Information Technology and Libraries

Volume 14 Number 2 ISSN 0730-9295

June 1995

Information Technology and Libraries (ISSN 0730-9295) is the official publication of the Library and Information Technology Association, a division of the American Library Association, 50 E. Huron St., Chicago, IL 60611; *Executive Director:* Linda J. Knutson. The journal is issued quarterly in March, June, September, and December.

Information Technology and Libraries publishes material related to all aspects of library and information technology. Some specific topics of interest are: Automated Bibliographic Control, AV Techniques, Communications Technology, Cable Systems, Computerized Information Processing, Data Management, Facsimile Applications, File Organization, Legal and Regulatory Matters, Library Networks, Storage and Retrieval Systems, Systems Analysis, and Video Technologies. *ITAL* welcomes unsolicited manuscripts. Submissions should follow the guidelines stated under "Instructions to Authors" on page 70 of the March 1995 issue. Manuscripts of articles, communications, and news items should be addressed to the editor: Thomas W. Leonhardt, *Information Technology and Libraries*, Bizzell Library, 401 W. Brooks, University of Oklahoma, Norman, OK 73019-0528; (405) 325-2611; Internet: leonhardt@aardvark.ucs.uoknor.edu. Copies of books for review should be addressed to: Susan B. Harrison, *ITAL Book Reviews*, The New York Public Library, 455 Fifth Ave., New York, NY 10016. Copies of software for review should be addressed to: George S. Machovec, *ITAL Software Reviews*, Colorado Association of Research Libraries, 3801 E. Florida Ave., Ste. 370, Denver, CO 80210-2542. ADVERTISING: Contact Stuart Foster, Ad Sales Manager, and Dolores LaPointe, Ad Coordinator, c/o *Choice*, 100 Riverview Center, Middletown, CT 06457; (203) 347-6933.

As a matter of policy, *Information Technology and Libraries*, as the scholarly organ of LITA, does not review LITA publications. Notice of new publications from LITA will generally be found in the "Other Recent Receipts" column following reviews.

Information Technology and Libraries is a perquisite of membership in the Library and Information Technology Association. Subscription price, $22.50, is included in membership dues. Nonmembers may subscribe for $50 per year in the U.S.; $55 in Canada, Mexico, Spain, and other PUAS countries; $60 in other foreign countries. Single copies, $15. Second-class postage paid at Chicago, Illinois, and at additional mailing offices. POSTMASTER: Send address changes to Information Technology and Libraries, 50 E. Huron St., Chicago, IL 60611.

Circulation and Production: ALA Production Services (Eileen Mahoney; Dianne Rooney; Bruce Frausto, and Donavan Vicha), American Library Association, 50 E. Huron St., Chicago, IL 60611.

Publication of material in *Information Technology and Libraries* does not constitute official endorsement by LITA or the ALA.

Abstracted in *Computer & Information Systems, Computing Reviews, Information Science Abstracts, Library & Information Science Abstracts, Referativnyi Zhurnal, Nauchnaya i Tekhnicheskaya Informatsiya, Otdyelnyi Vypusk,* and *Science Abstracts Publications.* Indexed in *CompuMath Citation Index, Computer Contents, Computer Literature Index, Current Contents/Health Services Administration, Current Contents/Social Behavioral Sciences, Current Index to Journals in Education, Education, Library Literature, Magazine Index, New Search,* and *Social Sciences Citation Index.* Microfilm copies available to subscribers from University Microfilms, Ann Arbor, Michigan.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences–Permanence of Paper for Printed Library Materials, ANSI 239.48-1992. ∞

Information Technology and Libraries

Volume 14, Number 2 | June 1995 | ISSN 0730-9295

SPECIAL ISSUE
Making GIS a Part of Library Service

Cover Design by Jim Lange. Interior Design by Dianne M. Rooney

Literacy Project, provides insight into the complexity of collection management and access services for digital spatial information at the Map and Geographical Information Center at the University of Connecticut Library. Both of these articles focus on the key issues libraries are confronting in their efforts to introduce GIS to their patrons and to make effective use of digital spatial data.

As the articles in this special issue demonstrate, GIS is an advanced, multimedia software that integrates a wide variety of technologies into a system for building and accessing digital libraries. These applications deal with large quantities of information in heterogeneous formats (images, text, tabular) and provide rapid access across distributed information from multiple platforms. These systems also provide a level of intelligence in information analysis and can be customized to meet evolving user needs.

The information about GIS and various related initiatives throughout the United States referenced in these papers shows that now is the moment of opportunity for information professionals to take bold and innovative steps to insure that programs are in place to meet the expanding information needs of our citizens. These papers also show that there is strong agreement that GIS initiatives are, in large part, a collaborative effort among a variety of information professionals and that each type of information professional has a unique contribution to make.

Note

1. National Research Council, *Promoting the National Spatial Data Infrastructure Through Partnerships* (Washington, D.C.: National Academy Press, 1994), 15.

Digital Spatial Libraries: A Context for Engineering and Library Collaboration

Kate Beard

At first glance, engineers and librarians would seem to be worlds apart in their interests. A revolution in information technology, however, is generating a convergence in interests and an environment for interesting and necessary collaborations. Increasingly, traditional analog library tools and mechanisms are being replaced by digital counterparts, some of which call for expertise beyond the domain of library science. The National Science Foundation (NSF) together with the National Aeronautics and Space Administration (NASA) and the Advanced Research Projects Agency (ARPA) have funded six digital library projects that are being carried out by multidisciplinary teams made up of librarians, engineers, computer scientists, and professionals from several other disciplines. This paper begins with a review of the technological innovations that have allowed the emergence of digital libraries and fostered new multidisciplinary collaborations. The paper identifies a few key areas where collaborations are occurring among librarians and engineers, and illustrates these with examples from digital spatial library projects at the University of California, Santa Barbara (Project Alexandria) and the University of Maine (the BASIN Project). Particular emphasis is placed on digital spatial libraries, which have more unusual requirements for information processing. Collaborations in support of digital spatial libraries involve librarians, map librarians, and professionals in a new field of engineering referred to as spatial information engineering.

Changing Information Technology: A Context for Collaboration

Technology is acknowledged as a powerful force in changing political and organizational dynamics and provoking realignments in group and individual behaviors. Information technology may be similarly described. Significant changes in information technology and communications infrastructures are affecting the behaviors of business, government, academia, and individual citizens. Some of these changing dynamics foster new and interesting collaborations, in some cases creating unexpected bedfellows. Recent examples include the shifting mergers and alliances among the traditional telephone companies; the cable television, cellular communications, and video distribution companies; and the entertainment and publishing industries. A less visible, but nonetheless auspicious, collaboration is the one forming between librarians and engineers.

Engineers and librarians are two groups that one generally does not associate as having collaborative interests. Librarians have been quite secure in their own world of collecting, cataloging, and distributing information in analog formats; and engineers have likewise been content in their domain. Spatial information engineers have traditionally been involved in geodesy, photogrammetry, surveying, cartography and, more recently, GIS. Enter information technology.

The tremendous rate of new developments in computing and telecommunications technology has forced a transition to new modes of operation for libraries. Increasingly, information is being created and offered in a digital format without ever migrating to a paper format. This transition to a digital environment has, in and of itself, provoked significant changes in the tools and roles of libraries. Another significant influence, however, has come from the development of the National Information Infrastructure (NII) (U.S. Congress 1993), which can now support the distribution of data to users, without the need for them to ever physically visit the library.

The NII consists of computers, software and databases, fax machines, local area networks, access networks, and regional and national networks embodying various technologies with speeds in the hundreds of gigabits per second (Kettinger 1994). Formally introduced by the High Performance Computing Act of 1991, which established the National Research and Education Network (NREN), the NII is being strengthened by the Clinton administration in an effort to build an information highway that will carry pictures, voice, video, and textual data anywhere in the United States. Collections available for distribution include materials from the Library of Congress, the National Library of Medicine, and Congressional Research Service databases. The number of accessible collections has continued to grow with the addition of global change datasets, climate and weather databases, human genome data, and other databases. Recent additions of wide-area information services (WAIS, gopher, WWW) have broadened the extent and

Kate Beard is Assistant Professor, Department of Surveying Engineering and National Center for Geographic Information and Analysis, University of Maine, Orono, Maine.

functionality of network communications, and the rate of data contributions is accelerating daily.

The evolving NII already provides a great deal of connectivity throughout the country. Between 1988 and 1992, NSFNET, the predecessor to the Internet, went from a T1 backbone (1.55 megabits per second) connecting 400 universities to a T3 (45 megabits per second) backbone connecting over 4,500 universities, businesses, libraries, government agencies, schools, and museums. Traffic increased during this period from 100 million packets per month to almost 12.5 billion packets per month (NSFNET 1993). By February 1993 this volume had again doubled.

In addition to the role that the rapid pace of technological improvements played in promoting conversion to digital libraries, further impetus came in 1993 in the form of congressional acts and executive orders. The Congress, in its Electronic Library Act of 1993 (S.626), and the president and vice president, in their report on "Technology for America's Economic Growth," called for the development of digital libraries. Subsequent white paper series (Fox 1993) and various NSF-sponsored workshops helped detail and define a research agenda for digital libraries, which were identified as a national challenge in the Information Infrastructure Technology Applications component of the U.S. High Performance Computing and Communications Program. In late 1993 NSF, NASA, and ARPA jointly sponsored a call for proposals on the topic of "digital libraries." The focus of this project was to dramatically advance the means to collect, store, and organize information in digital forms, and to make it available for searching, retrieval, and processing via communications networks in a user-friendly way. Six projects, centered at Carnegie-Mellon University, the University of California Berkeley, the University of California, Santa Barbara, Stanford University, the University of Illinois, and the University of Michigan, were funded under this initiative. Each of these projects involves substantial collaboration and interaction of librarians, computer scientists, engineers, and others.

The Emergence of Digital Spatial Libraries

Pressure for digital spatial libraries has come from numerous sources. Rapid growth in the volume of GIS applications and users has increased demand for spatial data and for the transmission of data over the networks. In particular, demand for better management and access to federally produced geospatial data provoked a call for a National Spatial Data Infrastructure (NSDI). The initiation of NSDI was specified in President Clinton's Executive Order 12906, issued in April of 1994. Conceived of as an umbrella of policies, standards, and procedures under which organizations and technologies would interact to foster efficient use, management, and production of geospatial data (FGDC 1994), NSDI is a vision for the future of a spatial data community that has broad implications for the underlying technology, user base, and necessary support services. Infrastructure services called for under NSDI include:

- collection of raw data and provision of value-added data;
- development of mechanisms to create, archive, and distribute the collected data and provide for creation of associated metadata;
- provision of capabilities to locate, browse, search, retrieve, preview, and transfer spatial data;
- production of well-defined, fixed products and ability to generate custom one-of-a-kind products; and
- ability to integrate, manipulate, augment, and generalize spatial data based on domain-specific needs (Mularz et al. 1995).

The Federal Geographic Data Committee (FGDC), which was charged with overseeing and facilitating the NSDI, has several ongoing efforts including development of standards, framework data, thematic datasets, and a National Geospatial Data Clearinghouse (Tosta 1994a). The clearinghouse is an effort to use the NII to facilitate access to spatial data. Through cooperative agreements, the FGDC has funded nine partnership grants totaling $225,000 for 1994. A new round of cooperative grants will be awarded in 1995. The 1994 cooperative agreement partnerships all include components of–or can be classed as–digital spatial library projects. Of these projects, three (Florida, Iowa, and Montana) have libraries as primary partners (for summaries of these projects see Tosta 1994b).

Only three years ago such digital library initiatives would have been impossible to carry out, and today construction of digital libraries is still intellectually and technically challenging. These challenges cannot be met by any one discipline acting alone. The library community is redefining its roles and responsibilities and moving actively to adapt (McClure, Moen, and Ryan 1994), but to be successful it must be prepared to join forces with engineers, computer scientists, cognitive scientists, educators, and others. Conversely, technological developments from computer science and engineering will not be fully successful without input from librarians. As some have suggested, the technological infrastructure that supports the Internet continues to grow at a much faster rate than our knowledge about how to use its resources (McClure 1994). To date, emphasis has been placed on delivery of information rather than on delivery of information services. Only information services that offer true user-friendly software, meet real needs, and make life easier will succeed. The expertise of librarians in managing and disseminating information, built over the centuries, must be merged with sound engineering to ensure success in digital library designs.

Areas of Collaboration

The components of a digital library include assemblages of electronic data, cataloging and indexing mechanisms, tools for locating, searching, and browsing data collections, mechanisms to retrieve and potentially process data from remote and distributed locations, and interface tools that allow tasks to be performed easily by non-expert users. Because each of these components requires expertise from several different disciplines, collaborations are the norm—as is evidenced by several existing projects (Fox et al. 1993). In the following section on Project Alexandria, three specific areas of collaborations are highlighted: data conversion, cataloging, and user interface design. Special consideration is given to collaborations in the development of digital spatial libraries, as these involve interesting "fringe" areas of library science (map librarians) and engineering (spatial information engineering).

Project Alexandria

One of the NSF-funded digital library projects focusing specifically on spatial information, Project Alexandria is a digital spatial library project representing a collaboration among map librarians, computer scientists, geographers, electrical engineers, and spatial information engineers (Smith et al. 1994). The project's long-term goal is to support not only library functions in traditional areas involving the use of digitized maps and images, but also to support library functions in areas that involve the use, for example, of photographs, medical images, or digital representations of art and architectural plans or illustrations in books and journals. While spatially-indexed information has traditionally been defined in a very narrow way—as collections of facts tied to specific locations on the surface of the earth, Project Alexandria will support collections from the whole spectrum of spatially indexed information.

In building a digital spatial data library, many sets of issues that normally arise in the context of analog digital libraries must be resolved. Maps, atlases, and images have posed major problems for conventional analog libraries. They require special forms of cataloging and indexing; their physical dimensions require special storage; and they often deteriorate rapidly with use. Because of these problems, maps and images are stored in special libraries, or are otherwise less than ideally accessible. Ideally, many of these storage and access issues should disappear in a digital world. Storage media are no longer incompatible, and indexing systems can be extended to include spatial queries. Retrieved information can be displayed or input to geographic information systems and used in various types of analysis. In reality, however, many highly specific technical impediments must be overcome in building a digital spatial data library. Such impediments include, for example:

- poor understanding of user requirements and lack of a sound basis for user interface design;
- the lack of appropriate models for both data and metadata;
- technical problems in supporting an appropriate browse capability for distributed access;
- the lack of appropriate spatial indexes and tiling systems to accommodate a wide variety of queries; and
- performance problems due to the large-volume characteristic of spatial information.

In Project Alexandria, each of these impediments is being addressed by multidisciplinary teams. Approaches to resolving some of these impediments are described under the three sections that follow. In some cases, representative examples of collaborations from other digital library projects are included.

Data Conversion

Historically, the task of collection development, regardless of format, belonged to librarians. Electronic library collection development, however, is not as simple as acquiring and organizing print and nonprint materials. One challenge for the digital library comes in converting large repositories of analog data into searchable electronic documents. Scanning is a straightforward solution for simple conversion of analog documents to images, but the addition of intelligence to the resulting images—the indexing of the images for efficient searching and providing efficient storage of images and efficient compression for browsing—involves input from electrical engineers, spatial information engineers, and computer scientists working with librarians.

In the Carnegie-Mellon University Informedia digital library project (Kanade et al. 1994), research is focusing on searching and extracting information from video. The project is using speech recognition to automatically capture and transcribe narrative and dialogue from video footage. Transcriptions are then time-aligned with the video. Natural language processors are applied to improve the transcriptions and identify topics and subtopics in transcript collections. Machine vision techniques are used automatically to identify segment boundaries using beginning or end points of shots, scenes, or conversations. This detailed conversion process is thus involving the input of several players from the specialties of speech recognition, natural language programming, and machine vision.

In Project Alexandria, the conversion of spatial analog products concerns extracting metadata simultaneously with extraction of data, with extracting image features for indexing, and with compressing images for efficient browsing.

On an analog map, both data and metadata are encoded in map symbols and in map sheet marginalia. Under current data conversion methods, typically only the spatial information is captured, while attribute and some topological information is neglected and added later by manual encoding. This process is being revised to encode attribute and topological relations at the time of data conversion.

The raw data of images (intensity/color/texture) are frequently not the information of primary interest for typical users, and thus storage of raw images is not the most efficient search format. To support content-based searches, image features need to be detected by appropriate preprocessing. Dynamic indexing is also called for. Two users of the same image may be interested in different information content, so a one-time static model of the image may be inadequate. The solution is both the definition of a set of generic images extracted at the time of storage and evolutionary dynamic indexing that links local user profiles with image processing tools to extract user-defined features.

A wavelet decomposition approach is being developed to support compression for browsing. This approach permits the decomposition of images into sub-images of coarser resolution, which can be subsequently recombined to perfectly reconstruct the original image. In this way an advantage in storage is achieved, since the need to store both high-resolution and coarse images is eliminated. Only sub-images are stored, which have the same memory requirements as high-resolution images. Defining an image feature is an application- and data-dependent task, so an important first step is to decide on the choice of image-level features.

In such data conversion efforts, librarians and map librarians can assist in identifying a pertinent thesaurus of features for indexing. Electrical engineers specializing in image and signal processing can assist with indexing and data compression, while spatial information engineers can help formalize map content for automated extraction of data and metadata.

Cataloging

Librarians have a long history of experience and training in cataloging methods. They have knowledge about how to organize information effectively for storage and retrieval and about appropriate connections that can be made between or among pieces of information. Traditional expressions of this knowledge and the traditional navigational tools used in libraries—such as the card catalog—are inadequate given the volume and formats of data now being made available over the Internet. Indexing of these electronic resources imposes new and challenging design constraints. Not only must indexing be efficient and effective from a librarian's information access and retrieval perspective, it must also satisfy requirements of effective database design and efficient software engineering.

Many existing library systems lack genuine database support for managing fast-growing collections of spatio-temporal and other datasets. Most existing catalog systems represent data in a flat relational view. In large part, data-handling difficulties result from the use of such existing file systems as a repository for these datasets. The varied contents, formats, and lineage, as well as the large size of these datasets, result in an unmanageable collection of files scattered over a network. The absence of database modeling and management further complicates the task of maintaining these datasets, many of which will not be static.

Spatially indexed data come in many different types and in a large variety of representations, so it is important to have an appropriate data model for organizing both the data and the metadata. Since spatial data have no standard format, the underlying data model for a catalog should be a flexible one. Map librarians, having a thorough knowledge of the range of formats and variety of representations of spatial data, are thus the essential domain experts who provide critical information for database design specifications.

The core of Project Alexandria's catalog system is a subdatabase of catalog information concerning the main items stored in the full database. The data stored in the catalog database include datasets organized into categories with four parts: a description, a tool box, source formats, and alternative formats. The "description" section includes metadata, abstract textual descriptions of the data, and reduced datasets or compressed images. The "toolbox" contains commonly used operations for manipulating the specific data types; "source formats" document the format of other data; and the "other formats" section allows storage of a user's commonly used formats (e.g., a user may store an ASCII format for transfer of data to a statistical package).

As documents are integrated into very large collections covering an entire scientific domain, links among the documents become increasingly important to help with searching and browsing. Librarians often have the expertise and experience to identify the important linkages between documents that should be incorporated within catalog entries. Design of the catalog component is thus a close collaboration between the map librarians, spatial engineers, and computer scientists who do the conceptual database design and the software engineers who are responsible for the implementation design.

User Interface Design

The design of a user interface is a crucial element for the success of any digital library. To date, user interfaces for electronic information retrieval have been given poor ratings by users (Fox et al. 1993). As digital libraries emerge it is important to improve user interfaces. Collaborative efforts can only improve the possibilities for success. Ideally, the interface should provide assistance like that offered by experienced librarians. Their knowledge in interpreting and responding to user requests provides vital input to the interface design effort.

Project Alexandria employs a three-tier approach to user interface design. The first step involves an investigation of various formal systems with well-defined pertinent objects, operations, and behaviors. In order to formalize the objects and operations that a user will need, designers must understand the problem domain and have sufficient knowledge of and practice with formalization tools (specification languages, algebras, etc.). This involves spatial information engineers working in close collaboration with map librarians. The second step investigates alternative interaction procedures that allow prospective users to perform intended manipulations. The specification of objects and operations identify core functionality around which different interaction techniques can be designed. In this phase the designer must be proficient in cognitive analysis, psychology, human computer interaction techniques, and graphic arts. These are the user interface designers. The third step is to implement the visualization and interaction designs on the specific platforms using their operating systems. This phase requires extensive expertise in user interface programming, graphics packages, and user interface management systems. These personnel are the user interface software engineers.

An important guiding principal in interface design is task-oriented access to electronic information (Fox et al. 1993). A starting point for task information is a focus on current information retrieval practices. For Project Alexandria, the map librarians outlined several typical request scenarios from different levels of users.

The map librarians noted that persons requesting spatial data always specify at least one parameter, and usually more. The library staff typically ask a basic set of questions on area, scale, subject, time, and format to make sure they steer persons to the correct items (e.g., "What area are you interested in?"; "What subjects should the item depict?"; "What time period should it portray?"; "Where is x located?" [x being a cultural or a physical feature]). In the case of spatial data, general users are often not familiar with key terms and concepts such as the meaning of scale. Queries such as, "Do you want a map at 1:10,000?" do not mean anything to the layperson, and in such cases assistance must be offered in a way that the user will understand—e.g., "How large an area should this show?" Such cues and responses typically provided by librarians need to be converted to equally responsive interfaces.

BASIN: Another Case Study Collaboration

The BASIN (Browsable hyper-Archive of Spatial Information on the Net) Project is a new project under way at the University of Maine to create a digital spatial data library. Involving collaboration amongst spatial information engineers, the library, and a subset of spatial data providers and users, the project combines the use of new technology—including high speed archival storage, telecommunications protocols, and an upgraded network interface (FDDI)—to provide access to digital spatial data collections. The objectives of the project are to:

- provide fast access to real (as opposed to fictitious) digital spatial data for instruction and student research projects;
- expose students to the wealth of data that is becoming available electronically;
- familiarize students with new communication technology so they can become both smart consumers and active participants in the new technology; and
- foster interdisciplinary collaboration through data sharing.

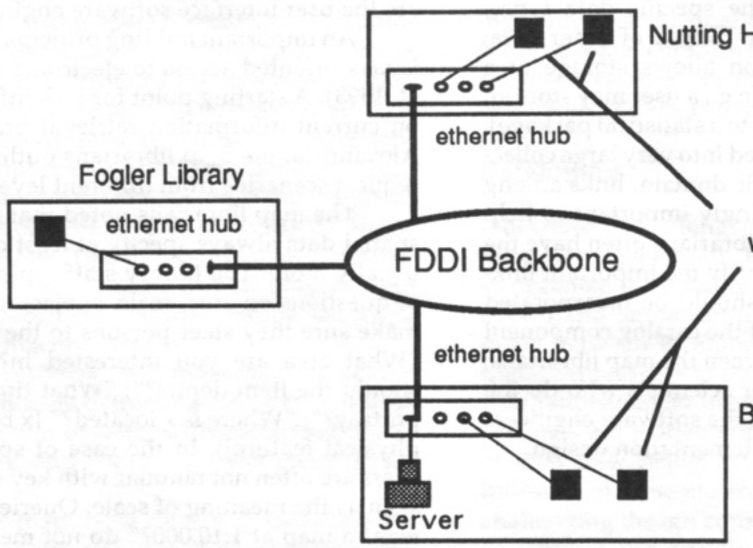


Figure 1
Schematic for BASIN

The motivation for BASIN was based on a number of factors. These include the fact that:

- several departments across campus have expressed a need and interest in access to spatial datasets;
- several departments have collected spatial data sets but have limited means for archiving and sharing these data sets;
- several Maine State agencies have compiled data sets of potential benefit for student instruction and research that are not, however, easily accessible or available in a timely manner to meet class assignment schedules; and
- several federal agencies are compiling and beginning to distribute electronically spatial datasets which students could make better use of through faster access and enhanced storage.

Specifically, the BASIN Project will consist of a server with RAID (Redundant Array of Inexpensive Disks) storage located in and managed by the Department of Surveying Engineering. Data in the BASIN archive will initially reside in fast-access hard-disk storage. As the archive grows, additional near-line storage facilities, such as a slower access/high capacity magneto/optical storage, may be added. Two switched Ethernet hubs will be provided to the BASIN archive to serve targeted student populations. Initially, the target populations will consist of students in surveying engineering, civil engineering, forest management, and wildlife, because these departments are most prepared to take immediate advantage of spatial data in their courses. A third switched Ethernet hub will be located in the Science and Engineering Center of Fogler Library to serve faculty and students at large. These hubs consist of 10 MB bandwidth between the archive and the workstations, allowing for a maximum transfer rate at Ethernet speed, which is unaffected by other local network traffic. Other departments—or any workstation on the network—may access BASIN using their current network connections. BASIN will employ two standard protocols: WAIS (Wide Area Information Server) and HTTP (HyperText Transfer Protocol) for indexing, browsing, searching, and data retrieval functions. The BASIN catalog (which describes the data archive) will use the Content Standard for Digital Geospatial Metadata, which was adopted as a federal standard in June of 1994. A schematic for BASIN is shown in figure 1.

In addition to providing dedicated access to the spatial data archive, the library plays a key role in this project in assisting in the creation of the BASIN catalog, and in connecting the spatial data catalog to the USMARC standards through the crosswalk (see Mangan in this issue). By establishing the crosswalk BASIN, entries will be accessible through the library's online catalog.

Conclusion

The projects and issues described above address many research issues fundamental to the development of true digital libraries that will fully exploit the potential of a digital world. It will be interesting to watch the progress of the NSF-funded digital library projects and evaluate the success of their collaborative ventures. These projects acknowledge that digital libraries involve more than simple conversion of library materials to digital formats, with straight transfer of analog tools such as page images of card catalogs. Instead, they reflect a fundamental rethinking of how information is captured, stored, and accessed. Fortunately, this rethinking is evolving as a shared endeavor among librarians, engineers, and many others.

Reference List

Federal Geographic Data Committee (FGDC). 1994. "The 1994 Plan for the National Spatial Data Infrastructure" (March).

Fox, E. (ed.). 1993. *Digital Libraries: A Sourcebook Report for the National Science Foundation.* Blacksburg, Va.: Virginia Polytechnic Institute and State University, Department of Computer Science.

Fox, E. A., D. Hix, L. T. Nowell, D. J. Brueni, W. C. Wake, and L. Heath and D. Rao. 1993. "Users, user interfaces, and objects: Envision, a digital library." *Journal of the American Society for Information Science* 44, no. 8: 480–91.

Kanade, T., M. Mauldin, R. Reddy, M. Sirbu, S. Stevens, D. Tygar, and H. Wactlar. 1994. *Proposal to the National Science Foundation: The Informedia Digital Video Library.* Carnegie-Mellon University.

Kettinger, W. J. 1994. "National infrastructure diffusion and the U.S. information superhighway." *Information and Management* 27: 357–68.

Mangan, E. 1995. "The making of a standard." *Information Technology and Libraries* 14, no. 2 (June 1995).

McClure, C. R. 1994. "Network literacy: A role for libraries." *Information Technology and Libraries* 13, no. 2 (June 1994) 115–25.

McClure, C. R., W. E. Moen, and J. Ryan. 1994. *Libraries and the Internet/NREN: Perspectives, Issues, and Opportunities.* Westport, Conn.: Meckler.

Mularz, D. E., H. G. Smith, and T. J. Mowbray. 1995. "Realizing the NSDI vision through a community architecture." *Geo Info Systems* 5: 1, 52–53.

NSFNET Packet. 1993. "NSFNET packet history." *Linkletter* 6: 1, 12.

Smith, T., M. Goodchild, S. K. Mitra, O. H. Ibarra, and J. Dozier. 1994. Proposal to the National Science Foundation: "The Alexandria Project: towards a distributed digital library with comprehensive services for images and spatially referenced information." University of California, Santa Barbara.

Tosta, N. 1994a. "Data data: The National Spatial Data Infrastructure: Where are we now?" *Geo Info Systems* 4, no. 1: 25–28.

Tosta, N. 1994b. "Data data: Partners." *Geo Info Systems* 4, no. 10: 24–25, 53.

U.S. Congress. House Committee on Science, Space and Technology. 1993. "National Information Infrastructure Act of 1993." Report to accompany H.R. 1757. Washington, D.C.: Government Printing Office.

Assisted Search for Knowledge (ASK): A Navigational Tool Set to Global Change Data and Information

Roberta Y. Rand

Traditionally, librarians and information management professionals have recognized the need to assist researchers in complex multidisciplinary research areas, and have developed and applied tools to assist in the storage and retrieval of information. In response to the vast amounts of research data and information now being collected, it is readily acknowledged that additional retrieval tools are needed. These additional tools can be viewed as "meta" tools, enabling and extending the full use of existing tools. This article describes a prototype online search system that will provide Internet and dial-up access to an array of databases related to environmental and demographic change. This "digital library" will link together diverse databases in different locations and will consist of unstructured full-text information, structured alphanumeric information, and satellite imagery and map information, and will deliver data to researchers in a form and context people can understand.

The United States Global Change Research Program (USGCRP) was established to observe, understand, and predict global change and to make the results of its research examining global change available for use in policy matters. The activities of the USGCRP are coordinated by the Committee on the Environment and Natural Resources Research (CENRR) (previously CEES, or the Global Change Committee on Earth and Environmental Sciences). Because research data and information are of fundamental importance in understanding and predicting global change, the 1992 "Global Change Data and Information Management Program Plan" was written and published to clearly state the collective commitment of the agencies participating in the USGCRP. In the program plan, participating agencies commit to work with each other, with academia, and with the international community to make it as easy as possible for researchers and others to access and use global change research data and information. It is toward this end that the federal agencies participating in the USGCRP are organizing the Global Change Data and Information System (GCDIS), which takes full advantage of the mission's resources and of the responsibilities of each agency, and links the services of the participating research data and information resources to each other and to users. These agencies realize it is critical to the success of the global change research agenda to have precise and accurate access to data and information.

The planning and organization of the GCDIS is being fully coordinated by a subcommittee of the CENRR, called the Interagency Working Group on Data Management for Global Change (IWG). The "Global Change Data and Information System Implementation Plan" builds on the broader program plan to define the construction of the GCDIS. The implementation plan states that the participating agencies will identify the vast array of data and information to be included, basing the criteria on the highest priority areas of interest, and will design and implement data and information services that are adequate to support the full breadth of the USGCRP.

To start this process, the agencies have initiated several pilot projects that are intended to broaden the scope of GCDIS. The Access Subgroup, one of the three subgroups of the IWG (the two others being the Library and Information Subgroup and the Contents Subgroup[1]), will coordinate the IWG activities necessary to develop the IWG access infrastructure. These activities include: providing mechanisms for the agencies to share access-related experiences and expertise, the development and application of standards and technology, demonstrations and pilot projects, mechanisms for feedback from the broad user community, and assessing the performance of the GCDIS access system. Several projects are underway to provide access to GCDIS: a gopher, a World Wide Web Browser, and the Assisted Search for Knowledge (ASK).

Developing a Prototype System: ASK

ASK builds on an earlier project, the "Thesaurus Project," [2] which was coordinated by the U.S. Department of Agriculture's National Agricultural Library (NAL). The objective of the "Thesaurus Project" was to establish a "proof of concept"—to show that traditional methods for search and retrieval currently employed could be greatly improved through available new technology. The intention was to:

> expand existing controlled vocabulary (keyword) capabilities at all levels by using a computerized, interactive, integrated knowledge base, that is, a semantic network combined with natural language understanding. This would be achieved by linking together existing distributed vocabularies and dictionaries, using keyword mapping, and by adopting other mechanisms to provide concept-based searching. The result would be enhanced access to multiple, distributed metadata directories and data collections, without ownership, using natural language queries. The success of this approach will be determined, in part, by the rate of development and the direction of emerging semantic networking technology.

Roberta Y. Rand is USDA Global Change Data and Information Management Coordinator, Information Systems Division, National Agricultural Library.

Global Change Data and Information System (GCDIS)

Participants:

Department of Agriculture (USDA)
Department of Commerce (DOC)
Department of Defense (DOD)
Department of Energy (DOE)
Department of Interior (DOI)
Department of State (DOS)
Environmental Protection Agency (EPA)
National Aeronautics and Space Administration (NASA)
National Science Foundation (NSF)

Databases:

Agencies' earth science data holdings, including satellite imagery, maps, text, and statistical data
Selected analyzed and assimilated datasets
Selected outputs from global change models
Published documents
Socioeconomic data necessary for study of human and medical dimensions of global change

Figure 1
GCDIS Participants and Databases

> The process would be accelerated by procuring currently available off-the-shelf, client/server software which adheres to common standards for this technology.

ConQuest Software, Inc., was the basis of the "Thesaurus Project." ConQuest has an existing commercially available text and image retrieval system that uses natural language processing techniques, word meaning processing, and concept-based information retrieval built from many electronically available dictionaries and thesauri.

The fundamental concept behind the ASK pilot project is to develop a prototype system that links databases diverse in format and content, while enabling users with different skills, needs, and access methods to obtain relevant information from these databases. This linkage will be done by using a natural language inquiry and a common user interface. The ASK project plan calls for the development of four prototypes to be delivered over the next twelve months. The first of these prototypes was delivered on January 10, 1995.

Prototype #1

A software testbed, this prototype demonstrates the key principles that are the essential technical foundation for continued successful development. It includes the ability to enter a command line query via the client software, which searches multiple databases over the Internet and provides a single merged response. The success of prototype #1 delivery is the baseline for providing increasing support to a broad cross-section of users who are concerned with access to and the analysis of global change data.

In addition, a User Concept of Operations (ConOps) is developing and will be reviewed at each stage of development by the ASK Users Working Group. This group represents experts on the identified user categories (researchers, policymakers, K-14, and the general public) and provides a mechanism to assess user needs for enhanced system development.

Prototype #2

Scheduled for availability on April 10, 1995, prototype #2 will: add a generic graphical user interface (GUI), be Z39.50 compliant, contain one or more additional knowledge bases (the National Institute of Health's Unified Medical Language System, the Defense Technical Information Center, and the NASA Thesaurus), and will be publicly available in Washington, D.C., on the Mall and in several Smithsonian locations, for Earth Day.

Several of the databases expected to be accessed by the ASK prototype system include significant geographic content. Therefore, a versatile capability for handling various types of geographic data, including digital imagery, is an important aspect of the program. Toward this end, E-Systems will provide OASIS (Open Architecture for Scientific Information Systems)—a software product that implements system-level building blocks based on common industry standards and commercially available platforms—for configuring systems to provide data management and graphic processing capabilities for spatial data. The modular software architecture and adherence to an open-systems concept fits the ASK concept and will become an integral component of ASK in prototype #2.

Prototype #3

Scheduled for availability in July 1995, prototype #3 will add a commercially-off-the-shelf (COTS) geographic information systems (GIS) capability, additional GUIs for multiple user classes, a consistent data presentation model, or non-native (non-ConQuest) search engine to demonstrate the capability to link to and effectively utilize the functionality of existing search engines of participating agencies.

Prototype #4

Finally, scheduled for availability in October 1995, prototype #4 will provide the ability to manage metadata and source selection using a knowledge base for simultaneous information access across multiple databases (including non-native search engines), and will include retrospective searching, real-time profiling, and on-disk (CD-ROM) product searching.

ASK is a team effort. ConQuest Software, Inc., which won the contract and is the lead company, is providing full-text search and retrieval engines and integration tools linking the databases together. ConQuest has teamed with E-Systems, which is acting as the project systems integrator and providing spatial data processing assistance. The other members of the team are: Infrastructures for Information, Inc., which is delivering capabilities for filtering and viewing diverse documents; WAIS, Inc., which is providing Internet communications protocols; and the University of California at Santa Barbara, the institution that is home to Project Alexandria.

Many of the ASK documents are available via FTP and are searchable using ConQuest software. It is also possible to access the prototypes over the Internet (see below).

For more information about the GCDIS-ASK project, send an e-mail message to the ASK listserv: ASK@circles.org. To subscribe to the ASK listserv, send an e-mail message to: majordomo@gaia.circles.org; leave the subject line blank, and type in the body of the record: "subscribe ASK [your name]."

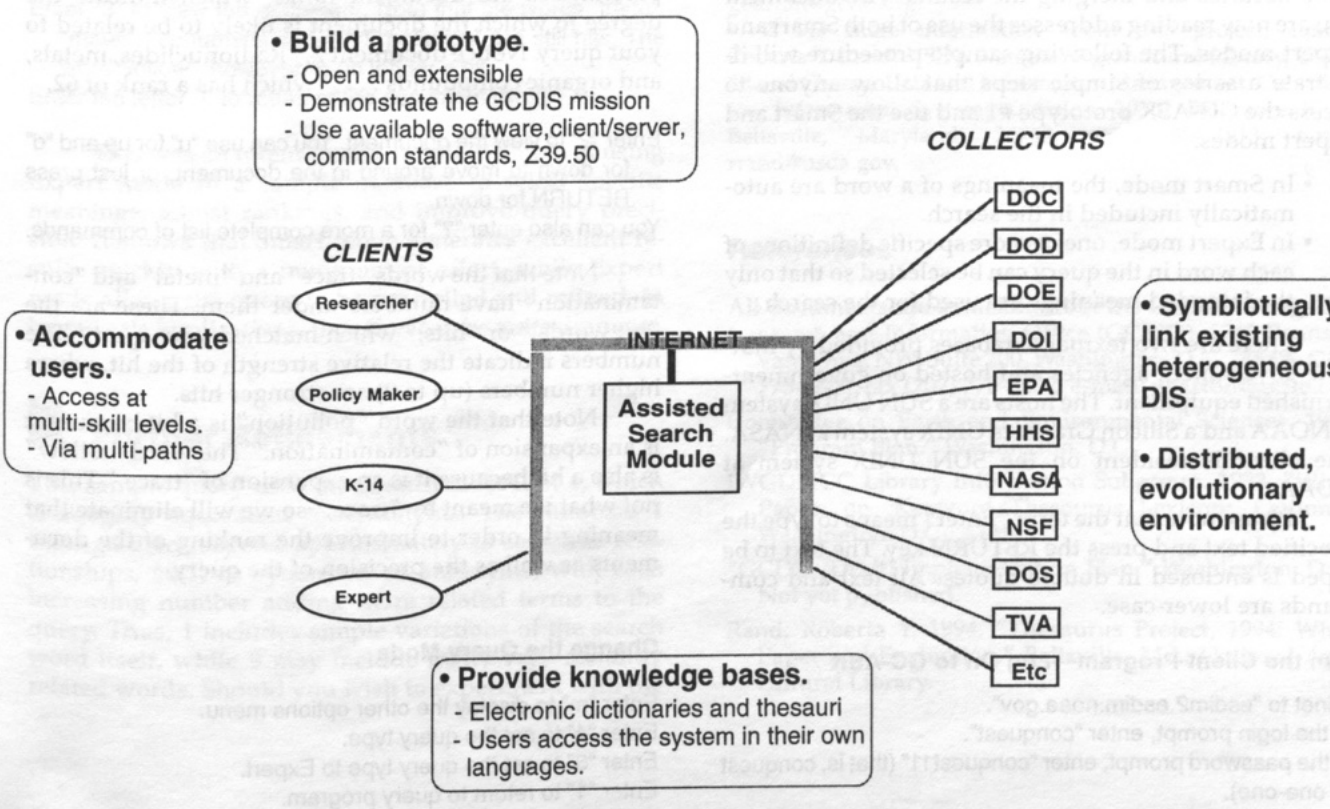


Figure 2
Overview of GC-ASK System

Smart and Expert Mode— Sample Procedure for GC-ASK Prototype #1

On January 10, 1995, prototype #1, the first of four GC-ASK prototypes, was successfully delivered. It is a software testbed that demonstrates key principles that form the essential technical underpinnings for continued development. It includes the ability to enter a single command-line query via the client software, which then searches multiple databases over the Internet and provides a single, merged relevance-ranked response.

The next major milestone, prototype #2, is scheduled for delivery on April 10, 1995. It will add a GUI, Z39.50 protocol compliance, one additional knowledge base, and additional databases. The success of prototype #1 has established a solid base for information support to a broad cross-section of users concerned with access to and analysis of global change data.

As part of the prototype #1 delivery, a sample procedure (or "script") is provided, which illustrates the fundamental process for searching across multiple remote libraries and merging the results. The document you are now reading addresses the use of both Smart and Expert modes. The following sample procedure will illustrate a series of simple steps that allow anyone to access the GC-ASK prototype #1 and use the Smart and Expert modes.

- In Smart mode, the meanings of a word are automatically included in the search.
- In Expert mode, one or more specific definitions of each word in the query can be selected so that only the intended meanings are used for the search.

There are two textual databases provided by multiple government agencies and hosted on government-furnished equipment. The hosts are a SUN UNIX system at NOAA and a Silicon Graphics UNIX system at NASA. The client is resident on the SUN UNIX system at NOAA.

Please note that the term "Enter" means to type the specified text and press the RETURN key. The text to be typed is enclosed in double quotes. All text and commands are lower-case.

Run the Client Program—Log On to GC-ASK

Telnet to "esdim2.esdim.noaa.gov".

At the login prompt, enter "conquest".

At the password prompt, enter "conquest11" (that is, conquest one-one).

You will briefly see some status information. This is part of the login procedure and has nothing to do with the prototype.

You will see some licensing information. Press RETURN to continue.

You will see a list of libraries (i.e., databases) that you can select. Number 1 (noaa_lib) is the library at NOAA, number 2 (nasa_lib) is the library at NASA.

Enter "1" to select the library at NOAA. An asterisk appears to indicate your selection.

Enter "2" to select the library at NASA. An asterisk appears to indicate your selection.

Enter "q" to quit.

After a few seconds you will see the "SMART>" prompt. This means that the system is ready to accept a query in SMART mode, which will automatically expand your query to include words that have similar meanings.

Query in Smart Mode

Enter "trace metal contamination".

After a few seconds, you will see a numbered list of documents returned by the server. The numbers in parentheses are document ranks, which indicate the degree to which the document is likely to be related to your query. Notice document 2, "Radionuclides, metals, and organic compounds . . . ," which has a rank of 62.

Enter "2" to view the document. You can use "u" for up and "d" for down to move around in the document, or just press RETURN for down.

You can also enter "?" for a more complete list of commands.

Note that the words "trace" and "metal" and "contamination" have numbers under them. These are the "hit words," or "hits," which matched your query. The numbers indicate the relative strength of the hit, where higher numbers (up to 9) are stronger hits.

Note that the word "pollution" is a hit because it is an expansion of "contamination." The word "follow" is also a hit because it is an expansion of "trace." This is not what we meant by "trace," so we will eliminate that meaning in order to improve the ranking of the documents as well as the precision of the query.

Change the Query Mode

Enter "m" to display the other options menu.

Enter "4" to set the query type.

Enter "3" to set the query type to Expert.

Enter "1" to return to query program.

You will see the "EXPERT>" prompt.

Query in Expert Mode

Enter "trace metal contamination". A word selection menu is displayed.

Enter "1". All available definitions and parts of speech for "trace" are displayed. Each definition with an asterisk will be used during the search.

Enter "2–13". Meanings 2 through 13 no longer display an asterisk and will not be included during the search. All occurrences of "trace" will be found, and meaning 1 will be the only meaning used to find words that are related to "trace".

Enter "q" to return to the word selection menu.

Enter "2" to select a definition of "metal".

Enter "2" to deselect definition 2.

Enter "q" to return to the word selection menu.

Enter "3" to select a definition of "contamination".

Enter "q" to return to the word selection menu. All meanings are still selected.

Enter "q" to execute the search. After a few seconds, you will see the numbered list of documents returned by the server. Notice that the documents have been re-ranked and that our "radionuclides" document is no longer #2. Its rank has been lowered to 60 and its position in the list is now #3.

Enter "3" to view the document. Note that the words "trace", "metal", "contamination", and "pollution" are still hits. The word "follow", however, is no longer a hit.

Enter the letter "l" to return to the document list.

This was an intentionally simple example of using Expert mode in a sample database to select specific meanings, adjust rankings, and improve query precision. It shows that Smart mode generates excellent results quickly with a minimum of effort, while Expert mode allows the query to be controlled and refined. In large-scale applications, the effect can be quite dramatic.

Further Refinements

This sample query used an expansion level of 4, which is roughly equivalent to synonyms. The numbers 1 through 9 correspond approximately to semantic relationships, such as synonyms or antonyms, with each increasing number adding more related terms to the query. Thus, 1 includes simple variations of the search word itself, while 9 may include many very distantly related words. Should you wish to experiment with the expansion level, you will find it in the other options menu, "m", option #3. Remember that entering "?" shows you the commands you can use while exploring. When you are finished, go to the logout instructions below.

Logout

Enter "q". The system asks if you want to exit the program. Enter "y".

Should you have any questions or comments about this document or the GC-ASK project, please contact Cheri Pender at 1-800-787-1715 or via e-mail at cpender@cq.com.

Notes

1. For the purposes of this paper, emphasis has been placed on the Access Subgroup. For more information about the Library Information Subgroup or the Contents Subgroup, contact the Global Change Research and Information Office (GCRIO) at 1747 Pennsylvania Ave, NW, Suite 200, Washington, D.C., 20006; (202) 775-6600; fax (202) 775-6622; e-mail gbarton@gcrio.org.

2. For more information about this project, contact Roberta Y. Rand, USDA National Agricultural Library, USDA Global Change Data and Information Management, Coordinator, Information Systems Division, 10301 Baltimore Blvd., Beltsville, Maryland, 20705; (301) 504-6813; e-mail rrand@usda.gov.

References

All documents are available from the Global Change Research and Information Office (GCRIO), 1747 Pennsylvania Ave, NW, Suite 200, Washington, D.C., 20006; (202) 775-6600; fax (202) 775-6622; e-mail: gbarton@gcrio.org.

Committee on Earth and Environmental Sciences. 1992. "Program Plan." Washington, D.C.

IWGDMGC Library Information Subgroup. 1993. "White Paper on Keyword/Thesaurus/Lexicon: Options." Washington, D.C.: Author.

"GCDIS (Draft) Implementation Plan." Washington, D.C.: Not yet published.

Rand, Roberta Y. 1994. "Thesaurus Project, 1994: White Paper and Evaluation." Beltsville, Md.: National Agricultural Library.

Accessing Spatial Data Online: Project Alexandria

Mary Lynette Larsgaard and Larry Carver

Project Alexandria, one of six National Science Foundation (NSF) digital library initiatives (DLI), has as its objective online access to spatial data in a distributed environment, with search options by areas of interest on a base map or by text, as the user prefers. The project, which is based at the University of California, Santa Barbara, has partners and participating agencies and libraries throughout the United States. The development of a rapid prototype system began in October 1994 and was completed by the end of February 1995. The testbed system will be operational by late 1997.

Map libraries have a reasonably lengthy history of map cataloging; the first map catalog at Harvard University is dated 1831 (Merrett 1976, 3). Extensive employment of map cataloging dates from after World War II, and the practice became especially common with the availability of the USMARC Map format on OCLC. Online shared cataloging brought substantial benefits to map libraries almost immediately by creating and broadening access to cartographic information. Nonetheless, for as long as map cataloging has gone on, it has had as its constant counterpart frustrations known only to the map librarian who, for instance, is asked questions at the map-sheet or air-photo-frame level, but whose collection is cataloged at the map series or flight level (this assumes that the map librarian has cataloging access to remote-sensing imagery in the first place).

Over the past two decades air photos have, for several reasons, seldom been cataloged in U.S. map libraries. Air photos held by any given library have tended to be local area photos for which no shared catalog record exists. Cataloging of this material most often requires time-consuming original cataloging, which frequently means that such items will be cataloged last, only after records that derive from shared cataloging or that will be used most heavily in a shared-cataloging environment have been cataloged. Added to the burden of the labor-intensive task of original cataloging is the challenge of describing a graphic object (such as a map) using only text. Even twenty years ago, exasperated rare-map catalogers, buried in USMARC note fields (5XX), were heard to mutter that the best idea was to photograph the map, put the photo in an aperture card with author/title/date/publishing information if any, and be done with it.

The Beginnings of Project Alexandria

Larry Carver, head of the Map and Imagery Laboratory (MIL) at the University of California, Santa Barbara's (UCSB) Davidson Library, noticed these problems in the late 1960s when he began working with the library's collection (all three map cases and two hundred maps of it). For many years, Carver's focus was developing the collection in the areas of greatest interest to the faculty—remote-sensing imagery and digital data; today, the collection encompasses 5.1 million items. In the mid-1980s, Carver began work on a grant proposal to make spatial data accessible at the sheet/frame level, using a map-search interface. The Research Libraries Group (RLG), of which UCSB is a member, became involved in Carver's project and in 1987 the Keck Foundation awarded RLG a grant to produce a design document for such a system (Bloch 1988; RLG 1989). For a variety of reasons, there was no further progress on the matter for several years.

ESRI and NSF

In the early summer of 1993, Carver met with Jack Dangermond, of the Environmental Research System Institute (ESRI), concerning the possibility of using geographic information systems (GIS), and specifically ESRI's ArcView and Arc/Info GIS softwares, as a method to access spatial data. In one sense, the idea under consideration turned GIS inside-out in that it made use of the attributes for the cataloging fields and the graphic representation for the browse file of the item or (space allowing) the actual spatial data in digital form. At their meeting, an agreement was reached between Carver and Dangermond whereby ESRI would provide software and staff time to work with MIL staff in setting up a prototype. An exploratory meeting was then held at MIL in September of 1993, with representatives from the Association of Research Libraries, ESRI, and the U.S. Geological Survey's Water Resource Division; a general plan of action was developed at that meeting.

At about the same time, the National Science Foundation (NSF) issued an RFP for the Digital Libraries Initiative (DLI). Carver consulted extensively with

Larry Carver is Head, and **Mary Lynette Larsgaard** is Assistant Head, Map and Imagery Laboratory at the Davidson Library, University of California, Santa Barbara.

various UCSB faculty about the initiative, and the outcome was the decision that UCSB would apply for the grant. A grant proposal was submitted in February 1994 with five co-principal investigators and a cast of twenty-two additional faculty and staff, also serving as co-principal investigators (Smith et al. 1994). Dr. Terence Smith, who holds a dual appointment in computer science and geography and is chair of the Computer Science Department, has served as the lead contact person and organizer for the project.

In August 1994 UCSB was notified that it had received one of the DLI grants—simultaneously fulfilling hopes and fears, and causing some of those who had worked on the grant proposal to recall the old saying, "Watch out what you ask for; you may get it." It had been intellectually stimulating and enjoyable to think about what needed to be done and how it could be done, but it was obviously going to require considerable work to come through with all the deliverables that the proposal had promised. Staff at UCSB officially began working on Project Alexandria on October 1, 1994; slated as a four-year project, it is scheduled to end in 1997. The development of a rapid prototype was the focus of activity from October 1994 through March of 1995. Present efforts are concentrated on the creation of a testbed system.

Project Alexandria: A Summary

Project Alexandria, which takes its name from the location of the famed library of antiquity, has as its overall objective to develop a digital library that provides quick, easy access to large, diverse collections of geo-referenced information, including maps, remote-sensing images, and pictorial and textual materials. A graphic user interface to the digital library will allow users to search in a manner that best suits them (via map, command line, or menu). The digital library will be a distributed database, permitting users with access to the Internet to view and retrieve materials, regardless of location. In fact, although items will be held at libraries dispersed throughout the United States, it will appear to users as if the materials are held at a single, local library site.

Alexandria will provide a full range of electronic library services, including an electronic reference desk and some forms of image processing (e.g., file conversion and compression), to assist users in taking full advantage of holdings. One of the project's long-term goals is to integrate access to more traditional text materials along with access to multimedia and digital materials. The client group will include any user of geo-referenced information—from K-12 students, to academic researchers, to members of the general public.

Key Members of the Project Team

Project Alexandria is a consortium of university, library, and industrial partners, centered at the University of California, Santa Barbara (UCSB). The UCSB component of the team includes faculty from the computer science, electrical engineering, and computer engineering departments, the Center for Remote Sensing and Environmental Optics, and the National Center for Geographic Information and Analysis (NCGIA), as well as staff from the Davidson Library's Map and Imagery Laboratory. Other university partners include faculty from NCGIA sites at the State University of New York, Buffalo (SUNY–Buffalo), and the University of Maine (Orono). Library partners include the Library of Congress, the University of California's Division of Library Automation, the library at SUNY–Buffalo, the library of the U.S. Geological Survey, and the St. Louis Public Library. Industrial partners at the time of writing are DEC (Digital Equipment Corporation), ESRI, ConQuest Software, and Xerox.

Testbed System—Four Library Components

During the first six months of the project, the Alexandria team built a prototype online system using primarily ESRI's ArcView software, and began developing detailed specifications for the main testbed system. For the remainder of the project, the main testbed system will be developed and tested, eventually taking the form of a distributed digital library with components at the previously mentioned library sites, as well as at other interested libraries and agencies. Facilities at each library site may include various combinations of four library components: (1) a graphical user interface with browse capabilities that support access to each of the services by textual and visual query languages; (2) a catalog component that provides rapid response to queries and produces a minimum of false drops; (3) an ingest component designed to incorporate new items into the distributed digital library, including such procedures as digitization, reformatting, and automatic extraction of catalog information; and (4) a storage component that provides for high-speed access to large collections. The set of components implemented at any one site will depend upon the needs and abilities of the site.

The graphical user interface is in many ways the most crucial of the four components. While a variety of interfaces will be available, each must be simple and intuitive in order to meet diverse user requirements. The interface will support text-based and visual-based query languages. An especially challenging part of the project is content-based searching; that is, searching a digital geo-referenced object (such as a scanned air photo) for certain features. The interface will also support browsing and visualization of information (with databases transmitted upon user request within the confines of copyright law).

The core of the digital library is an electronic catalog, which the system searches for items requested by users. The catalog contains both a database of information about the library collections and efficient mechanisms for searching the database. This metadata will comply with USMARC and with the Federal Geographic Data Committee's (FGDC) recently issued Content Standard for Digital Geospatial Metadata (Federal Geographic Data Committee 1994). Metadata includes bibliographic information, the contents of various geo-referenced objects (at as many layers as is necessary for user access—e.g., grandparent, parent, and child levels), and thumbnail versions of images and maps to make browsing easy and convenient.

The ingest component provides for loading into the digital library not only metadata but digital forms of geospatial data. Participating libraries will be encouraged to first load data that is unique or held by only a very few sites. An important part of the project plan for database creation is a hierarchical decomposition of digital images, based on the use of a technique known as "wavelet transformation." After undergoing this transformation, a digitized image may be represented as a set of images at different levels of resolution in such a way that the different sub-images, when "added up," result in the original image. Given the enormous size of images, this technique is very important for storage, as the image is archived only once, rather than several times at different levels of resolution. The sub-images are used for browsing and for extracting information about the contents of an image.

In such an ambitious and massive undertaking, there are any number of research matters to be addressed. However, the overarching issues are:

- *Graphical user interface design:* Evaluation of all aspects by many different user groups is essential. Project team members are especially interested in finding out how users search for spatial data in an online environment, particularly users (e.g., K-12) who have had relatively little access to large spatial-data collections.
- *Network aspects:* Speedy, reliable transmission of information between the library and its users is vital.
- *Standards:* There are a substantial number of standards in many different areas that must be followed, from description to transmission of databases.
- *Scalability:* To be successful, Alexandria must be scalable and extensible; that is, it must continue to work effectively, from a user's point of view, as it grows in size, and it must be capable of being extended to any geo-referenced collection.

For further details on Alexandria, see the Mosaic home page (http://alexandria.sdc.ucsb.edu), which has a copy of the grant proposal and provides detail about the above matters.

Progress Report: The First Six Months of Alexandria

Throughout the months of effort put into the prototype, there was considerable activity in many areas, including numerous meetings (held more or less monthly) with topics covered ranging from a presentation by Cliff Lynch (University of California Division of Library Automation) on Z39.50 to intense status-and-discussion meetings chaired by Dr. Terence Smith. Members of the executive board also met with partners and prospective partners. The following summary focuses on the development of a rapid prototype, an element essential to the progress of Alexandria. For details beyond this summary, see Frew (1995).

October 1994: Planning and Hiring Staff

The plan for the prototype was to follow the course of action set out during early discussions with ESRI and to have the prototype completed by March 1, 1995—a deadline that was then changed to February 24 to coincide with a an NSF site visit to determine whether the project was making sufficient progress to justify continued funding. While the project did not formally begin until October 1, UCSB project staff began working in August and September to hire project staff and initiate work on the rapid prototype. Two computer engineers—one full-time, to work on metadata database modeling, and one two-thirds time, to incorporate metadata and spatial datesets into ArcView—were in place by early October. Two half-time geography graduate students were also hired to ingest metadata and digital data into the prototype. By mid-October, an administrative assistant was in place to set up Alexandria offices (starting with empty rooms) in Phelps Hall on the UCSB campus. A principal engineer and a senior engineer were hired by the end of the month. Meanwhile, the executive board and the rapid-prototype (RP) team were each meeting weekly. The primary focuses for the month were:

- building a list of spatial datasets to be used in the project, focusing on remote-sensing imagery and maps of Santa Barbara County, California, and especially on the area of the campus, for which the holdings of MIL are most extensive;
- getting all hardware and software in MIL and Phelps set up;
- orienting engineers with Sybase and ArcView, and working on structure of metadata schema; and
- training graduate students on ArcView and scanner software.

November 1994: Creating the Rapid Prototype

RP team members concentrated on having the first formal demonstration of the rapid prototype ready for a December 9 visit by the executive board of the NCGIA. Major tasks for the attainment of this goal were the construction of:

- *Metadata:* The Alexandria Metadata Schema (about 50 minimum fields) was constructed primarily from the FGDC standards document, with additional fields taken as needed from USMARC. Since the FGDC standard is intended for application to digital spatial data only, and because some analog data was to be ingested, metadata was derived for those datasets (29) selected for the December 9 demonstration;
- *Spatial data:* ArcView was successfully connected to Sybase and bounding-coordinate rectangles were constructed within ArcView; scanning of air photos began;
- *Timelines:* The project schedule was set up, responsibilities were assigned, and a script for the demo was written by RP team leader, James Frew. The executive board named team leaders and teams were formed.

December 1994: Expanding the Rapid Prototype

All metadata and spatial data were ingested into ArcView and were ready for testing one week before the December 9 visit. The demonstration went very well—the system crashed only once, at the very beginning, and otherwise worked impeccably. After the demonstration, the RP team began planning for the next NSF site visit—February 24, 1995—and set goals for the month of January: determine additional data types to be added; determine additional metadata fields (another 30, primarily for more detailed information on aerial photographs); and develop the ingest component and, importantly, public-user interfaces. Other teams, including the image processing team and the library team, began meeting regularly.

January 1995: Determining Storage Requirements and Load Procedures

Air-photo coverage for the general UCSB campus area was selected, as were several geological and ecological maps. The age of photos ranged from 1928 to the early 1990s and were chosen at ten-year intervals. The additional 30 metadata fields were added to the structure, and work began on the ingest and public-user interfaces, using TCL/TK software. Ingest of a 450,000-record metadata set began; the metadata set was made up of frame-level records for NASA/Ames flights from the early 1970s to the early 1990s. The loading of this metadata dataset required 150MB to store unindexed; it also helped establish a procedure for ingesting such datasets.

February 1995: Developing the User Interface

The RP team aimed to have all metadata and spatial data in the prototype by February 17, allowing a week for testing prior to the NSF site visit. A beta test demonstration was held February 20 during the first meeting of Project Alexandria's advisory board. Work on the public-user interface, creation of browse files and metadata, and scanning items were primary tasks during the first weeks of the month. Toward the end of the month, work began on establishing a World Wide Web (WWW) presence.

Lessons Learned

From a library technical-processing point of view, the most important task deserving of effort for the testbed is the need to use computer software and hardware to derive bounding coordinates. It will come as no news to catalogers that metadata creation is a time-consuming process and that determining bounding coordinates—essential for a digital library focused on geo-reference information—is easily the most time-consuming part of spatial-data cataloging, especially for large-scale (e.g., 1:6,000) aerial photographs. For the prototype, after considerable wear and tear on the dispositions of the two staff primarily responsible for creating metadata, a method of using ERDAS software evolved. A geography graduate student digitized enough of the 1:24,000-scale U.S. Geological Survey topographic quadrangle (Goleta), within whose bounds the aerial photographs were all located, to derive coordinates. Alexandria needs to develop an ingest front-end to a map database upon which the cataloger may "place" an item being cataloged. The software then determines the coordinates and enters them into the appropriate fields in the metadata workform. Another technical-processing task that is already being addressed is the full-scale implementation of a joint FGDC-USMARC metadata schema, far beyond the roughly 80 fields (literally in the thousands) that were employed for the February datasets.

Of course, there will be many more developments over the next three or so years—a considerable number of which will be far beyond this author's expertise—that will continue to make Alexandria an important, exciting, and interesting project. It will take the extensive knowledge and hard work of persons with many different capabilities and from many different disciplines to make Alexandria a reality. Readers are encouraged to consult the Project Alexandria Mosaic homepage for future developments.

Bibliography

Bloch, Cecil. 1988. *Geo-Referenced Information Network External Design*. 2 vols. Mountain View, Calif.: Research Libraries Group.

Federal Geographic Data Committee. 1994. "Content Standard for Digital Geospatial Metadata" (June 8). Washington, D.C.: Federal Geographic Data Committee.

Frew, James. 1995. "The Alexandria rapid prototype: Building a digital spatial data library." Paper submitted for the 1995 Advances in Digital Libraries Conference (http://alexandria.sdc.ucsb.edu).

Merrett, Christopher. 1976. *Map Cataloguing and Classification: A Comparison of Approaches*. Occasional publication series, no. 7. Sheffield, England: University of Sheffield Postgraduate School of Librarianship and Information Science.

Research Libraries Group. 1989. "RLG enters new sphere with geoinformation project." *News*, no. 19: 3–9.

Smith, T., Goodchild, M., Mitra, S. K., Ibarra, O. H., and Dozier, J. 1994. Proposal to the National Science Foundation: "The Alexandria Project: Towards a distributed digital library with comprehensive services for images and spatially referenced information." University of California, Santa Barbara.

The Making of a Standard

Elizabeth U. Mangan

This article describes how the Content Standard for Digital Geospatial Metadata was developed and how it relates to USMARC Format for Bibliographic Data to provide the ability to communicate and access descriptions for digital spatial datasets.

Origins

The Federal Geographic Data Committee (FGDC) was established by the Office of Management and Budget (via circular A-16) in October 1990 to promote the coordinated development, use, sharing, and dissemination of geographic data. This interagency committee, which is chaired by the Secretary of the Interior, has representatives from the Department of Agriculture, Department of Commerce, Department of Defense, Department of Energy, Department of Housing and Urban Development, Department of the Interior, Department of State, Department of Transportation, Environmental Protection Agency, Federal Emergency Management Agency, Library of Congress, National Aeronautics and Space Administration, National Archives and Records Administration, and the Tennessee Valley Authority. Other federal agencies participate on various FGDC subcommittees and working groups.

In June 1992 the FGDC hosted a forum to discuss what information would be needed to describe a digital dataset so that numerous data-collecting agencies could coordinate and share spatial data. During these deliberations the participants agreed on the necessity for a standard. The FGDC accepted the offer of ASTM Section 18.01.05 to develop a draft standard to define the required content information.[1] The ASTM draft was slightly revised by the FGDC before it was then offered for public review from October 1992 to April 1993 as the "Content Standard for Spatial Metadata."

This review period generated extensive comments from a wide range of potential digital data producers and users. The comments were summarized, resulting in a packet almost two inches thick, and were presented to the members of the FGDC standards working group for consideration. This group held an intense four-day session in April 1993 to consider each comment that had been submitted. For nearly each data element present in the circulated draft, comments ranged from "Expand it and include more detail" to "Get rid of it!"

Global Issues or Assumptions

The first step in the analysis of the comments the working group received was to establish some global issues or assumptions to serve as a framework in revising the standard. These assumptions were that:

- The standard was to be considered an independent document defining the data needs of the spatial data community and was not linked to any specific implementation, although two implementation methods were proposed–the Spatial Data Transfer Standard (SDTS or FIPS 173) for the transfer of data, and USMARC for access to the metadata in a catalog environment;
- While the standard should provide for uniform description of spatial data independent of the form or media, the current standard would be intended to cover only digital forms; any effort to cover nondigital forms would be deferred since expanding the standard to cover nondigital forms would require an extensive effort;
- The standard was intended to encompass all means of describing locations—geo-referenced coordinates, coordinates whose relationship to the earth is unknown, street addresses, mile markers, and indirect positional references through objects which have a known location—but would also include ancillary datasets important to spatial analysis but not spatially referenced; and
- The term *dataset*, which represents the foundation for the standard, should be defined simply as "a group of related data," and that what constitutes a specific dataset should be left to the provider of an individual dataset.

Additionally, the working group determined that the standard, and therefore the data elements included in the standard, must be sufficient to support four activities: availability, defined as the information needed to determine what data exists for a given geographic area; fitness-for-use, defined as the information needed to determine if a dataset meets a specific need; access, defined as the information needed to acquire an identified dataset; and transfer, defined as the information needed to process and use a dataset. While these characteristics form a continuum through which a user moves, via a variety of choices, to determine what data is available, evaluate its usefulness, determine how to obtain it, and learn how it can be used, the order in

Elizabeth U. Mangan is Head, Data Prep and Files Maintenance, Geography and Map Division, Library of Congress.

which specific data elements are evaluated and the relative importance of any particular data element will not be the same for all users. Therefore, each data element was measured against these criteria and, in order to be part of the standard, had to fulfill at least one of the supported activities: availability, fitness-for-use, access, and/or transfer.

Keeping these assumptions in mind, the working group considered the comments received during the review period and analyzed the appropriateness of each data element in the standard. The working group also deliberated over the issue of which of the data elements should be identified as "mandatory" (must be included) and which could be "optional" (may be included). During the discussions, an additional category—"required"—was defined similar to mandatory but for which the values *unknown* or *not applicable* were valid. During subsequent revisions, these categories were changed to coincide with the USMARC-equivalent terms "mandatory," "mandatory if applicable," and "optional."

A revised draft was circulated in July 1993 for further review and testing. After this six-month testing period, the working group again revised the document based on comments received and on the results of the testing; they then submitted it for review by the FGDC coordination group. The final draft version, renamed "Content Standard for Digital Geospatial Metadata," was then produced and was approved by the FGDC on June 8, 1994.

Purpose of the Standard

The standard specifies the information content of metadata for a dataset of digital geospatial data. Metadata are information about the content, quality, condition, and other characteristics of data.[2] The purpose of the standard is to provide a common set of terminology and definitions for the documentation of digital geospatial data.[3]

The geospatial metadata standard was developed to document geospatial data acquired or developed by federal government applications and programs, although the FGDC invites and encourages organizations and persons from state, local, and tribal governments, the private sector, and nonprofit organizations to use the standard to document their geospatial data.[4] The major uses of this information are: to maintain an organization's internal investment in geospatial data; to provide information about an organization's data holdings to access catalogs, clearing houses, and brokerages; and to provide information needed to process and interpret data received through a transfer from an external source.

As stipulated in Executive Order 12096, signed by President Clinton on April 11, 1994, "beginning nine months from the date of the order, each [executive branch] agency shall document all new geospatial data it collects or produces, either directly or indirectly, using the standard under development by the FGDC, and make that standardized documentation electronically accessible . . . . Within one year of the date of this order, [executive branch] agencies shall adopt a schedule . . . for documenting, to the extent practicable, geospatial data previously collected or produced, either directly or indirectly, and making the data documentation electronically accessible."[5] Since map libraries have traditionally depended heavily on federal agencies for spatial information, they need to follow closely the evolution and application of this standard.

Implementations

The metadata standard is not intended to specify the means by which metadata information is organized in a computer system, nor does it specify the means by which this information is transmitted or communicated to the user. These decisions were guided by the variety of methods of organizing data in a computer, the differences between data providers in describing their data holdings because of varying institutional and technical capabilities, and the rapidly developing methods of providing information on the Internet for different purposes.

The FGDC is, however, pursuing the Spatial Data Transfer Standard (SDTS) for transfer, and USMARC for retrieval or searching access, as the proposed implementation methods for the geospatial metadata standard. These existing standards were selected because the SDTS is the federal government's standard for the transfer of geospatial data and is already a Federal Information Processing Standard (FIPS 173), while USMARC is the widely accepted cataloging-data standard in the library community, thus providing an existing infrastructure of libraries to disseminate information to the public.

Once the geospatial metadata standard was finalized, it was compared to the USMARC Format for Bibliographic Data to determine the extent to which the metadata data elements were compatible with existing appropriate USMARC fields. While most of what might be considered traditional bibliographic elements included in the standard were easily correlated with USMARC, the data elements dealing with specialized information related to digital data—such as the method of storage, the mathematical information, the requirements to make use of the dataset, and information on how to obtain a copy—had no corresponding fields. Therefore, two proposals to modify USMARC to accommodate the metadata standard were prepared and presented to the interdivisional MARBI Committee[6] at the American Library Association's (ALA) Annual Conference in June 1994. These proposals suggested modifying five existing fields by adding new subfields and (in some cases) defining a new indicator, adding two fields from the USMARC Format for Community Information, and defining six new fields. After much discussion, MARBI approved changes to five fields (with some modifications to the proposals), the addition of the fields from the community information format, and the addition of six new fields. The new fields were approved in a provisional status.

The fields that were modified were: Coded Cartographic Mathematical Data (034), Source of Acquisition (037), Cartographic Mathematical Data (255), Security Classification Control (355), and Nonspecific Relationship Entry (787). Those added from the Community Information Format were: Hours (301) and Address (270); and new fields created were: Geospatial Reference Data (342), Planar Coordinate Data (343), Digital Graphic Representation (352), Data Quality Note (514), Entity and Attribute Information Note (551), and Source of Data Entry (786).

In order to facilitate the use of USMARC in communicating metadata information, a crosswalk to correlate the geospatial metadata data element to the appropriate USMARC field or subfield was prepared (see figure 1, pp. 102–110). As you will notice, the portion of the metadata standard dealing with dial-up access is not yet supported by USMARC. The proposal to support these data elements was presented to MARBI and accepted during the ALA Midwinter Meeting in February 1995, so the crosswalk will soon be complete.

Notes

1. Formerly the American Society for Testing and Materials.
2. Federal Geographic Data Committee, "Content Standard for Digital Geospatial Metadata" (Washington, D.C.: Federal Geographic Data Committee, 1994), 1.
3. Federal Geographic Data Committee, v.
4. Federal Geographic Data Committee, vi.
5. Executive Order 12096, Sec. 3(b).
6. MARBI, the Machine-Readable Bibliographic Information Committee, is made up of members from the Association for Library Collections & Technical Services, Library and Information Technology Association, and Reference and Adult Services Division.

Bibliography

Federal Geographic Data Committee. 1994. "Content Standard for Digital Geospatial Metadata." (June 8). Washington, D.C.: Federal Geographic Data Committee.

U.S. Department of Commerce. 1992. Spatial Data Transfer Standard (SDTS). (Federal Information Processing Standard 173.) Washington, D.C.: Department of Commerce, National Institute of Standards and Technology.

USMARC Format for Bibliographic Data. 1994. Washington, D.C.: Library of Congress, Network Development and MARC Standards Office.

USMARC Format for Community Information. 1993. Washington, D.C.: Library of Congress, Network Development and MARC Standards Office.

CSDGM		USMARC	
line number	Data Element		TAG
1	**Identification Information**	**	**
1.1	Citation	*see Section 8: Citation Information: Item*	**
1.2	Description	**	**
1.2.1	Abstract	Summary, etc. note. Summary, etc.	520
1.2.2	Purpose	Summary, etc. note. Expansion of summary, etc.	500
1.2.3	Supplemental Information	General note	500
1.3	Time Period of Content	*see Section 9: Time period information*	
1.3.1	Currentness Reference	General note	500
1.4	Status	**	**
1.4.1	Progress	Action note. Status	583‡l
1.4.2	Maintenance and Update Frequency	Current publication frequency	310‡a
1.5	Spatial Domain	**	**
1.5.1	Bounding Coordinates	**	**
1.5.1.1	West Bounding Coordinate	Cartographic mathematical data. Statement of coordinates	255‡c
		Coded cartographic mathematical data. coordinate - westernmost longitude	034‡d
1.5.1.2	East Bounding Coordinate	Cartographic mathematical data. Statement of coordinates	255‡c
		Coded cartographic mathematical data. coordinate - easternmost longitude	034‡e
1.5.1.3	North Bounding Coordinate	Cartographic mathematical data. Statement of coordinates	255‡c
		Coded cartographic mathematical data. coordinate - northernmost latitude	034‡f
1.5.1.4	South Bounding Coordinate	Cartographic mathematical data. Statement of coordinates	255‡c
		Coded cartographic mathematical data. coordinate - southernmost latitude	034‡g
1.5.2	Data Set G-Polygon	**	**
1.5.2.1	Data Set G-Polygon Outer G-Ring	Cartographic mathematical data. Outer G-ring coordinate pairs	255‡f
1.5.2.1.1	G-Ring Latitude	G-ring latitude	034‡s
1.5.2.1.2	G-Ring Longitude	G-ring longitude	034‡t
1.5.2.2	Data Set G-Polygon Exclusion G-Ring	Cartographic mathematical data. Exclusion G-ring coordinate pairs	255‡g
		G-ring latitude & G-ring longitude	034‡s & ‡t
1.6	Keywords	**	**
1.6.1	Theme	**	**
1.6.1.1	Theme Keyword Thesaurus	Topical subject thesaurus	650‡2
1.6.1.2	Theme Keyword	Topical subject	650‡a
1.6.2	Place	**	**
1.6.2.1	Place Keyword Thesaurus	Geographic subject thesaurus	651‡2
1.6.2.2	Place Keyword	Geographic subject	651‡a
1.6.3	Stratum	**	**
1.6.3.1	Stratum Keyword Thesaurus	Geographic subject thesaurus	651‡2
1.6.3.2	Stratum Keyword	Geographic subject	651‡a
1.6.4	Temporal	**	**
1.6.4.1	Temporal Keyword Thesaurus	Topical subject thesaurus	650‡2
1.6.4.2	Temporal Keyword	Topical subject	650‡a
1.7	Access Contraints	Restrictions on access note. Terms governing access	506‡a
1.8	Use Constraints	Terms governing use and reproduction note. Terms governing use and reproduction	540‡a

** Compound data element; made up of parts

Figure 1. Crosswalk

CSDGM		USMARC	
line number	Data Element		TAG
1.9	Point of Contact	*see Section 10: Contact Information*	
1.10	Browse Graphic	**	**
1.10.1	Browse Graphic File Name	Electronic location and access. Electronic name	856‡f
1.10.2	Browse Graphic File Description	Electronic location and access. Public note	856‡z
1.10.3	Browse Graphic File Type	Electronic location and access. Public note	856‡z
1.11	Data Set Credits	Added entries: personal	700
		Added entries: corporate	710
1.12	Security Information	**	**
1.12.1	Security Classification System	Security classification control. System name	355‡e
1.12.2	Security Classification	Security classification control. Security classification	355‡a
1.12.3	Security Handling Description	Security classification control. Handling instructions	355‡b
1.13	Native Data Set Environment	Systems details note	538
1.14	Cross Reference	*see Section 8: Citation Information: Related work*	
2	**Data Quality Information**	**	**
2.1	Attribute Accuracy	**	**
2.1.1	Attribute Accuracy Report	Data quality note. Attribute accuracy report	514‡a
2.1.2	Quantitative Attribute Accuracy Assessment	**	**
2.1.2.1	Attribute Accuracy Value	Data quality note. Attribute accuracy value	514‡b
2.1.2.2	Attribute Accuracy Explanation	Data quality note. Attribute accuracy explanation	514‡c
2.2	Logical Consistency Report	Data quality note. Logical consistency	514‡d
2.3	Completeness Report	Data quality note. Completeness report	514‡e
2.4	Positional Accuracy	**	**
2.4.1	Horizontal Positional Accuracy	**	**
2.4.1.1	Horizontal Positional Accuracy Report	Data quality note. Horizontal position accuracy report	514‡f
2.4.1.2	Quantitative Horizontal Positional Accuracy Assessment	**	**
2.4.1.2.1	Horizontal Positional Accuracy Value	Data quality note. Horizontal position accuracy value	514‡g
2.4.1.2.2	Horizontal Positional Accuracy Explanation	Data quality note. Horizontal position accuracy explanation	514‡h
2.4.2	Vertical Positional Accuracy	**	**
2.4.2.1	Vertical Positional Accuracy Report	Data quality note. Vertical positional accuracy report	514‡i
2.4.2.2	Quantitative Vertical Positional Accuracy Assessment	**	**
2.4.2.2.1	Vertical Positional Accuracy Value	Data quality note. Vertical positional accuracy value	514‡j
2.4.2.2.2	Vertical Positional Accuracy Explanation	Data quality note. Vertical positional accuracy explanation	514‡k
2.5	Lineage	**	**
2.5.1	Source Information	**	**
2.5.1.1	Source Citation	*see Section 8: Citation Information: Source*	
2.5.1.2	Source Scale Denominator	Source of data entry. Material specific details	786‡m
2.5.1.3	Type of Source Media	Source of data entry. Physical description	786‡h
2.5.1.4	Source Time Period of Content	Source of data entry. Period of content	786‡j
2.5.1.4.1	Currentness Reference	Source of data entry. Period of content	786‡j
2.5.1.5	Source Citation Abbreviation	Source of data entry. Abbreviated title	786‡p
2.5.1.6	Source Contribution	Source of data entry. Source contribution	786‡v

** Compound data element; made up of parts

Figure 1, continued. Crosswalk

CSDGM		USMARC	
line number	Data Element		TAG
2.5.2	Process Step	**	**
2.5.2.1	Process Description	Action note. Process description	583‡a
2.5.2.2	Source Used Citation Abbreviation	Action note. Action identification	583‡b
2.5.2.3	Process Date	Action note. Time of action	583‡c
2.5.2.4	Process Time	Action note. Time of action	583‡c
2.5.2..5	Source Produced Citation Abbreviation	Action note. Action identification	583‡b
2.5.2.6	Process Contact	Action note. Action agent	583‡k
2.6	Cloud Cover	Data quality note. Cloud cover	514‡m
3	**Spatial Data Organization Information**	**	**
3.1	Indirect Spatial Reference	Digital graphic representation. Indirect reference description	352‡i
3.2	Direct Spatial Reference Method	Digital graphic representation. Direct reference method	352‡a
3.3	Point and Vector Object Information	**	**
3.3.1	SDTS Terms Description	**	**
3.3.1.1	SDTS Point and Vector Object Type	Digital graphic representation. Object type	352‡b
3.3.1.2	Point and Vector Object Count	Digital graphic representation. Object count	352‡c
3.3.2	VPF Terms Description	**	**
3.3.2.1	VPF Topology Level	Digital graphic representation. VPF topology level	352‡g
3.3.2.2	VPF Point and Vector Object Type	Digital graphic representation. Object type	352‡b
3.4	Raster Object Information	**	**
3.4.1	Raster Object Type	Digital graphic representation. Object count	352‡b
3.4.2	Row Count	Digital graphic representation. Row count	352‡d
3.4.3	Column Count	Digital graphic representation. Column count	352‡e
3.4.4	Vertical Count	Digital graphic representation. Vertical count	352‡f
4	**Spatial Reference Information**	**	**
4.1	Horizontal Coordinate System Definition	**	**
4.1.1	Geographic	**	**
4.1.1.1	Latitude Resolution	Geospatial reference data. Latitude resolution	342‡c
4.1.1.2	Longitude Resolution	Geospatial reference data. Longitude resolution	342‡d
4.1.1.3	Geographic Coordinate Units	Geospatial reference data. Coordinate or distance unit	342‡b
4.1.2	Planar	**	**
4.1.2.1	Map Projection	**	**
4.1.2.1.1	Map Projection Name	Geospatial reference data. Name	342‡a
4.1.2.1.2	Map Projection Parameters	**	**
4.1.2.1.2.1	Standard Parallel	Geospatial reference data. Standard parallel or oblique line latitude	342‡e
4.1.2.1.2.2	Longitude of Central Meridian	Geospatial reference data. Longitude of central meridian or projection center	342‡g
4.1.2.1.2.3	Latitude of Projection Origin	Geospatial reference data. Latitude of projection origin or projection center	342‡h
4.1.2.1.2.4	False Easting	Geospatial reference data. False easting	342‡i
4.1.2.1.2.5	False Northing	Geospatial reference data. False northing	342‡j
4.1.2.1.2.6	Scale Factor at Equator	Geospatial reference data. Scale factor	342‡k
4.1.2.1.2.7	Height of Perspective Point Above Surface	Geospatial reference data. Height of persepective point above surface	342‡l
4.1.2.1.2.8	Longitude of Projection Center	Geospatial reference data. Longitude of central meridian or projection center	342‡g
4.1.2.1.2.9	Latitude of Projection Center	Geospatial reference data. Latitude of projection origin or projection center	342‡h
4.1.2.1.2.10	Scale Factor at Center Line	Geospatial reference data. Scale factor	342‡k

** Compound data element; made up of parts

Figure 1, continued. Crosswalk

CSDGM		USMARC	
line number	Data Element		TAG
4.1.2.1.2.11	Oblique Line Azimuth	**	**
4.1.2.1.2.11.1	Azimuthal Angle	Geospatial reference data. Azimuthal angle	342‡m
4.1.2.1.2.11.2	Azimuth Measure Point Longitude	Geospatial reference data. Azimuth measure point longitude or straight vertical longitude from pole	342‡n
4.1.2.1.2.12	Oblique Line Point	**	**
4.1.2.1.2.12.1	Oblique Line Latitude	Geospatial reference data. Standard parallel or oblique line latitude	342‡e
4.1.2.1.2.12.2	Oblique Line Longitude	Geospatial reference data. Oblique line longitude	342‡f
4.1.2.1.2.13	Straight Vertical Longitude from Pole	Geospatial reference data. Azimuth measure point longitude or straight vertical longitude from pole	342‡n
4.1.2.1.2.14	Scale Factor at Projection Origin	Geospatial reference data. Scale factor	342‡k
4.1.2.1.2.15	Landsat Number	Geospatial reference data. Landsat number and path number	342‡o
4.1.2.1.2.16	Path Number	Geospatial reference data. Landsat number and path number	342‡o
4.1.2.1.2.17	Scale Factor at Central Meridian	Geospatial reference data. Scale factor	342‡k
4.1.2.1.3	Other Projection's Definition	Geospatial reference data. Local planar, local, or other projection or grid description	342‡v
4.1.2.2	Grid Coordinate System	**	**
4.1.2.2.1	Grid Coordinate System Name	Geospatial reference data. Name	342‡a
4.1.2.2.2	Universal Transverse Mercator	**	**
4.1.2.2.2.1	UTM Zone Identifier	Geospatial reference data. Zone identifier	342‡p
4.1.2.2.3	Universal Polar Stereographic	**	**
4.1.2.2.3.1	UPS Zone Identifier	Geospatial reference data. Zone identifier	342‡p
4.1.2.2.4	State Plane Coordinate System	**	**
4.1.2.2.4.1	SPCS Zone Identifier	Geospatial reference data. Zone identifier	342‡p
4.1.2.2.5	ARC Coordinate System	**	**
4.1.2.2.5.1	ARC System Zone Identifier	Geospatial reference data. Zone identifier	342‡p
4.1.2.2.6	Other Grid System's Definition	Geospatial reference data. Local planar, local, or other projection or grid description	342‡v
4.1.2.3	Local Planar	**	**
4.1.2.3.1	Local Planar Description	Geospatial reference data. Local planar, local, or other projection or grid description	342‡v
4.1.2.3.2	Planar Georeference Information	Geospatial reference data. Local planar or local georeference information	342‡w
4.1.2.4	Planar Coordinate Information	**	**
4.1.2.4.1	Planar Coordinate Encoding Method	Planar coordinate data. Planar coordinate encoding level	343‡a
4.1.2.4.2	Coordinates Representation	**	**
4.1.2.4.2.1	Abscissa Resolution	Planar coordinate data. Abscissa resolution	343‡c
4.1.2.4.2.2	Ordinate Resolution	Planar coordinate data. Ordinate resolution	343‡d
4.1.2.4.3	Distance and Bearing Representation	**	**
4.1.2.4.3.1	Distance Resolution	Planar coordinate data. Distance resolution	343‡e
4.1.2.4.3.2	Bearing Resolution	Planar coordinate data. Bearing resolution	343‡f
4.1.2.4.3.3	Bearing Units	Planar coordinate data. Bearing units	343‡g
4.1.2.4.3.4	Bearing Reference Direction	Planar coordinate data. Bearing reference direction	343‡h
4.1.2.4.3.5	Bearing Reference Meridian	Planar coordinate data. Bearing reference meridian	343‡i
4.1.2.4.4	Planar Distance Units	Planar coordinate data. Planar distance units	343‡b
4.1.3	Local	**	**
4.1.3.1	Local Description	Geospatial reference data. Local planar, local, or other projection or grid description	342‡v
4.1.3.2	Local Georeference Information	Geospatial reference data. Local planar or local georeference information	342‡w

** Compound data element; made up of parts

Figure 1, continued. Crosswalk

CSDGM		USMARC	
line number	Data Element		TAG
4.1.4	Geodetic Model	**	**
4.1.4.1	Horizontal Datum Name	Geospatial reference data. Names	342‡a
4.1.4.2	Ellipsoid Name	Geospatial reference data. Ellipsoid name	342‡q
4.1.4.3	Semi-major Axis	Geospatial reference data. Semi-major axis	342‡r
4.1.4.4	Denominator of Flattening Ratio	Geospatial reference data. Denominator of flattening ratio	342‡s
4.2	Vertical Coordinate System Definition	**	**
4.2.1	Altitude System Definition	**	**
4.2.1.1	Altitude Datum Name	Geospatial reference data. Name	342‡a
4.2.1.2	Altitude Resolution	Geospatial reference data. Vertical resolution	342‡t
4.2.1.3	Altitude Distance Units	Geospatial reference data. Coordinate or distance units	342‡b
4.2.1.4	Altitude Encoding Method	Geospatial reference data. Vertical encoding level	342‡u
4.2.2	Depth System Definition	**	**
4.2.2.1	Depth Datum Name	Geospatial reference data. Name	342‡a
4.2.2.2	Depth Resolution	Geospatial reference data. Vertical resolution	342‡t
4.2.2.3	Depth Distance Units	Geospatial reference data. Coordinate or distance units	342‡b
4.2.2.4	Depth Encoding Method	Geospatial reference data. Vertical encoding level	342‡u
5	**Entity and Attribute Information**	**	**
5.1	Detailed Description	**	**
5.1.1	Entity Type	**	**
5.1.1.1	Entity Type Label	Entity and attribute information note. Entity type label	551‡a
5.1.1.2	Entity Type Definition	Entity and attribute information note. Entity type definition and source	551‡b
5.1.1.3	Entity Type Definition Source	Entity and attribute information note. Entity type definition and source	551‡b
5.1.2	Attribute	**	**
5.1.2.1	Attribute Label	Entity and attribute information note. Attribute label	551‡c
5.1.2.2	Attribute Definition	Entity and attribute information note. Attribute definition and source	551‡d
5.1.2.3	Attribute Definition Source	Entity and attribute information note. Attribute definition and source	551‡d
5.1.2.4	Attribute Domain Values	**	**
5.1.2.4.1	Enumerated Domain	**	**
5.1.2.4.1.1	Enumerated Domain Value	Entity and attribute information note. Enumerated domain value	551‡e
5.1.2.4.1.2	Enumerated Domain Value Definition	Entity and attribute information note. Enumerated domain value definition and source	551‡f
5.1.2.4.1.3	Enumerated Domain Value Definition Source	Entity and attribute information note. Enumerated domain value definition and source	551‡f
5.1.2.4.2	Range Domain	**	**
5.1.2.4.2.1	Range Domain Minimum	Entity and attribute information note. Range domain minimum and maximum	551‡g
5.1.2.4.2.2	Range Domain Maximum	Entity and attribute information note. Range domain minimum and maximum	551‡g
5.1.2.4.3	Codeset Domain	**	**
5.1.2.4.3.1	Codeset Name	Entity and attribute information note. Codeset name and source	551‡h
5.1.2.4.3.2	Codeset Source	Entity and attribute information note. Codeset name and source	551‡h
5.1.2.4.4	Unrepresentable Domain	Entity and attribute information note. Unrepresentable domain	551‡i
5.1.2.5	Attribute Units of Measurement	Entity and attribute information note. Atttribute units of measurement and resolution	551‡j
5.1.2.6	Attribute Measurement Resolution	Entity and attribute information note. Atttribute units of measurement and resolution	551‡j
5.1.2.7	Beginning Date of Attribute Values	Entity and attribute information note. Beginning date and ending date of attribute value	551‡k
5.1.2.8	Ending Date of Attribute Values	Entity and attribute information note. Beginning date and ending date of attribute value	551‡k

** Compound data element; made up of parts

Figure 1, continued. Crosswalk

CSDGM line number	CSDGM Data Element	USMARC	TAG
5.1.2.9	Attribute Value Accuracy Information	**	**
5.1.2.9.1	Attribute Value Accuracy	Entity and attribute information note. Attribute value accuracy	551‡l
5.1.2.9.2	Attribute Value Accuracy Explanation	Entity and attribute information note. Attribute value accuracy explanation	551‡m
5.1.2.10	Attribute Measurement Frequency	Entity and attribute information note. Attribute measurement frequency	551‡n
5.2	Overview Description	**	**
5.2.1	Entity and Attribute Overview	Entity and attribute information note. Entity and attribute overview	551‡o
5.2.2	Entity and Attribute Detail Citation	Entity and attribute information note. Entity and attribute detail citation	551‡p
6	**Distribution Information**	**	**
6.1	Distributor	*see Section 10: Contact Information*	
6.2	Resource Description	Source of acquisition. Note	037‡n
6.3	Distribution Liability	Electronic location and access. Public note	856‡z
6.4	Standard Order process	**	**
6.4.1	Non-digital Form	Source of acquisition. Form of issue	037‡f
6.4.2	Digital Form	**	**
6.4.2.1	Digital Transfer Information	**	**
6.4.2.1.1	Format Name	Source of acquisition. Additional format characteristics	037‡g
6.4.2.1.2	Format Version Number	Source of acquisition. Additional format characteristics	037‡g
6.4.2.1.3	Format Version Date	Source of acquisition. Additional format characteristics	037‡g
6.4.2.1.4	Format Specification	Source of acquisition. Note	037‡n
6.4.2.1.5	Format Information Content	Source of acquisition. Note	037‡n
6.4.2.1.6	File Decompression Technique	Source of acquisition. Note	037‡h
6.4.2.1.7	Transfer Size	Source of acquisition. Additional format characteristics	037‡g
6.4.2.2	Digital Transfer Option	**	**
6.4.2.2.1	Online Option	**	**
6.4.2.2.1.1	Computer Contact Information	Electronic location and access. Instructions	856‡i
6.4.2.2.1.1.1	Network Address	Electronic location and access. IP address	856‡b
6.4.2.2.1.1.1.1	Network Resource Name	Electronic location and access. Electronic name	856‡f
6.4.2.2.1.1.2	Dialup Instructions	**	**
6.4.2.2.1.1.2.1	Lowest BPS		
6.4.2.2.1.1.2.2	Highest BPS		
6.4.2.2.1.1.2.3	Number DataBits		
6.4.2.2.1.1.2.4	Number StopBits		
6.4.2.2.1.1.2.5	Parity		
6.4.2.2.1.1.2.6	Compression Support		
6.4.2.2.1.1.2.7	Dialup Telephone		
6.4.2.2.1.1.2.8	Dialup File Name	Electronic location and access. Electronic name	856‡f
6.4.2.2.1.2	Access Instructions	Electronic location and access. Instructions	856‡i
6.4.2.2.1.3	Online Computer and Operating System	Electronic location and access. Operating system	856‡o

** Compound data element; made up of parts

Figure 1, continued. Crosswalk

CSDGM		USMARC	
line number	Data Element		TAG
6.4.2.2.2	Offline Options	**	**
6.4.2.2.2.1	Offline Media	Source of acquisition. Form of issue	037‡f
6.4.2.2.2.2	Recording Capacity	**	**
6.4.2.2.2.2.1	Recording Density	Source of acquisition. Additional format characteristics	037‡g
6.4.2.2.2.2.2	Recording Density Units	Source of acquisition. Additional format characteristics	037‡g
6.4.2.2.2.3	Recording Format	Source of acquisition. Additional format characteristics	037‡g
6.4.2.2.2.4	Compatibility Information	Source of acquisition. Note	037‡n
6.4.3	Fees	Source of acquisition. Terms of availability	037‡c
6.4.4	Ordering Instructions	Source of acquisition. Note	037‡n
6.4.5	Turnaround	Source of acquisition. Terms of availability	037‡c
6.5	Custom Order Process	Source of acquisition. Note	037‡n
6.6	Technical Prerequisites	Source of acquisition. Note	037‡n
6.7	Available Time Period	Source of acquisition. Terms of availability	037‡c
7	**Metadata Reference Information**	**	**
7.1	Metadata Date	Action note. Time of action	583‡c
7.2	Metadata Review Date	Action note. Public note	583‡z
7.3	Metadata Future Review Date	Action note. Public note	583‡z
7.4	Metadata Contact	Action note. Action agent	583‡k
7.5	Metadata Standard Name	Action note. Authorization	583‡f
7.6	Metadata Standard Version	Action note. Authorization	583‡f
7.7	Metadata Time Convention	Action note. Public note	583‡z
7.8	Metadata Access Constraints	Restrictions on access note. Materials specified & Terms governing access	506‡3 & ‡a
7.9	Metadata Use Constraints	Terms governing use and reproduction note. Materials specified & Terms governing use and reproducti	540‡3 & ‡a
7.10	Metadata Security Information	**	**
7.10.1	Metadata Security Classification System	Security classification control. System name	355‡e
7.10.2	Metadata Security Classification	Security classification control. Security classification	355‡a
7.10.3	Metadata Security Handling Description	Security classification control. Handling instructions	355‡b
8	**Citation Information: Item**	**	**
8.1	Originator	Added entries: Personal Added entries: Corporate	700 710
8.2	Publication Date	Publication, distribution, etc. Date of publication	260‡c
8.3	Publication Time	Publication, distribution, etc. Date of publication	260‡c
8.4	Title	Title	245‡a
8.5	Edition	Edition	250
8.6	Geospatial Data Presentation Form	Physical description. Extent	300‡a
8.7	Series Information	**	**
8.7.1	Series Name	Series statement. Title	440‡a
8.7.2	Issue Identification	Series statement. Vol. no./sequential designation	440‡v
8.8	Publication Information	**	**
8.8.1	Publication Place	Publication, distribution, etc. Place of publication	260‡a
8.8.2	Publisher	Publication, distribution, etc. Publisher	260‡b
8.9	Other Citation Details	General note	500
8.10	Online Linkage	Electronic location and access. Uniform resource locator	856‡u

** Compound data element; made up of parts

Figure 1, continued. Crosswalk

CSDGM line number	CSDGM Data Element	USMARC	TAG
8.11	Larger Work	Host item entry	773
8	**Citation Information: Source**	**	**
8.1	Originator	Source of data entry. Main entry heading	786‡a
8.2	Publication Date	Source of data entry. Place, publisher, and date of publication	786‡d
8.3	Publication Time	Source of data entry. Place, publisher, and date of publication	786‡d
8.4	Title	Source of data entry. Title	786‡t
8.5	Edition	Source of data entry. Edition	786‡b
8.6	Geospatial Data Presentation Form	Source of data entry. Physical description	786‡h
8.7	Series Information	**	**
8.7.1	Series Name	Source of data entry. Series data for related item	786‡k
8.7.2	Issue Identification	Source of data entry. Series data for related item	786‡k
8.8	Publication Information	**	**
8.8.1	Publication Place	Source of data entry. Place, publisher, and date of publication	786‡d
8.8.2	Publisher	Source of data entry. Place, publisher, and date of publication	786‡d
8.9	Other Citation Details	Source of data entry. Note about source	786‡n
8.10	Online Linkage	Source of data entry. Note about source	786‡n
8.11	Larger Work	Source of data entry. Note about source	786‡n
8	**Citation Information: Related work**	**	**
8.1	Originator	Nonspecific relationship entry. Main entry heading	787‡a
8.2	Publication Date	Nonspecific relationship entry. Place, publisher, and date of publication	787‡d
8.3	Publication Time	Nonspecific relationship entry. Place, publisher, and date of publication	787‡d
8.4	Title	Nonspecific relationship entry. Title	787‡t
8.5	Edition	Nonspecific relationship entry. Edition	787‡b
8.6	Geospatial Data Presentation Form	Nonspecific relationship entry. Physical description	787‡h
8.7	Series Information	**	**
8.7.1	Series Name	Nonspecific relationship entry. Series data for related item	787‡k
8.7.2	Issue Identification	Nonspecific relationship entry. Series data for related item	787‡k
8.8	Publication Information	**	**
8.8.1	Publication Place	Nonspecific relationship entry. Place, publisher, and date of publication	787‡d
8.8.2	Publisher	Nonspecific relationship entry. Place, publisher, and date of publication	787‡d
8.9	Other Citation Details	Nonspecific relationship entry. Note about source	787‡n
8.10	Online Linkage	Nonspecific relationship entry. Note about source	787‡n
8.11	Larger Work	Nonspecific relationship entry. Note about source	787‡n
9	**Time Period Information**	**	**
9.1	Single Date/Time	**	**
9.1.1	Calendar Date	Time period of content. Formatted 9999 B.C. through A.D. time period	045‡b
9.1.2	Time of Day		
9.20	Multiple Dates/Times	Time period of content. Formatted 9999 B.C. through A.D. time period	045‡b
9.30	Ranges of Dates/Times	**	**
9.3.1	Beginning Date	Time period of content. Formatted 9999 B.C. through A.D. time period	045‡b
9.3.2	Beginning Time		
9.3.3	Ending Date	Time period of content. Formatted 9999 B.C. through A.D. time period	045‡b
9.3.4	Ending Time		

** Compound data element; made up of parts

Figure 1, continued. Crosswalk

CSDGM		USMARC	
line number	Data Element		TAG
10	**Contact Information**	**	**
10.1	Contact Person Primary	**	**
10.1.1	Contact Person	Address. Contact	270‡p
10.1.2	Contact Organization	Address. Title of contact	270‡q
10.2	Contact Organization Primary	Address. Contact	270‡p
10.3	Contact Position	Address. Title of contact	270‡q
10.4	Contact Address	**	**
10.4.1	Address Type	Address. Type of address	270‡i
10.4.2	Address	Address. Address	270‡a
10.4.3	City	Address. City	270‡b
10.4.4	State or Province	Address. State or province	270‡c
10.4.5	Postal Code	Address. Country	270‡e
10.4.6	Country	Address. Postal code	270‡d
10.5	Contact Voice Telephone	Address. Telephone number	270‡k
10.6	Contact TDD/TTY Telephone	Address. TDD or TTY number	270‡n
10.7	Contact Facsimile Telephone	Address. Fsx number	270‡l
10.8	Contact Electronic Mail Address	Address. Electronic mail address	270‡m
10.9	Hours of Service	Address. Hours contact available	270‡r
10.10	Contact Instructions	Address. Public note	270‡z

** Compound data element; made up of parts

Figure 1, continued. Crosswalk

GIS and Research Libraries: One Perspective

Nancy M. Cline and
Prudence S. Adler

The Association of Research Libraries (ARL), in partnership with members of the GIS community, manages the ARL GIS Literacy Project. The project seeks to introduce, educate, and equip librarians with skills needed to provide access to spatially referenced data. The project was formulated in a manner that permits each library to design a program unique to that institution's needs. The program at Pennsylvania State University is illustrative of how one library has grappled with the introduction of GIS and is an example of an initiative that is flexible enough to meet the changing demands on campus and beyond, while utilizing the project as an opportunity to experiment with networked-based services.

In June of 1992, the Association of Research Libraries (ARL), in partnership with members of the geographic information systems (GIS) community, initiated the ARL GIS Literacy Project. The multiphased project seeks to introduce, educate, and equip librarians with skills needed to provide access to spatially referenced data in all formats and to provide effective access to selected electronic information resources in library collections. Collaboration with others in the public and private sectors is a key element of the project and has been instrumental to the successful integration of GIS services and resources into libraries.

The increasing reliance upon GIS by multiple communities, including government agencies at all levels and members of the academic and research community, signaled the need for research librarians to become effective users of GIS. It is the mission of ARL and research libraries "to promote equitable access to, and effective use of recorded knowledge in support of teaching, research, scholarship, and community service." Thus research libraries are constantly challenged with the introduction of new services and formats, and with the design of training programs that will assist in the effective integration of new programs and services into research libraries. GIS is one such service. The amount and nature of spatial data, as well as its use by a growing and diverse array of campus departments and other users, presents opportunities for libraries to rethink current practice and to do so in an environment conducive to research, education, and public access.

The goals of the ARL GIS Literacy Project were designed to meet the current needs of libraries and users while addressing the changes that libraries are facing during a time of experimentation, transition, and transformation to increasingly networked-based services. These goals include:

- the introduction of GIS to a variety of libraries (e.g., public, state-based, and academic and university libraries in public and private institutions) to address diverse user information needs;
- the development of a team of GIS professionals within the research library community willing to lend time and expertise to applications, user training, and education programs related to GIS;
- the encouragement of connections among federal, state, and local GIS users and information;
- the promotion of research, education, and the public right-to-know through improved access to government information;
- the initiation of library projects to explore new applications of spatially referenced data and to evaluate the introduction of these services in research libraries; and
- the implementation of programs to allow institutions that have invested in networking capabilities to leverage the sharing of resources via networks.

The ARL GIS Literacy Project provides a framework and a forum for research libraries to introduce, experiment with, and engage in GIS activities. ARL, in cooperation with GIS vendors and users, solicits donations of GIS software and data, organizes regular training sessions for project participants, sponsors an electronic mail list, and works with government agencies on GIS programs and related issues. Financial support and expertise have been provided by GIS vendors, GIS practitioners in the public and private sectors, and foundations.

Although originally envisioned as a one-year project for twenty-five research libraries, well over seventy libraries (divided into two phases) are participating, and a third GIS program has been launched with twenty-eight Canadian research libraries. The ARL project was formulated in a manner that permitted each library to design a program customized to its own needs and that would integrate GIS services locally. Indeed, no one model has emerged, with each participant designing programs unique to their institutional needs. One common element to those libraries with more active programs is the ongoing commitment of resources, including personnel, by library management. A second success factor

Nancy Cline is Dean, University Libraries, Pennsylvania State University; **Prudence Adler** is Assistant Executive Director, Federal Relations and Information Policy, Association of Research Libraries.

relates to the project's third goal—the encouragement of connections between federal, state, and local GIS users and information. This goal was, from the outset, seen as critical to each institution's success with the project, particularly in viewing GIS services as a campus-based resource versus a department-based resource. In the past year, we have witnessed a significant increase in campus-based licensing arrangements for GIS software and data, with the library participating as a key player. It is now very evident that GIS activities will be a central public service function in the years ahead.

The GIS program at the Pennsylvania State University is illustrative of how one library has grappled with the introduction of GIS in an environment where there is a broad range of information needs and where there was already a significant investment in GIS services in a variety of other campus departments. The Penn State library's program provides an excellent example of how one research library has established an initiative that is flexible enough to meet the constantly changing demands on campus and beyond, while at the same time utilizing the ARL project as an opportunity to experiment with new networked-based services. Finally, a strong and evolving GIS program has emerged because of the library administration's substantial commitment to devoting the needed resources to this endeavor—including, most importantly, personnel.

Penn State University—A Perspective on GIS

Pennsylvania State University is a state-related, multicampus university. It is the land-grant university of Pennsylvania; and its mission is to provide teaching, research, and public service of the highest quality to the people of the commonwealth, the nation, and the world. The university has approximately 70,000 students and more than 4,000 full-time faculty distributed across twenty-three campus locations.

The University Libraries, with collections of over 3.3 million cataloged items and significant collections of maps, archives, government documents, microforms, audio visual materials, and other items, provides centralized support to meeting the academic program needs across all campuses. Access to the libraries' holdings is through LIAS, the Libraries Information Access System; and there is increasing emphasis on acquiring electronic resources that can be made available on the network. Any opportunity to incorporate additional computer databases, especially ones that can be offered on the network, is given serious consideration since it enables the university to provide a more equitable information environment at all its campuses. Suggestions come from librarians with collection development assignments, from students and faculty, from consortia partners, from publishers, and from other channels.

Penn State has a strong commitment to strategic planning, which is closely tied to budget allocations. In 1992, in response to declining appropriations from the state, the university sharpened the focus of its planning. All units were required to cut their budgets by 10 percent, taking the reductions over a three-year period. The university then reallocated a percentage of funding to each unit, to be used for new or enhanced programs of strategic value to the institution. Important new initiatives were thereby funded out of our own resources, since new sources of funds were practically nonexistent. It was in this context that the University Libraries began to explore a role with geographic information systems.

GIS: A New Initiative

When ARL embarked upon the ARL GIS Literacy Project, this new direction offered Penn State an ideal opportunity to assess the potential impact of GIS for an academic research library. There were, however, many factors to be weighed before deciding to participate in the project. For Penn State, as for many other research libraries, the early 1990s marked a period of diminished budgets, and there were some who believed that undertaking a major new initiative was not a fiscally sound decision. GIS, unlike bibliographic databases, could not be easily fitted into the decision-making scheme, and it was hard for some librarians to envision why GIS might be important to anyone other than geographers and map aficionados. In addition, GIS did not fit the typical library paradigm of bibliographic data in word-based retrieval systems; it also seemed to require additional steps in helping users actually work with data. GIS went beyond the traditional levels of reference service in an academic library context. Some wondered if we could afford such levels of support. However, others—including many documents librarians—had an early grasp of the vast amounts of data that might become available from federal agencies and saw GIS as a key initiative in providing service to users. From my perspective as dean of the University Libraries, and as a member of the ARL board, I had developed a personal understanding of the potential value of participating in the project, but also knew that if we were to succeed as a partner in the project it would take the commitment of people close to the user services.

After careful consideration, Penn State agreed to "sign on." The major goal of the ARL project—to introduce, educate, and equip depository librarians with the skills needed to provide access to spatially referenced data in multiple formats and provide effective access to selected federal electronic information resources in depository collections—was a natural fit for a library that had made a significant investment in its federal depository designation for about three-quarters of a century. If the geographic resources being produced by the federal government were moving to electronic format, we had to be positioned to work with those resources and to assist patrons in accessing them. The ARL project seemed to offer a bridge to those future information environments in which we intended to be a leading player. As a multicampus university, we hoped to provide a capability that would extend across many university locations.

Also important was the fact that Penn State University had a strong geography department and that many librarians felt confident that the faculty and students would derive benefit from the libraries' involvement in the GIS project. We were later to discover that our clientele for GIS was much more diverse, coming from many different departments, and that many of the geographers were already using more sophisticated software than the project offered. At the university's University Park campus, our largest campus and the hub for university administration, we identified a strong core of faculty who wanted to expand the use of digital cartographic information.

Participation in the ARL GIS project was a strategic choice. The value of embarking on it was not to achieve short-term gains. Rather, it was envisioned as a long-term investment that would position the university's libraries to work effectively with a major set of resources—spatial data—and to do so in a networked environment.

The University Libraries at Penn State proposed to ARL that, if chosen as a participant, staff would "develop a joint team of library and teaching faculty who would develop both GIS skills and expertise with government files. We envisioned that the librarians would incorporate GIS products into our daily reference service and include demonstrations and discussions of GIS software in classes given to students using federal data." We intended to draw on a pool of experience that already existed in the geography department's GIS lab. This lab used software from the Environmental Research System Institute (ESRI), and lab personnel were willing to help us in developing skills among the libraries' staff and faculty.

Evolving a GIS Agenda: A Collaborative Effort

Building on the computer skills and the technological and network infrastructure that existed in the geography department, the Center for Academic Computing, and the University Libraries, staff would experiment with "developing and sharing GIS applications in a networked environment." We predicted that as "more GIS based files are distributed by GPO, this interest on campus will only grow exponentially. It is essential that we develop in libraries the skills to deal with this information format." A short time later, we are still discovering how true those words were—now, more than ever, research libraries must develop capabilities to work with spatial data and devise ways of collecting and sharing vast sets of data. User expectations are growing rapidly in areas relating to GIS.

At Penn State, we chose initially to focus on the GPO/depository aspect of GIS and agreed to have the head of the documents and maps section of the library serve as the lead librarian on the ARL project. By the fall of 1992, we had installed equipment in the documents and maps section and soon thereafter installed similar hardware in the Earth and Mineral Sciences Library (the branch library serving the geography department). The plan was to train staff who worked in these libraries to use GIS/ArcView. The complement of employees who were trained included part-time students, staff, and librarians. Although we had decided upon the people to participate, we still had only tentative ideas about how the service would emerge, what data might be involved, and where we might be heading with the project.

As an initial step, we began to promote GIS around the University Park campus, beginning with a series of demonstrations. Throughout the 1993/94 academic year we had a small complement of librarians and staff who could work comfortably with GIS, using ArcView. This included the librarians from the Earth and Mineral Sciences Library, the maps librarian, and the head of documents and maps section. Some of the staff in these units were also mastering ArcInfo and assisting in the use of more sophisticated GIS resources. Relationships were slowly building with users of GIS and, on a steady basis, there were new users wanting to learn to use specific items. Nevertheless, it somehow seemed that the GIS initiative was not taking off in the manner we had anticipated.

In January 1994 a key faculty member from the geography department suggested that we work together to develop a GIS support position. Due to budget reductions, neither the department nor the libraries had sufficient funding to establish a full-time position alone, but by pooling our funding we thought we could establish the position. Moving into the somewhat murky waters of a shared position, we nonetheless proceeded to recruit a technology associate whose responsibilities were to manage the geography department's Advanced Geographic Information Systems Laboratory and to manage affiliated GIS computing facilities in the libraries. In addition, the person was to provide training and technical support services for both areas. As is often the case, the person hired into the new position helped to define it. In this instance, we saw a marked increase in expectations, generated by his enthusiasm for teaching others about GIS.

Within a year the position evolved and, as expectations grew, we soon recognized the need to provide greater stability for the program. A library faculty position (for a three-year period of time) was established. The position was defined with responsibility to provide academic leadership to the GIS initiatives, to work with appropriate faculty and staff in the libraries, to develop GIS collections and services for the university and its multiple constituencies, and to develop instructional and outreach programs for these resources. This position would continue to be shaped by the university's changing needs and the directions in which the ARL partnership evolved.

With a full-time position dedicated to GIS, the incumbent had the freedom to focus full energy on assessing user needs, developing training and educational components, and defining strategies to capitalize upon the software and data resources received through the ARL project and the depository program. One of the most important gains was that we now had a person who could move between the various library units, spend time in the academic departments, talk to people in the computer lab, and thereby begin to develop a strong network of people—people skilled in the uses of GIS, people eager to learn more about GIS, people with specialties, people who had connections to other organizations. This essential interaction among people began to propel the GIS agenda and it soon appeared that the University Libraries would serve as the hub for activities relating to GIS within the university.

A conference was developed on the topic of digital spatial libraries, bringing to the University Park campus two leading experts: Nancy Tosta, from the U.S. Geological Survey; and Jack Dangermond, from Environmental Systems Research Institute, Inc. In addition, a panel of experts was assembled to respond to the keynote speakers. The panel included a faculty member, a documents/map librarian, a consultant, and the executive director of a national association of geographers. Attendees of the conference included students, faculty, librarians, and library staff. New interpersonal relationships were established from the conference and a whole new level of visibility was achieved.

Conclusion

In reflecting on the first few years, it is important to note that we have learned how critical it is to foster human interactions, or personal networking, to accomplish change. Having the expertise on staff was an important step, training programs were essential, and installing hardware and software were required steps. However, in seeking to realize the potential of such a powerful new capability as GIS, we underestimated the need to break out of traditional thinking and to reach into new areas. Adding the new position was a key in enabling progress. It not only added to the number of people working with GIS, it was a position that functioned with a degree of freedom to go out and make new friends, find strategic alliances, locate data resources, and promote GIS in research and teaching programs. We realized it was not fair to assume that the skills needed to work with GIS would be naturally existent within the library staff. GIS requires an understanding of computing and the ability to work with visual representation of data, in addition to the knowledge and skills typically found in libraries relating to the organization of data, knowledge of information retrieval systems, reference services, and collection development. To implement GIS as a strategic direction for a library, requires a commitment to developing this special combination of strengths, either in individuals or in teams.

The changes of the past year have helped modify our approach to GIS literacy. In considering what impact GIS will have on the learning environment in a university, we remain committed to the University Libraries serving a central role. At Penn State, the libraries are respected for their neutral role, serving the breadth of the university and respecting the diverse information needs of students, faculty, staff, and the public.

We envision providing a coordinated spatial data handling infrastructure for the university. This infrastructure will enhance the accessibility, communication, and use of spatially referenced data to support the full range of educational programs. We will create a supporting environment for the training and use of the technology, offering courses, seminars, or other forms of instruction. This environment will span academic disciplines and will accommodate undergraduate and graduate students, faculty, staff, and independent learners. The University Libraries intends to assist novice or nontechnical users and expects that more sophisticated users will receive detailed training through their departments.

The University Libraries have taken a lead in working with other organizations, as well as with units within the university, to provide a broad-use spatial data handling capability. It is to libraries that users continue to turn for data, for resources to assist in the interpretation of data, and for other materials to augment the data. The segue to spatial data is logical since libraries are experienced with collection development, cataloging, access, and preservation issues. In many instances, we find that others have a "use-it-once-and-forget-it" approach to spatial data, which will make it highly difficult to sustain research activities over time. Planned collection of data resources, with commitments to resource-sharing among institutions, would be beneficial to many constituencies. At present there is risk of substantial duplication of effort, even within a single institution. The library can use its position of neutrality to coordinate a program for defining the handling of digital spatial data. A critical part of this is to develop procedures for cataloging the existing data in order to reduce redundant data collection, overcome fragmentation of data coverage, and improve efficiency for users.

The growing quantity of digital spatial data is phenomenal, and more geographic information is available today in digital format than on paper. For some users this means greater empowerment, but for others there is a risk of isolation. Libraries remain a location to which any users can turn for access to this information.

Terabytes of digital spatial data are being produced daily. We cannot wait for digital library concepts to mature; instead, it is important to be a part of shaping the future through participation in projects such as the ARL GIS Literacy Project. At the same time, we need to develop pragmatic plans of how to build, maintain, and access spatial libraries if educational institutions are to adequately serve the growing information needs of students and faculty. It is hard to predict how GIS will evolve. The participation of academic and research libraries should ensure attention to several key issues: where will the data exist, how do users access data, how do we teach users to make effective use of spatial information, and how do we archive or preserve data for long-term usage in research?

At Penn State we have made a beginning. This is a strategic initiative and we will see many changes as our plans evolve. What is clear is that there is a logical and important role for the University Libraries to play in defining the opportunities for research and instruction using geographic information systems and spatial data resources.

Identifying Issues and Concerns: The University of Connecticut's MAGIC—A Case Study

Patrick McGlamery

This paper will outline the various issues and concerns of providing spatial data to a research community that have been encountered by the Map and Geographic Information Center (MAGIC) at the University of Connecticut's Homer Babbidge Library. The situation at the University of Connecticut is presumably little different from that of many other large state institutions of higher education. The university's library began developing its digital spatial data collection in 1987. Looking at the experiences at the University of Connecticut over the past eight years can provide a case study of the issues and concerns of migrating a spatial data collection from the paper domain of maps to the digital domain of geographic information systems (GIS).

History

At the 1991 annual meeting of the North American Cartographic Information Specialists (NACIS) in Milwaukee, Patrick McGlamery, map librarian at the Homer Babbidge Library, gathered together librarians and geographers to discuss the issues and challenges of dealing with digital spatial data in libraries. Five critical areas of interest emerged during that dialogue: collecting/archiving, cataloging/indexing, networking, distributing, and education.

In February 1992, McGlamery was invited to attend the National Center for Geographic Information and Analysis' (NCGIA) Research Initiative 9: "Institutions Sharing Geographic Information" in San Diego. For that session he wrote a paper, "Libraries as Institutions for Sharing," wherein he expanded on the agenda developed in Milwaukee. Those conceptual issues have become the agenda for the development of the Homer Babbidge Library's Map and Geographic Information Center.

Collecting/Archiving

The problems of the collection and long-term storage and preservation of digital information, especially digital cartographic and database files with spatially referenced fields, need to be addressed. In the evolving arena of digital information, hardware, software, and data formats and standards are of crucial importance. The problems of the nonstandard "publishing" of electronic information, the resultant problems of identification and procurement, as well as adherence to standard formats (as they are developed) demand attention.

> Libraries, especially research libraries, store large quantities of material . . . and have for centuries. Problems of space, prioritized use, varieties of media format and preservation of information are continuing concerns. For example, libraries have confronted the problems of preservation and conservation of materials with national and international strategies. These strategies have provided initiatives for insisting materials be published on acid-free paper, nationally coordinated comprehensive microfilming projects of brittle books among others. (McGlamery 1995)

The MAGIC Case Study

The collection development of MAGIC supports the research and teaching needs of the University of Connecticut. With digital data, as with analog data, our collection development policy is determined by area, scale, "language," and theme. The collection has been limited to data collection that is in the public domain. (Years ago we purchased street data for Connecticut from MapInfo and realized, belatedly, the limitations of the proprietary one-copy-one-machine model in a networked university environment). The primary collection interest of MAGIC is Connecticut. The secondary interest is New England, especially those features that directly affect Connecticut, such as Long Island Sound, the Connecticut River, and the transportation corridor between New York and Boston. The third interest is the United States and gross administrative divisions of the world's nations.

The Homer Babbidge Library's Map and Geographic Information Center is collecting its primary interest areas at a scale of 1:100,000 or greater, which means TIGER (Topologically Integrated, Geographically Encoded and Referenced), 1:100,000 Digital Line Graphs (DLG), 1:25,000 data and "40 feet to the inch" data of select areas. TIGER, of course, is distributed through the federal depository program on CD-ROM, as are the 1:100,000 DLGs. These data either have been or will be subsequently extracted from the CD-ROMs and converted to supported proprietary formats or the "languages" described below. 1:25,000 data are being acquired from the state's Department of Environmental Protection, and the "40 feet" data are being acquired from the Metropolitan District Commission in Hartford.

Patrick McGlamery is Map Librarian, Map and Geographic Information Center, Homer Babbidge Library, University of Connecticut.

MAGIC is collecting its secondary interest area at a scale of 1:100,000 to 1:1,000,000. These data will be taken from the 1:100,000, 1:250,000 Land Use/Land Cover, 30 Second Arc, and the Digital Chart of the World.

The tertiary interest areas will be a public domain dataset of the counties and states of the United States and the world. We have the data and are planning to convert the City/County Data Book going back to the 1940s from FORTRAN to dBase with geo-links, but that will be a project with separate funding. Congressional districts are now in the TIGER format, though not converted, and there is interest here in converting those—again, pending resources. At present the two ArcUSA disks provide good data, although the problem with them is their proprietary nature. When the term "language" is used, it refers to the proprietary file format; not whether it is in English, French, or German, but whether it is ArcInfo, MapInfo, or ASCII.

One philosophical underpinning is that "a library provides access to information, not simply data." A major part of the collection development process in the Map and Geographic Information Center consists of processing the data into supported formats, something that is fairly controversial. By converting the data, is it being manipulated beyond what librarians should be doing? Is the information feed being streamlined too much? In addition, MAGIC is offering programs such as MapInfo for Windows, MapInfo for DOS, ArcView and Idrisi to campus users across the network.

Collecting the data has historically not well-served the users coming to the map library. They could get the data, but more often than not were frustrated by their inability to get at the information in the data. Since software and data have been mounted on the MAGIC server, the issues have become the "learning curve," connectivity, and the user's own computer—all simpler to deal with than converting data.

Data Conversion

Converting data, then, has become part of the collection building process for MAGIC. The initial problems were determining what formats to support, then finding conversion software. Three packages that have been successful are TIGER to MapInfo Translator (TMT) 5.3 and, more recently, ArcInfo/MapInfo (AIMI) 2.1 and MapInfo's ArcLink 3.0. Neither of the last two programs is particularly easy to use, as they have their quirks and tend to have to be nursed, but both do the job after a fashion. Other conversion packages are in development.

Thematic Data

Thematic data are going to take us to places librarians have never thought of and will push the limits of our ability to manage internal communication and cooperation. Thematic data are data that can be geo-referenced. Census data, instances of Lyme disease, soil types, pottery shard attributes, automobile accidents—all this and more, as long as it is tied to the earth's surface, but not as maps—are attribute data that can be mapped. Do we collect it? Do we have a responsibility to it? As long as geo-coding attributes are maintained and the data are cataloged and made available, we do not have to worry about it. With that said, MAGIC is extracting data from CD-ROMs.

> **MAGIC: Collecting/Archiving Issues**
>
> Developing and maintaining collection development policies for digital data is as imperative and effective as it is for analog data.
>
> Differentiating between the research and reference collections is important.
>
> Determining where the research collection is to be housed—data integrity; just-in-time or just-in-case; fresh data or canned data.
>
> Preservation of data collections through consistent backup, proper storage, and refreshing of magnetic media is the role of the library.

Cataloging/Indexing

Navigating through electronic information is confounded by the sheer volume of information, its dynamic and unpublished nature, and its abstract, "virtual" dimension. As more sophisticated descriptive cataloging techniques are developed, standardized and made available, item level cataloging needs to focus less on the distribution media and more on the file and, indeed, field level. The multidimensional layering process of GIS analysis demands ever-increasing refinement of the descriptive art of the cataloger.

Libraries have developed, over the centuries, sophisticated processes of bibliographic (and cartographic) control and access through the development of the descriptive and subject catalog record. These processes have culminated in the past few years with the establishment of an international database of MARC records. MARC formats exist for books, maps, serials and, among others, computer files. The complex nature of GIS data is being described using Spatial MetaData standards. These standards, developed by the Federal Geographic Data

```
OCLC: 31399360
1      040   UCW ≠c UCW
2      043   n-us-ct
3      090   G/3783/F3/P3/1993/U5/MAP_WIN
4      099   MAGIC Fileserver G/3783/F3/P3/1993/U5/MAP_WIN
5      049   UCWC
6    110 1   United States. ≠b Bureau of the Census.
7  245 10    TIGER/Line 1992; the coast-to-coast digital map data base ≠g [computer files]
             ≠p Railroad Line Features. ≠p Fairfield County, Connecticut.
8      256   Computer data (9 files : 176, 22380, 131, 3864, 28160, 15872, 308, 88, 1536 bytes)
             and documentation (2 files).
9      260   Washington, D.C. : ≠b The Bureau, ≠c 1993 ≠e (Storrs, Conn. : Homer Babbidge
             Library, ≠g 1994).
10     538   Data in MapInfo for Windows format in \MAP_WIN. Data issued compressed into
             RAIL.ZIP using PKZIP utility.
11     522   TIGER\Line feature data for Fairfield County, Connecticut.
12     500   Accompanied by geospatial metadata and GIF image files
13     500   Data has been extracted from TIGER/Line 1992 to include only TIGER Line Feature,
             Class B, Railroad data. Includes address ranges for geo-coding and address matching.
14  505 0    Bethel—Bridgeport—Brookfield—Danbury—Darie—Easton—Fairfield—Greenwich—Monroe—New
             Canaa—New Fairfield—Newton—Norwalk—Redding—Ridgefield—Shelton—Sherman—Stamford—
             Stratford —Trumbull—Weston—Westport—Wilton towns.
15  651 0    Fairfield County (Conn.) ≠x Railroads ≠x Computer files.
16  651 0    Fairfield County (Conn.) ≠x Census, 1990 ≠x Computer files.
17  71022    Homer Babbidge Library. ≠b Map and Geographic Information Center.
18  852 0    ≠b magic.lib.uconn.edu ≠c G/3783/F3/P3/1993/U5/map_win
```

Figure 1
Example of a Record

Committee, are compatible with the MARC format and crosswalks are being created. (McGlamery 1995)

The MAGIC Case Study

MAGIC's holdings are described using MARC for the Homer Babbidge Library's OPAC, HOMER. Describing library materials for bibliographic control and access has been a primary activity of libraries for over a hundred years. Larsgaard's *Map Librarianship* has an excellent overview of map cataloging, including its pros and cons. MAGIC has been cataloging and classifying the data files mounted on the server. The MARC computer files format and the G Schedule Classification Scheme are being used. Figure 1 is an example of a record.

The 099 field uses an adapted DOS subdirectory hierarchy to emulate an LC classification number. The G Schedule was chosen because, though the information being cataloged is not a map, it is geographic material.

The 256 field is important to users because they are going to have to live with the size of these files. The data are compressed, "zipped," but when they are unzipped the user needs to understand the impact on their storage device. I have created a separate record for each format the data are stored in; MapInfo for Windows, MapInfo for DOS, ArcView (ArcInfo PC), and ArcInfo Interchange format (E00). Most of the record stays the same, except for the call number, the 256 field, and the 538 field.

The 505 field is searchable and allows the user to search by each of the towns in the county. Towns in Connecticut are the smallest administrative unit.

Finally, the 852 field provides location/call number, in this case the IP address of the server and the call number. With increased access to remote OPACs and through OCLC, users from around the world looking for Connecticut information are able to find the record, get the call number, and, connecting to MAGIC, retrieve the data.

The MARC record is how libraries have handled bibliographic control. It is not necessarily the best way to retrieve spatial or cartographic information, but is simply the most expedient in the library environment. Other descriptive formats are on the horizon. The Spatial MetaData (SMD) Standards or "Content Standard for Digital Geospatial Metadata," developed by the standards working group of the Federal Geographic Data Committee, is a parallel record that is being designed and implemented by the data community. SMD will provide a standard description of nontextual electronic data. A crosswalk to allow the librarian to transfer

MAGIC: Cataloging/Indexing Issues

MARC is only one of a number of text-based systems for locating spatial data. WAIS is another. The developing technologies of the National Spatial Data InfraStructure are evolving. Graphic interfaces to spatial data will become more effective than text-based interfaces.

MAGIC: Networking Issues

Network capacity for large numeric/graphic data files is limited. As the rest of the information world moves toward image data, such as documents and pictures, the spatial data provider can only benefit.

As difficult as it is for libraries to become equipped, it is more difficult for the user. Libraries need to become the advocate for the user's connectivity. sidebar

the SMD record into MARC is in development (see"The Making of a Standard" by Elizabeth Mangan, p. 99).

Networking

Large numeric databases, which are the basis of digital spatial data and cartographic formats, reside most comfortably and are processed most efficiently on intensive (or mainframe type) computers. Spatial data, on the other hand, are used most effectively on the personal computer and workstation. This dilemma is resolved, in part, by complex networking. Remote storage of information, ready distribution of data, and appropriate linkages that connect available information need to be studied and addressed.

> Libraries have a long history of aggressively utilizing computer networks. Bibliographic utilities, such as OCLC, have existed and flourished for decades, distributing bibliographic citations and library holdings nationally and internationally. These systems have allowed for a high level of information sharing among national and international library communities. More recently the Internet has allowed for interactive perusal of research libraries' bibliographic databases such as the Library of Congress, MELVYL, and even the University of Connecticut's HOMER, from remote sites throughout the world. (McGlamery 1995)

The MAGIC Case Study

Magic.lib.uconn.edu is the IP address of the fileserver at the University of Connecticut's Map and Geographic Information Center. It is a major step in the development of a virtual library at the university. The MAGIC fileserver was purchased as a stand-alone machine, with the idea that the library might be able to find the money to purchase LAN software. It has come together in a piecemeal fashion. In 1992 a Tangent 486DX EISA processor with 16 MB of RAM and a RAID (Redundant Array of Inexpensive Drives) SCSI device was purchased. It was a 33 MHz machine with two 676 MB hard drives, a 500 MB tape backup and a 16/4 Mbit TokenRing Card. The cost of the system was approximately $8,000. In January 1995 the motherboard was replaced to boost the CPU to a 90 MHz Pentium and 32 MB of RAM was added for an additional $2,500. Three 1.05 GB hard drives were installed for $2,600. The ten-user Novell site license cost $1,500. All machines in the library are networked. Though map libraries should make a strong commitment to make maps on map-sized paper, the center has an HP 560 Color Inkjet printer. There is a sense that the printing of maps by federal, state, and local government agencies will dry up in five years or so. (An irony is that, with so much spatial information, hard copy is difficult if not impossible to obtain. This can already be seen if we choose to look. The ESRI User Conferences produced "ARC/INFO Maps," a good example of maps that are plotted, but not printed. The U.S. Bureau of the Census' tract maps are another example; five to ten copies are made where five hundred would have at one time been printed. Plotters have a press-run of one copy.) In the not-too-distant future, large-format plotters will be necessary in order to supply our users with current mapping. Currently the electrostatic color plotter is the plotter of choice, but it is expensive.

Distribution

Data compilation continues to be the most difficult aspect of modern map production. The voracious appetite for information presented by GIS analysis and the large, complex nature of numeric data files and databases are evolving rapidly. Distributing electronic information in the appropriate file formats, with documentation, will streamline geographic analysis.

Modern libraries have developed extensive and complex distribution systems. Copying machines, microform copying, document delivery, and faxing are as much a part of those systems as are local, state, regional, and national interlibrary loan networks. (McGlamery 1995)

The MAGIC Case Study

At the University of Connecticut, MAGIC's holdings are made accessible to the user community in the Map and Geographic Information Center and the library (reference desks), on the university's WAN backbone, and via FTP to the world. A library provides access to information, not simply data. MAGIC supports four GIS and computer mapping software packages—ArcView, MapInfo for Windows, and MapInfo for DOS, and IDRISI—in addition to a growing variety of application packages such as MapExpert, AutoRoute, and PCGlobe. Access to spatial information includes the performance and maintenance of file conversion for supported formats.

Because of the multiple dimensions of the spatial data, access to digital spatial information is a complex endeavor. On a paper map, what you see is what you get. Digital data, especially spatial data, can be portrayed in a number of formats: as a map, a database, a report, or a spreadsheet. The map can be on a screen, in a file, as an image, or on paper. The library needs to set the limits regarding what access it can supported—the TIGER discs in a shoebox, or the extracted TIGER data available via the Internet (see the University of Virginia's World Wide Web offering at http://www.lib.virginia.edu/socsci/ for example). As mentioned above in the section on collection development, MAGIC extracts and stores all material for the state of Connecticut on its server. This includes TIGER, census attribute data, the State DEP, Digital Orthophotos, and so on. The vector data are in four proprietary formats: MapInfo for Windows, MapInfo for DOS, ArcView (or ArcInfo PC) and ArcInfo Interchange (E00); attribute data are in dBase 3.

MAGIC: Distribution Issues

Data and information are two very different products. Libraries must deliver information. Therefore data should be processed to the point where the user can easily get to the information.

The storage facilitites of libraries and users are different. As much as possible, data should be reduced to its most usable size. It is easier for the user to aggregate up than to synthesize down.

These are the formats most frequently used by the extended Connecticut GIS community. As the user community was being defined, storing the data in ASCII format was considered. That would transfer the burden of conversion to the user, not a large task for the "data elite," but an insurmountable one for the emerging GIS user. In addition to the impediment of data conversion, there is the need for software. For the librarian the goal is to get the user to the information as efficiently and quickly as possible; therefore both impediments need to be addressed.

Education

User education is a considerable hurdle to the effective use of GIS and spatial data. GIS is a whole new way of dealing with information—with database management, image processing, and spatial analysis. College level, secondary, and elementary curricula are necessary. Staff education about the use, storage, extraction, and retrieval of multidimensional files is an issue which should be addressed in library schools.

Libraries have offered bibliographic instruction beginning with the use of the card catalog in elementary schools to navigating the Internet with Mosaic. Map librarians are used to the user who needs to be helped into the graphics format of maps, and readily supply interpretive legends, advice on scale and projections. The standardization of the MLS curricula and the degree have established a consistent professional foundation. This foundation has been the underpinning to shared cataloging, interlibrary loan and collection development. (McGlamery 1995)

The MAGIC Case Study

Only in the past year has MAGIC begun to study the need for education and to strategize solutions. The first consideration is how to get the user logged onto MAGIC. A number of "how to" devices have been developed to introduce the user to MapInfo for Windows or DOS and to help them get logged on to the most general reference tool, MapUser. These really are the most simple aids. The next level of instruction has been a mixture of education and training. A "how to" device on digitizing a map was also developed, a difficult process, but one that is frequently done in the map library for novices. This past fall a MAGIC seminar series was begun. Once a month a speaker was invited to lecture on an aspect of GIS; topics ranged from a general overview of GIS, to statistical analysis of spatial data by a staff member of the state's Office of Policy and Management,

MAGIC: Education Issues

It is difficult to know the limits of training the user. Certainly training them to get at the information, and perhaps even to interpret the information at some basic level, but certainly not to serve as a GIS instructor, is advisable.

to a discussion of census mapping. This spring two seminars are scheduled, one on cartographic presentation and one on geo-coding. The goal of the series is to introduce faculty, staff, and graduate students to concepts in GIS. Attendance has been good, but most importantly varied. Primary attendees have tended to be the emerging GIS users in the business, nursing, economics, and computer science departments. The computer scientists seem naturally to gravitate to large datasets!

The ARL GIS Project

Since 1992 several opportunities have advanced the efforts of librarians to provide services to spatial data. The ARL GIS Literacy Project has been fulfilling its goals of developing a core of GIS professionals in the library community willing to lend time and expertise to new applications, user training, and GIS education programs. The ARL GIS Literacy Project has stimulated and encouraged the establishment of "centers of excellence" for GIS and it has promoted research, education, information, and the public right-to-know through improved access to government information. In addition the project has initiated library projects to explore new applications of spatially referenced data and has introduced GIS to public, academic, research, state, rural, and large urban libraries to address the diverse information needs of disparate user communities. (ARL 1994)

At the Homer Babbidge Library, the ARL GIS Literacy Project moved the library director to set aside a portion of money for the purchase of a sophisticated computer and to send the map librarian to a two-day workshop on GIS and ArcView software. This attention from the library's administration has provided a necessary impetus in the support and development of MAGIC.

Summary

Spatial information in research libraries has made great strides in the past three to four years. Truly, the librarian responsible for spatial data is riding turbulent seas. In order to manage the amount and variety of spatial data available today, a clear agenda is followed at the Homer Babbidge Library Map and Geographic Information Center. Issues of collecting/archiving, networking, cataloging/indexing, distributing, and education keep the role of the library in clear perspective.

References

Association of Research Libraries. GIS Literacy Project, 1992. Washington, D.C.: ARL.

Environmental Systems Research Institute, Inc. 1990. *Understanding GIS: The ARC/INFO Method.* Redlands, Calif.: Author.

Larsgaard, Mary Lynette. 1987. *Map Librarianship: An Introduction.* 2d ed. Littleton, Colo.: Libraries Unlimited.

McGlamery, Patrick. 1995. "Libraries as Institutions for Sharing." In *Sharing Spatial Information; Proceedings of the 9th Specialists' Session*, ed. Harlan Onsrud and Gerry Rushton.

Tomlin, C. Dana. 1990. *Geographic Information Systems and Cartographic Modeling.* Englewood Cliffs, N.J.: Prentice-Hall.

Key Terms

ARC/INFO A full-featured geographic information system (GIS) from ESRI.

ArcView Desktop browse and query software from ESRI with an advanced graphical user interface (GUI) and object-oriented scripting language.

bandwidth A measure of the volume of data that can flow through a communications link. Image data tend to exist as large data sets, thus moving image data sets from one computer to another requires high bandwidth or performance will be slowed.

browse image A low-resolution image that displays quickly on computer screens. Because they are typically of 512 x 512 pixels or less, they do not provide detailed information, but are useful for quick display.

Contents Standard (CSSM) The Contents Standard for Spatial Metadata. A document produced by the Federal Geographic Data Committee that describes spatial metadata. Approved June, 1994.

coordinate A set of numbers that designates location in a given coordinate reference system, such as x,y in a planar coordinate system or an x,y,z in a three-dimensional coordinate system. A coordinate represents a location on the earth's surface relative to other locations.

coordinate system A reference system for defining points on the earth's surface. A coordinate system can be planar (flat) or nonplanar (spherical). The Cartesian system (x,y) is an example of a planar coordinate system. Latitude and longitude (expressed as angles) is an example of a nonplanar coordinate system. A common coordinate system is used to spatially register geographic data for the same area.

data set A named collection of logically related data items arranged in a prescribed manner.

EOS The Earth Observing System. An effort to study the earth as system while tracking long-term changes on a global scale. EOS, a mission of the National Aeronautics and Space Administration, will produce petabytes of satellite image data.

FGDC The U.S. Federal Geographic Data Committee. Composed of representatives of several federal agencies, the FGDC has the lead role in defining spatial metadata standards, which it describes in the Contents Standards for Spatial Metadata.

geo-reference To establish the relationship of one data set to another through reference to common locations in both data sets. Geo-referencing requires that data be placed in a common coordinate system.

GIS Geographic Information System. An organzied collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information. Complex spatial analysis and geographic data processing is possible with a GIS that would be diffiuclt, time-consuming, or inpracticable otherwise.

Landsat A series of satellites that produce images of the earth.

metadata Metadata is information about the content, quality, condition and other characteristics of information kept in a database. Metadata is software accessible, and therefore is vital in the development of advanced database systems, such as a digital spatial library.

petabyte A measure of data size. One petabyte is equivalent to 1,000 terabytes.

quadrangle A four-sided figure, bounded by parallels of latitude and meridians of longitude, used as an area's unit in mapping. A well-known maps series is the USGS 7.5-minute topographic quadrangle. Each map in this map series covers 7.5 degrees of latitude and longitude and provides basic earth information such as elevation, hydrography, vegetation, and cultural features such as roads and buildings.

satellite image A picture of the earth taken from an earth-orbital satellite. Satellite images may be produced photographically, or by on-board scanners.

scale The extent of reduction needed to display a representation of the earth's surface on a map. A statement of a measure on the map and the equivalent measure on the earth's surface, often expressed as a representative fraction of distance, such as 1:24,000 (one unit of distance on the map represents 24,000 of the same units of distance on the earth). Map scale can also be expressed as a statement of equivalence using different units: for example, 1 inch = 1 mile or 1 inch = 2,000 feet.

Scale can be used as a measure of viewable detail; small scale implies less detail is visible, large scale implies more detail is visible. Thus, scale can be used to control display; as scale increases (becomes larger and more "zoomed in") more detail can be displayed without overcrowding the screen display.

STDS Spatial Data Transfer Standard. A federal standard designed to support the transfer of different types of geographic and cartographic spatial data. A wide variety of raster and vector data types, models, and structures, as well as associated attribute data, can be transferred between dissimilar systems using STDS. Also known as federal information processing standards (FIPS) 173.

terabyte A measure of data size. A terabyte of data is equivalent to 1,000 gigabytes of data or 1,000,000 megabytes of data.

vector A coordinate-based spatial data structure comonly used to represent geographic features. For example, a linear feature is represented as an ordered list of x,y coordinates.

Selected terms and definitions from the Glossary in the ESRI White Paper Series, GIS Approach to Digital Spatial Libraries, supplied courtesy of Environmental Systems Research Institute, Inc. (ESRI)

Electronic Information Delivery: Ensuring Quality and Value

Ed. Reva Basch. Brookfield, Vt.: Gower. 1995. 264p. $79.95 (ISBN 0-566-07567-9).

The fourteen articles in this collection were written by a diverse set of professionals representing various sectors of the information industry as well as some of its more sophisticated customers, end users, and intermediaries. The contributors include information consultants, online and CD-ROM publishers and vendors, library school faculty, and information specialists from corporations, government agencies, library associations, and the legal profession.

Unlikely as it may seem, there is consensus among the contributors on fundamental issues: they all agree on the basic criteria for judging the quality of database and information products and services; they all believe that information products should be subject to the same types of performance measures as other kinds of goods and services (quality is defined by the customer); and they all feel that an ongoing commitment to information quality management is essential.

The ten quality criteria championed by the Southern California Online Users Group continue to have positive ripple effects in the information accuracy/error rate; accessibility/ease of use; integration; output; documentation; customer support and training; and value-to-cost ratio. The reality for information consumers, however, is that databases vary tremendously with regard to quality as measured by these criteria. As several contributors advise, *caveat searcher*—let the searcher beware.

The editor has done an exemplary job of selecting relevant, readable, complementary, and useful articles. There are historical perspectives that address the effects of legacy print and mainframe systems and the persistent lack of real standardization ("'the chicken-and-egg' question: do you standardize the query or the data? If the query, how do you get everybody to do it the same way?"). There are descriptions of TQM (Total Quality Management) in an aerospace database production environment, in a corporate information services setting, and in scattered European projects. Barbara Quint's answer to the question of quality in mediated online services is communicated in the title of her article, "Better Searching through Better Searchers." Carol Tenopir's article focuses on "absolute quality factors," what she considers most important in defining and making a quality database. Peter Jacso provides model test searches to explore CD-ROM databases for incompleteness, inaccuracies, and inconsistency. There is a chilling review of the legal liabilities of independent information professionals (whose best defenses include competence, modest claims, and disclaimers). The collection concludes with three papers that report on the information quality movement in Finland, the United Kingdom, and Europe in general.

Although expensive, this title is recommended for library schools and large research collections. It is essential reading for database developers, producers, and vendors. As many of the contributors remark, the issues addressed here are becoming even more critical as commercial and noncommercial electronic information delivery via the Internet proliferates.—*Eddy Hogan, California State University, Sacramento*

The Elements of Information Gathering: A Guide for Technical Communicators, Scientists, and Engineers

By Donald E. Zimmerman and Michel Lynn Muraski. Phoenix, Ariz.: Oryx. 1995. 242p. paper, $19.95 (ISBN 0-89774-800-X).

A comprehensive overview of effective ways to gather the information needed for various kinds of research, this volume is more than adequate for people unfamiliar with libraries and other data resources. A strength of the book is its emphasis on analysis and evaluation of the information being gathered, from judging the quality of an expert to evaluating the quality of a particular journal. This accent on developing critical thinking skills is commendable. One weakness (at least from a librarian's point of view) is a lack of emphasis on librarians as an important resource in libraries. While not ignored altogether, the librarian's role in saving time and frustration is underplayed; as its subtitle implies, the book is aimed at researchers and students working in institutions that almost certainly have large libraries staffed by professionals just waiting to be asked!

The book is well organized into three parts. The first, consisting of chapters 1 through 6, is titled "Gathering and Exploring Information Sources." This provides a thorough overview of the research process—how to find information using scientific methodology, problem-solving methodology, and traditional library sources. There are separate chapters for printed and electronic resources. Appendix A provides a brief introduction to searching library catalogs and other databases through the Internet with just enough information to get started. The last chapter in this section discusses a variety of ways to remain organized throughout the research process. How to take notes and what content to include are discussed, as well as ethical issues such as how and why to avoid plagiarism. Appendix B discusses four computer programs to help manage citations: Papyrus, Reference Manager, WP Citation, and Pro-Cite. However, an important criterion—price—for deciding which to purchase is omitted. The prices of these programs vary significantly, and while they understandably might be left out due to changeability and other reasons, even ballpark figures would have been useful.

The second part of the book is titled "Planning and Conducting Interviews and Surveys." Chapters 7 and 8 provide a good background in techniques for interviewing individuals and groups and ways to assess information gathered this way. Chapters 9 through 13 review survey methodology, from planning to administering to evaluating data.

Part three, titled "Exploring Advanced Research and Evaluation Methodologies" (chapters 14 through 16), furnishes brief introductions to usability testing for new products, ethnography and case studies, and other advanced research designs.

Useful diagrams, tables, and an index are included. The book is recommended for special and academic libraries.—*Donna E. Cromer, University of New Mexico*

Emerging Communities: Integrating Networked Information into Library Services

Ed. Ann P. Bishop. Papers presented at the 1993 Clinic on Library Applications of Data Processing, April 4–6, 1993. Champaign, Ill.: Publications Office, Graduate School of Library and Information Science, University of Illinois at Urbana-Campaign, 1994. 304p. $30 (ISBN 0-87845-094-7).

The contents of the papers in this compilation range from broad conceptual ideas to the specifics of designing library gateways or integrating network services into libraries. Many of the ideas presented in this volume illustrate the authors' hopes for libraries' roles in the future. In the two years since the conference, many of the predicted developments have occurred, while some have changed dramatically. Although the details of some of the articles may no longer be relevant, the community relations and planning processes described in several of the articles continue to be valid. Examples include the use of networked information in a wide range of libraries, including K–12, corporate, and academic libraries.

The articles on integrating computer networks with academic libraries and the roles of librarians in the emerging era of networked information are among the collection's best. As a collection of edited transcripts of conference presentations, the book may be less useful to individuals working out the technical details of bringing networked information to libraries than to library administrators promoting technology to their communities. The articles on the design and evaluation of electronic gateways and online library catalogs might be useful to individuals designing or customizing a computer interface.

Much has changed in networked information since this conference was held. For example, the introduction notes that NCSA (National Center for Supercomputing Applications) Mosaic was demonstrated at the conference by the designers from NCSA at the University of Illinois. The introduction of the World Wide Web (WWW) has changed readers' perspectives on two-year-old Internet discussions. Even though such details may be out of date, time has only added credibility to the validity of the presenters' visions of where libraries should be headed. This nicely edited collection is readable for librarians at all levels of technical ability and will appeal to administrators as well as practitioners.—*Connie V. Dowell and Steve Bischof, Connecticut College Library*

From A to Z39.50: A Networking Primer

By James J. Michael and Mark Hinnebusch. Westport, Conn.: Mecklermedia. 1995. 166p. paper, $35 (ISBN 0-88736-766-6).

"Z.39.50, [the] information retrieval protocol . . . , will dramatically change how the information industry does business," say the authors—both experienced and very knowledgeable automated library system developers. Michael is a vice-president of Data Research Associates, with primary responsibility for standards conformity—especially Z39.50. Hinnebusch is a network administrator with the Florida Center for Library Automation and serves as chair of the Z39.50 Implementation Group. The discussion is not limited to Z39.50, however; over 80 percent of it deals with OSI, TCP/IP, and local area networks. While Z39.50 was originally designed as an OSI (open system interconnection) reference model protocol, it has been even more widely used with TCP/IP (the protocol used on the Internet) and local area networks.

The "network revolution," which Michael describes in the opening chapter, requires tools to access the great variety of databases available. OPACs represent only a small fraction of these resources. All of the networking required today to access information—some 95 percent of which is not in printed form, according to Michael—makes networking standards imperative. Of these, the most important is Z39.50, a protocol that offers user "interoperability," the availability of a search-and-display strategy with which the user is familiar. With Z39.50, it is not necessary to learn the search language of the target database; instead one uses the search language of one's own system. The authors stress that while Z39.50 was originally designed for retrieving bibliographic records, it now extends to abstracts, full text, and images.

This book is recommended for those with little time and a basic knowledge of library automation. Chapter 12 by Hinnebusch is particularly useful as an overview of Z39.50; however, readers should be aware that it was probably written in mid-1994—before the balloting on version 3 of the protocol was completed and work on version 4 began. The treatment on OSI, TCP/IP, and LANs, while extensive, is no better than adequate. There are scores of books and articles that are as good or better.

While the index is very good, there is no glossary and no bibliography. Many terms are not adequately defined in context; therefore, the reader should have a good data processing dictionary handy. Since most of the useful literature on Z39.50 is found in the professional journals, the lack of a bibliography means that a literature search will be required for anyone wishing more details on the protocol.

The layout, typography, and quality of editing are good, although one still wonders whether any book other than a fine edition warrants a price of $.21 per page.—*Richard W. Boss, Information Systems Consultants Inc.*

Library Information Technology and Networks

By Audrey N. Grosch. New York: Marcel Dekker, 1994. 384p. $150 (ISBN 0-8247-3971-7).

Originally intended as a revision of Stephen R. Salmon's *Library Automation Systems*, Grosch's book expands the scope to include the development of networks as an important factor in library systems. The introductory chapters offer a thoroughly researched overview of the early development of library systems. The role of bibliographic networks and cooperative initiatives are also covered.

Current developments in networking, both local and global, are especially well synthesized despite the fact that at the time of publication, Grosch noted that there were "still relatively few commercial business sites" on the Internet (p. 157). Oh, the difference a few months make in Internet-land! Still, this chapter provides an excellent description of not only the development of the Internet, but also how it actually functions.

The NISO regulating standards, so important for effective systems planning, are also explained in an uncomplicated, narrative style. Hints of future developments such as the Computer Interchange of Museum Information (CIMI) are brief yet tantalizing.

Other chapters cover the library systems marketplace and the factors involved in systems evaluation, procurement, and enhancement. Besides outlining the fundamental questions of platforms and anticipated use, these chapters include an overview of the frills system vendors provide or fail to provide. Noticeably absent, for example, is Notis's ill-conceived Horizon initiative. New concurrent "multiuser" systems appear to favor UNIX platforms and are still very much in the beta stage, but Grosch offers a good overview of these possibilities as well. More interesting is Grosch's grasp of the philosophical debate over the issue of proprietorship vs. non-proprietorship of database resources that individual will need to resolve with the growth of their own systems.

Finally, there are more frills to consider: PC-based client software purchased separately to work in tandem with the library system. Once again, Grosch provides an excellent synthesis of leading packages.

Grosch's book is remarkable for its thorough research, international scope, explicit detail, and value to the professional community with a price to match. It should be considered a standard text for systems planners and students alike.—*Mary Hemmings, University of Calgary Law Library*

Multimedia Technologies for Training: An Introduction

By Ann E. Barron and Gary W. Orwig. Englewood, Colo.: Libraries Unlimited. 1994. 225p. $29 (ISBN 1-56308-262-4).

This practical and inspiring guide should serve a wide audience. Both authors have solid reputations in the training field and they have put together a book that is easy enough for the novice to understand yet provides sufficient detail for the seasoned trainer to use as a refresher while picking up new ideas and techniques. Also, the book need not be read from beginning to end; each chapter can stand alone and therefore serves well the reader who may only want to learn about video technology or local area networks.

The text is upbeat and positive, continually urging the reader to ask questions, read articles and books about the technology, join organizations and network, and seek out others who share his or her interests.

The book opens with a short chapter covering the long-term benefits of using multimedia technology to train persons in business, industry, and higher education and emphasizing careful planning and development. Subsequent chapters cover different technologies: CD-ROM, video, digital audio, telecommunications, teleconferencing and distance education, development software for training applications, and simulations and virtual reality.

Each chapter begins with an introduction, reviews positive and negative elements, and details specific applications with numerous diagrams and easy-to-understand graphics. The chapters are packed with practical information and provide manufacturer/vendors' names, addresses, and phone numbers, and "800" numbers when known. The chapter conclusion is followed by references, a list of resources, and recommended readings on that topic. The book ends with an extensive glossary of terms with short, clear definitions and an index for quick location of topics.

This introductory work lays excellent groundwork for delving further into specific multimedia technology topics of interest. Although the field of technology changes very rapidly and some of the names and addresses will be out-of-date before long, this work will serve as a useful introduction for some time. The book is highly recommended for those who need to know just a little about general technologies and those who want to know a lot about multimedia technology with an emphasis on training applications.—*Claudette S. Hagle, University of Dallas*

Navigating the Networks: Proceedings of the ASIS Mid-Year Meeting, Portland, Oregon, May 21–25, 1994

By Deborah Lines Andersen, Thomas J. Galvin, and Mark D. Giguere. Medford, N.J.: Learned Information. 1994. 257p. paper, $29.95 (ISBN 0-988734-85-7).

This publication is based on the plenary and contributed papers, as well as papers and summaries of panels, presented at the midyear meeting of ASIS in 1994. In the preface, Pat Molholt indicates that although there are still technical challenges to be met with electronic information, this conference focuses on "the politics, the sociology, the legal issues that, along with the interfaces, gateways, protocols, and wires, make networking the change agent it has become" (p. 2).

The paper by Ann Bishop describes and evaluates the Blacksburg Electronic Village (BEV) pilot study in which Virginia Tech, the C & P Telephone Company of Virginia, and the town of Blacksburg joined together to provide electronic information and communication services to the community, including businesses, libraries, schools, local government, and individuals. Ray R. Larson's plenary session paper reports on the design and development of a network-based electronic library at the University of California at Berkeley for Computer Science technical reports.

Two of the papers discuss rural public libraries and network access. Charles McClure's paper is a summary of an assessment of the impact of providing Internet access to rural public libraries in upstate New York in a successful project named Project GAIN, (Global Access Information Network), with suggestions on how other libraries can do the same. Jackie Shane's paper describes how rural communities are linked to networks through academic institutions, commercial services such as Compuserve, and community freenets.

Philip Doty's paper, "Planning for and Evaluating an Internet Connection," explores some of the social effects of networking technologies that are beyond purely technical and economic problems, and suggests some strategies for overcoming them. Two UC Berkeley graduate student papers that also address the social impact of networked communications are included: one discusses virtual communities; the other, the impact of information technology on Latin America.

Two papers describe interesting ongoing research projects and the methodologies used but report no results since the studies were not complete at the time. Carolyn Frost and Joseph Jones plan to test gopher searching by designing classification schemes using Dewey, Library of Congress, and an existing gopher. Since conventional methods of organizing information, such as library catalogs, do not exist for the Internet, gopher is an attempt to provide organization and structure, although there are few rules or standards for gopher construction. H. Julene Butler's paper describes the ongoing project that will study the attitudes of researchers from the science and social science disciplines toward electronic publication and determine whether the electronic journal is a viable channel for formal scholarly communication.

Thomas H. Martin presents a brief but thoughtful discussion of how free public services and commercial services that charge fees can coexist peacefully on the Internet and not jeopardize the value of the Internet to scholars. Judith Weedman looks at how humanists engage in informal scholarly communication and explores the use of the Humanist listserv.

Three of the contributed papers by Michael Buckland, Gregory Leazer, and Richard Smiraglia present fairly technical discussions of retrieval systems in networked environments. They address ways of making online catalogs more powerful to the user, not just automated traditional card catalogs.

This publication includes the conference program at the beginning and an author/presenter index and subject index at the end. One shortcoming is that it does not contain all of the speaker presentations and panel discussions that occurred at the conference, although some of these might have been interesting. As can be seen from the nearly universally informative papers included, they cover a variety of topics, some fairly technical, on networks. The book is recommended only for libraries interested in a book on the broad scope of networking.—*Marilyn L. Hankel, University of New Orleans Library*

Notes for Music Catalogers: Examples Illustrating AACR2 in the Online Bibliographic Record

By Ralph Hartsock. Lake Crystal, Minn.: Soldier Creek, 1994. 355p. paper, $40 (ISBN 0-93699-6633).

Music catalogers need this book on their desks. Like the two other publications in the Soldier Creek Music Series (*Music Subject Headings* by Perry Bratcher and Jennifer Smith and *Music Coding and Tagging* by Jay Weitz) it presents a wealth of information about the creation of notes in MARC records in a handy manual.

The organization of the book is clear and easy to understand. It is also thorough. Each chapter presents the AACR2 rule governing the note along with Library of Congress Rule Interpretations and Music Cataloging Decisions affecting its application. The relevant note is then illustrated by numerous bibliographic record examples gleaned from the OCLC database, all of which are Library of Congress records. These examples highlight the relevant note in boldfaced type. Because AACR2 leaves the creation of notes largely up to the cataloger's judgement, notes are often created by inexperienced persons who tend to be too informative, on the one hand, or unclear, on the other. This book offers examples of notes formulated by decisions based not only on the rules but on sound judgement.

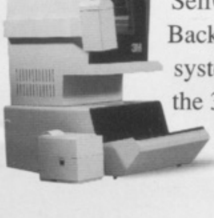
Of course, the book should never be used as a final source, and local practice will dictate how formalized some notes need to be. In case of doubt, only AACR2, the Library of Congress *Rule Interpretations*, and the *Music Cataloging Decisions* can help resolve problems. But many problems about how to avoid verbosity to the point of tediousness and terseness at the expense of clarity can be avoided by consulting this book.

Like any book, this work is not without flaws. It would have benefited from better proofreading, particularly when the reader seems to be invited to beat romanization?? of the title proper into parentheses under uniform titles (p. 70 LCRI: 1.7B4. Variation in title, subparagraph number 2). These mechanical details are few, however, and increase the value of the book for its fun.—*Richard D. Claypool, The New York Public Library*

It's called the 3M Materials Flow Management System. And it can boost your library's productivity *dramatically.*

It starts with processing: bar coding your materials in standardized locations and protecting them with 3M Tattle-Tape security strips.

Next comes checkout. Nothing has streamlined this part of materials flow more than the 3M SelfCheck System. Backed by 3M detection systems at your exits, the 3M SelfCheck System lets patrons check out their own materials *without compromising security.*



Then there's check-in, which used to take many labor-intensive steps. The new 3M Staff Workstation combines these into one, quick, ergonomic procedure.

3M Materials Flow Management: There's no more efficient route from Point A to Point A.

For more information, call us at 1-800-888-1889 ext. 63 or Email jmnelson@mmm.com. In Canada, call (519) 451-2500.

3M *Innovation*