NRC committee also recognized a severe shortage of adequate repository space and the importance of quality indexing and the need for the national metadata catalog. In particular, the NRC report recommends the funding and operation of at least three regional geoscience data repositories, built onto existing repositories if possible. Likewise, the report recognized that the major long-term costs in data preservation is the transport of cores, cuttings, and paper logs between facilities, and applying appropriate indexing to the data holdings. We recognize that the NGDRS program to date has been addressing most of the needs noted in the NRC report, and our steering committee has been evaluating the appropriate priorities of the effort relative the report's recommendations.



The Board on Earth Sciences and Resources has raised support for this study from various agencies and private organizations, including the US Department of Energy, US Geological Survey, National Science Foundation, Smithsonian, AAPG, POSC, and AGI.

### ***NGDRS Steering Committee***

AGI hosted a meeting of the NGDRS Steering Committee on October 2, 2002, along with other geoscience community representatives interested in geoscience preservation, in Houston, Texas at the new Texas Bureau of Economic Geology Geoscience Research Center. The primary focus of this meeting was to identify the key next steps in the NGDRS program and how it needs to focus in the future with regards to the results of the NRC report.

Fifty people attended the meeting, representing all facets of the geoscience community. The key actions identified from this meeting was the establishment of task groups to focus on engaging industry, government, and academia in the data preservation process. Additionally a task group was identified to work on issues of GeoTrek and indexing issues. Additionally, the group agreed to continue their work during a data preservation meeting at the Geological Society of American Annual Meeting in Denver on October 26, 2002.

### **NGDRS Steering Committee Meeting Attendees**

Allen Britton	CoreLab
Allen Hittelman	NOAA
Bill Kempner	ChevronTexaco
Bill Krebs	Energy and Geoscience Institute
Camille Lupton	OmniLaboratories
Cathy Allen	Railroad Commission of Texas
Charles Warrner	ExxonMobil
Charles Weiner	Texas Crude
Christopher Keane	AGI
Christopher Sorlien	Univ. of California - Santa Barbara
Dan Francis	Cal State Long Beach
David Applegate	AGI
David Archer	POSC
David Bouguet	ExxonMobil
David Lehman	ExxonMobil
Don Clarke	City of Long Beach
George Dellagiarino	MMS
Greg Wahlman	BP America
James Parr	Anadarko Petroleum
Jan Van Sant	AGI Foundation
Jim Handschy	ConocoPhillips
John Firth	Ocean Drilling Program
John Steinmetz	Indiana Geological Survey
Larry Knauer	ChevronTexaco
Larry Woodfork	AGI Past President
Linda Gunderson	US Geological Survey
Marcus Milling	AGI
Mike Padgett	EEX
Mike Partridge	Railroad Commission of Texas

Paul Cutler	National Research Council
Richard Bajura	National Research Center for Coal and Energy
Richard Lahann	ConocoPhillips
Richard Lane	National Science Foundation
Rob Finley	Illinois Geological Survey
Robert Graebner	AGI
Robert Merrill	Samson
Robert Schafer	Cattle River Minerals
Robert Shafer	C&M Storage
Robert Sneider	Sneider Oil and Gas
Scott Tinker	Texas BEG
Sherilyn Williams-Stroud	ChevronTexaco
Shirley Dutton	Texas BEG
Skip Rhodes	El Paso Production
Susan Longacre	ChevronTexaco
Tammy Dickinson	National Research Council
Tony Walton	University of Kansas/DOSECC
Walter Snyder	National Science Foundation
Wayne Ahr	Texas A&M
William Lawson	Dept. of Energy

The current actions planned for the task groups have been outlined by the Metadata Catalog group. During the Digital Mapping Technology meeting at the US Geological Survey in the Spring of 2003, the NGDRS Metadata Catalog task group will meet with representatives of the State Geological Surveys, the USGS, and data repository managers to initiate standardization of minimum metadata requirements for cores, cuttings, logs, and seismic data. Additionally, discussions will also be held to propose a standard citation format for use of geoscience data in peer-reviewed journal articles. It will be the responsibility of this group to subsequently work with the AGI and the various geoscience societies to implement and foster the use of data citation.

### ***Discussions with DOSECC***

Discussions with DOSECC began in the second quarter of 1999. DOSECC is a consortium of 48 universities and research laboratories that 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 mobile floating drill rig for coring of lake bottoms, such as the Great Salt Lake. DOSECC recognizes the long term scientific 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, DOSECC is leasing a smaller lot within the same property now for storing cores and equipment in sea containers. Depending on the success of commercializing Toole, DOSECC and AGI are keeping the option of a facility in Utah open.

### ***Data Transfer Status***

#### **NGDRS Cores and Cuttings Transfers**

<b>Source</b>	<b>Liner Ft. of Cores &amp; cuttings</b>
Unocal	1,109,016
Chevron (Cores)	934,157
Chevron (Cuttings)	10,038,898
Shell	1,350,000
BP/Amoco	1,200,000
Altura	255,000
<b>Total To Date</b>	<b>14,887,071</b>

#### **Recent NGDRS Well Log Transfers**

<b>Source</b>	<b>Well Logs</b>
Legacy Well Log Library	180,000

#### **NGDRS Paleontological Transfers**

<b>Source</b>	<b>Section Equivalent (ft)</b>
Chevron	43,200,000

#### **NGDRS Seismic Data Transfers**

<b>Source</b>	<b>Line-miles</b>
Phillips	2000

#### **Public Data Integrated into the NGDRS**

<b>Source</b>	<b>Boxes/Logs</b>
Texas BEG cores	100,656
Texas BEG well logs	87,772
Texas RRC logs	552,524
Alabama cores	1,091
Oklahoma cores	4,604
MMS logs	44,455
US Geological Survey cores	370,000
<b>Total</b>	<b>1,161,102</b>

### ***Total Priority Data Transfers Relative to Phase I Targets***

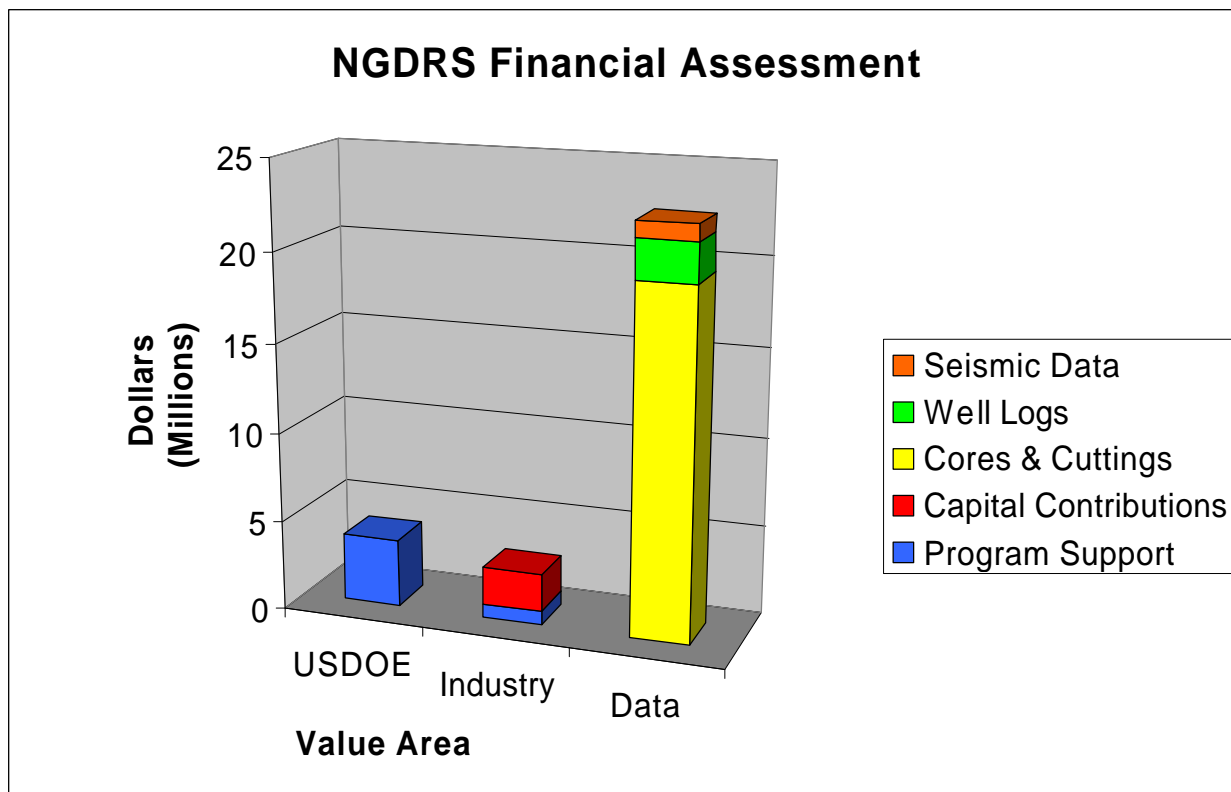
To date for the NGDRS project, nearly three-quarters of the volume of identified at-risk cores and cuttings have been transferred into the public domain. The NGDRS steering committee established priority to cores and cutting data given their particular economic risk, and thus to date, that data type has been the primary transfer target. A test case regarding seismic transfers was performed, but given the vast volumes of data that needs to be transferred and converted, current funding levels preclude major initiatives into that area.

<b>Data Type</b>	<b>Phase I Target</b>	<b>Phase III Transfers</b>	<b>Percent Completion</b>
Cores & Cuttings	17.5 M liner feet	14.9 M liner feet	85%
Seismic Data	100 M line-miles	961 line-miles	<1%
Paleo Data	Not quantified	43.2 M section ft.	>100%

### ***Assessing the Return on Investment of the NGDRS***

The American Geological Institute has begun a preliminary assessment of the financial dynamics of geoscience data, including the value of contributions, either in data or financial support, by the various stakeholders in the NGDRS. This effort is greatly facilitated by the recent survey of minimum total available geoscience data conducted by the National Research Council Committee of the Preservation of Geoscience Data and Collections.

<b>Data Type</b>	<b>Min. Total Volume</b>	<b>At-Risk Data Volume</b>	<b>Data Transferred To Date</b>	<b>Total National Value of Data</b>	<b>Value of Transferred Data</b>	<b>Percent of At-Risk Data Transferred</b>	<b>Percentage of Total Data Transferred</b>
Cores & Cuttings	55M ft	17.5 M ft	14.9M ft	\$73M	\$19.9M	85%	27%
Seismic Data	357M line-miles	100M line-miles	961 line-miles	\$3.57B	\$960,000	<1%	<1%
Well Logs	46M logs	7.1M logs	530,000 logs	\$184M	\$2.1M	7.5%	1%



**Figure 1. Financial Assessment of the NGDRS**

To date, the National Geoscience Data Repository System has received \$3.8 million in financial support for the US Department of Energy, \$2.87 million in financial support from industry, and \$22.3 million in data contributions. Additionally, financial and facility contributions to the Texas Bureau of Economic Geology of \$14 million in 1996 and 2002 transfers from Shell and BP also represents a derivative of the NGDRS project. The total cost-share investment by industry of \$25.2 million represents over a 6.5 to 1 return on investment by the US Department of Energy, even without attempting to place future use value on the data or attributing the University of Texas/Shell Oil transaction.

#### ***Legacy Log Library Well Logs***

The Legacy Log Library has contributed the well log records for the Texas Railroad Commission Districts 1-6 to the NGDRS and are now being integrated into the well log records of the Texas Bureau of Economic Geology. As these logs are integrated into the Bureau metadata catalog, they will be made available via GeoTrek.

#### ***Unocal/Spirit Energy Cores and Cuttings***

Unocal's onshore cores and cuttings are now cataloged in the NGDRS metadata catalog. The metadata is house at the American Geological Institute in its database servers. This transfer covered 2082 core records, representing 1,109,016 linear feet of core from across the nation.

Quality control by AGI allowed the inclusion of 1198 core records into the metadata catalog. Additional work on the data has determined the geolocating of the additional 884 core records is not possible given the existing metadata. However, the data is included in the system for those queries not dependent on geographic location. Unocal continues to store their core and cutting holdings 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.

### ***Unocal/Spirit Energy Utah Core Transfer***

The NGDRS is assisting in the transfer of Unocal's Utah cores and cuttings from Shulenburg, 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, need to be listed in GeoTrek. In March, the Utah Geological Survey sent a copy of its metadata to the American Geological Institute for review for integration into GeoTrek. The review is ongoing and integration of the Unocal Utah cores is expected in 2001.

### ***Chevron Cores and Cuttings***

Chevron transferred its metadata catalog of over 180,000 core and cuttings records to the NGDRS. These records represent 934,157 feet of cores, over 10M feet of cuttings, 14M washed paleo sample bags, 41,942 paleontology slides, and 56,621 oil samples.

Similar to the arrangement by Unocal, Chevron is maintaining the cores and cuttings at the C&M Storage facility. However, all of the cores and cuttings in the database are now released to the NGDRS for inclusion in the metadata catalog. The metadata records of the Chevron cores and cuttings are undergoing quality control at this time. The data is not extensively geocoded, so translations from TRS coordinates to latitude and longitude need to be performed. Full integration of the data into the metadata catalog, including geolocation of the records will occur in 2001.

### ***Altura Midland Core Facility***

Altura has transferred ownership of some 85,000 boxes of core and cuttings to the Bureau of Economic Geology at the University of Texas at Austin in 2000. The construction of a new repository in Midland was completed and physical movement of core and cuttings boxes began in October 2000. The metadata records for the Altura core was processed and integrated into the BEG's metadata catalog under direction of the NGDRS. The consolidated BEG catalog will be reintegrated into GeoTrek.

### ***Phillips Seismic Tapes***

Phillips Petroleum has transferred selected seismic holdings for the Santa Barbara Channel in California to the NGDRS. AGI completed a pilot project to evaluate the feasibility and costs for digitizing and transcribing the analog data to current media and format. The data was stored on 1-inch analog tapes, for which there are few known working readers. A selected number of tapes,

representative of the Santa Barbara channel were transcribed and processed to check for validity. The processed seismic lines demonstrated excellent quality and provide a new set of data for the geoscience community to use. Copies of the tapes are available on request for private sector and academic researchers..

### ***Marathon Oil Cores and Cuttings***

Marathon Oil approached AGI concerning the contribution of their cores and cuttings to the NGDRS from the Littleton, CO facility. That facility has been slated for closure, and their Permian Basin cores and cuttings were also in danger of being discarded. After initial discussions with Marathon representatives, the company decided to contract with C&M Storage to hold and maintain their data holdings in Schulenberg, TX. Discussions are ongoing regarding the incorporation of non-proprietary holdings of Marathon into GeoTrek using similar arrangements as those with Chevron and Unocal. Approximately 100,000 boxes of core are at issue in these discussions.

### ***Texaco/Chevron Midland Proposal***

Texaco's Midland operations had approached AGI concerning the development of a Permian Basin-wide core facility, including identification of non-proprietary cores and cuttings for release to the public. Following a meeting in October 2000, both Texaco and Chevron indicated a strong desire to move forward expeditiously with this process. However, with the merger between Texaco and Chevron no further activity has occurred on this effort. At this time, AGI does not expect any activity regarding this issue in the near future.

### ***Houston Geoscience Research Center Contribution***

BP has made a contribution of the former Amoco core repository on West Little York Road in Houston, Texas to the Texas Bureau of Economic Geology. With the transfer of the facility, the Texas BEG has also been provided additional land, rights to the proceeds from adjacent land sales, an initial endowment investment, along with 400,000 boxes of non-proprietary cores and cuttings previously held at that facility. With the addition of this substantial volume of new space, the severe space shortage has been temporarily relieved within the Gulf Coast region, and in particular, within Texas. BP is in similar negotiations regarding other former Amoco facilities which may further alleviate the immediate space issues for core storage.

### ***Operating the Metadata Catalog***

The operation of the metadata catalog continued during the second half of 2002. The following databases are currently available on the metadata catalog:

- 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
- Unocal Onshore Cores and Cuttings

- Chevron Onshore Cores, Cuttings, and Paleontological samples
- US Geological Survey Cores and Cuttings

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

	<b>1998</b>	<b>1999</b>	<b>2000</b>	<b>2001</b>	<b>YTD 2002</b>	<b>Total</b>
<b>NGDRS Website Hits</b>	30,911	61,152	48,656	64,746	69,018	274,483
<b>Unique Visitors to NGDRS</b>	1,331	4,336	6,218	7,900	4,805	24,590

### ***Redesign of GeoTrek***

GeoTrek, the metadata catalog for the NGDRS was redesigned in 2001. This redesign revolved around 2 areas: backend remote database integration and an improved user interface.

The back-end remote database change was completed and implemented with the integration of the Unocal cores. The system now is capable of querying remote database systems over the Internet, allowing data holders to control the availability and extent of data accessible through the metadata catalog. Additionally, it allows transparent integration of additional information, such as core disposition, core photos, analytical data, etc., to be readily accessible by the end-user.

Additionally, in August 2001, the entirety of the GeoTrek code was ported to Linux and interfaced with a number of Open Source database systems. This effort was targeted to ensure the greatest viability of the metadata catalog regardless of funding levels. This has allowed GeoTrek to now be run on lower cost servers and against a wider range of databases, including several without annual service licenses. It is expected that these arrangements will assist in bringing additional public data holders into the NGDRS fold.

A new user interface system has been identified in the Open Source community that will neatly tie into the GeoTrek metadata catalog. A concerted effort will be engaged to integrate this widely used interface, which has the benefit of being well-tested on a number of browser technologies, into the end-users GeoTrek experience.

A major obstacle facing GeoTrek over the next year results from the fact that Microsoft no longer supports Java “out-of-the-box” with its web browsers and operating systems. Given the reliance of GeoTrek’s front-end to Java, as well as MapServer’s most interactive front-end on Java, concern is raised about the need to explore other strategies for user interfaces. In general, AGI expects the lack of default Java support to be a major hurdle within the next 12 months as a number of users do not have dedicated IT staffs to address these issues.

### ***International Geoscience Data Repository Meeting***

The National Geoscience Data Repository System was represented during the fourth meeting of the National Geoscience Data Repositories, held in Stavanger, Norway. This meeting provided an opportunity for representatives of the various National Data Repositories to meet and discuss issues facing them, strategies for data management and funding, and to assess differences in data preservation policies around the world. A summary report of the meeting was published in the June 2002 issue of *Geotimes*.



The fundamental issue facing all National Data Repositories is the issue of valid and high quality metadata. Metadata appears to be a universal problem with most data providers not documenting their data acquisitions adequately, leading to the disposal of a large amount of data throughout the world. However, most countries appear to have systems in place to encourage metadata improvement, either through legislative threat or incentives by the NDR providing enhanced data management to the acquiring entities.

Universally, National Data Repositories are focusing on well logs and seismic data, in particular the digitization of well logs to reduce paper costs is the single largest effort occurring throughout the world. Cores and cuttings are also of high interest, and generally seen as the detailed data which the well logs will lead researchers to. Most NDR programs are funded either directly or indirectly from royalties and fees collected from petroleum production, while providing time-limited confidentiality of selected data and free access to public data for educational purposes.

## Conclusions

The NGDRS has seen a distinctive increase in activity and interest, particularly since October 2000. This new spike in activity is the result of increased activities in the petroleum sector, including new funding to examine infrastructure issues facing many of the companies over the long-term. Over the past several years, the petroleum industry has been focused on short-term issues and cost-savings. However, with increased activities and continued industry consolidation, longer time horizons have reemerged.

Despite a lack of available public repository space, the NGDRS has press ahead in coordinating transfers both to existing facilities and to virtualize some transfers, whereby previously private data is made public. This has resulted in the NGDRS realizing transfer of 85% of targeted cores and cuttings, with over 14M linear feet of cores and cuttings now in the public sector. Additionally, large-scale transfers of seismic data have been evaluated, but bases on the recommendation of the NGDRS steering committee, cores have been given priority because of the vast scale of the seismic data problem relative to the available funding. The rapidly changing industry conditions have required that the primary core and cuttings preservation strategy evolve as well. To date, the NGDRS project has realized a 6.5 to 1 return on investment of Department of Energy funds.

AGI remains actively involved in realizing the recommendations of the National Research Council regarding the preservation of geoscience data. The GeoTrek metadata catalog system continues to operate. Porting of GeoTrek to Linux and MySQL was completed in August 2001 and now efforts are underway to examine the use of MapServer and other open source software to further enhance the metadata catalog.

## **References**

None applicable.