

Geological Data and Collections in Peril: Case Example in Georgia

By Mark D. Cocker

Georgia Geologic Survey (former)
Watershed Protection Branch (currently)
Georgia Environmental Protection Division
19 Martin Luther King, Jr., Dr., SW
Atlanta, Georgia 30338
Telephone: (404) 657-6131
e-mail: mark_cocker@mail.dnr.state.ga.us

INTRODUCTION

A report by the National Research Council Committee (NRC) on the Preservation of Geoscience Data and Collections investigated types of geoscience data and collections, their estimated volume, and factors that threaten loss or degradation of these data and collections (Musser, 2003). Types of data and collections included drill core, cuttings, thin sections, washed residues, well logs, fossils, minerals, rocks, surface geophysical surveys, scout tickets, and chemical analyses. The emphasis of the National Research Council Committee appeared to be on the vast amount of data collected by hydrocarbon and mineral exploration companies in the western U.S., as well as data collected and stored by relatively well-funded, mainly western, state geological surveys and the USGS. The Committee's recommendations focused on preservation of the data and collections and value-added functions such as documentation and outreach (Musser, 2003). Rocks, whether they are drill core, cuttings, fossils or hand specimens, represent the most voluminous (and heavy) portion of the data and collections. Rocks are, in some ways, the cheapest to preserve, and in other ways, probably the most costly to preserve and document.

A significant amount of other geoscience data that wasn't addressed in the NRC study includes field notes, maps, photographs, and publications, as well as data that are difficult to quantify (e.g. institutional memory). Some state geological surveys may depend more heavily on these types of data, which require smaller expenditures to acquire or maintain than, for example, drill core. State geological surveys are more likely to have more focused geological data and collections pertinent to their respective state, and should have the greatest interest in preservation and documentation of those data, as well as promotion of these through outreach programs. However,

when a state geological survey suddenly ceases to exist, will there be any stewardship of the data and collections?

STATUS OF THE GEORGIA GEOLOGIC SURVEY

At the end of 2004, after serving the State of Georgia for 115 years, the Georgia Geologic Survey (GGS) was abruptly terminated. The State Geologist retired in August of 2005. A small handful of geologists continue to work on geologic problems and mapping within the Regulatory Support Program, Watershed Protection Branch, Environmental Protection Division, Georgia Department of Natural Resources. No organization or individual has been charged with ownership of the GGS's geoscience data or collections. The threat of non-stewardship and perhaps permanent loss of a significant amount of scientific and historical data and collections pertaining to the State of Georgia is real. Permanent loss of institutional memory is highly probable.

Although the collection of geoscience data had increased significantly in Georgia during the past 30 years, the GGS's management exhibited little interest and committed inadequate funding, time and personnel to the organization, documentation, preservation and storage of these geoscience data and collections. Prior to termination of the GGS, guidelines had not been established to organize, document, manage, preserve and store these data and collections, as well as new data and additions to the collections that continue to accumulate. Maintenance and updating of digital data and media have not been addressed. A few geology programs continue, e.g., the National Cooperative Geologic Mapping Program's STATEMAP component, and these accumulate new data, such as field observations, photographs, maps, and core logs, and new collections, such as drill core.

TYPES OF GEOLOGICAL DATA AND COLLECTIONS

GGs's geological data and collections are basically the same as most governmental geological surveys. These data and collections include written field observations in notebooks and maps, well logs, petrographic and XRD analyses, geochemical analyses, geophysical well logs, photographic records (film and digital), drill core, cuttings, washed residues, minerals, rocks and fossils. Mapping projects conducted under the STATEMAP program offer a good example of the breadth of these collections. These projects have allowed the field geologist to observe the geology and geological relations that describe outcrops, roadcuts, and mine exposures and to record these observations by a variety of methods. During the course of mapping 7.5' quadrangles, 300 to 500 sites per quadrangle may be examined and recorded. During the GGS mapping program, these sites are continuously added to a GIS outcrop database, which now contains more than 6000 sites. The database currently includes the quad number, the site number, an interpretation of the geological unit (i.e., the formation symbol), as well as a coded shade-set number for map plotting purposes. (Note: at the beginning of this GGS mapping program, publication of digital geological database examples was rather limited, and the year-to-year continuity of the program was not established, so only the minimum amount of geological data was entered into the database. Future tasks may include coding of site descriptions and perhaps linking of digital photos to each outcrop). With this database, outcrops can be quickly plotted on a topographic base map either as a hard copy or on-screen. Contacts for an interpretive geologic map are then digitized on-screen in relation to the outcrops. Outcrops are assigned a slightly darker color shade and can be plotted relative to the interpretive geology. Relative, size, shape and distribution of outcrops are also apparent on the outcrop coverage. The addition or linking of digital photographs can document lithology, sedimentary structures, alteration, mineralization, structural deformation, and geological hazards.

Another aspect of the STATEMAP program involves shallow core drilling. This drilling is invaluable in areas where outcrops are poor to non-existent. Because of equipment limitations, hole depth is limited to 50 feet. Sites are selected mainly for the opportunity of locating geologic contacts, and core is logged principally for lithology and contacts. In addition to the written core logs, digital photographs of the core are taken with a 2.5-foot scale marked in inches. These images are clipped, and the core is reconstructed into 10-foot lengths by digitally pasting the images end-to-end.

A part of the product produced by the GGS for the STATEMAP program consists of a geological report that includes descriptions of the formations, structure,

mineralization, aquifers, and geologic hazards. Selected annotated photographs of outcrops document observations in the report. Descriptions of new core are also included. Because the file size of a digitally reconstructed core hole is on the order of several tens of megabytes, current hardware and software cannot handle these files in a text document, and so they are not included in the published STATEMAP product.

Older data, maps and manuscripts exist only in hard copy paper or mylar formats. More recently, data, maps and manuscripts were compiled or created digitally and stored on a variety of evolving formats and media. A significant amount of data exists only in the form of hard copy publications. Some of the newer publications are available only on CD ROMs. Without a management plan and support, will those publications on CDs be readable in 10 or 20 years?

The GGS lacked a management plan to develop consistent data-recording methodologies and store and preserve that data. Over a period of many years, diverse types of geologic data were collected and recorded by numerous staff members with different education and experience levels, employing a variety of evolving techniques, tools and media. In addition, geologists were not required to provide copies of the data to the technical files. This resulted in the actual physical loss of unpublished data or misplacement of data files when staff members left the GGS.

During the past 17 years, digital technology advanced slowly within the GGS. In 1988, one personal computer was available to a staff of approximately 40. Computers were gradually acquired over the next seven years so that eventually the entire staff had access to a personal computer. Data storage was problematic, with inadequate hard drives and a policy that limited the number of available diskettes. With no linkage to a common server, file sharing was difficult. Reusing diskettes was a management policy as diskettes were "expensive" and long term data storage was a foreign concept. Even as technology advanced and file sizes grew rapidly, only one CD writer was made available to the entire staff of the GGS. Files from a PC were transferred to a server and then to another PC where the CD writer was installed. This procedure remains as computers and other related hardware have not been updated since 1999. As with the hard copy data files, the GGS did not develop a strategy for planning how data were to be stored, backed-up, or archived. Software acquisition and software training were neglected by GGS management, with few staff members advancing beyond basic word processing and spreadsheet computer literacy.

Migration of data to newer formats is vital, as technology continues to advance and older technology is no longer supported. A change in software approximately 5 years ago resulted in many data files becoming inaccessible or corrupted. Over the course of 10 to 12 years, numerous document files were created with one particular

word processing software, with data files and accompanying graphs prepared with that company's spreadsheet software. Manual entering of large amounts of data into spreadsheets represented a considerable investment of time. A change in software vendors by the State of Georgia resulted in the removal of the previous software, installation of another company's software, and the resultant loss or corruption of a significant amount of digital data.

ARCHIVING DATA FILES

The archiving of data files and collections is a critical function of a geoscience organization. A collection of data files should be easily searchable and accessible. Most of the GGS's data files are referred to as the technical files; these are housed mainly in standard file drawers and flat files (for maps). Despite a recent, multi-year attempt to develop a digital catalog of the technical files, the most effective search technique remains the manual method. The present digital catalog database is an alphabetical file listing and is not searchable by key words, topics, authors, dates, or subject areas. This digital catalog was developed by people with no technical background, and no input from the geologic staff was considered. Recently, a compilation of drill hole data for a selected depth interval in a selected multi-county area required a month-long manual search of file drawers to find and retrieve logs from five different locations. The existence or location of some drill hole logs remains unknown.

PUBLICATIONS

GGS publications should also be regarded as data sources, as these publications contain data unavailable anywhere else. Ideally, data would be archived in a data repository. Depending on the author or the reviewers, some or all of the collected data may be included in the publication. Some publications, e.g. maps, may be compilations of new and older published and unpublished data from a variety of sources. As these publications are data sources, they should also be documented, preserved and made available for access by other geologists and customers. Other manuscripts and maps were at various stages of completion when the GGS was terminated and continue to be published. Still other geologic projects, including STATEMAP mapping, will continue to produce more data and publications.

Publication, sales, and preservation of the publication inventories, require an agency to commit funds, sales staff and space. Documentation is especially important to the customer, in order to search for, and find, what they need. Traditionally, a geological survey's publications are documented in a catalog. The GGS's annual catalog of publications consists of a simple sequential (mainly chronological) listing by type of publication, i.e. bulletin, open-file

report, hydrologic atlas, etc., and by title, author and date. The publications are not arranged by subject matter, e.g. economic geology, or other logical method to quickly find a publication of interest. An annotated bibliography could provide more pertinent information regarding the publications. More recently the GGS produced an on-line catalog, but it is just a digital version of the hard-copy catalog without a key word search. Even the most recent catalog of GGS publications is far from complete, e.g., it does not indicate the existence of 26 geologic maps and 8 open-file reports completed during seven years of GGS participation in STATEMAP mapping. Customers searching the catalog would not be aware of these publications, and the staff servicing the customers would probably also not be aware of these publications. The GGS has not funded either new publications or reprints of older publications that have been sold out. As publications begin to be sold out, at what point will termination of the GGS affect accessibility and availability of their publications and data?

STORAGE OF COLLECTIONS

Drill core and cuttings, petrologic, mineral and fossil collections belonging to the GGS have been stored in a non-climate controlled warehouse in Atlanta, GA. Many core and cuttings boxes are up to 40 to 60 years old and have suffered the effects of high temperatures, high humidity, dust from nearby industrial activities, and neglect. Because of poor lighting, security issues, access, air-conditioning and heating, and the lack of other basic facilities, the warehouse has never served as a research facility. Project files (data), mylar originals of published maps (required for reprints), office files, rare and historic USGS Professional Papers and Bulletins, etc., bound professional periodicals, and excess older GGS publications also have been semi-permanently stored at the warehouse. Deterioration of materials and data over time has been inevitable.

Prior to the termination of the GGS, an unknown quantity of drill core and cuttings, rock, mineral and fossil collections, maps, project files, equipment, and GGS publications were discarded, as a result of a lack of interest and understanding by decision-makers regarding the present and future value of the data and collections.

INSTITUTIONAL MEMORY

The institutional memory of a geoscience organization consists of: the undocumented experiences, observations and interpretations that are accumulated by an organization's personnel mainly during the course of their field and laboratory work, conversations with colleagues both within and outside their organization, knowledge gained at professional meetings, and reading or knowledge of pertinent published literature and unpublished

or “gray” literature. Institutional memory also includes other types of information and knowledge such as road or property access, new roadcut or other excavation exposures outside of one’s current study area, and professional contacts outside the agency (e.g., consultants and industry geologists, who may have little or no publication record). A discussion of what constitutes institutional memory is open-ended, but essentially it is undocumented knowledge and expertise that, in order for the organization to survive and flourish, can (and must) be passed on to other personnel.

During the past 25 years at the GGS, an unknown and immeasurable amount of institutional memory was permanently lost, as experienced geological personnel were reassigned, retired, or moved on to new employment. Nearly all of the GGS geologists who were reassigned or acquired employment with other Georgia state agencies have retired or are within a few years of retirement. In this author’s experience, the institutional memory of former staff generally fades rapidly with time. Currently, the two remaining GGS geologists have about 10 years to retirement age with no new or potential opportunities for new geological staff to pass on this institutional memory.

RECOMMENDATIONS FOR PRESERVATION AND INCREASED AVAILABILITY OF GEOSCIENCE DATA AND COLLECTIONS IN GEORGIA

These recommendations may be specific to Georgia because of the current circumstances, but may serve as a guide if other state geological surveys risk termination:

- Restore a GGS that has a legislative mandate to collect, preserve, document, and disseminate geoscience data and collections.
- Adequately fund and staff a geological survey to be able to achieve that mandate.
- Construct a climate-controlled core warehouse facility with a permanent, geologically-trained staff.
- Provide training for new and current employees and encourage retention of employees or emeritus employees to preserve and transfer institutional memory.
- Scan older publications, maps, and data files and make them available in a digital format.
- Update older digital files to current standards and maintain newer digital files with newer digital formats and technological advances.
- Develop a digital on-line catalog of GGS publications, geoscience files, and archival maps, etc., such as that developed by the North Carolina Geological Survey (available at <http://www.geology.enr.state.nc.us/>).

REFERENCES

- Musser, L.R., 2003, Preservation of Geoscience Data and collections, in Soller, D.R., ed., *Digital Mapping Techniques '03—Workshop Proceedings*: U.S. Geological Survey Open File Report 03-471, p. 195-196, available at <http://pubs.usgs.gov/of/2003/of03-471/musser/index.html>.
- National Research Council, Committee on the Preservation of Geoscience Data and Collections, 2002, *Geoscience Data and Collections—National Resources in Peril*: Washington, D.C., National Academy Press, 107 p., available at <http://www.nap.edu/catalog/10348.html>.