

GEOLOGICAL SURVEY OF ALABAMA

Berry H. (Nick) Tew, Jr.
State Geologist

GEOLOGIC INVESTIGATIONS PROGRAM

**INITIATION OF METADATA DEVELOPMENT
FOR GEOLOGICAL AND GEOPHYSICAL DATA
AT THE GEOLOGICAL SURVEY OF ALABAMA:
FINAL REPORT FOR FISCAL YEAR 2008**

Open-File Report 0913

by
W. Edward Osborne and Sandy M. Ebersole

Prepared in partial fulfillment of Grant Award No. 08HQGR0118 with the
U.S. Geological Survey

Tuscaloosa, Alabama
2009

CONTENTS

	Page
Introduction.....	1
Purpose and methods	2
Goals	2
Results	2
Discussion.....	4

TABLE

1. Comparison of the goals and results for the Fiscal Year 2008 project	3
---	---

**INITIATION OF METADATA DEVELOPMENT FOR GEOLOGICAL AND
GEOPHYSICAL DATA AT THE GEOLOGICAL SURVEY OF ALABAMA:
FINAL REPORT FOR FISCAL YEAR 2008**

BY W. EDWARD OSBORNE AND SANDY M. EBERSOLE,
PRINCIPAL INVESTIGATORS

INTRODUCTION

The Geological Survey of Alabama (GSA) maintains a core and sample storage warehouse in the Mary Harmon Bryant Special Collections Facility on the University of Alabama (UA) campus in Tuscaloosa, Alabama. The warehouse contains about 12,000 square feet of shelved space. GSA is closely affiliated with the State Oil and Gas Board of Alabama (OGB), which requires that cores and well cuttings be submitted and archived, and the samples are maintained by the GSA in its facility. Alabama is a major oil and gas producing state, ranking 11th nationally in natural gas production and 15th in liquid petroleum production. Consequently, GSA maintains an extensive collection of samples from oil and gas wells. The warehouse contains processed well cuttings from nearly 4,000 oil and gas wells and cores from more than 3,300 oil and gas wells. In addition, the facility houses about 450 industrial cores, 244 vibracores, nearly 2,800 sets of water well cuttings, and more than 1,500 coal samples. In total, more than 67,000 boxes of samples are stored in the GSA warehouse.

The GSA was founded in 1848 and, since then, GSA geologists have made extensive collections of Paleozoic, Mesozoic, and Cenozoic fossils from numerous world-class sites that exist in Alabama. More than 182,000 fossils occupy 69 cabinets and are housed in the basement of Walter B. Jones Hall (the main GSA building on the UA campus) as well as in the warehouse in the Special Collections Facility.

During Fiscal Year 2007, GSA conducted an inventory of its geological and geophysical collections for the National Geological and Geophysical Data Preservation Program (NGGDPP). The results of the inventory were entered into the online survey at the completion of the project. For Fiscal Year 2008, GSA initiated development of site-specific metadata records for individual samples in its collections. Because GSA collections are extensive, metadata preparation in Fiscal Year 2008 was limited to a significant part of the collections in the core and sample warehouse and to fossils of Paleozoic age in the paleontology collection. These metadata have been uploaded to the National Digital Catalog.

PURPOSE AND METHODS

The purpose of this project is to generate site-specific metadata records that describe, at the individual sample level, the contents of a significant part of the geological collections at the GSA. Where possible, the data were extracted from existing databases. For example, information on core and well cuttings taken from wells that were formally permitted by the OGB were extracted from the OGB well database, which was made available to GSA for this project. Construction of some records (for example, non-permitted industrial well samples) required compilation of information from both digital and paper sources. Completion of all records for washed oil and gas well cuttings, washed water well cuttings, and industrial core included a physical inventory of samples in the GSA warehouse. Construction of metadata records for specimens in the paleontology collection required examination of the labels accompanying the individual fossils in the cabinets of the collection.

GOALS

At the time the Fiscal Year 2008 proposal was submitted, the GSA had not completed the inventory of its collections. Therefore, it was not possible to provide an exact number of samples for which metadata were to be written. As a first approximation, we anticipated creating metadata for 4,000 sets of washed oil and gas well cuttings, 1,935 oil and gas well cores, 450 industrial minerals cores, and 2,700 sets of cuttings from water wells (table 1). Because the well sample washing program has been discontinued, GSA is currently discarding a large part of its collection of unwashed oil and gas well cuttings. Metadata needs be written for those sets of unprocessed cuttings that are retained, once the disposal process is complete. Unwashed well cuttings retained probably will not exceed 50 sets. In total, GSA proposed to generate metadata records for about 8,900 sets of samples currently stored in the GSA warehouse. GSA estimated its paleontology collection contains about 18,900 fossils of Paleozoic age, each of which requires a metadata record (table 1). Consequently, the goal of the Fiscal Year 2008 project was to generate more than 28,000 metadata records.

RESULTS

The actual number of metadata records completed during the Fiscal Year 2008 project is compared with the number of records proposed (goals) in table 1. Discrepancies between the actual numbers and goals are discussed below.

Table 1. Comparison of the goals and results for the Fiscal Year 2008 project.

Type of sample	Goals	Completed
Washed oil and gas well cuttings	4,000	3,605
Unwashed oil and gas well cuttings	50	0
Oil and gas cores	1,935	3,034
Industrial minerals cores	450	454
Water well cuttings	2,700	2,771
Fossils of Paleozoic age	18,900	4,687 records (17,027 total specimens)
Total	28,035	14,551

The actual number of washed oil and gas well cuttings stored in our warehouse is less than estimated (table 1). Also, records for unwashed well cuttings were not created, because the GSA has not completed its disposal process. In contrast, the number of records for oil and gas cores greatly exceeded the goal, primarily because many wells have multiple sets of cores. The actual number of records for industrial cores and water well cuttings is very close to the goals. The most significant difference between projected and actual records generated is in the paleontology collection. The number of metadata records for fossils is 4,687; however, the number of specimens is 17,027. Multiple specimens are often grouped within a specimen tray (especially when the specimens are the same species and/or collected from the same locality on the same date) and assigned a single identifier for the set (one metadata record for multiple specimens). The difference between the 17,027 specimens and the original estimation of 18,900 lies with the discovery made during the project that one of the Paleozoic cabinets is filled with fossils on loan from an outside source and are not property of GSA. Efforts are being made to return these fossils to their owner.

The metadata records for the GSA collections were compiled in Microsoft Excel files. The files were formatted according to the instructions posted at http://datapreservation.usgs.gov/docs/NGGDPP_MetadataPreparation.pdf. The files were saved in pipe-delimited csv format and were uploaded to the National Digital Catalog at <http://my.usgs.gov/csc/nggdpp/upload>.

DISCUSSION

The process of constructing metadata records for samples stored in the GSA warehouse resulted in additional benefits. Cross-checking the records in the OGB database against the actual samples stored in the warehouse revealed numerous inconsistencies in the OGB database. Because of the data generated during this project, these problems can now be corrected. The process of “groundtruthing” the records for samples stored in the GSA warehouse allowed the Co-Principal Investigator (Osborne) to closely inspect the collections and identify storage problems. Similarly, the physical inspection of paper records in the paleontology collection allowed the collection manager and Co-Principal Investigator (Ebersole) to examine the Paleozoic part of the paleontology collection in great detail, revealing storage and preservation deficiencies and identification of Paleozoic specimens on loan from an outside source. Storage problems in both the GSA warehouse and paleontology collection will be addressed in future NGGDPP Program Years.

GEOLOGICAL SURVEY OF ALABAMA

P.O. Box 869999
420 Hackberry Lane
Tuscaloosa, Alabama 35486-6999
205/349-2852

Berry H. (Nick) Tew, Jr., State Geologist

A list of the printed publications by the Geological Survey of Alabama can be obtained from the Publications Sales Office (205/247-3636) or through our web site at <http://www.gsa.state.al.us/>.

E-mail: pubsales@gsa.state.al.us

The Geological Survey of Alabama (GSA) makes every effort to collect, provide, and maintain accurate and complete information. However, data acquisition and research are ongoing activities of GSA, and interpretations may be revised as new data are acquired. Therefore, all information made available to the public by GSA should be viewed in that context. Neither the GSA nor any employee thereof makes any warranty, expressed or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed in this report. Conclusions drawn or actions taken on the basis of these data and information are the sole responsibility of the user.

As a recipient of Federal financial assistance from the U.S. Department of the Interior, the GSA prohibits discrimination on the basis of race, color, national origin, age, or disability in its programs or activities. Discrimination on the basis of sex is prohibited in federally assisted GSA education programs. If anyone believes that he or she has been discriminated against in any of the GSA's programs or activities, including its employment practices, the individual may contact the U.S. Geological Survey, U.S. Department of the Interior, Washington, D.C. 20240.

AN EQUAL OPPORTUNITY EMPLOYER

Serving Alabama since 1848