NGGDPP Final Technical Report – FY 2014
Award G14AS00026

Name of State Geological Survey: Kentucky Geological Survey

Project Title: Kentucky Geological Survey FY2014 Geologic Data Preservation Project

Term Covered: September 1, 2014 – August 31, 2015

Principal Investigator(s): Douglas Curl
Kentucky Geological Survey
228 MMRB
University of Kentucky
Lexington, KY 40506-0107
859-323-0519, 859-257-1147 (fax), doug@uky.edu

Submittal Date of Final Technical Report: 1/8/2015

Abstract:
KGS is legislatively mandated to archive and make available oil and gas records for wells permitted through the Kentucky Division of Oil and Gas. Records are transmitted from the Division to the KGS Lexington office where they are scanned and prepared for online access. The KGS western Kentucky office maintains a separate file of historical well records that has been compiled from a variety of sources. This collection is known to contain some original documents not in the main office, as well as working copies that have valuable annotations that have been made over the years. The only access to these unique documents is by visual inspection of the original documents at the western Kentucky office. KGS has lacked resources to make significant progress with the comparison of the two collections to identify unique documents to scan and append to the online repository. There are estimated to be 80,000 document folders that need review, and this is likely to constitute a multi-year effort.

The KGS 2013-14 NGGDPP project started this effort and was successful in completing the scanning and preparation of 15% of this oil and gas documents collection (40,100 pages of records from 12,384 folders). Most of these newly scanned records are now available for viewing online through the KGS oil and gas data search. We proposed to continue preserving this collecting by employing students both in the western Kentucky and the Lexington office to further this significant effort. The proven workflow used was to review, scan, and rebuild the online file as a single process so the work will have immediate benefits. As soon as they are processed, the scanned documents are made available online through the Kentucky Geological Survey oil and gas well search (http://kgs.uky.edu/kgsweb/DataSearching/OilGas/OGSearch.asp).

Objectives:
1. Review 20,000 folders of oil and gas records for documents missing from the Kentucky Oil and Gas Repository, and add digital images to the existing online resource.

2. Update the KGS collections folder in the USGS ScienceBase catalog to include recently identified collections, and update metadata records for select collections that have changed since originally entered.

Results:

1. Approximately 8650 oil and gas record folders with approximately 60,000 pages of records were reviewed for documents missing from the Kentucky Oil and Gas Repository. This number of reviewed folders fell well short of the objective for review of 20,000 folders, which was due to several factors:
    - A staffing change of the scanning personnel occurred in the middle of the project (Jan 2015), which resulted in a slowdown due to the requisite training and it taking about a month to get up to the same speed as the previous staff person.
    - The scanner that was used in the KGS western Kentucky office began to fail during the 2nd quarter of the project, and dramatically slowed down the scanning process. KGS replaced the scanner at the end of April 2015, which improved the scanning process, scan quality and time taken. Also, a number of longer log images that were scanned with the failing scanner had to be rescanned with the newer scanner, because upon a later review of the images online, it was discovered that the failing scanner had produced problems with these longer images. 4685 log images had to be rescanned due to this issue.
    - During the process of reviewing documents that were previously scanned and online, the scanning staff found either low quality or unreadable scans (black pages) from previous scanning efforts, or found documents that needed to be updated with newer documents found in the record folders (typically, these were electronic logs with useful notations that were not in the original scans of the same logs). A decision was made to re-scan these documents to either fix the quality or update the documents so that the documents archive had the best available information.
    - The folders are organized in cabinets which are organized geographically. During this project year, the scanning process has moved into areas of higher oil and gas development which both has more wells, but also more documents per well. Some of these documents from these higher production areas include longer electronic logs, which take longer to scan. In the previous phase of this project (Data Preservation FY2013 - G13AP00090), there were approximately 3.2 pages/well record. During this phase, that ratio increased to 4.7 pages/well record. When there are more pages per well, there are also more filing errors that the scanning staff must resolve. This includes finding physical pages filed with the incorrect well or finding documents from two record folders that need to be combined. This process will also slow down the scanning process, but improves the integrity of these preserved documents.

    During the well record folder review, missing documents were scanned at the KGS western Kentucky office, which resulted in scanning of 31,517 pages of documents for 6,756 oil and gas well records (this includes 4378 pages that cannot be posted online due to confidentiality requirements). Above the 6756 reviewed oil and gas well records,

approximately 1900 well records were found to have low quality or unreadable scanned documents that were rescanned. This scanning process also resulted in 779 new oil and gas record data records added to the KGS oil and gas database. These scanned documents and new records are now available online through the KGS oil and gas well search.

2. KGS has a web application which generates an XML file of metadata from the most current data in the database for each collection. The intent of this application is for a user to download all the metadata of current items for a collection into the National Digital Catalog (NDC) XML format. The metadata can then be submitted to the NDC via the NDC Dashboard. For KGS items, wholesale replacement of all metadata for each collection is required on each upload for a collection. This is because data may change for existing items, and to ensure that all items are added to the collection.

   A new XML metadata was created for oil and gas well records and uploaded to the NDC. An additional 7274 metadata items were added to the catalog for oil and gas well records, which is a total of 147,674 metadata items for this collection.

Future work:
Scanning of missing documents from the Kentucky Oil and Gas repository is an on-going process, and KGS is currently funded for the FY2015 through the Data Preservation program to continue this work. KGS is anticipates reviewing and scanning missing documents from approximately 15,000 well records during this phase of the project.