



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

July 22, 2011

U. S. Geological Survey
Office of Acquisition and Grants
Frances Pierce, NGGDPP Grant Program Manager

RE: Grant #DOI – USGS G10AP00120, titled “Development of Metadata for a Portion of Mud Logs at the Oklahoma Geological Survey.”

Dear Ms. Pierce:

Attached is a copy of the Final Technical Report for Grant No. DOI – USGS G10AP00120. 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

**DEVELOPMENT OF METADATA
FOR A PORTION OF MUD LOGS
AT THE OKLAHOMA GEOLOGICAL SURVEY**

AWARD NO. USGS G10AP00120

FINAL TECHNICAL REPORT

Co-Principal Investigators:

G. Randy Keller, OGS	Email: grkeller@ou.edu	405-325-7968
Jane L. Weber, OGS	Email: jlweber@ou.edu	405-325-7331

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, 2010
Grant Completion Date: July 31, 2011
Submittal Date of Final Technical Report: July 22, 2011

TABLE OF CONTENTS

TABLE OF CONTENTS.....	2
ABSTRACT	3
INTRODUCTION.....	4
APPROACH TO WORK	5
Organization of Logs.....	5
Cataloging of Logs	5
Mapping of Data to Metadata Elements	5
Example of Excel Metadata Template Record	6
Uploading of Metadata to National Digital Catalog	7
Example of CSV File Record	7
ACCOMPLISHMENTS RELATIVE TO PROJECT GOAL.....	8

DEVELOPMENT OF METADATA FOR A PORTION OF MUD LOGS AT THE OKLAHOMA GEOLOGICAL SURVEY

ABSTRACT

The Oklahoma Geological Survey (OGS) cataloged and created sample-specific metadata for 880 mud logs and submitted that metadata for inclusion in the National Digital Catalog. The mud logs had been filed in Township-Range-Section order in file cabinets but no card file or list existed of logs in the collection.

The first attempt to catalog the logs, starting with those filed under Township North-Range East of the Indian Meridian, revealed 1) many logs were misfiled; 2) material other than mud logs was mixed with the logs; and 3) there were multiple copies of many logs. The entire collection was then examined and reorganized, correcting filing errors, removing extraneous material, and culling replicates. Concentrating on mud logs from the northeast sector of Oklahoma, the logs were cataloged and described on an Excel spreadsheet.

Following an OGS procedure developed for a previous National Geological and Geophysical Data Preservation Program (NGGDPP) grant, metadata were created in an Excel template that contained all 7 elements required by the NGGDPP plus 3 elements listed as optional. The metadata file was populated by mapping mud log properties to one of the elements in the template using Copy and Paste Special functions. The resulting metadata were then extracted to a Comma Separated Value (CSV) file format using the pipe character (|) as a record delimiter.

The CSV file was uploaded through the NGGDPP web interface at <http://my.usgs.gov/csc/nggdpp/upload> , successfully validated on the first try, and loaded into the National Catalog

DEVELOPMENT OF METADATA FOR A PORTION OF MUD LOGS AT THE OKLAHOMA GEOLOGICAL SURVEY

INTRODUCTION

The Oklahoma Geological Survey (OGS) is responsible for maintaining an estimated several million items of geological and geophysical data related to the natural resources of Oklahoma. In FY2007, as part of the National Geological and Geophysical Data Preservation Program (NGGDPP), OGS identified 26 data collections to list in the National Digital Catalog(NDC). Only a few of those collections have been individually cataloged. Many are at some stage of being sorted and organized to enable manual or digital searches; the remaining ones are not yet in any useable form. A long-term goal of OGS is to catalog and/or computerize all its data items for the benefit of the public.

For its FY2010 NGGDPP grant, OGS proposed to catalog and prepare sample-specific metadata for a subset of mud logs and then submit that metadata for inclusion in the National Catalog. Mud logs are sometimes referred to as gas logs, gas detector logs or, for older versions, drillers' logs. The subset chosen was an estimated 600 logs from the northeastern sector of Oklahoma. That region is identified as having coordinates Township North and Range East of the Indian Meridian. There was no list, hardcopy or digital, or card file of the logs; but they were sorted somewhat by Township-Range-Section location in file cabinets. The methodology for creating the metadata, extracting a Comma Separated Value (CSV) file from an Excel spreadsheet, was used previously with our FY2009 NGGDPP grant for rock core samples. One of the purposes of that project was to serve as a prototype experience for metadata development for any subsequent well-related dataset.

APPROACH TO WORK

Organization of logs: The mud log collection was supposedly arranged in Township-Range-Section order in several file cabinets. However, it was soon discovered that 1) a substantial number of logs had been misfiled; 2) material other than mud logs, such as drill stem tests, electric logs, and core reports, was mixed with the mud logs; and 3) there was duplication of logs, with sometimes as many as 13 replicates present. It became apparent that the original plan to handle only mud logs from the northeast sector had to be revised. Instead, the entire collection first had to be examined and reorganized. Student help was especially useful in the refiling effort but not as adept at recognizing extraneous log material or replicate logs.

Cataloging of logs: The front page of each log was Xeroxed to produce a sheet of log header information. It was more convenient for data entry to read from individual sheets than from whole logs to build the basic database in an Excel spreadsheet. When depth information was absent in the header, it was obtained later by inspecting the original log trace. Several data elements normally used to characterize well-related data often did not appear anywhere on the mud log: API number, elevation, quarter section, date, latitude, and longitude. If missing, these data were sought from the Natural Resources Information System (NRIS) database at Oil-Law, IHS Energy data, or, if only spatial coordinates were required, the Spatial Calculator at the University of Oklahoma's Center for Spatial Analysis (CSA). NRIS- and CSA-derived latitude and longitude values are based on the NAD83 datum standard and Topographic Mapping Company's landgrid. It is recognized that IHS Energy values and those taken directly from mud log headers may be based on a different datum and/or landgrid.

The resulting database, containing 27 fields of information, was designed to accommodate OGS purposes. Data were quality-control checked by comparing County Name vs. API Number and making sure latitude and longitude coordinates fell in the appropriate sector of Oklahoma.

Mapping of data to metadata elements: The approach for this aspect of the project closely followed the one developed for our FY2009 NGGDPP grant. The product was to be a CSV file of metadata. To preclude possible problems associated with commas used in descriptive fields, the default comma record delimiter (List Separator) on the work computer was replaced with the pipe character (|) before work was begun.

The NGGDPP program named 7 required metadata elements and several optional elements. We decided which of our fields of sample properties to use in populating the 7 required elements plus 3 of the optional elements. The match-ups made were:

NDC Metadata Element**OGS Database Fields Used**

<i>collectionID</i> (req.)	(Assigned by NCGDPP. 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 NCGDPP. 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

To combine more than one column of sample data into one metadata element, a formula with 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. During the mapping process, 7 records were excluded due to incomplete data.

An example of the first two records of the completed metadata template file is shown here, divided at the end of the line for display purposes. The first record contains the names of the metadata elements, required and optional, corresponding to the data presented in the second record.

Example of Excel Metadata Template Record

collectionID	title	abstract	alternateTitle	verticalExtent	coordinates
1088928	Mud log from Well: STATE 1-22 Operator: VEGA ENERGY CO	This mud log sample shows drilling rate and lithology from 200 ft to 4200 ft at an elevation of 4663 ft in CIMARRON County, OKLAHOMA.	API Number: 3502520765	ft,4200,200	-102.935987, 36.800097

alternateGeometry	supplementalInformation	datasetReferenceDate	dataType
Public Land Survey System TOWNSHIP 04 NORTH RANGE 01 EAST of the Cimarron Meridian SECTION 22	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/ .	19870203	Lithology Log

Uploading of metadata to National Digital Catalog: After the desired samples were mapped to the Excel metadata template, the metadata file was saved as a CSV file. An example of the first two records of the CSV file is shown below. As described above, the first record contains the names of the metadata elements corresponding to the data presented in subsequent lines. Note that the pipe symbol (|) is being used as a delimiter to separate metadata elements. Bolded items were not bolded in the uploaded file but are bolded here to illustrate which information was derived from our database.

Example of CSV File Record

collectionID|title|abstract|alternateTitle|verticalExtent|coordinates|alternateGeometry|supplementalInformation|datasetReferenceDate|dataType
1088928|Mud log from Well: **STATE 1-22** Operator: **VEGA ENERGY CO**|This mud log sample shows drilling rate and lithology from **200** ft to **4200** ft at an elevation of **4663** ft in **CIMARRON** County, OKLAHOMA.|API Number: **3502520765**|ft,**4200,200**|-**102.935987, 36.800097**|Public Land Survey System TOWNSHIP **04 NORTH** RANGE **01 EAST** of the Cimarron Meridian SECTION **22**|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/>.| **19870203**|Lithology Log

We uploaded a CSV file containing 880 site-specific metadata records for mud logs through the interface at <http://my.usgs.gov/csc/nggdpp/upload>. This file included:

Northeast OK records	330
Southeast OK records	68
Southwest OK records	113
Panhandle OK records	369

Because the original collection was more disorganized than we had imagined, our estimate of 600 samples from the northeast was in error. Therefore, we added samples from other regions as time and resources permitted to make up for the shortfall. The upload produced no validation errors.

ACCOMPLISHMENTS RELATIVE TO PROJECT GOAL

The goal for this project, to provide sample-specific metadata to the National Digital Catalog while cataloging a subset of the mud log collection maintained by OGS, was achieved. In the process, the accessibility and accuracy of our mud log collection were greatly improved. Reorganizing the file led to removal of extraneous material and replicate logs; it also caused us to discover misplaced logs and to note and correct misleading typographical errors on the logs. The database of mud log header information begun will eventually be made available on our website, enabling interested persons to determine whether or not we have logs they need without having to make a trip to the OPIC facility.

A separate outcome of the project was that it successfully demonstrated the application and value of the data handling procedure developed for our FY2009 NNGDPP grant. A lot of time was saved by being able to follow that step-by-step methodology to create metadata for another data collection related to drill holes.

The proposed goal referred to creating metadata for about 600 samples, the number of northeastern Oklahoma mud logs initially estimated to be in the collection. When the actual number turned out to be far less, we included logs from other regions as we finished processing them. At that point our aim became to develop metadata for a minimum of 600 logs with the personnel we had and within the time frame of the project. We were able to exceed that number; we submitted metadata for a total of 880 sample records.