

**Kansas Geological Survey
Open File Report 96-45**

Digital Geologic Map Data Capture:

A summary of the October 1996 meeting of the
Digital Geologic Map Standards Committee
and Working Group on Data Capture
American Association of State Geologists
and United States Geological Survey

Including the report,
National Geologic Map Database Project, Summary of FY 1997 Project Plan
by David Sollars, USGS

prepared by

Jorgina A. Ross, Manager
Automated Cartography Group

October 31, 1996

Disclaimer: The Kansas Geological Survey does not guarantee this document to be free from errors or inaccuracies and disclaims any responsibility or liability for interpretations based on the data used in the production of this document or decisions based thereon. This report is intended to make results of research available at the earliest possible date, but is not intended to constitute final or formal publication.

INTRODUCTION

The Working Group on Data Capture of the Digital Geologic Map Standards Committee established by the American Association of State Geologists and the United States Geological Survey met in Denver, Colorado, in conjunction with the annual meeting of the Geological Society of America.

During the GSA convention, the Kentucky Geological Survey and Ohio Geological Survey each provided demonstrations of their techniques for capture of digital geologic map data. Other state surveys (notably Illinois) provided computer demonstrations of their digital geologic map data and its applications. The Kansas Geological Survey provided demonstrations of the accessibility and use of digital geologic data through on-line connection to the KGS home page on the Internet. These demonstrations provided opportunities for informal discussions among members of the Working Group on Data Capture.

A formal meeting of the Digital Geologic Map Standards Committee and its various working groups was held on the evening of October 29, 1996. The Committee meeting and meeting of the Working Group on Data Capture were chaired by David Sollars, USGS, project chief for the National Geologic Map Database Project. The following meeting notes provide an informal report on the discussions at this meeting. The official minutes of the meeting will be included when they are made available by the chairman. Also included in this open-file report is a copy of the report, *National Geologic Map Database Project, Summary of FY 1997 Project Plan*, distributed by David Sollars.

MEETING NOTES

The October meeting of the Digital Geologic Map Standards Committee was convened at 7:00PM, October 29, 1996, at the Hyatt Regency Hotel in Denver. David Sollars, USGS, chaired the meeting.

David Sollars provided an overview of the distributed database design currently anticipated for the National Geologic Map Database (NGMDB). He envisions a system which provides user access to a searchable catalog of maps and data (both paper and digital) with links to data archives throughout the nation (USGS, state geological surveys, libraries, etc.). A prototype of this system is currently located on the world wide web at:

<http://wwwflag.wr.usgs.gov/ngmdb/>

Sollars emphasized that the current project is not an effort to impose mapping requirements on state surveys and other organizations from above. The objective is to develop a flexible system which facilitates access to all geologic map data while providing the necessary information to users (or potential users) of the data to insure appropriate use of the data.

Sollars distributed a copy of the attached report, *National Geologic Map Database Project, Summary of FY 1997 Project Plan*," and briefly highlighted its content.

Of particular interest to the Kansas Geological Survey in the initial development phase of the NGMDB is Section III. (*Populating the catalog database*) on page 2 of the report. This activity, which involves creation of bibliographic records for published

maps in all formats, is of direct relevance to the Kansas Geological Survey's librarian and the ongoing development of the *Bibliography of Kansas Geology*. Also of immediate interest are Section V.D. (page 4) relating to initial state geological survey participation in the implementation of NGMDB links to state survey web pages; and Section VI.B. (page 5) providing for links to state geologic maps (Geologic Map of Kansas: KGS Map Series M-23).

The committee meeting was then opened for informal discussions among participants before breaking into separate meetings of working groups. At this point numerous representatives of other state surveys asked for explanations of the techniques used at the Kansas Survey. There was an extended discussion with the Kentucky Survey's programmer/systems analyst for GIS development.

The techniques used by the Kentucky Survey involve scanning of stable-base mylars which have been photo processed in some way to remove all but the geologic contacts on Kentucky's geologic quadrangles. (Kentucky is in the apparently enviable position of having already produced published geologic maps at 1:24,000 for each of the USGS 1:24,000 topographic quadrangles covering the state.) The scanned image is then vectorized. In a process described as "lengthy," staff members then clean and annotate the resulting line work. The time required to capture the geologic information in digital form and produce a computer generated plot file for a map of a single quadrangle amounts to about five weeks.

Apparently believing that the Kansas Survey used similar techniques, the staff member from Kentucky remarked that there were probably still some surveys that used the "old" manual digitizing technology. At best, he was surprised when informed that the Kansas Survey used manual digitizing as the basis of its automated cartography operations. When shown a completed, 1:24,000 scale, geologic map of the Gladstone quadrangle in Chase County, derived from R. C. Moore et al, (1951), and informed that the entire process (see steps 1 - 4 on pages 3 and 4 of the following report by Ross) took about five days, he was even more surprised.

James Cobb, Kentucky's Assistant State Geologist, subsequently expressed his concern that the procedures which they were using were very costly and time consuming. Dr. Cobb had a sample of the end product for one of their computer generated geologic quadrangles. A fair assessment of the product would suggest that while containing the basic information elements desired for such a product, the presentation of that information was rather primitive.

One important fact became apparent from the discussion period. To some extent, each state survey had some unique concerns and needs in terms of desirable end products. In that regard it should be expected that different concerns should exist regarding the type of information to be captured in digital form and the extent of quality assurance efforts associated with the process. On the other hand it appeared that much benefit could be achieved from a more extensive exchange of information regarding results of various data capture techniques.

At this point, individual working groups began their independent meetings. The following notes relate to the meeting of the Working Group on Data Capture.

Principle participants in the discussions of this group were:

David Sollars, USGS (Reston), Chairman
Patrick Muffler, USGS (Menlo Park)
Rob Krumm, Illinois Geological Survey
Barbara Stiff, Illinois Geological Survey
Gina Ross, Kansas Geological Survey
David Collins, Kansas Geological Survey
James Cobb, Kentucky Geological Survey
C. R. Berquist, Division of Mineral Resources, Commonwealth of Virginia
Gayle McColloch, West Virginia Geological & Economic Survey

The report, *Compilation of Digital Geologic Map Data at the Kansas Geological Survey*, (Ross, 1996) was distributed to the Working Group on Data Capture. At the request of the chairman, plans were made for a meeting, sponsored by the working group, to be held in late January or early February at which participants would present reports summarizing critical aspects of their data capture operations. The meeting is tentatively planned for St. Louis, although there was some discussion of the possibility of having the meeting at a state survey (with interest in either Illinois or Kansas). The chairman expects that attendance would be around 100.

The critical aspects of data capture to be covered in the reports (each about 5 pages with a couple of illustrations, to be compiled beforehand into a USGS open-file report) included:

Relation of data capture operations to mission, client needs, etc.
Descriptions of source documents (scale, date, detail, etc.)
Quality/accuracy assessment of source documents
Summary of data capture techniques
Quality control procedures
Hardware, software and materials requirements
Staff requirements
Time and dollar costs
Description of final digital products

The chairman emphasized that while other working groups had objectives which involved delivery of a specific product such as a 'standards' document, the objective of the Data Capture group was to inform and provide the guidance of practical experience to the other working groups. It was also suggested that some surveys (such as Kansas) with a long history of computer mapping might present other reports at the winter meeting which would review changes in perspective with time and changing technology.

There was considerable discussion regarding the adequacy of information (metadata) carried along with digital geologic data to advise potential users of the quality of the data and help insure appropriate use (and avoid inappropriate use) of the data. The chairman commented on several occasions that the Working Group on Metadata (Tom Mettillie represents KGS on that working group) is too small and needs broader input.

In a wrap up session of the Digital Geologic Map Standards Committee other working groups reported plans for similar meetings. Chairmen of different working groups

indicated they would be notifying participants of future plans by e-mail. There was considerable interest in the issue of appropriate information content of geologic maps, nomenclature, symbology, etc. David Sollars indicated that the current list of working groups was not fixed, and that other groups would probably be formed as work proceeded. In particular, he indicated the need to wrap up efforts which have apparently been going on for some time to establish a recommended (i.e., USGS approved?) set of geologic map symbols and recommended colors or patterns to be used in representation of different age rock units.

There was much discussion before the meeting concluded of the need to keep proceedings of the committee and working groups as open and inclusive as possible to avoid adverse reaction due to lack of understanding of the committee objectives. Minutes of the meeting, including minutes of working group sessions are to be placed on a home page to be established by David Sollars and his staff.

The meeting adjourned at about 10:00PM.

References

Ross, J.A., 1996, Compilation of Digital Geologic Map Data at the Kansas Geological Survey, Open File Report 96-58, Kansas Geological Survey, 30 pp.

National Geologic Map Database Project Summary of FY 1997 Project Plan

Background

Objective: Develop the National Geologic Map Database (NGMDB) as a comprehensive reference tool and data management system for spatial geoscience information in paper and digital form.

Strategy: This project has been designed as a set of related tasks that will develop, over time, a NGMDB with increasing complexity and utility. This is being accomplished through a network of geoscientists and librarians committed to supporting the objectives of the NGMDB. The first phase of this Project is the building of a comprehensive catalog of all available paper and digital maps, and book publications containing maps, that adhere to the earth-science themes specified in the National Geologic Mapping Act of 1992. The second phase is the identification, documentation, and archiving of existing digital maps, according to a simple, low-level set of format standards. The third phase will be the development of National geologic map coverage at a variety of map scales (primarily 1:100,000-scale), according to a more complex set of content and format specifications that are standardized through general agreement among all partners in the NGMDB (especially including the Association of American State Geologists, or AASG, and Geologic Division programs). During the development of these phases of the NGMDB, extensive work will be conducted to develop a Web interface and search engine and to continually improve it, and to develop the data management and administrative protocols necessary to ensure that the NGMDB will function efficiently in the future.

I. Project management

The project chief (Dave Soller) will continue to provide the primary input to NGMDB design, and will coordinate among the various project participants to ensure that project objectives are met. The state geological surveys are primary collaborators in the NGMDB; Tom Berg (State Geologist of Ohio and Chair of the AASG Digital Geologic Mapping Committee) is the principal liaison to the NGMDB project.

II. Software development and data management

This activity will be conducted by a team of six persons (Ed Pfeiffer, Chris Isbell, Dennis McMacken, Jim Mathews, Alex Acosta, and Janet Barrett) residing in various administrative units in Flagstaff. For FY 1997, their primary objectives will include revision of the NGMDB Web site and assessment of new graphical search techniques that will permit enhanced customer utility, adopting a new database management software (Oracle), designing the map catalog data structure and managing its records, and beginning to implement links to other servers so that the NGMDB can become, as intended, a distributed system.

III. Populating the catalog database

The first phase of this Project is the building of a comprehensive catalog of all available paper and digital maps, and book publications containing maps, that adhere to the earth-science themes specified in the National Geologic Mapping Act of 1992. The USGS library has agreed to build the map catalog with assistance from map indexers hired by this project to support the primary catalog compiler in Menlo Park, CA.

The Project Chief, the three regional librarians, and the software development team in Flagstaff have decided that the map catalog will include a small subset of metadata for each paper and digital map. These metadata (perhaps better described as bibliographic listings) will provide sufficient information to perform geographic and topical searches of the catalog database.

Data for the map catalog will be largely compiled in Menlo Park, under the guidance of the regional librarian, Nancy Blair. To support this activity, Nancy has written a draft instruction manual for persons creating the bibliographic records. Entry of records received from other contributors, in other USGS libraries and in State Geological Surveys and Universities, will be checked for conformity and accuracy in Menlo Park. At regular intervals, these entries will be transferred to the NGMDB server in Flagstaff.

The library will create bibliographic records for published maps in all formats. Work will be prioritized as follows: USGS publications, State Geological Surveys and related state agencies, geological organizations, other published sources of earth-science mapping, and non-published maps (e.g., theses). In FY97, the library intends to complete the instruction manual, complete work on all USGS publications, work with State Geological Survey librarians to develop the bibliographic listings for their publications, and, to the extent possible, conduct work on the three lower-priority categories listed above.

IV. Development of standards

A variety of standards are needed to facilitate the building of a NGMDB. Through extensive discussions and preliminary work with the AASG's Digital Geologic Mapping Committee, each Division program (through their appointed Program representative to the NGMDB), the Federal Geographic Data Committee's (FGDC) Standards Working Group, and others (including the Geological Survey of Canada), this project now serves as the focal point for the development of geologic map standards. At a USGS/AASG meeting in St. Louis in 8/96, attended by members of the NGMDB project, the AASG Digital Geologic Mapping Committee, the USGS National Cooperative Geologic Mapping Program, the USGS Mineral Resource Surveys Program, and the Geological Survey of Canada, five standards working groups were begun (items B-F, below). Standards activities are listed below. Working group members are listed in the Participants section attached to the end of the document.

A. Digital cartographic standards.

USGS Open-file Report 95-525 (authored by Mitch Reynolds) was technically reviewed in 1996 under an effort coordinated by Dave Soller for the FGDC's Geologic Data Subcommittee. In FY97, review comments will help guide the revision of that document for submission as FGDC draft cartographic standards for geologic maps. Document revision will occur in early FY97, with submittal to the FGDC planned for 1997. Concurrently, GIS users and the regional Publications Groups will be polled to determine if these standards should be converted to formats useful in GIS

and graphics illustration software widely used in the Division. Dave Soller will coordinate this on behalf of the FGDC.

B. Metadata.

A USGS/AASG working group has been formed to develop guidance and tools for creating metadata. The group is chaired by Peter Schweitzer (USGS). As stated in the minutes of the 8/96 St. Louis meeting, this is the group's charge:

- 1) examine the FGDC metadata content standard for adequacy
- 2) examine how to implement metadata, in a standard format for geologic map information
- 3) establish guidelines that interpret the meaning of the metadata elements, and provide examples
- 4) bring into the discussion the states not represented at the meeting
- 5) for maps destined for the NGMDB map catalog, develop a metadata form.

C. Guidelines for digital geologic map distribution.

A USGS/AASG working group has been formed to develop standards for a common suite of digital map files and ancillary files and a common file "packaging" (e.g., all files contained in a compressed UNIX tar file). This is designed to aid the user by improving the commonality of digital publications served on the Internet by the various NGMDB cooperators. The group is chaired by Todd Fitzgibbon (USGS). This is the group's charge:

- 1) convene a meeting where a draft document which lays out guidelines for the release of geologic map data is written
- 2) release guidelines document as a USGS Open-file report (by Sept. 30, 1997).

D. Data capture.

A USGS/AASG working group has been formed to promote coordination and progress in digital mapping techniques. The group is chaired by Dave Soller (USGS). As stated in the minutes of the 8/96 St. Louis meeting, this group's charge is to:

- 1) provide guidance on data-capture techniques through facilitation of information exchange
- 2) conduct workshops to share experiences and insight and to evaluate technology. For example, a formal workshop on digital mapping techniques will be held in late winter, 1997; broad participation from the USGS and AASG will be sought.

E. Spatial accuracy.

A USGS/AASG working group has been formed to investigate issues of spatial accuracy on geologic maps. The group probably will be chaired by Dick Berg (Illinois State Geological Survey). As stated in the minutes of the 8/96 St. Louis meeting, this group's charge is to:

- 1) produce a general-interest publication that explains the accuracy of geologic maps
- 2) investigate technologies to improve spatial accuracy of geologic maps
- 3) look at issue of disclaimers on spatial accuracy, and the use and misuse of map data.

F. Data model.

A USGS/AASG working group has been formed to develop standards for digital geologic map content and data structure. The group is co-chaired by Todd Fitzgibbon and Gary Raines (both USGS). The group has the general mission to identify the "core" elements common to all geologic maps and legends, and to build a data model that can be implemented by all cooperators

in the NGMDB. This is a complex task requiring extensive research and development; preliminary results will be available in 1997, but the entire model may require 2 years for full development.

V. Populating the NGMDB with existing digital data, according to interim, minimal standards

The second phase of this Project is the identification, documentation, and archiving of existing digital maps, according to a simple, low-level set of format standards. Digital data currently exists in a wide variety of formats. Some is published and is of high quality whereas some data sets were used solely as an interim step during production of a conventional paper map. To address the project's second phase, the Project will:

A) Identify digital geologic map data produced by the Geologic Division, and evaluate that data for linking to, or serving through, the NGMDB. A project member from the Eastern Mapping Team (Steve Schindler), the Central Mapping Team (Ron Wahl), and the Western Mapping Team (Todd Fitzgibbon) will work in collaboration with the regional publications groups and, with Soller's assistance, in collaboration with the six Division program representatives to the NGMDB.

B) Work with the regional publications groups to develop a network of computers to archive and serve the data, in support of both the NGMDB and, more generally, in support of USGS interests in serving published data. Fitzgibbon, Wahl, and Schindler will work on this task. Work with the six program representatives to achieve this goal collaboratively and to link the NGMDB to databases currently maintained by those programs; Soller will assist by coordinating with the program representatives.

C) For map data to be made accessible through the NGMDB, standard metadata and ancillary files must be created according to pending guidelines for digital map distribution (see activity IV-C, above). Fitzgibbon, Wahl, and Schindler will coordinate the preparation of digital maps (identified in activity V-A) for serving through the NGMDB.

D) To enable the NGMDB to serve digital data held by state geological surveys, Soller and the AASG's Digital Geologic Mapping Committee will identify selected states that will work with the NGMDB project in the initial implementation of a database link to state Web pages and data servers. This link will require the standards or guidelines for serving digital data, developed in collaboration through the working group described in activity IV-C. As technical issues are explored and addressed through this initial implementation, additional states can then be linked to the NGMDB.

VI. Development of standardized, National geologic map coverage

The third phase of this Project will be the development of national geologic map coverage at a variety of map scales (primarily 1:100,000-scale), according to a set of content and format specifications that are standardized through general agreement among all partners in the NGMDB (especially including other Division programs and the AASG). As noted in activity IV, above, these standards will be under development in FY97. While standards are being developed, this Project will begin to create small- and intermediate-scale digital geologic maps of regional and National scope; standards will be applied to these data when applicable. The NGMDB will undertake the following activities in FY97:

A) 1:100,000-scale geologic map of the Nation

This map will become the primary digital geologic map of national scope built and maintained by the NGMDB. Work will be conducted primarily by Dave Soller and Susan Price. This activity will begin by developing an inventory of all available paper and digital geologic maps at this scale. At the end of FY97, an interim report and map showing the status of 1:100,000-scale geologic mapping will be produced. A prioritized conversion of paper maps to digital form will then be undertaken, in coordination with the AASG and other Division programs. Upon widespread acceptance of a standard geologic map data model, all digital quadrangle maps will be converted to this model. The potential for developing a single coherent national map from these quadrangles will then be discussed among NGMDB cooperators.

B) State geologic maps

These maps are a valuable reference tool for the States and the USGS, and should be made available as digital files. The state maps will be inventoried and assessed in a process similar to the plan for the 1:100,000-scale maps; the work will be conducted primarily by Dave Soller and Susan Price.

B) Small-scale regional and national geologic maps

Selected small-scale maps of large regions or the Nation (e.g., at 1:1,000,000-scale and smaller) should be developed into digital map files for regional analysis and planning, and for general education. Currently, only the King-Beikman map is available. The following maps will be addressed by the Project in FY97:

- 1) *Thickness and character of Quaternary sediments in the glaciated U.S. east of the Rocky Mountains (1:1,000,000-scale)*. This map is in final preparation for conventional printing as USGS Map I-1970 A-D. Digital files used to print the map have been documented with FGDC-standard metadata. 230 map files and assorted images and text are being prepared for publication on CDROM as USGS DDS-38, and for serving through the NGMDB. Patricia Packard and Dave Soller will continue on this effort, which should conclude by early 1997.
- 2) *Quaternary Atlas of the United States (1:1,000,000-scale)*. Roughly 3/5ths of this series has been compiled and printed conventionally. Compilation of the remaining quadrangles will continue under a separate project. The NGMDB will focus on digital conversion of the printed maps and on developing interim products that may include a CDROM of map images and data. Chuck Bush will have primary responsibility for this effort. Patricia Packard and Dave Soller will provide technical assistance.
- 3) *Geologic Map of North America (1:5,000,000-scale)*. This map in four quadrants is nearly compiled, and work has begun to convert it to digital format.

VII. National Paleontologic Database

The National Paleontological Database (NPD) and the Geological Names Lexicon will be built, by Bruce Wardlaw and Nancy Stamm (USGS). The prototype for both databases has been designed, and work in FY97 will concentrate on: 1) scanning and editing E&R documents for inclusion in the NPD; 2) accessing data on the CD-ROM "GNULEX" (Lexicon) and converting it to format for the modern Lexicon; 3) revising the database structure for the NPD and Lexicon as needed; and 4) setting up the databases on a Web server and developing links to the NGMDB.

VIII. National Geochronologic Database

In Fall, 1996, John Sutter will convene a meeting of selected Division geochronologists, to develop a plan for an on-line database. The existing database on CDROM will be evaluated, to determine whether the database design is sufficiently robust and modern to serve the needs of geochronologists. It is anticipated that in FY97 the appropriate database structure, search protocols, and presentation software will be determined, and a plan for populating the database will begin to be invoked.

IX. Linking geochemical and geophysical databases to the NGMDB

As noted in activity V, Soller will coordinate with the six program representatives to the NGMDB. Appropriate databases will be identified and linked to the NGMDB, thereby addressing requirements of the National Geologic Mapping Act.

X. Status of geologic mapping activities

Knowledge of mapping activities across the Nation's geoscience community ("who is mapping where") is important both to the Division and for inquiries into the availability of geologic maps. The NGMDB will provide this service in support of its general intention to provide users with access to all available information on geologic maps. John Sutter has begun this effort by contacting the leaders of STATEMAP, EDMAP, and FEDMAP projects for information on those projects. A database containing status of mapping information will be linked to digital quadrangle index maps for the Nation, enabling a user to search the database graphically or textually.

For further information, contact the project chief (Dave Soller, USGS, 703-648-6907).

Participants in standards working groups (*NOTE: this is the initial list of participants from the St. Louis meeting; additional participants will be added as the working groups develop their plans*):

Metadata Working Group

Peter Schweitzer, USGS (Chair)
Greg Hermann, New Jersey GS
Rob Krumm, Illinois GS
Tom Metille, Kansas GS

Guidelines for digital geologic map distribution Working Group

Todd Fitzgibbon, USGS (Chair)
Ian Duncan, Virginia GS
Gene Ellis, USGS
Ron Hess, Nevada GS
Bernie Hoyer

Data capture Working Group

Dave Soller, USGS (Chair)
Warren Anderson, Kentucky GS
Elizabeth Campbell, Virginia GS
Rob Krumm, Illinois GS
Gina Ross, Kansas GS
Dave Wagner, California GS
Tom Whitfield, Pennsylvania GS

Spatial accuracy Working Group

Dick Berg, Illinois GS (Chair)
Warren Anderson, Kentucky GS
John Davis, Kansas GS
James McDonald, Ohio GS
Gary Raines, USGS
Jay Raney, Texas GS
Bill Shilts, Illinois GS
Dave Wagner, California GS

Data model Working Group

Todd Fitzgibbon, USGS (co-Chair)
Gary Raines, USGS (co-Chair)
Boyan Brodaric, Geol. Survey of Canada
Jim Cobb, Kentucky GS
Greg Hermann, New Jersey GS
Bruce Johnson, USGS
James McDonald, Ohio GS
Bill Shilts, Illinois GS
Ron Wahl, USGS