# Award Number G09AP00092

# National Geological and Geophysical Data Preservation Program

# Preservation of Geologic Data and Collections in Illinois: Compilation, Documentation and Planning

# Final Technical Report

**Fiscal Year 2009**

**7/1/2009-6/30/2010**


Ronald Klass
Information Management Coordinator
Tel. 217-244-5585
Fax 217-333-2830
rklass@illinois.edu

Zakaria Lasemi
Senior Geologist and Head, Industrial Minerals and Resource Economics Section
Interim Director, Earth Resources Center
Tel. 217-244-6944
Fax 217-333-2830
lasemi@isgs.illinois.edu

**Illinois State Geological Survey**
**Institute of Natural Resource Sustainability**
**University of Illinois**
**615 East Peabody Drive**
**Champaign, IL 61820**


**August 23, 2010**

## Abstract

This report summarizes the work performed for the Illinois State Geological Survey FY09 NGGDPP project.

### Collection Metadata

- Metadata records were created for 2 additional collections.
- Metadata records for 3 collections not completed in FY08 were substantially augmented.
- Metadata records for other existing collections were updated.
- All metadata records were uploaded to the National Catalog.

### Collections Inventory

- An additional 4 collections were identified and entered into the collection inventory. A total of 40 collections are now listed in the National Catalog.

### Data Preservation Techniques Workshop

- Three staff members attended the Data Preservation Techniques Workshop July 14-15, 2009 hosted by the Indiana Geological Survey and sponsored by USGS NGGDPP and Association of American State Geologists.

## Collection Metadata

### Albert V. Carozzi Collection (Hand Samples)

Dr. Carozzi was a distinguished professor of Geology at the University of Illinois from 1955 until his retirement in 1989, amassing this considerable worldwide collection over his tenure. The collection represents many years of research, especially in the area of petroleum exploration and projects completed by many of his Masters and Ph.D. students. Metadata records for 13,856 specimens were created and uploaded to the National Catalog.

While over half of the specimen localities are within the United States, many are in Europe, South America, and other locations worldwide. Student help was used to research and assign a coordinate to each specimen based on available information. In most cases internet services such as the USGS Geographic Names database were used to obtain a reasonably precise coordinate for each locality. For samples used in theses, the thesis was reviewed to obtain locality and coordinate information.

Coordinates for approximately 2000 specimens could not be estimated for various reasons. We plan to do additional work in FY10 (time allowing) to obtain coordinates for as many additional specimens as practical.

### Weller Collection (Paleontological Samples)

This collection consists of mostly Pennsylvanian and Mississippian samples primarily collected and cataloged by M. Weller from 1926-1942. Most of Weller's numerous (approximately 75,000) specimens are in the main Paleontological samples collection and thus are not cataloged. This collection represents his cataloged specimens; many are type or figured starting with specimens from Weller's first published paleontological paper in 1929. Metadata records for 4001 paleontological samples in the Weller Collection were created and uploaded to the National Catalog.

### Geologic Samples Library Paleontological Samples, Silurian subset

Our main Paleontological samples collection consists of approximately 150,000 specimens. Until this project none were cataloged, but in general the samples are physically arranged chronostratigraphically (by geologic system and then by stratigraphic unit within the system). In FY08 we began creating metadata for the Silurian subset of this collection, ending FY08 with metadata for 11,086 samples from 460 localities (roughly half of the entire Silurian subset).

For FY09, we decided to complete the metadata for the Silurian portion of the collection, using the same student who had become quite proficient at the task. We also completed numerous sets of "problem" specimens that required additional work or research and thus had been set aside. The completed Silurian subset was uploaded to the National Catalog and consists of 25,924 samples from 982 localities.

Many metadata sample locations do not yet have a precise latitude/longitude point, generally for one of two reasons: first, some do not list a section, but rather name a specific quarry or locality. Staff members expressed a strong preference that they (rather than the student) derive detailed location information because they are familiar with most of the localities. However until then, the point information for most of these samples is approximate (county or state centroid). The derivation and precision of each sample location is noted in the alternate geometry section of the XML data.

**Lead-Zinc mining district borehole records (including Geochemical data)**

These paper records and mine maps were donated to us after the last operating Lead-Zinc mining company in Illinois closed.

In FY08 a student began organizing the data and maps, and entered location and formation data from a total of 452 boreholes into our Enterprise well and borehole database, 372 of which included geochemical analysis data for Lead, Zinc, and Iron at regularly spaced depths in the borehole and which were used to create metadata records for the geochemical data collection. However the location data were imprecise, based on PLSS sections and not suitable for precise mapping.

For FY09, the student continued sorting and organizing the numerous mine maps. A subset of mine maps were chosen to be scanned and then georeferenced in ArcGIS so that more precise locations could be assigned to the boreholes. This required significant effort and interpretation to determine valid reference points but it allowed us to significantly improve the precision of the location of virtually all boreholes previously entered. Revised information for 372 boreholes with geochemical analysis was uploaded to the National Catalog. Additional information, maps, and data remain to be sorted through and processed; we may continue work on this project if and when resources become available.

While working on this project, we discovered a set of historic logbooks containing similar data that previously were unknown to current staff. These logbooks contain data for boreholes from several Lead-Zinc mining companies drilled primarily from the 1940’s through 1960’s. A detailed review and inventory of the logbooks was made to facilitate future metadata creation. Approximately 1800 boreholes are represented, most have precise location information and many contain geochemical analyses (and/or estimates) at various depths. We also determined that while these additional boreholes are in the same general areas as the current NGGDPP project Lead-Zinc project data, it is unlikely that any of the logbook data duplicate boreholes in the NGGDPP project. We plan to create metadata for the logbook boreholes if and when resources become available.

### Electrical Earth Resistivity Survey records

We have a collection of Electrical Earth Resistivity Survey data, reports and maps going back to the 1930's. In general these were used to determine sites with better potential for water wells in a survey area.

In FY08 we created metadata for sites that were cataloged in historical paper summary forms. These forms cover from the 1930's through approximately 2000. In order to capture the last 10 years of surveys as well as some earlier surveys missing from the historical summaries, we then began systematically examining each of the paper-based files. Since the paper files were being systematically reviewed we decided to also scan them at the same time. In FY08 when the project ended we had scanned and cataloged additional data from 6 of Illinois's 102 counties.

The Geophysicist in charge of this collection found the FY08 NGGDPP work so valuable that he has used some of his staff's time as well as used other funds to hire a summer intern to continue the work of systematically reviewing and scanning the remaining county files as well as adding previously non-cataloged projects to the metadata database under the direction of the PI. As of mid-summer 2010, an additional 14 counties (making a total of 20) have been scanned; corresponding metadata for 358 locations was created and uploaded to the National Catalog.

### Updates to existing metadata

In addition to the collections previously described, the following collections were updated to include new data and data revisions. With these updates, all of our collections with metadata in the National Catalog are current as of mid-2010.

- Rock Cores
- Rock Cuttings

## Collections Inventory

In FY09 we inventoried 4 additional collections not previously inventoried in FY08:

- Geochemical data (historic)
- Fluorspar District Mine Borehole Records (historic)
- Clay Mineralogy Data (historic)
- Clay Mineralogy Data (current)

The information on these collections was provided to the National Catalog where a total of 40 collections from the Illinois State Geological Survey are currently listed.

## Data Preservation Techniques Workshop

Staff members D. Keefer, R. Klass, and D. Mikulic attended the Data Preservation Techniques Workshop July 14-15, 2009 hosted by the Indiana Geological Survey and sponsored by USGS NGGDPP and Association of American State Geologists. We feel this event was an excellent opportunity to learn and share data preservation techniques and best management practices. We appreciate the NGGDPP setting up and providing funding for this workshop, and look forward to other workshops in the future.

## Comparison of Proposal Goals with Project Accomplishments

The FY09 proposal lists the following specific objectives:

> 1) Continue the compilation of existing metadata and the creation of new metadata for identified collections
> 2) continue the uploading of metadata on ISGS collections into the National Catalog and
> 3) produce a Final Technical Report that summarizes the results of the project.

It should be noted that at the time the FY09 proposal was due, we were only 4 months into the FY08 project and had not as of yet fully completed and analyzed our collection inventory in order to prioritize specific collections for metadata creation. For these reasons the FY09 proposal did not identify specific collections for which to create metadata. However, by the start of FY09 we had identified a number of collections for which to create metadata; we also decided to continue creating and/or refining metadata on some FY08 collections.

This final report, itself a deliverable, summarizes the work done in FY09. To recap, we created new metadata for 2 collections, and created significant additional or revised metadata for 3 collections first started in FY08. These accomplishments meet the goals of the FY09 proposal.

## Conclusion

We believe the collection metadata generated and uploaded for this project represents an appropriate and successful use of the resources provided under the contract. We are continuing metadata creation for FY2010. We also plan to investigate and potentially use OGC WMS services to deliver metadata to the National Catalog in the future.

## Acknowledgements

Several staff members at the Illinois State Geological Survey participated in this project:

- **Jared Freiburg** helped coordinate scanning and georectifing of the Lead-Zinc mining district mine maps.
- **Don Keefer** advised on many aspects of the project.
- **Don Mikulic** selected and coordinated the preparation of the paleontological samples for metadata creation.
- **Rod Norby** (retired) assisted with paleontological-related issues.
- **Tim Young** assisted with Electrical Earth Resistivity report issues.