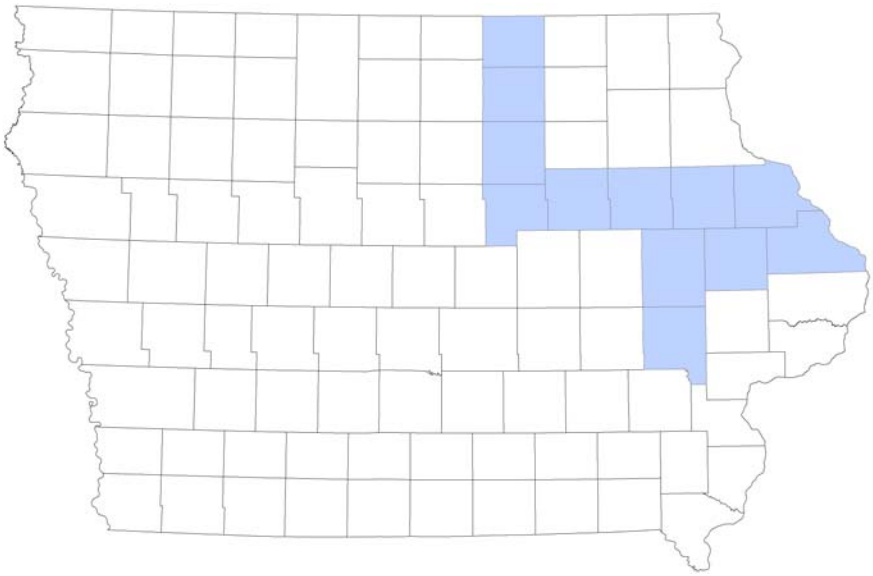
Digital Infrastructure and Metadata for Geologic Collections, East-Central Iowa Focus



Iowa Geological Survey Technical Report
National Geological and Geophysical Data Preservation Program
United States Geological Survey Award Number G13AP00101.
Term Covered by Award: 09/15/2013 - 09/14/2014
Submitted: 09/09/2014

Robert D. Libra, Principal Investigator
Iowa Department of Natural Resources
Wallace State Office Building
502 East 9th Street
Des Moines, Iowa 50319-0034
(515) 725-3408
Robert.Libra@dnr.iowa.gov

Jason A. Vogelgesang
Iowa Geological Survey
IIHR – Hydroscience & Engineering
The University of Iowa
300 Trowbridge Hall
Iowa City, Iowa 52242-1585
(319) 335-4231
Jason-Vogelgesang@uiowa.edu

Abstract

This report documents work performed by the Iowa Geological Survey for the National Geological and Geophysical Data Preservation Program (NGGDPP) under award number G13AP00101 for 2013. The project focused on preserving data in a twelve county east-central region of Iowa. Grant objectives of creating and/or updating digital infrastructure and creating metadata were met. The majority of paper documents preserved to electronic format included well logs and paper documents. Metadata for these and other documents were created and uploaded to the National Digital Catalog.

Background

The IGS was established as the Iowa Geological Survey (IGS) in 1892, but existed discontinuously since 1855. Systematic collection of samples and records for wells began in the 1930s. While much of the data have been converted to digital records, large portions of the data are still held as paper documents and files in printed reports, maps, field notes, etc. Many of the paper documents are deteriorating, often poorly maintained, and difficult to access efficiently for research, resource or environmental assessment, among other uses. Conversion to digital formats and addition to the GEOSAM and other appropriate document storage locations alleviates these shortcomings to a large degree.

In early 2014, the Iowa Geological Survey transitioned away from the Iowa Department of Natural Resources to IIHR-Hydroscience and Engineering, at the University of Iowa. The IGS, now a part of the University of Iowa (UI), remains the state's lead earth and water science. IGS staff completed this report under the direction of Robert Libra, State Geologist of Iowa and NGGDPP Principal Investigator, who is employed with the Iowa Department of Natural Resources.

Figure 1. Well records and other document storage the Oakdale Iowa Geological Survey Building.

The IGS occupies two buildings on UI campuses. The main office is located in Trowbridge Hall on the central UI campus. The Oakdale Iowa Geological Survey (OIGS) building is on the west edge of the University of Iowa Research Park campus. The building houses the IGS' Rock Library, the

repository for driller's logs, well cores, well chips, and rock specimens from exposures throughout Iowa. Additionally, the building provides storage for the bulk of the paper files that document the physical collections (Figure 1). Included are a wide variety of older paper files and archived maps.

The majority of the most critical and widely used geologic information for the state is electronically held in the IGS Geologic Sample Database (GEOSAM). Access is free and available to the general public and is provided through a web-based interface (http://ifis.iowafloodcenter.org/ifis/more/geosam), which also provides tools to query the database and view the electronic documents. Geologic information accessible via GEOSAM includes: site identifiers and location, well construction information, geologic logs (lithologic and/or stratigraphic), and water production information. In early 2014, GEOSAM's web-based interface was redesigned to allow users to locate sites geographically. Further enhancements to the interface are planned including the addition of more text search fields. GEOSAM and GEOSAM-compatible databases form the backbone of the current IGS data preservation system and its ongoing preservation plan.

Established 2013 Grant Activities

IGS included the following in its 2013 NGGDPP grant proposal:

- Establish locations; enter well, geologic, and hydrogeologic information as needed for pertinent collections not held in GEOSAM
    - Focus on twelve counties in the east-central part of the State
    - Regional and statewide collections of unrecorded well logs, maps, and other items will be included
    - Migrate collected data to GEOSAM
    - Develop and submit NGGDPP-compliant metadata for these additions
- Enter or update GEOSAM records as needed
    - Well log data previously not added
    - Geologic log, well construction, aquifer information
    - Additional information on existing well, cuttings, core records
- Modify GEOSAM by developing tables/views to manage additional types of geologic information as part of an IGS long term data preservation strategy
- Submit all developed metadata to the National Catalog
- Supply a complete update of Iowa data to the National Catalog
- Create electronic versions of prioritized statewide publications
    - For publications that only exist in paper format
    - Develop metadata for these additions
- Develop process for scanning paper geophysical logs and scan 10 logs
- Develop metadata for a collection of historic coal mine township and underground mine maps
- Develop metadata for a collection of IGS Annual Reports; create index of electronic file locations
- Submit all developed metadata to the National Catalog
- Supply a complete update of Iowa data to the National Catalog

The project focused efforts on capturing data from a twelve county region of east-central Iowa, when applicable, to continue our geographic data preservation approach. Prior grants have focused on other regions of the state, working systematically from the southwest to the southeast in a clockwise manner.

Results of 2013 Grant Activities

NGGDPP Metadata Development

The IGS manages a total of 10 collections defined for NGGDPP. Metadata records were created and submitted for all items associated with a geographic location, including rock cores, rock cuttings, well logs, field notes, maps, and paper reports. Items contained in GEOSAM, including rock cores, rock cuttings, field notes, and well logs, have been assigned a point location and horizontal accuracy as part of GEOSAM entry. In the case of maps and paper reports, centroid coordinates were assigned.

Metadata were prepared in CSV format for all of the collections submitted as part of this project. The metadata were generated using views and queries from SQL Server databases. The CSV format was chosen as the metadata creators have more experience with that format.

Metadata preparation was carried out as follows:

1) SQL server tables, views, and custom functions needed to produce the necessary text strings were developed.
2) Queries were developed to select the data from the views prepared in step one to a temporary file.
3) The temporary file was imported into Excel and validated to ensure National Catalog compatibility.
4) The Excel file was saved to CSV format with pipe characters as delimiters and the file was uploaded to the National Digital Catalog website.
5) The final step was to request NGGDPP staff to load the data into the catalog, replacing previously submitted data sets.

The NGGDPP metadata elements, along with a description of the data items provided for each element, are listed in Table 1. In instances where required data elements were not available, metadata records were not prepared, e.g. if latitude and longitude had not been stored for a data item, the item was omitted from the metadata. If changes/updates were not made to a collection, metadata was not submitted. In this case previously uploaded metadata are still current.

The IGS metadata supplies data for the optional items in some cases. For most collections, optional items were either populated for all records or no records. In other cases, the optional items were populated where the data to be supplied was available.

Table 1. Description of metadata elements developed from IGS data.

Tag/Data Element	Description of data supplied
collectionID	Collection ID from previously defined Iowa NGGDPP collections.
title	Name of collection and IGS identifier for the item. Untitled documents were given a descriptive title.
abstract	General descriptive information about the site and/or data item including identifier and site type; other information, depending on the collection being documented, includes total depth, sample depth, completion date, collection date, aquifer, and author.
dataType	NGGDPP catalog supplied values.
supplementalInformation	Statement concerning archive location of samples/logs and/or a URL where additional information may be accessed, if available.
coordinates	Longitude and latitude in decimal degrees (NAD83) derived from coordinates stored in GEOSAM. Centroid coordinates are provided for maps and paper reports.
onlineResource (optional element)	Hyperlinks are provided where available.
datasetReferenceDate	The date the metadata for the collection was prepared

Additions to IGS Collections for East-Central Iowa

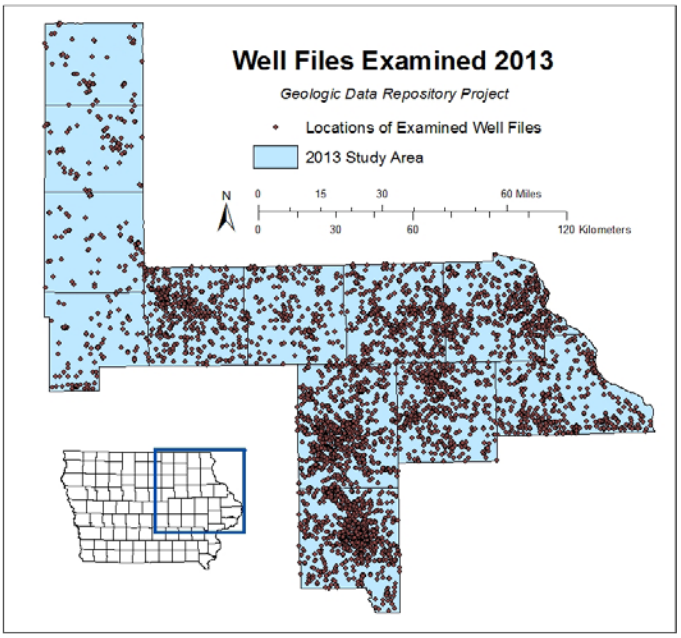


Figure 2. East-central Iowa study area showing the location of 5,969 well sites (of 4,152 well folders examined) with relevant information scanned and added to GEOSAM.

Previous NGGDPP grants have allowed the IGS to focus geologic data preservation efforts in the northwest, southwest, south-central, north-central, and northeast areas of the state. This year's work follows that regional approach and focuses on twelve counties in east-central Iowa (Figure 2). While the vast majority of preserved data lie within the east-central focused area, some data plot outside of the twelve county region.

A total of 5,969 well file folders from east-central Iowa were examined (not including archived well logs described below). While GEOSAM contained the basic information on each of these records, additional information existed only in paper format. Relevant information from 4,152 folders (Figure 2) was scanned and attached as PDFs to the existing well data in GEOSAM.

Archived Well Logs Scanned and Entered in GEOSAM

The transition of the IGS to the University of Iowa initiated a major inventory and organization of archived material. As such, a number of boxes of driller's logs were discovered. An investigation into these records disclosed that these logs had never been entered into the GEOSAM database. As a part of this NGGDPP grant, approximately 1,500 driller's logs were scanned and entered into GEOSAM. This time consuming process involved looking through historic and current landowner data, locating the well as best possible, entering all applicable information to GEOSAM, and scanning and uploading the log to GEOSAM. The logs contain valuable geologic, hydrogeologic, well construction, and location data that is now accessible electronically.

Preservation and Cataloging of Paper Reports

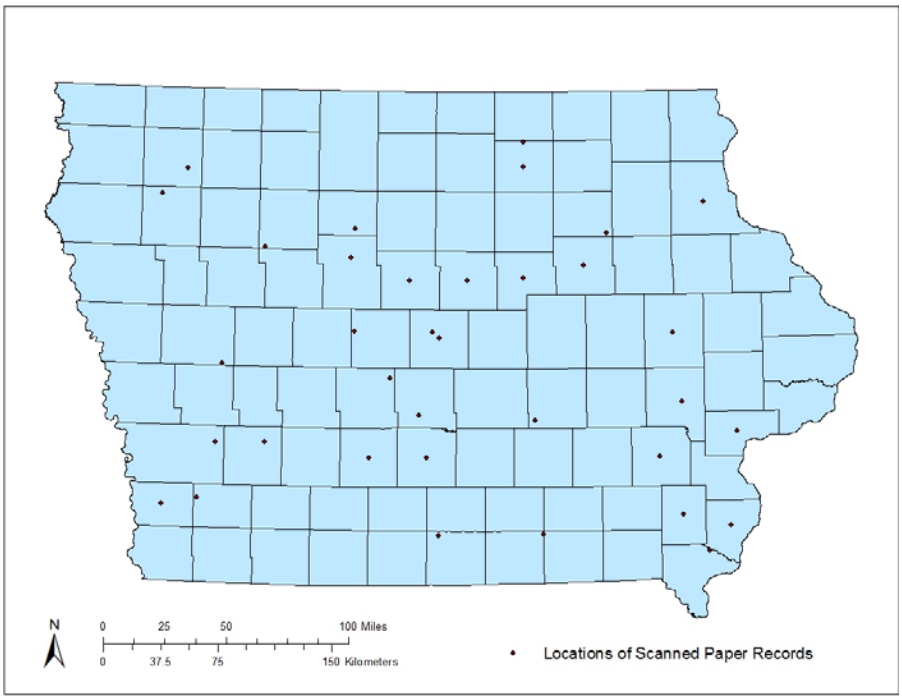


Figure 3. Map of Iowa showing the locations assigned to 82 paper reports that were scanned as part of year's work.

Throughout its history, the IGS has generated an extensive collection of paper reports documenting geologic studies. Reports include projects ranging in scope from statewide to narrowly-focused in geographic and subject area. In recent years, newly published IGS reports are created as electronic documents or as electronic and paper documents. The electronic versions are publicly available as PDF downloads from the IGS website (http://s-iihr34.iihr.uiowa.edu/publications/search). However, older reports exist as paper documents only. In some cases the reports have been superseded by newer work, but older reports represent foundation work and are important from a historical perspective. Prominent among this group of documents are the Iowa Geological Survey Annual Reports V. 1-38 were published from 1893 through 1941. The reports include 239 separate chapters. The reports were scanned by the University of Iowa Libraries in 2011 and made available at: http://ir.uiowa.edu/igsar/. Metadata were submitted this year for these publications and an indexed table was created.

Remaining paper-only publications were prioritized for digital conversion. Prioritized and preserved paper report categories included: 31 Open File Reports, 26 Iowa Geology publications, 12 Miscellaneous Publications, 9 Ground Water Resources publications, 2 Water Supply Bulletins, 1 Annual Report Separate, and 1 Aeromagnetic Survey publication. A total of 4,487 pages were scanned during this project year. Scanned documents were added to the IGS publication website (link above). Metadata were developed and submitted for these paper documents.

Metadata Records Added to the National Catalog

Metadata describing the information held in GEOSAM, as well as other geologic information from Iowa, were loaded into the National Digital Catalog in September 2014. The number of data items included in the metadata is summarized by collection as follows:

Table 2. Summary of new metadata provided to NGGDPP catalog.

GEOSAM Metadata Records Submitted **East-central Iowa Focus*	**New Records 2013**	**Total Records Uploaded**	**Approximate Collection Total**
Rock Cores	5	1,291	1,300*
Rock Cuttings	193	36,615	39,245
Hand Samples	0	258	425
Thin Sections	0	1,865	5,200
Well Logs (driller's and strip)	5,730	81,469	85,000
Paper Reports	82	319	886
Maps	473	1,549	2,000
Geochemical Data	0	20,022	20,022
Photographs	0	0	20,000
Field Notes	27	4,199	8,000
Total	**6,510**	**147,587**	**182,069**

**Metadata numbers for rock cores are less than prior years. It was discovered that metadata were being generated for every record containing a core log. Current numbers reflect only physical rock cores in our collection.*

Geophysical Logs

The IGS has seen an increase in the demand for digitally accessible geophysical logs, particularly from the oil and gas industry. As such, we proposed to develop a process to scan electrical geophysical logs. The logs typically measure 8.5" in width and can be several feet in length. The IGS has experienced issues saving these files due to their disproportional length and width. We discovered a local printing company had experience in resolving the size issues we faced and used that company's scanning services to scan 41 logs. The IGS decided that utilizing the printing company scanning services was more cost-effective than purchasing the correct equipment and software to scan the logs in-house.

GESOAM updates

2013 project goals included modifying GEOSAM by developing tables and/or views to manage additional types of geologic data. With GEOSAM serving as the IGS' main geologic data storage database, ongoing modifications for efficient data retrieval are necessary. During the 2013 year, new tables and views were created to better manage and retrieve geologic data. One such example is a table that was created to link GEOSAM data to wells used in the Iowa Department of Natural Resources Source-water Protection program (SWP). The table allows the SWP program to quickly gather geologic data pertaining to their work in protecting groundwater resources.

Maps

Part of this year's goals involved the creation of metadata for a collection of approximately 200 maps including historic maps showing coal mines compiled by the Iowa State Mine Inspector's office in the 1960s, underground mine layout maps, and miscellaneous geologic maps. The metadata for these maps was submitted in December 2013. In addition, many of these maps were added to the Iowa Department of Natural Resources' Natural Resources Geographic Information Systems (NRGIS) online library (https://programs.iowadnr.gov/nrgislibx/).

2013 Award Summary and 2014 Award Plans

Grant objectives of creating and/or updating digital infrastructure and creating metadata were met. This project year focused on preserving data within a twelve county region in east-central Iowa. The majority of paper documents preserved to electronic format included well logs and paper documents. Metadata for these and other documents were created and uploaded to the National Digital Catalog.

Previous NGGDPP grants allowed the IGS to convert valuable geologic data to digital format in the northwest, southwest, south-central, north-central, northeast, and now east-central regions of the state. Proposed NGGDPP work for 2014 follows and completes this regional approach, focusing on thirteen counties in the southeast region of the state. After completion of the 2014 NGGDPP project, data in all geographic regions of the state will have experienced focused preservation efforts.