



Oklahoma Geological Survey  
THE UNIVERSITY OF OKLAHOMA  
MEWBOURNE COLLEGE OF EARTH & ENERGY

July 3, 2012

**U. S. Geological Survey**  
**Office of Acquisition and Grants**  
Betty Adrian, NNGDPP Grant Program Manager

RE: Grant #DOI – USGS G11AP20183, titled “Preparation of Metadata for the Oklahoma Mud Log Collection at the Oklahoma Geological Survey.”

Dear Ms. Adrian:

Attached is a copy of the Final Technical Report for Grant No. DOI – USGS G11AP20183. This report fulfills one of the deliverable requirements under the referenced grant.

We very much appreciate the support the U. S. Geological Survey has provided us for conducting this project. If you have any questions or need additional information, please let us know.

Thank you.

Sincerely,

G. Randy Keller  
Director  
Oklahoma Geological Survey

Jane Weber  
Database Manager  
Oklahoma Geological Survey

Attachment

c: Margaret Eastman, USGS Contracting Officer  
Leslie Flenniken, OU Grants Specialist

**PREPARATION OF METADATA  
FOR THE OKLAHOMA MUD LOG COLLECTION  
AT THE OKLAHOMA GEOLOGICAL SURVEY**

**AWARD NO. USGS G11AP20183**

**FINAL TECHNICAL REPORT**

Co-Principal Investigators:

G. Randy Keller, OGS	Email: <a href="mailto:grkeller@ou.edu">grkeller@ou.edu</a>	405-325-7968
Jane L. Weber, OGS	Email: <a href="mailto:jlweber@ou.edu">jlweber@ou.edu</a>	405-325-5830

Oklahoma Geological Survey, University of Oklahoma  
100 E. Boyd Street, Room N-131  
Norman, Oklahoma 73019-0628  
Phone: 405-325-3031  
Fax: 405-325-7069

Grant Initiation Date: August 1, 2011  
Grant Completion Date: July 31, 2012  
Submittal Date of Final Technical Report: July 3, 2012

## TABLE OF CONTENTS

<b>TABLE OF CONTENTS.....</b>	<b>2</b>
<b>ABSTRACT .....</b>	<b>3</b>
<b>INTRODUCTION.....</b>	<b>4</b>
<b>APPROACH TO WORK .....</b>	<b>5</b>
<b>Cataloging Logs.....</b>	<b>5</b>
<b>Mapping Data to Metadata Elements.....</b>	<b>5</b>
<b>Figure 1. Mapping OGS Sample Data to NDC Metadata             Elements .....</b>	<b>6</b>
<b>Figure 2. Example of Excel Metadata Template File.....</b>	<b>6</b>
<b>Figure 3. Example of CSV File Output .....</b>	<b>7</b>
<b>Uploading Metadata to National Digital Catalog .....</b>	<b>7</b>
<b>ACCOMPLISHMENTS RELATIVE TO PROJECT GOAL.....</b>	<b>9</b>

# PREPARATION OF METADATA FOR THE OKLAHOMA MUD LOG COLLECTION AT THE OKLAHOMA GEOLOGICAL SURVEY

## ABSTRACT

The Oklahoma Geological Survey completed the second year of a two-year project to build a database of Oklahoma mud logs maintained among its collections and to prepare and submit sample-specific metadata for those logs to the National Digital Catalog. FY2010 efforts, which concentrated on the northeast, southeast, southwest, and panhandle sectors of the state, had produced 880 metadata records. In FY2011, 202 additional records for those regions plus 2003 records for the northwest region were processed. A total of 3085 Oklahoma mud logs have been submitted for inclusion in the National Catalog.

The methodology used to catalog and describe the logs and then create metadata from that information followed the basic protocol originally developed for our FY2009 National Geological and Geophysical Data Preservation Program (NGGDPP) grant and utilized for our FY2010 grant. Data from logs were entered manually into an Excel spreadsheet. Recognized sources of Oklahoma well information were consulted to supply additional or missing information. Columnar data from the Excel spreadsheet were combined and mapped to the 7 required metadata elements plus 3 optional elements of the NGGDPP metadata template file. The resulting metadata were then extracted to a Comma Separated Value (CSV) file format.

Initial attempts to upload the CSV file via the NGGDPP web interface encountered problems related to maintenance changes in progress. With assistance from Rick Brown in Rolla, we eventually uploaded the CSV file through <http://my.usgs.gov/csc/nggdpp/upload>. The validation process generated one record error. Nevertheless, a request was made to load the file in the Catalog.

# **PREPARATION OF METADATA FOR THE OKLAHOMA MUD LOG COLLECTION AT THE OKLAHOMA GEOLOGICAL SURVEY**

## ***INTRODUCTION***

In FY2007 the Oklahoma Geological Survey (OGS) identified 26 geological and geophysical data collections to enter in the National Digital Catalog as part of the National Geological and Geophysical Data Preservation Program (NGGDPP). A long-term goal of OGS is to catalog and/or computerize these data sets for the benefit of the public. About half of the collections have been individually cataloged. Others are in the process of being sorted and organized, while the remaining collections are not yet in a form easily used by the public.

One OGS collection of particular use to the petroleum industry is mud logs, sometimes referred to as gas logs, gas detector logs, or drillers' logs. With FY2010 NGGDPP funding, OGS examined and sorted all the mud logs in its collection at that time and then developed metadata for a portion of them. For FY2011 OGS finished both the cataloging and creating metadata for all its mud logs. Starting with a sorted set of logs this year, OGS applied the same metadata creation procedure used for the FY2010 NGGDPP project, which had mirrored the methodology originally developed for our FY2009 grant for rock core samples. The mud log project was spread over two years due to the amount of work involved and the limited number of personnel available.

## **APPROACH TO WORK**

**Cataloging logs:** The FY2011 project was a continuation of work performed under our FY2010 grant. For FY2010 we developed site-specific metadata for 880 mud logs from the northeast, southeast, southwest, and panhandle regions of Oklahoma. In FY2011 we processed mud logs from the northwest sector plus logs from other sectors that had been added to the collection after the FY2010 metadata file was submitted.

In brief, the cataloging procedure consisted of:

1. Xerox front page of log to produce convenient sheet containing well header information
2. Enter well header data into existing Excel spreadsheet database containing 28 fields designed to accommodate OGS needs
3. Acquire and fill in missing data elements
  - a. For depth values inspect original log trace
  - b. For API number, elevation, quarter section location, date, latitude and/or longitude, seek information from the Natural Resources Information System (NRIS) database (originally built by OGS but now at Oil-Law) or IHS Energy commercial database. If only spatial coordinates are required, use the Spatial Calculator at the University of Oklahoma's Center for Spatial Analysis (CSA). NOTE: NRIS- and CSA-derived latitude and longitude values are based on the NAD83 datum standard and Topographic Mapping Company's landgrid. IHS Energy values and those taken directly from mud logs may be based on a different datum and/or landgrid.
4. Perform a quality-control check of the data by comparing County Name vs. API Number and making sure latitude and longitude coordinates fall in the correct sector of Oklahoma.

**Mapping data to metadata elements:** Since a deliverable for the project was to be a CSV file of metadata, before work began the default List Separator on the work computer was changed from a comma (,) to the pipe character (|). This precluded possible problems arising from commas used in descriptive fields.

Mapping mud log sample properties listed in the OGS database to metadata elements required for the National Digital Catalog followed the match-ups outlined last year as shown in Figure 1. This approach resulted in populating not only the 7 required metadata elements but also 3 of the optional elements.

To combine more than one field of sample data into one metadata element, formulaic script was written for the first sample record and copied to all subsequent records. The resulting column of metadata was then copied but only its "values" were pasted into the Excel metadata template file. An example of the first two records of a completed metadata template file is shown in Figure 2, divided at the end of the line for display purposes. The first record contains the names of the metadata elements, required and

optional; the second record shows the OGS sample information corresponding to each of those metadata elements.

**Figure 1. Mapping OGS Sample Data to NDC Metadata Elements**

<b><u>NDC Metadata Element</u></b>	<b><u>OGS Database Fields Used</u></b>
<i>collectionID</i> (req.)	(Assigned by NGGDPP. Same for all records.)
<i>title</i> (req.)	Lease, Well Number, Operator
<i>abstract</i> (req.)	Top, Bottom, Elevation, County
<i>dataType</i> (req.)	(Controlled by NGGDPP. Same for all records.)
<i>supplementalInformation</i> (req.)	Not derived from dataset. Same for all records.
<i>coordinates</i> (req.)	Latitude, Longitude
<i>datasetReferenceDate</i> (req.)	Log Start (Spud if Log Start not given)
<i>alternateTitle</i> (opt.)	API Number
<i>alternateGeometry</i> (opt.)	Township Number and Direction, Range Number and Direction, Section Name of meridian included to prevent confusion between samples east of Indian Meridian and those east of Cimarron Meridian
<i>verticalExtent</i> (opt.)	Top, Bottom Also included under <i>abstract</i> to enhance description of resource

**Figure 2. Example of Excel Metadata Template File**

<b>collectionID</b>	<b>title</b>	<b>abstract</b>	<b>alternateTitle</b>	<b>verticalExtent</b>	<b>coordinates</b>
1088928	Mud log from Well: SHUMAN FARM 1 Operator: IREX CORP	This mud log sample shows drilling rate and lithology from 4500 ft to 6600 ft at an elevation of 1982 ft in HARPER County, OKLAHOMA.	API Number: 3505921074	ft, 6600, 4500	-99.853957, 36.949249

<b>alternateGeometry</b>	<b>supplementalInformation</b>	<b>datasetReferenceDate</b>	<b>dataType</b>
Public Land Survey System TOWNSHIP 29 NORTH RANGE 25 WEST of the Indian Meridian SECTION 36	Contact the manager of the Oklahoma Geological Survey's Oklahoma Petroleum Information Center at 405-325-3031 to access the sample material free of charge. Copying fees apply. Additional information is available at <a href="http://www.ogs.ou.edu/">http://www.ogs.ou.edu/</a> .	19820102	Lithology Log

The completed Excel metadata template file was saved as a CSV file. Figure 3 shows an example of output from a CSV file, using the same two records chosen for the illustration in Figure 2. Note the use of the pipe symbol (|) as a delimiter to separate metadata elements

### Figure 3. Example of CSV File Output

collectionID|title|abstract|alternateTitle|verticalExtent|coordinates|alternateGeometry|supplementalInformation|datasetReferenceDate|dataType  
 |Mud log from Well: **SHUMAN FARM 1** Operator: **IREX CORP**|This mud log sample shows drilling rate and lithology from **4500** ft to **6600** ft at an elevation of **1982** ft in **HARPER** County, OKLAHOMA.|API Number: **3505921074**|ft, **6600, 4500**|-**99.853957, 36.949249** |Public Land Survey System TOWNSHIP **29 NORTH** RANGE **25 WEST** of the Indian Meridian SECTION **36**|Contact the manager of the Oklahoma Geological Survey's Oklahoma Petroleum Information Center at 405-325-3031 to access the sample material free of charge. Copying fees apply. Additional information is available at <http://www.ogs.ou.edu/>.|**19820102**|Lithology Log

*Note:* Certain items are bolded here only to highlight which information was derived from our database. They were not bolded in the CSV file.

**Uploading metadata to National Digital Catalog:** We encountered major difficulties when we first tried to upload the CSV sample-specific metadata file. The web interface at <http://ndc.sciencebase.gov> was “Down for maintenance” and the original interface at <http://my.usgs.gov/csc/nggdpp/upload> was in a state of transition. We were neither informed that changes were being made nor sent instructions as to what to do in the meantime. Only with help from Rick Brown in Rolla were we eventually able to upload the CSV file. At the time of upload, the changes being made were not intuitive, perhaps because the programming was not yet complete. Our file was intended as a replacement file, as it included all the mud log records submitted for FY2010 plus the records added this year. However, steps for how to achieve that action were not specified at the time of the upload.

The uploaded file contains:

	<u>From FY2010</u>	<u>New</u>	<u>Total</u>
Northeast OK records	337	61	398
Southeast OK records	68	20	88
Southwest OK records	113	8	121
Panhandle OK records	369	106	475
Northwest OK records	—	<u>2003</u>	<u>2003</u>
Sub-totals	887*	2198	
Grand TOTAL submitted in FY2011			<b><u>3085</u></b>

\* Includes 7 records excluded from FY2010 submission due to incomplete data at that time.

A total of 26 mud logs were excluded from the new records added due to missing or incomplete data:

No lithology given	14
No API Number found	10
No Elevation data found	2

The upload produced one validation error, apparently the result of an unwanted comma. It would be helpful if the validation process could somehow flag problem records – up to a maximum of 5, for example – when an error is detected. That would help the sender identify where a correction is needed.

In spite of the one error, we requested that the file be loaded into the Catalog. Later we verified the presence of our data at <http://myusgs.gov/csc/nggdpp/state/ok>.

## ***ACCOMPLISHMENTS RELATIVE TO PROJECT GOAL***

The primary goal of the project, to finish cataloging and creating sample-specific metadata for all Oklahoma mud logs in the OGS collection, was accomplished. More than 2000 logs from the NW sector of Oklahoma, plus 195 logs from other areas found recently in stored material, were processed. Built upon last year's submitted file of 880 records, an updated and expanded CSV metadata file containing 3085 records was uploaded as a replacement file to the National Digital Catalog. This constituted our main deliverable of the project.

A separate outcome of the project was completion of an Excel spreadsheet of Oklahoma mud logs suitable for publishing as a downloadable file on the OGS web site. This information enables interested persons to determine whether or not we have logs they need without having to make a trip to the Data Library.

Completing the project encompassed tasks from 3 Phases of our long-term data preservation approach: (a) process individual collections; (b) prepare electronic catalog of items in each collection; and (c) prepare metadata for each collection. With the added attention now given to the easily-accessible mud logs, we are able to look toward accomplishing the final Phase of data preservation — provide scanned images of individual items online. Efforts in that direction have begun, as we are asking users who scan the logs to provide us with copies of their digital images, thereby leveraging their efforts and resources to help us. We have also acquired a log scanner able to handle the narrower logs and plan to generate digital mud log images as time and personnel allow.

Finally, not to be overlooked as an accomplishment, was the chance to employ data handling and metadata preparation methodology we developed during a previous NCGDPP project, thereby using our limited resources more efficiently.