Final Technical Report

Prepared for the
National Geological and Geophysical Data Preservation Program

Generating New and Updated Metadata Records, and Digitizing Documents in the Tennessee Geological Survey's Coal Exploration, Zinc Mining, and Mineral Resources Collections from September 1, 2013 to August 31, 2014

United States Geological Survey
Award No. G13AP00089

Principal Investigator
Ronald P. Zurawski

Tennessee Department of Environment and Conservation
Tennessee Geological Survey
William R. Snodgrass TN Tower
312 Rosa L. Parks Ave., 12th Floor
Nashville, TN 37243
Telephone: 615-532-1502
Fax: 615-532-0199
E-mail: Ronald.Zurawski@tn.gov

November 26, 2014

Abstract

For FY2013 the National Geological and Geophysical Data Preservation Program awarded the Tennessee Geological Survey funds to continue the process of producing metadata records for the National Digital Catalog and to continue the conversion of its paper records into digital format. The proposed work was divided into three projects: (1) to digitally capture all of the remaining documents in the Zinc Mining Maps and Reports Collection (P1326); (2) to digitally capture and create metadata records for 1000 out of the remaining 1500 documents in the Coal Exploration Core Hole Logs Collection (P1334); and (3) to digitally capture and create metadata records for the mineral test hole files, which are part of the Mineral Resources Collection (P1328). We unfortunately report here that only the first two projects were successfully completed within the grant award period. A total of 1072 Coal Exploration Core Hole Log metadata records were compiled and the documents associated with these records have been converted to digital format scans that are safely stored, duplicated, and available for distribution as Portable Document or Tagged Image Format files. In addition, a total of 1035 maps, cross-sections, reports, and drill logs contained in the Zinc Mining Maps and Reports Collection have been scanned and the digital files are safely stored, duplicated, and available for distribution available as Portable Document or Tagged Image Format files. We also updated the Coal Exploration Core Hole Logs and Zinc Mining Maps and Reports collection metadata to indicate which documents are now available in digital format. The revised files have been uploaded and the records harvested onto the United States Geological Survey's ScienceBase website.

Introduction

The importance of cataloging Tennessee Geological Survey (TGS) collections and digitizing the documents in them continues to increase as potential users expect to have the ability to remotely query collection records and to acquire a document in a useful electronic format. Although a slow and somewhat tedious process, which has enlisted staff geologists that are trained to catalog geologic data in the field, the value added to the mission of TGS cannot be overstated. It is a goal worth striving for because the outcome helps ensure the preservation of historic documents, as well as decrease the staff time and the expense required to fulfill requests for information contained in these collections. Through the financial assistance provided by the U.S. Geological Survey (USGS) National Geologic and Geophysical Data Preservation Program (NGGDPP) TGS continues to make steps towards reaching that goal.

Currently and for the foreseeable future, TGS is unable to provide the on-line presence needed for its collections of geologic data. As a result, companies, government agencies, and the general public are unaware of the abundant site specific geologic data available in them. The goals for this year included completing the final phase of a multiyear effort to digitize all of the documents in the Zinc Mining Maps and Reports Collection; to make further steps to compile the metadata and digitize documents in the Coal Exploration Core Hole Logs Collection; and to begin compiling the metadata records and digitizing documents in the Mineral Resources Collection. By incorporating the newly compiled and/or updated metadata records for these collections into the National Digital Catalog (NGD), TGS is able to expand its on-line presence and provide the public with the means to search for, locate, and evaluate the type of information that is available in these collections.

Unfortunately, TGS was unable to complete one its project goals for this year. The effort to compile the metadata records and digitize documents in the Mineral Resources Collection was the responsibility of staff in TGS's Nashville office. Although the approximately 1000 mineral test hole files that are part of the Mineral Resources Collection were retrieved from TGS's warehouse in Nashville during the late spring of 2014, no metadata records were created, nor were any of the drilling data, site location maps, or detailed geologic logs scanned in fulfillment of this part on the FY2013 NGGDPP award. This was due, in part, for the following reasons:

1- As part of a major departmental reorganization during the latter part of 2013, the large format in-house continuous-feed copier/scanner that TGS planned to use for scanning the site location maps and geologic logs was transferred from TGS's office to a departmental satellite office as part of a move to consolidate oil and gas records and related equipment within the department's oil and gas regulatory program. Unfortunately, due to budgetary and other issues, the copier/scanner was not placed back into operation until the summer of 2014.
2- By the time the copier/scanner was once again available for use, TGS staff was obligated to complete more critical contractual obligations that had the same due date as the FY2013 NGGDPP project.

We still believe that preservation of these mineral test hole files and making them available through the National Digital Catalog is a goal worth pursuing. We therefore intend to resubmit this project in response to a future NGGDPP program announcement. In the near future we will be acquiring a replacement large format copier/scanner so that we will once again have in-house scanning capability for the site location maps and geologic logs that are part of this collection.

Work Completed

The two broad objectives for our NGGDPP FY2013 projects were to compile metadata in a format suitable for input into the NDC and to convert documents into a digital format.

The first objective was related to the Coal Exploration Core Hole Logs Collection. The metadata for a total of 1072 records was completed, which surpassed the target of 1000 originally proposed. The metadata file for the Coal Exploration Core Hole Logs Collection was created in Microsoft Excel and saved in .csv file format using the pipe (|) record delimiter character as a field separator for seamless uploading. The file was uploaded onto the USGS ScienceBase (NGGDPP website) and the records have been successfully harvested by USGS staff.

In addition to providing the metadata records for the Coal Exploration Core Hole Logs Collection, the following outcomes were realized as a result of completing this objective:

(1) We continued to review and better organize the material in the Coal Exploration Core Hole Logs Collection;

(2) An in-house Excel database started last year was populated with the new records. The importance of this database is that it includes additional information not included in the NDC to assist TGS staff with the searching and sorting of records, as well as the storage location of each document cataloged to date;

(3) The existing Excel database also feeds directly into a GIS (ArcMap) project to better visualize the geographic location and query the documents in the collection.

The second objective successfully completed this year was to continue the conversion of paper, Mylar, and linen documents into a digital format. In order to fulfill the proposed objective we were required to:

(1) digitally capture 1000 documents in the Coal Exploration Core Hole Logs Collection (a total of 1072 documents were completed) and;

(2) digitally capture the remaining documents in the Zinc Mining Maps and Reports Collection (a total of 1035 documents were scanned representing 950 records);

All documents larger than 11x17 inches were scanned at 400 dpi resolution in either grayscale or RGB color, uncompressed tagged image file format (.tiff) on a large format scanner. Page-sized and some tabloid sized documents were scanned in Portable Document File (PDF) format at 300 dpi resolution in either grayscale, black and white, or RGB color format on a Canon ImageRunner. Adobe Acrobat software was used to create a second digital file in PDF format for all files originally scanned in .tiff format. When necessary, free software called GIMP (GNU Image Manipulation Program), was used to crop, rotate, and digitally enhance the scanned documents.

The TGS staff involved in the project built upon lessons learned from the previous year's document scanning and digitizing effort. A new computer roll-out helped reduce the processing time to complete many of the steps in the digitizing process and prior experience resulted in a smoother workflow from document retrieval, to scanning, to digital enhancement of the documents in the collections. This was also true with regard to file naming, duplicating, storing, and all the accounting that goes along with verifying the completed tasks as a particular document moves through the process.

In addition to converting documents in the Coal Exploration Core Hole Logs Collection and Zinc Mining Maps and Reports Collection into a digital format, the following results were realized by working on this objective:

(1) An additional 45 records were added to the Zinc Mining Maps and Reports Collection. These documents have always been a part of the collection, but as our familiarity with this collection continues to improve we are able to understand and decipher relationships between documents that were once a poorly identified. The metadata for this collection has now grown to 3144 records.

(2) We updated the metadata records to indicate which documents are now available in digital format in each of the collections.

(3) We uploaded the .csv format files for the updated collections on the USGS ScienceBase website, revised the metadata describing the collections, and requested that the old versions of the collections be removed from the website.

Tennessee Geological Survey Staff Participation

A number of TGS personnel were involved in completing the work for this year's NGGDPP project. Chief Geologist Dr. Peter Lemiszki in TGS's Knoxville Field Office was responsible for scanning, image processing, and updating the metadata file for 1035 documents in the Zinc Mining Maps and Reports Collection. He was also responsible for completing the Final Technical Report and submitting the datasets onto the USGS ScienceBase website. Geologist Barry Miller in TGS's Knoxville Field Office was responsible for scanning, image processing, and creating the metadata files for the Coal Exploration Core Logs Collection. Secretary Becky Hawkins in the Knoxville Field Office assisted Pete and Barry with the scanning of page-size to tabloid-size documents from both collections.

Publications

Currently there are no plans to publish this information in any other form than what is available on the USGS ScienceBase website.