# TABLE OF CONTENTS

## Articles

**GIS: Its Impact on Library Services** ................................................................. page 5  
By David Cobb

**The Harvard Geospatial "Liboratory"** ............................................................. page 9  
By Thomas M. Parris

**Would You Mind Showing Me a Map of Amazon.Com? Cartographic and Geographic Reference in a Digital Age: A Canadian View** .......................................................... page 19  
By James Boxall

**GIS in Libraries at the Millennium: A European Perspective** .......................... page 27  
By Nick Millea

**Toward a Catalog for the Millennium: Digital Geospatial Metadata and Data in the Alexandria Digital Library** .......................................................... page 29  
By Mary Larsgaard

**Digital Historic Aerial Photographs: A Development Description** ................. page 37  
By Jenny Marie Johnson

## Et Cetera

**Book Review** ........................................................................................................ page 48

**Calendar** ................................................................................................................ page 2

**Editorial** .................................................................................................................. page 3

**Index to Advertisers** ............................................................................................. page 2
CALENDAR

ALA Midwinter Meeting:
San Antonio January 14-19, 2000

International Miami Map Fair:
Miami February 5-6, 2000

WAML Spring Meeting:
Edmonton May 31-June 4, 2000

SLA Annual Conference:
Philadelphia June 10-15, 2000

ESRI User Conference:
San Diego June 26-30, 2000

ALA Annual Meeting:
Chicago July 6-13, 2000

IFLA General Conference:
Jerusalem August 12-18, 2000

Meeting and exhibit announcements should be sent to the Editor.

ADVERTISING STATEMENT
Meridian accepts advertising of products or services as it improves communication between vendor and buyer. Meridian will adhere to all ethical and commonly accepted advertising practices and reserves the right to reject any advertisement deemed not relevant or consistent with the goals of the Map and Geography Round Table. Inquiries should be addressed to Donna Koepp, Government Documents Library, 6001 Malott, University of Kansas, Lawrence, KS 66045-2800.

SUBSCRIPTIONS
Meridian is published twice yearly. To subscribe, or to change an address, please write to Christine E. Kollen, Map Collection, University of Arizona Library, Tucson, AZ 85721. Subscription rates are $20.00 for individuals; $25.00 for institutions. Add $5.00 for foreign subscriptions. Individuals must prepay, institutions may be billed. All foreign subscriptions must be paid in U.S. dollars. Make checks payable to ALA/MAGERT.

PURCHASING BACK COPIES
We welcome your orders for back issues of Meridian. They may be obtained for $10.00 each ($12.00 foreign). Make checks payable to ALA/MAGERT. Send requests to Jim Coombs, Maps Library, Southwest Missouri State University, 901 S. National, #175, Springfield, MO 65804.

INDEX TO ADVERTISERS
Imago Mundi.......................... page 18
MAGERT............................... page 50
Washington Map Society ...... page 4
Editorial

We have all seen people in airports, airplanes, trains, and wherever with their laptops opened, and they are clicking away at some important project or maybe just playing solitaire. Similarly, when we begin to schedule meetings with colleagues, it is not unusual for several to pull out their palm pilots to check their daily commitments. And the Internet continues to add ca. 1 million new addresses daily around the world.

We are besieged by technology, its computers, cell phones, and faxes that constantly remind us that we have transfixed our lives to a more instant world – a world that feeds us far more information than we, as individuals, will ever be able to absorb. The question is, must we? I believe that libraries, even map libraries, will remain as future information sources and will continue to store millions of books, millions of maps, and yes, millions of digits for us to refer to in the future. This is what libraries have done throughout the ages and will continue to do despite the skeptics’ thoughts that the Internet is the ‘new library.’

As map libraries continue to collect information, it is important that they reflect the world that they now are a part of, and that means including digital information alongside their traditional paper map collection. This digital information will combine data (such as census information) with various geospatial software programs that will allow map libraries to create parcelled datasets for specific geographical areas, produce graphs and charts, and produce really nice maps – in digital or paper form.

We are no longer at the end of the information line as in the past when the map appeared from a government agency or commercial publisher. We are becoming an integral piece of the map information community as we are collecting digital data and creating data displays and maps for people on demand. This issue of Meridian approaches the end of a rapidly changing millenium and sets the stage for where libraries may be going with geographical information. It was important that we looked at this internationally, and I am very appreciative for Nick Millea and Jim Boxall’s pieces. This issue also marks my last as Editor. It has been a fun run, and it is impossible to thank all of those who have helped bring these last few issues to you, but they know who they are. It is now time for someone else to take over the reins and take Meridian even higher. I thank many of you for your kind comments, support, and for reading what I believe have been some very interesting issues on very interesting topics.

David Cobb
Harvard University
Fall 1999
WASHINGTON MAP SOCIETY
WASHINGTON, D.C. USA

The Portolan
Journal published three times per year with original articles, meeting summaries, book reviews, cartographic bibliography, and notices of meetings and events of interest. Sent worldwide. (Past contents list at web site below.) Contact: Mr. T. Sander, Washington Map Society, PO Box 10793, Burke, VA 22009-0793

Local and Regional Meetings and Field Trips: Nine per year, normally at Geography and Map Division, Library of Congress. Also at New York, Baltimore, Richmond, Williamsburg, and more.

Scholarly Writing Award: Annual Ristow Prize for Cartographic History and Map Librarianship. Contact: Dr. John Docktor, 150 South Strathcona Drive, York, PA 17403-3833

Membership Contact: Mr. Bert Johnson, 2101 Huntington Ave., Alexandria, VA 22303-1547

Web Site: http://www.cyberia.com/pages/jdocktor/washmap.htm

John Carter Brown Library Research Fellowships

The John Carter Brown Library will award approximately twenty-five short- and long-term Research Fellowships for the year June 1, 2000 - May 31, 2001. Short-term fellowships are available for periods of two to four months and carry a stipend of $1,200 per month. These fellowships are open to foreign nationals as well as to U.S. citizens who are engaged in pre-and post-doctoral, or independent research. Graduate students must have passed their preliminary or general examinations at the time of application. Long-term fellowships, primarily funded by the National Endowment for the Humanities (NEH) and the Andrew W. Mellon Foundation, are typically for five to nine months and carry a stipend of $2,800 per month. Recipients of long-term fellowships may not be engaged in graduate work and ordinarily must be U.S. citizens or have resided in the U.S. for the three years immediately preceding the term of the fellowship.

It should be noted that the Library's holdings are concentrated on the history of the Western Hemisphere during the colonial period (ca. 1492 to ca. 1825), emphasizing the European discovery, exploration, settlement, and development of the Americas, the indigenous response to the European conquest, the African contribution to the development of the hemisphere, and all aspects of European relations with the New World, including the impact of the New World on the Old. Research proposed by fellowship applicants must be suited to the holdings of the Library. All fellows are expected to relocate to Providence and be in continuous residence at the Library for the entire term of the fellowship.

Several short-term fellowships have thematic restrictions: the Jeannette D. Black Memorial Fellowship in the history of cartography; Center for New World Comparative Studies Fellowships for research in the comparative history of the colonial Americas; the Alexander O. Victor Memorial Fellowship in early maritime history; the Ruth and Lincoln Ekstrom Fellowship in the history of women and the family in the Americas; the William Reese Company Fellowship in bibliography and the history of printing; and the Touro National Heritage Fellowship for research on some aspect of the Jewish experience in the New World before 1825. Maria Llena Cassiet Fellowships are restricted to scholars who are permanent residents of countries in Spanish America.

The application deadline for fellowships for 2000-2001 is January 15, 2000. For application forms and fuller information, write to: Director, John Carter Brown Library, Box 1894, Providence, RI 02912. Tel.: 401-863-2725. Fax: 401-863-3477. L-mail: JCBL_Fellowships@brown.edu. Web site: http://www.JCBL.org
GIS: Its Impact on Library Services

David Cobb
Map Librarian
Harvard College Library
Cambridge, MA

How much does it cost to offer GIS services? This is usually the first question that I am often asked by librarians and administrators when they visit our collection. Unfortunately, cost is only one aspect of GIS services, and the impact of GIS services cannot simply be defined by cost, as it is more complex. In reality, GIS is a reflection of our increasingly complex information society. Like so many other software solutions, GIS allows users to view information geographically and far more quickly and efficiently than we could with paper maps. In other words, GIS allows the user to view large datasets (i.e. Census information), apply them to real world geography (i.e. boundaries), and then analyze their geo-spatial relationships.

GIS services allow libraries to view an increasing diversity of digital data geographically. Libraries of all types and sizes are witnessing a growth in their data collections whether it be from general acquisitions or from government deposit. The majority of this data is georeferenced, meaning that various GIS software products can use it. By not offering GIS services, libraries are restricting the use of their datasets and not allowing their users to take full advantage of the analysis and display of spatial data.

With this in mind I wish to offer some opinions and predictions on the impact of GIS services for libraries. Only time will tell whether the predictions have truth and the opinions are based on my experience with managing GIS services in a library environment for nearly 10 years – almost an eternity for GIS technology. More importantly, I hope that some of these ideas will generate discussion and urge others to share their ideas.

Map libraries, and map librarians, have traditionally concerned themselves with a collection of paper maps. They worried when USGS cancelled a particular series; they quickly acquired Rand McNally’s latest publication; and they followed with great interest the topographic map series produced by foreign countries. Today, and even more so in the future, the emphasis will be on data, not publications. This means that we are relying less and less on the paper maps produced by the parties above and are following their digital data products. As the digital data becomes more available, it is important that we then collect comparable digital geographic boundaries to match the data. The levels of geography that we may have taken for granted on paper maps will require separate layers in the developing GIS service systems. This changing emphasis to data will have its impact on our future staff as it will become incumbent that database management skills will be required of ALL...
The future medium, if not the current, is going to be the Internet and web services. Certain state agencies are much farther ahead than others, and this unevenness will stay with us for some time.

...I believe our changing world will bring an increased interest in geography that will naturally lead to an interest in GIS among disciplines from anthropology to zoology.

Staff, especially those working in collections dominated by government data.

The future medium, if not the current, is going to be the Internet and web services. We use the Internet almost solely for much of our communication via email and are using it increasingly to share data. What this means is that we are going to see less person to person reference service as we rely on email reference and build an increasing array of web links on our departmental websites. While government publication collections witnessed the transition from paper to CDs, I believe we are now going to witness a transition from CDs to WEB-based services. The recent demise of NTIS may be only the first such casualty as this change takes place. I also believe there will be less reliance on federal data. State publications were once the forgotten publications in most collections, were published in limited numbers, and difficult to identify and acquire. More states are now offering digital data on their agency websites, and researchers are finding that there is far more data collected and disseminated by state governments that will never appear in our decennial census for example. More importantly, the data is often collected and updated annually, allowing it to have far greater time analysis impact than data collected decennially. This will also lead to the obvious repercussions of uneven quantity and quality. Certain state agencies are much farther ahead than others are, and this unevenness will stay with us for some time. In addition, some states are concerned with sharing their data, and others are not.

In the midst of this great change will be our challenge to manage the digital information flow. This means that it will be our responsibility to identify the most useful resources and those that are simply duplicative or not very useful to our clientele. We must also plan for the change from analog reference to digital reference and yet not lose sight of the value of our historical paper collections. As digital collections become more commonplace, we must consider the changes to our acquisitions policies as we will begin to provide access to information that we do not own. Similarly, it is important that we, as geospatial experts, provide the necessary links between data and geography in order that spatial analysis becomes available.

While it is always difficult, if not impossible, to predict change, we have many prognosticators appearing as the millennium changes. I too have some ideas as to where GIS services in libraries may be in the next five years, and we might have some fun looking back at some conference in 2004!

First, I believe our changing world will bring an increased interest in geography that will naturally lead to an interest in GIS among disciplines from anthropology to zoology. I also believe that the majority of social science research will be using GIS by 2004, but few will know it because the services will be transparent on search engines and library servers by that time. Actually, I believe our success can be measured by this transparency as we will have made the technology far easier to use, and the manipulation of geographic data and boundaries will be simpler than today.

Most developed countries will have moved to complete digital map production by this time. I suspect that paper map depositories will continue to exist and paper map products will continue to be sold and distributed. However, all maps that are printed by national agencies will come from computer databases. The advantages to such a system are many:

- individual data layers will be available
- ease of access and manipulation
- ease of copying
Custom mapping will be more accessible. There will be many disadvantages as well:
- further division of the information rich vs. information poor
- preservation of information concerns may go unheeded
- further restricted government distribution possible
- challenge of copyrighted information

As we transition to an increased reliance on digital information, we are already seeing an increase in networked services which centralize storage and delivery of information. We are beginning to see many library services move to 7/24 modes meaning they are available seven days a week and 24 hours a day. Additionally, individual library identities are beginning to fade as centralized services create an overall information umbrella under which libraries come together to better serve a diverse clientele. Similarly, and perhaps more revolutionary, will be the merging of academic disciplines electronically and intellectually. The development of environmental research programs, for example, are bringing together scholars from diverse fields such as chemistry, law, and architecture. Furthermore, we will see an emergence of digital geospatial libraries that will not only offer access to geographical information but will also offer a variety of online services as well.

What other changes might we expect? A larger number of atlases are going to be electronic only. The new National Atlas of the United States (http://www-atlas.usgs.gov/) would not be available today if it had to be printed in traditional paper format. The web and CD-ROMs offer an efficiency and cost savings that will allow for an increasing number of titles to appear in this format. We can also expect this technology to produce ‘What If’ scenarios. Planners, researchers, and students will be using GIS to plan the consequences of floods, earthquakes, business site locations, highway changes, etc. Our users are also going to expect our map collections to house and service digital data. I believe there would have been very few of us in the late 1980s that were planning to house the 1990 Census in the map collection, and I know there are many collections that have the 1990 Census now and are eagerly awaiting the 2000 data. Interestingly, in spite of the fact that there will be fewer and fewer printed government publications and printed publications in general, I believe there will be an increased use of historical print items. This turn of events can be explained by the increasing array of digital index information that provides far more information for our users than when they were just using the old card catalog and the Reader’s Guide. Therefore, our users are finding citations to far more historical information, including maps, and such use will increase accordingly.

The increased interest in geography will be matched by the emphasis on what I call “Geography’s Three W’s:” (1) Where We’ve Been, (2) Where We Are, and (3) Where Are We Going? Maps seem to add some romance and mystery to historical time studies and are being used increasingly to set the stage for historical studies, and we can magnify this even more with genealogical research. Where We Are involves everything from a current street map answering just that question to establishing a base line for environmental quality using tree cover vs. parking lot layers in an urban GIS system. The ‘What If’ scenarios and trip planning cover the last of the three W’s.

Custom on-demand mapping, once considered to be very specialized, will become very common as it is now in the Harvard Map...
Collection and other map collections around the United States. Fewer research libraries are waiting for, or dependent upon, the U.S. Bureau of the Census or the U.S. Geological Survey for their latest map. Today, those custom maps (i.e. the GE-50 series) are being produced in map collections on demand. Services for scanned images are also going to be expected, and that demand will soon expand to include same size color similar to our oversize black and white copying. These demands and services will also attract an increasing amount of donor support for digital services. The Library of Congress, strangely enough, has led the way in attracting such support, and its excellent web services are a direct result. By 2004 all libraries will be on the web, and a significant majority will offer web services. In spite of these positives there will be an increasing competition with libraries to provide information which is already apparent.

It usually takes several years for most people to realize what they have missed. As a result, I can see a backlash demand for government services and information. While people have been demanding less of government, believing that it has grown too much, there will be an equal demand soon that government is now not providing enough. There will also be numerous choices for geo-information, running the full range from all forms of government to commercial and public sources. For libraries, the emphasis will be less on collecting while a greater emphasis on access to information will be the norm. Such changes in our map collection services are going to attract new groups of users and include many people that have never known of the map collection. Similarly, we as a group of information professionals are going to be forming new partnerships with disseminators of digital information, state and local agencies for data, and with faculty and other professionals to advise on GIS services.

This new millennium may not quite be the calamity or joyous event that some are predicting; however, I predict that it will be a significant point for many map collections. Many have resisted the transition from their secure paper collections to an increasing diversity of electronic mapping services in the 1990s. To paraphrase a popular show: "Resistance will be futile" soon after the millennium changes. It will be a rocky road for some, but it is a road that we will all be going down. Get ready, the millennium is coming, are you ready? And try to remember this mantra and you will be alright: "Our Collection Drives Our Technology; Our Technology Does Not Drive Our Collection!" We must always remember the history of our collections and place our current policies in their proper perspective. We also must realize that times are changing and that it is not in our best interests to be left behind. This is a new millennium and have you considered how you are going to react to it?
Introduction

Once the province of highly specialized workgroups within departments of geography, geospatial analysis methods have now achieved a level of maturity. These methods have been adopted as routine tools across a broad range of disciplines throughout the natural sciences, the social sciences, and the humanities – much the way use of statistical analysis methods spread in the middle of the 20th century. Geospatial techniques are now used throughout the academic community to study topics such as congressional districting, urban planning, environmental change, public health, and the origins of “sacred places.” These analyses integrate information from sources as diverse as satellite time-series imagery, demographic and economic censuses, and historical fire insurance maps.

The Geospatial “Liboratory” is an effort underway at Harvard University designed to alleviate the most common challenges users face when they embark upon a geospatial analysis project: finding interesting data, obtaining that data in a usable form, learning to use new data analysis tools, and accessing appropriate computing platforms. Our goal is to provide a combination of library and analytical laboratory services that will enable students, faculty, and other members of the Harvard University community to perform meaningful geospatial analyses within the strict time requirements of a problem set, term paper, or real-world issue. Toward this end, the Liboratory will provide multiple levels of service for different types of users, ranging from novices with straightforward projects to experts who need state of the art electronic cartography and geospatial analysis.

This concept represents a major departure from business as usual for both libraries and research groups that make extensive use of geospatial information resources. In the past, libraries have been reluctant to acquire and manage data due to cost, technology, and service issues. Similarly, individual research projects have been disinclined to develop good geospatial data catalogs, long-term management, and services that would enable others in the university community to learn about and use relevant data. The Liboratory concept addresses both sets of issues.

The basic architecture of the Liboratory is illustrated in Figure 1. It consists of five major components:

- Our existing on-line public access catalog known as HOLLIS – the Harvard On-Line Library Information System;
- An enhanced Geospatial Information Resources Catalog;
- A geospatial data repository;
- A set of web-based networked exploratory analysis services for simple visualization and manipulation of geospatial datasets; and,
- A distributed laboratory environment for advanced analysis of geospatial data objects.

The remainder of this article is...
We wish to promote serendipitous discovery of geospatial information resources...

divided into three sections. The first describes the basic architecture of the Liboratory in more detail. The second discusses our plan to implement the system using an iterative process of development, test, and deployment. The third describes some of the management challenges involved in implementing an information system that crosses traditional management boundaries.

**Liboratory Architecture**

A major goal of the Liboratory is to improve intellectual access to geospatial information resources, regardless of format. There are two important, but somewhat contradictory, aspects of this goal. First, we wish to promote serendipitous discovery of geospatial information resources to patrons who do not necessarily approach the library with an explicit interest in geospatial information resources such as maps, satellite imagery, or GIS datasets. Second, we wish to provide patrons with more specific queries the ability to perform precise searches for particular geographic regions of interest, scales, projections, and themes. These users want to quickly hone their searches to the specific map sheet, satellite image, or GIS data file that meets their precise needs. The Liboratory will use our existing on-line public access catalog—HOLLIS—to serve the needs of the first set of users. A more specialized Geospatial Information Resources Catalog will serve the needs of the second set of users.

![Figure 1: Architectural Components of the Harvard Geospatial Liboratory](image)
HOLLIS
To promote serendipitous discovery we will add a single high-level description for each geospatial information resource using current MARC/AACR2 cataloging standards to HOLLIS. This will enable users to discover geospatial information resources using familiar author, subject, title, and keyword searches. When found, patrons will be able to use the information in these records to:

- View a more detailed description of the resource as presented by the specialized Geospatial Information Resources Catalog;
- Visualize the dataset using the web-based exploratory analysis services (if the cataloged resource has been deposited in the repository in one of several supported formats);
- Download the dataset from the repository (if the cataloged object has been deposited in the repository); and,
- Download the enhanced metadata record as an XML document (for automated retrieval).

Geospatial Information Resources Catalog
While on-line public access catalog systems such as HOLLIS provide excellent starting points for serendipitous discovery, several essential functions are missing when it comes to supporting more precise queries. In some cases, the fields are not defined by MARC/AACR2 (e.g., acquisition date). In other cases, the supporting database and interface functions are not available in the online public access catalog system (e.g., the ability to index bounding polygons, or render locator and index maps). The Geospatial Information Resources Catalog will provide the metadata, database, and interface functions to:

- Allow users to specify a geographic region of interest, and search for information resources that describe that area. The region of interest may be described by direct entry of latitude and longitude coordinates (rare), by drawing on an interactive base map, or with the aid of a gazetteer that will translate a place name to geographic coordinates. When a specific record is displayed, the catalog will automatically generate a locator map that shows the patron's region of interest and the bounding polygon of the map or dataset on top of a relevant base map.
- Describe and index the thematic decomposition of each resource. Many maps and geospatial datasets describe multiple phenomena. In order to evaluate the utility of a given map or dataset for a particular application, users need access to descriptions of each theme and its corresponding legend or classification.
- Describe and index the geographic decomposition of each resource. Many maps and geospatial datasets are issued in series or "works in part" in which each element describes a subset of the overall geographic area (e.g., a map or dataset for each county of a state). Users asking for information about a precise geographic region of interest will be directed to the individual components that best describe their particular region of interest instead of the dataset as a whole. Conversely, the catalog will generate index maps when patrons view records for the series as a whole without having specified a region of interest.
- Describe and index the geographic unit of observation. One of the critical factors that users must assess before integrating data from multiple sources is whether the geographic units of observation match. For instance, users will be able to limit their search to data that is organized "by county," "by watershed," "by census tract," or "by pixel."
- Describe and index the extent, scale, and projection. Maps and geospatial datasets are constructed...
The resulting metadata will be stored in a relational database server that supports geospatial indexing and queries.

The repository will provide a mechanism for authenticating users in the context of the specified object's licensing regime.

For many users, there is a large gap between knowing a particular geospatial information resource exists and making use of it by acquiring it as a complex digital object.

Geospatial Data Repository

Ultimately, users want access to maps and data. In some cases (e.g., paper maps), the resources described in the catalog will not be in electronic format. In these cases, users will be directed to the holding library and a call number. When the object does exist in electronic format, it will live in the Geospatial Data Repository. The repository will provide:

- Secure storage (e.g., secure “off-site” backups and routine monitoring of data integrity).
- A mechanism for uniquely identifying (or “addressing”) each object.
- A mechanism for authenticating users in the context of the specified object’s licensing regime.
- A mechanism for delivering the objects to users and applications via a variety protocols (e.g., http, ftp, nfs, nfs, and ipx).
- “Virtual carrels” for regular users of the distributed laboratory environment. These carrels will allow patrons to save work in progress and resume multi-session projects from any workstation in the distributed laboratory environment.

Networked Exploratory Analysis Services

For many users, there is a large gap between knowing a particular geospatial information resource exists and making use of it by acquiring it as a complex digital object. The Laboratory will fill this gap by offering users the ability to perform simple visual analyses of most geospatial datasets via a web-based exploratory analysis environment. For some users, this service will provide the end product they need for their research (e.g., a GIF image of a map for insertion into a term paper or journal article). Users with more advanced analysis agendas will use this service to assess the applicability of various datasets (e.g., does this dataset represent the features I need for my analysis?). Within the context of a single geospatial data object, users will be able to:

- pan, scroll and zoom through the dataset;
- select which data “layers” (or variables) to display;
- interactively select geographic regions of interest within the dataset;
subset, convert formats, and download selected portions of the dataset; and,
launch new queries based on selections.

The Harvard Map Collection’s Massachusetts Electronic Atlas represents a prototype of this service that runs on a single (albeit large) geospatial data object.

Distributed Laboratory Environment
For users with more extensive analysis needs, the Laboratory will provide access to localized analysis facilities with full-fledged analysis products such as ArcView, ArcInfo, ERDAS Imagine, and robust non-commercial tools. The path between the data repository and the analysis product will be preprogrammed, so that patrons can focus on the substance of their analysis instead of mundane technical issues of moving files from one location to another and translating data from one format to another. These facilities will be located in both library and laboratory settings. Local and remote reference staff will assist users with the data and the analysis systems. Users who have tasks that are especially complex or that require large resources will find solutions here. This combination of centralized data, query, and map servers with powerful managed clients will enable users at Harvard to create state of the art maps and geospatial analyses.

There will be three “point of presence” categories for the distributed laboratory environment:

- Supported points of presence are public workspaces equipped with a geospatial data analysis workstation and supported by on-site staff that can assist users with the mechanics of using the data and software.

Figure 2: Components of the Geospatial Information Resources Catalog
These service points significantly expand the number of workstations and will enable users that have achieved a level of self-sufficiency to work on their own.

We have chosen to adopt a highly iterative process to develop, test, deploy, and enhance the system.

The task of building the repository has two major components: constructing and maintaining the platform and populating it with data.

These workstations will have direct access to data via a shared file system. These service points will be housed in existing library and laboratory settings and supported by trained reference librarians. In addition, these supported points of presence will provide access to specialized peripherals (e.g., scanners, digitizers, plotters). Likely supported points of presence include the Harvard Map Collection, the Government Documents Reading Room, the Cabot Science Library, the Kummel Library of the Geological Sciences, and the Frances Loeb Library of the Graduate School of Design.

-Controlled points of presence are essentially the same as supported points of presence. The primary difference is that there will be no on-site staff support with special knowledge of geographic analysis. On-site support will be limited to making sure the machines work and helping users find and launch applications. These service points significantly expand the number of workstations and will enable users that have achieved a level of self-sufficiency to work on their own. In addition, these points will provide resources and software for more complex tasks than are feasible through the simple web-based interface. Likely controlled points of presence include the computer laboratories operated by the Faculty of Arts and Sciences and the Graduate School of Design.

-Unsupported points of presence are workstations owned and maintained by individuals or laboratories with which the Liboratory has no formal relationship. Some intensive users will inevitably want direct access to the Liboratory from the office or laboratory setting. For these people, the Liboratory will provide basic connection instructions and recommended software configurations. Other than addressing issues of clarity or accuracy, no additional support will be offered to these users.

Implementation

The Liboratory is clearly an ambitious project with many technical components. As a result, we have chosen to adopt a highly iterative process to develop, test, deploy, and enhance the system. Our hope is that this approach will enable the Liboratory to: provide early access to useful, albeit incomplete, functionality so we can adjust our designs based on real experience instead of hypothesis; and quickly adapt to new technologies as they become available.

Geospatial Data Repository

The task of building the repository has two major components: constructing and maintaining the platform and populating it with data. The platform for the repository will be developed and released in the following distinct versions:

-V1.0 Layout and Organization
Workspace: This is a preliminary version designed to provide space for organization, layout, and description of an initial set of approximately 12 large, high-demand geospatial datasets. Access will be limited to Liboratory. Network access to these files will be provided using NTFS – the Windows NT networked file system.

-V2.0 Read Only Access: This version allows limited, read-only public access to fully processed datasets. Access will be limited to public workstations in the Map Collection via NTFS. A minor release (V2.1) will grant read-only access to reference librarians enrolled in the reference training program. A second minor release (V2.2) will grant read-only access to public workstations installed in libraries hosting supported points of presence. A third minor release (V2.3) will support access from controlled workstations. A fourth minor release (V2.4) will support file delivery via the web and ftp. In all cases, users will store data in a temporary workspace on the client...
workstation. Users will be advised that files left in this directory may be deleted at any time and encouraged to find their own persistent file space. Automated scripts will remove files older than a specified period of time (e.g., 2 days). More frequent deletions may be performed as necessary.

-V3.0 Virtual Carrels: This version will add support for “virtual carrels.” As with V2.x, initial access to virtual carrels will only be granted to Liboratory staff and reference librarians enrolled in the reference training program. Subsequent minor releases will expand the number of service points. V3.1 will grant virtual carrel access to supported points of presence. V3.2 will grant virtual carrel access to controlled workstations. Note that virtual carrels only provide time-limited file storage. They do not save any user specific information about the login environment (e.g., default application settings).

The contents of the geospatial repository will continuously grow over time as additional data is acquired and processed. The target contents for each major repository release are summarized in Figure 3.

### Geospatial Information Resources Catalog

The Geospatial Information Resources Catalog has five major components: the design of the metadata structure, authoring the metadata itself, the metadata server, the technical services interface, and the public services interface.

### Metadata

An XML DTD will define the structure of the enhanced geospatial metadata. A companion document will describe the semantics of each metadata field and its relationship to existing MARC and FGDC fields. These documents will be released in the following major versions:

- V1.0 MARC/FGDC Compatibility: This version will define the minimal superstructure required to embody the FGDC metadata standard and the relevant components of the MARC standard. It will be based largely upon the work already performed by the University of California at Santa Barbara Alexandria Project.
- V2.0 Enhanced Intellectual, Analytic and Structural: This version

The contents of the geospatial repository will continuously grow over time as additional data is acquired and processed.

### FIGURE 3

<table>
<thead>
<tr>
<th>Major Version</th>
<th>Master Datasets</th>
<th>Analytic Datasets</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V2.0</td>
<td>12</td>
<td>300</td>
<td>10GB</td>
</tr>
<tr>
<td>V3.0</td>
<td>25</td>
<td>500</td>
<td>20GB</td>
</tr>
</tbody>
</table>

No public access will be granted to objects stored in the geospatial data repository until a minimal level of metadata exists.

A companion document will describe the semantics of each metadata field and its relationship to existing MARC and FGDC fields.
will add the necessary structures to describe the full breadth of intellectual attributes, structural attributes (e.g. file formats) and spatial and thematic decomposition of datasets.

V3.0 Administrative: This version will add the necessary structures to describe administrative information such as access rights.

In the early phases of the effort, the production of geographic metadata will necessarily iterate with the design of the metadata structure. As a result of this iterative process, it may be necessary to retrospectively upgrade metadata authored early in the effort as the structure evolves. As the metadata structure matures, production will become routine. No public access will be granted to objects stored in the geospatial data repository until a minimal level of metadata exists. Hence, the production targets for metadata match the targets for major repository releases shown above.

Metadata Server
The metadata server will hold and provide both public and technical services access to the metadata. It will be produced in two distinct versions.

V1.0 Individual Files: This version is intended to store and provide browse only access to a small number of metadata descriptions (up to approximately 20 master datasets) as stand-alone XML files and their "public display" equivalents (generated from the XML by scripts). Access will be restricted to Liboratory staff. Subsequent minor releases will enable public read-only access to the "public display" versions via HTTP and NTFS.

V2.0 Database Server: As the project begins to scale-up, technical services and public users will need the ability to search the available metadata. This version will store all Liboratory related metadata in a relational database management system with geospatial search operators (e.g., Oracle Server and ESRI Spatial Data Engine).

Technical Services Interface
The technical services interface is the mechanism through which catalogers and collection managers interact with the metadata server. It will be released in two distinct versions to match the major versions of the metadata server.

V1.0 Authoring Tool and Conversion: This version will simply enable catalogers to author metadata that conforms to the structure defined above. A revision control system (e.g., RCS) will be used to store the resulting collection of XML files.

V2.0 Database Management Utilities: This version builds upon V1.0 by adding a set of utilities that:
1) load XML documents into the metadata server, 2) export XML documents from the metadata server; 3) import and export MARC and FGDC metadata descriptions into skeletal XML equivalents; 4) maintain corresponding metadata in HOLLIS; and 5) manage the overall quality of the metadata collection.

End User Interface
The end user interface provides public access to the metadata server. It will be released in two distinct versions to match the major versions of the metadata server. Subsequent releases will add additional functionality (e.g., exploratory analysis, and visual summaries of result sets). However, the specifics of these versions have yet to be planned in detail.

V1.0 Browse Only Web Site and "Read Me" files: This version will help users navigate the metadata before the database version of the metadata server is available. A small program will convert collection of XML documents into their "public display" equivalents (formatted html files) for public use via a web server.

V2.0 Search and Retrieval: This version will let users search the
geospatial metadata collection by geography in addition to traditional author, title, subject, and keyword searches. When the results of a search are stored in the repository, users will be able to acquire the object via standard file delivery protocols.

**Distributed Laboratory Environment**

The Liboratory will deploy three classes of workstations: Staff, Supported, and Controlled as defined earlier. The deployment of these workstations will be timed to match the release of the repository versions outlined above. This task involves designing, documenting and testing the configuration of each class of workstation before it is due to be deployed.

**Training**

Successful implementation, deployment and operation of the Liboratory will involve a significant amount of staff and patron training. This training can be broken down, by audience, into four major categories:

- Staff involved in implementing the Liboratory will be sent to appropriate technology training as part of our existing professional development program.
- Technical services librarians will be trained in the metadata structure, XML authoring tools, and metadata management utilities. This training will be provided internally by Liboratory staff.
- Reference librarians will be trained in basic geospatial information concepts, the end-user interface, and related application software (e.g., ArcView).
- End-users will be trained in basic geospatial information concepts, the end-user interface, and related application software (e.g., ArcView).

Liboratory staff and reference librarians will provide this training within the context of the existing library instruction programs.

**Management Challenges**

While the technology described above represents a significant challenge, it is but one of the many challenges faced by the Liboratory. Managing a project that cuts across faculty, departmental, and library boundaries at a large, highly decentralized institution such as Harvard represents a significant challenge.

Managing a project that cuts across faculty, departmental, and library boundaries at a large, highly decentralized institution such as Harvard represents a significant challenge.

The Implementation Team is responsible for designing, developing, testing, deploying, and maintaining the technology.

Liboratory staff and reference librarians will provide this training within the context of the existing library instruction programs.
active in acquiring geospatial information resources that will set policies for Laboratory content and quality control.

· **Technical Services Task Force:** A group of technical services librarians with responsibility for geospatial information resources that will review and provide advice regarding the metadata design and technical service interfaces.

· **Public Services Task Force:** A group of public services librarians with responsibility for geospatial information services that will review and provide advice regarding the end-user interfaces.

· **Systems Administration Task Force:** A group of systems administrators that will provide advice regarding systems interfaces and coordinate the operation and maintenance of servers and clients.

**Conclusion**

The Harvard Geospatial Laboratory is an ambitious project that is still in an embryonic phase. As the effort matures, we hope to work with other libraries toward the construction of a national geospatial metadata cataloging utility along the lines of OCLC.

1. This article is based upon work and numerous discussions with Harvard colleagues Micah Altman, Bonnie Burns, David Cobb, Jim Coleman, Dale Flecker, Diane Garner, John Howard, Robin Wendler, and others. However, the information presented in this article is the sole responsibility of the author, and does not represent the official position of Harvard University. Correspondence should be addressed to tparris@fas.harvard.edu.


"Would you mind showing me a map of Amazon.com?"
Cartographic and Geographic Reference in a Digital Age:
A Canadian View
James Boxall*
Map Curator, Dalhousie University and
President of the Association of Canadian Map Libraries and Archives

Wouldn’t it be nice if someone could find, download and use a simple street map of Halifax, Nova Scotia? Wouldn’t it be grand if the person was not charged several hundred dollars for that data? Sadly, it isn’t possible. The data exists. The infrastructure and software is in place. However, roads are not in the public domain - at least the data about where those streets are has been determined to be non-public. In Canada, it is possible to place copyright restrictions on streets, contours, place names and a host of geographic features. And if crown copyright doesn’t stop you, cost-recovery certainly will.

The above is an oversimplification of a tremendously confusing and difficult situation. It is possible to georeference the roads of Halifax from a remotely sensed image or to use an existing base map to ‘manually’ digitize centerlines, thus creating the road network. But who wants to do that for every city and town in Canada when the data exists? Hence the tongue-in-cheek title to this article. It may be easier to answer that question and find Amazon.com than it would be to wander through the maze of geographic information (GI) policy and regulations in Canada. Perhaps, instead of Amazon.com, GI will become available through “canadagov.com”?

This article attempts to briefly explore several issues and trends that may impact what we do and how we do it. Everything presented is related to the growth in use of GIS and related computing and communication technologies within libraries. In the Canadian context, the term geomatics is used as an all-encompassing idea. To talk about map collections without GIS would be like talking about libraries without the Internet. Also, many of these issues have not been covered in recent special issues of journals devoted to GIS and librarianship, some of which are referenced at the end of this article. One cannot talk about reference or collections, or any other functions of Canadian map collections without getting mired down in the issues surrounding information policy in Canada. And so a greater amount of this article attempts to explain, at least from one very biased perspective, why things in Canada are so unique. There is some irony of course: GIS technology was invented in Canada, by Canadians and for Canadians.

A Political Landscape
Not only do we need to contend with the internal policies of our collections and libraries, but recent trends suggest that during the next century we will be even more preoccupied with policy issues.

*Views expressed herein are those of the author and do not necessarily reflect those of his employer or the Association of Canadian Map Libraries and Archives.
Governments, courts, and policy institutions are having difficulty seeing old concepts in the light of new technologies. Most nations seeking to create spatial data infrastructures are doing so within the context of government rationalisation and commercialisation. Canada is just entering the second phase of developing the Canadian Geospatial Data Infrastructure.

...
are most often government led and usually exclude anything but the most general level involvement of those outside government.

Digital Access in Canada

There are a number of issues that converge or overlap in terms of the print and digital world. Certainly crown copyright and cost-recovery are two such issues. Also, it should be recognised that most Canadian libraries are attempting to grapple with the problems of computing and communications infrastructures. Such concerns are not related to the technology per se but to the actual buying and upgrading of infrastructure. While funding for higher education has increased in the US, it has decreased in Canada. According to some sources, the endowment funds of Princeton are greater than the budget of the province of Nova Scotia. Therefore, any issues we attempt to deal with in Canada should be placed in a context of regional or continental disparity.

Some agencies and groups have even gone so far as to use Freedom of Information (FOI) legislation to gain access to information priced outside of the reach of most library or non-profit budgets. It is sad to note that many map librarians are spending more time and energy worrying about and following updates related to copyright, cost-recovery and other access issues than they spend actually helping patrons and users. However, one must always take care in painting such a negative picture. The tendency is to descend into a 'whine and complain' modus operandi. The map librarians, archivists and curators in Canada are diligently continuing their traditional efforts to maintain and enhance their collections. However, the atmosphere is clearly one that requires additional, almost Herculean, energy to keep pace with those who are placing barriers in the way of access.

Perhaps it is also an aspect of 'Canadian culture' (an oxymoron most people outside Canada find rather quaint), but we tend to always err on the side of apology. For example, while we vehemently oppose and criticise the policies that deny us access to the information we need and have a right to, at the same time we go to great lengths to make certain not to offend those we are being critical of. This is the great Canadian way. As I often state to counterparts in the United States, "you fought a war against the monarchy; we didn't." This has the net effect of creating a climate of slow progress and change. Some of the rationale for not getting rid of crown copyright does make sense -- it allows for a quickly recognisable imprimatur for quality control and authenticity. However, some of it makes no sense at all.

Since 1993, Industry Canada established a funding stream for "SchoolNet," an online network to place content and connectivity into schools and the broader educational sector. It is perhaps ironic that the federal agency that has the mandate to deal with providing content is an "industrial development" department.

Established in 1993, Canada's SchoolNet is designed to promote the effective use of information technology amongst Canadians by helping all Canadian public schools and libraries connect to the Internet. SchoolNet is a collaborative initiative between provincial and territorial governments, universities and colleges, education associations, the telecommunication industry and other private sector representatives. (www.schoolnet.ca/home/info/about.html)

Because SchoolNet is the "only game in town," libraries that are seeking funding for digital projects are required to conform to the policies and regulations surrounding the program. Again, the policies of cost-recovery and crown copyright still apply to SchoolNet, although cost-recovery is only an issue if one desires more than the level or detail of information.
necessary. Although SchoolNet is an exceptionally useful system for the k-12 sector (and maybe only k-10), it does not meet the access needs of higher education or the private sector. Those groups are being channelled to "commercialised" systems.

Policy (as opposed to politics?)

On March 12, 1998, the federal government set into motion a review of both the National Library and National Archives of Canada. The "English review" as it came to be known (named after its chair, Dr. John English), was to find areas of common concern and opportunities to reduce overlap and create more on-line access to materials in and about Canada. Suffice it to say that much of the input or consultation within that review made mention of the lack of digital access and lack of digitization going on in Canada relative to other nations. An example that this author gave to that review had to do with accessing an 1879 panoramic map of Halifax, Nova Scotia on-line through the Library of Congress "American Memory Project" (http://memory.loc.gov/ammem/pmhtml/panhome.html).

While the access to and inclusion of the Canadian maps by the Library of Congress is appreciated, most Canadian map librarians wonder why such innovative actions did not take place in Canada.

There are projects underway in Canada, and the National Library and National Archives are working very hard to put more content on-line. There is even a Canadian Initiative for Digital Libraries (http://www.nlc-bnc.ca/cidl), but such actions are rather slow to get started and usually face a long period of committee building and consensus reaching rather than actual project work, fund raising or collaboration. Additionally, there have been several reviews and reports recently regarding the depository system for libraries. These can be read via the DSI web site under the "reports" heading at http://dsp-psd.pwgsc.gc.ca.

On March 5, 1999, the province of British Columbia put out a call to find a private sector partner to operate LandData BC. According to the press release, this is part of the government’s efforts to “streamline business and cut red tape and confirmation of the ‘innovative supplier/expert consumer’ relationship model between private sector service providers and government.” The release goes on to suggest that the government “is committed to making its information readily accessible to the public through LandData BC,” and it is “confident that this course of action will result in the best possible model for service delivery” (http://www.env.gov.bc.ca/srmb/landdata.htm).

Under the current climate of cost-recovery, it is clear that there is a de-emphasising of depository programs in favour of cost-recovery efforts. The success or failure of this change was best outlined by Andrew Hubbertz (1999) who showed that in one case the projected sales and “recovery” was far below what was anticipated. His cited case of an ice atlas that was once a depository item which is now sold for $1000 is almost humorous. The atlas was not bought by one academic library.

GI Access

The following examples highlight how access to GI is currently being governed and viewed within the governmental sector. Again, this is critical to the Canadian situation and unfortunately will seem rather unusual to those with little experience in trying to access or use GI from Canada.

On March 5, 1999, the Province of British Columbia put out a call to find a private sector partner to operate LandData BC. According to the press release, this is part of the government’s efforts to “streamline business and cut red tape and a confirmation of the ‘innovative supplier/expert consumer’ relationship model between private sector service providers and government.” The release goes on to suggest that the government “is committed to making its information readily accessible to the public through LandData BC,” and it is “confident that this course of action will result in the best possible model for service delivery” (http://www.env.gov.bc.ca/srmb/landdata.htm).

After viewing the above service, one also finds such key terms or phrases as:

"Indexes of Air Photo operations for 1993 through 1998 are now officially available for download, at no charge, through LandData BC. For LandData BC visitors who are not yet registered customers, it will be necessary to apply for an Account (also via our home page) before placing orders through LandData BC."
Most readers will note the fact that the service provider seems proud that they are providing an index for free. And if you are not a customer you will need to apply for credit or be approved. Even though there is a move to privatise and commercialise these services, the similarities to real private sector entities are less than ample. How many Wal-Mart or Sears stores require you to apply for credit before being allowed to look at their catalogue or buy some clothes?

The situation in Nova Scotia is slightly different, and in some ways is improving. With the development of GeoNOVA (Nova Scotia Geographic Information) there has been a policy shift towards cost-recovery, while at the same time trying to support public access.

As the following policy statement related to GI dissemination attempts to explain:

"Government may donate its Geographic Information to educational and public libraries for reference use only. Government departments may provide Geographic Information to educational and public libraries at no cost, provided that the libraries treat the information as reference material, available to their clients in view only form. This will provide wider exposure and access to the information, encourage research, and provide more direct experience to both students and researchers with the use of Canadian information. The usual license and other documentation must also accompany such dissemination of information." (http://www.gov.ns.ca/home/dhma/pub/idsp.htm#83.2)

While these types of policy statements can be supportive or helpful for Canadian libraries, one still needs to recall that requests for data have to be made, data has to be understood and training developed, and a license has to be signed. There have been few if any studies done to quantify the costs associated with managing such restrictive and/or bureaucratic GI dissemination systems.

The main agency responsible for GI in Canada is Natural Resources Canada. Through Natural Resources divisions came the development of CEONet and other "components" of a Canadian Geospatial Data Infrastructure. Overseeing these developments, in part, is a CGDI Secretariat and an Inter-Agency Committee for Geomatics (IAGC).

It was in 1996 that IAGC developed and released a white paper on the “Barriers to the use of Geomatics Data.” This paper discussed many of the concerns raised here and even expressed some fairly blunt (from a government perspective) honesty about cost-recovery, copyright and other barriers to the use of GI data. The simple fact that the paper had the word “use” in the title goes far to reassuring map librarians that progress can be made towards bettering access to GI. Sadly, there have been other developments outside of or parallel to the IAGC which has moved more “open access” policy directions to the background.

As an example of recent developments in access and “commercialisation,” I offer the following direct quote from CEONet which is the government sponsored "clearinghouse" of earth observation information. This example is instructive for two reasons. First, the system it creates can be simplified using simple and widely available technology, such as a simple secure ftp site. Secondly, the scenario below would suggest some kind of environmental hazard or emergency is in process, which should mean that there should be no need for commercialised data sources and requestor systems such as CEONet.

On notification of a flooding event from Environment Canada, the VAS [value added service] links to the RSI acquisition service which is established as a 'favourite' link in their CEONet client interface. An acquisi-

Even though there is a move to privatise and commercialise these services, the similarities to real private sector entities are less than ample.
The system it creates can be simplified using simple and widely available technology.

The primary goal was to get research data into the hands of researchers and students within the higher education sector.

The rule of thumb is "don't download it all."

Reference and Access in Canada - examples

DLI The Data Liberation Initiative was established in 1994 as a cooperative project between Statistics Canada (the national census agency) and the research community (represented by a number of academic societies). The primary goal was to get research data into the hands of researchers and students within the higher education sector. It was designed as a means to decrease the total costs associated with buying the data at market prices. In practice, this has meant that larger institutions pay a yearly subscription fee which allows a download of all the data made available. Herein lies the rub. The rule of thumb is “don’t download it all.” Also, the system is designed so that information that is supposed to be made available within the depository program shouldn’t be found on the DLI ftp site. This ensures that DSP materials get the widest audience, which includes the whole public, while DLI materials get into the hands of the institutional license holders.

It is critical to remember that DLI is a licensing arrangement similar to any other commercial venture. The researchers license the data and promise not to create any commercial or competitive products from the data. The cost ranges from $3,000 to $12,000 per site license and any data or results are for educational purposes and non-commercial. The general impression one gets from listserv discussions is that (a) DLI is seen as a positive activity and (b) other federal and provincial agencies are keeping a very close eye on how DLI works and if it is renewed beyond the first five year project phase.

As a user of the DLI "system," this author's view is that it needs to be greatly improved. First of all, the data remains in very raw form and requires much time, energy and money on the part of the data receiving institution. Secondly, there is an undue burden placed on the data user to make absolutely certain that all uses and copies of data are kept under strict control. This doesn’t even take place for commercial information. It may be an unrealistic expectation for users to gain a very clear appreciation for the strictness of crown copyright. Most users may view the data as “theirs,”
and so may choose to ignore the letter of the agreement. Perhaps DLI will prove one important thing – it costs more to try to recoup part of the costs than it does to simply release the information. In fact, it may be worse than such a simplistic notion as that because a wide release policy may actually help spur economic growth and research activity, whereas a more strict dissemination and cost-recovery process will do the opposite.

**GCLI** Partly in relation to DLI, but mostly as a result of efforts to mirror the successful ARL GIS Literacy project, a Canadian version of the project was started in 1995. Canadian libraries, in partnership with the Canadian Association of Research Libraries, ARL, and ESRI Canada developed the GIS in Canadian Libraries Initiative (GCLI). The GCLI experience in Canada has been similar to that of US libraries with the ARL project. Some have moved very quickly and far in developing GIS services in libraries as a result of the ARL project, while others have moved more cautiously or (for whatever reasons) have been unsuccessful. In Canada, libraries have been striving to implement GIS and related data services. Often, as is the case at the University of Toronto (http://utcat.library.utoronto.ca:8002/01map.html#010108) and Carleton University (http://www.library.carleton.ca/magic/), the data services and project components of DLI are closely or seamlessly tied to GCLI.

Implementing GIS in Canadian libraries is still a very new experience. McGill University has been the most "planned" case. The map collection within the geography department was revamped under a long-term plan to center the renewed service point around GIS and geospatial data services (http://www.geom.mcgill.ca/keeslib/welcome.html). McGill now has the Walter Hitschfeld Geographic Information Centre. Brock University has also had success from another perspective through the development of an on-line atlas (http://www.brocku.ca/maplibrary/census/census.html) using data and project "momentum" (a little competition goes a long way).

Again, it is vital to note that these developments are within a context that, for the most part, does not allow libraries to develop the required expertise to enhance service or consider planning for the longer-term impact of GIS on themselves and their users. In the US, libraries felt pressure to come to grips with GI and GIS because of the quantity of data made available freely or inexpensively. In Canada, map libraries see only a trickle, at best, of data, and what data is available is normally restricted in use by the crown or made too expensive by government agencies. Also, the use of examples above (or the exclusion of others) is problematic in the specific case of this author. Again, the case in Canada is varied, but generally any lack of success can almost always be traced to the problems associated with the dissemination of GIS in Canada.

**Added Concerns**

Job advertisements over the last year suggest that nearly 20 positions in North America required a map and/or geography librarian, or a candidate with "extensive GIS knowledge." This does not include opportunities for map and geography librarians outside the mainstream. With a background in GIS, the job market is wide open and will remain so for at least the next five years. Mobility of new map library employees has increased. It may be a generalisation, but a career in map librarianship may no longer be the norm. Also, the first decade of the next century will see a period of increased retirements within the educational sector. All indications are that people need more education not less. Another large group of youth is making its way through the school and higher education systems. The key question for map...
and geography librarianship is whether our positions will be filled, and if filled, what will be the background of the people. There seems to be little active planning to nurture future map librarians or to seek out ways to make certain map librarianship is viewed as a specialist field.

Add into the above mixture of issues the continual changes in technology, user demands and expectations, data sources, the need for training, and shrinking budgets, and suddenly one has a situation that seems unsustainable. At the very least the situation is time consuming and stressful. It should be possible to point to a large number of positive Canadian examples of excellent work in GIS and GI within the libraries. While it can be said that those staffing the map libraries are doing their best and answering above and beyond the call of duty, it sadly cannot be said that the progress desired by all is being achieved. It would be interesting to see how we Canadian map librarians, curators and archivists would respond to a US data access system. I feel we would do better than most! Now if our governments could be convinced of that......

In addition to the URL’s given within the text of this article, the following have provided the basis for this work:


It may be a generalisation, but a career in map librarianship may no longer be the norm.

While it can be said that those staffing the map libraries are doing their best and answering above and beyond the call of duty, it sadly cannot be said that the progress desired by all is being achieved.

Again, the case in Canada is varied, but generally any lack of success can almost always be traced to the problems associated with the dissemination of GI in Canada.
GIS in Libraries at the Millennium:  
A European Perspective

Nick Millea*  
Map Librarian  
Bodleian Library

An initial perspective on the role of GIS in European map libraries was aired by Jan Smits, Map Curator at the Royal Library in The Hague at the 1998 Ligue des Bibliothèques Européennes de Recherche (LIBER) Groupe des Cartothecaires conference in Kraków when he summised that:

"... most map curators, with the exception maybe of those in the United Kingdom and Scandinavia, do not want to learn a technology which will be predominant in the future spatial library."

This article will endeavour to justify or dispute these observations by citing primarily British examples of the inroads GIS is making into the map library.

An intriguing email was posted to lis-maps, the UK-based map librarianship listserv by the University of Leicester’s John Castleford in July 1998, when he invited map librarians to partake in “a little debate.” It was suggested that university map libraries may well be in decline... while simultaneously demanding who will be supplying readers with the customised mapping they are inevitably going to be demanding.

The response to the Castleford challenge was resoundingly silent—a few isolated comments, but nothing particularly substantial, tending to reflect the overall inactivity which is seemingly prevalent in the United Kingdom. Examples of GIS being harnessed by British map libraries are rare.

At the Bodleian Library in Oxford, the Map Room has recently invested in MapInfo Professional, with staff currently undergoing training before launching GIS to readers later in 1999. This course of action has been the direct result of student demand.

A more structured GIS approach has been employed by the map library at Cambridge’s British Antarctic Survey, which has taken on a holistic approach, whereby mapmaking and geographical information are all closely linked. Under the aegis of the Mapping and Geographic Information Centre (MAGIC), part of the mapping division’s function is to provide maps to order from the Antarctic Digital Database (which is Antarctica’s equivalent to Ordnance Survey (OS) digital data in Britain). This runs parallel to surveying, mapmaking, topographic survey and the provision of more traditional printed map and aerial photography information. MAGIC operates with “geographic information,” which may be in various forms, assisting readers in their access and use of these data.

*With assistance from A. Paul R. Cooper and Sarah Dobson (British Antarctic Survey), and Barbara Morris (Data Library, Edinburgh University Library.)
A final British example is the Digimap Project which began on 1 October 1997 based in the Data Library of the University of Edinburgh through which staff and students at six British universities (Aberdeen, Edinburgh, Glasgow, Newcastle, Oxford and Reading) have been able to gain Internet access to OS digital mapping. Details are available at: http://digimap.ed.ac.uk/.

This has been made possible through a partnership arrangement with OS, and funding from the Electronic Libraries Programme (eLib) of the Joint Information System Committee (JISC) of the Higher Education Funding Councils. Digimap registered users (some 500 in the first year) have access to a substantial pool (around 3.2%) of the OS National Topographic Database. Digimap has also begun to resolve many of the problems associated with the provision of online viewing, plotting and down-loading of digital map data, including the far-reaching implications for map librarians. Plans for a national service to go online in September 1999 are well advanced.

Three examples of the inroads GIS is making into the map libraries of continental Europe can be seen in Germany, Sweden and Switzerland.

Wolfgang Crom of the Württembergische Landesbibliothek in Stuttgart has incorporated web technology, to devise a system of “hot” indexes detailing a chronological breakdown of map series holdings within the Library, an example of which can be viewed at: http://www.wlb-stuttgart.de/~www/referate/kartograph/bwtk50i.htm.

This index details holdings of the Baden-Württemberg 1:50,000 Topographische Karte. Such a system builds in an extra dimension to the map collection, providing the map user with an instant history of the institution’s holdings of a particular sheet. Indexes covering a variety of series for the southwest of Germany and parts of Switzerland have been incorporated into the Library’s portfolio of services.

Similar technology has been employed at the ETH-Bibliothek in Zürich, where Jürg Bühler has created an “electronic” map library, which can be accessed at: http://www.ethbib.ethz.ch/ks/karten_e.html.

Digital maps were incorporated into the collection from 1993, the Library well aware of its obligation to make electronic media available and to employee GIS to manipulate digital maps for the benefit of readers. In order to enhance user-friendliness, the ETH-Bibliothek has selected Adobe Photoshop, Adobe Illustrator and ArcView GIS (ESRI) as versatile support software for GIS.

The Royal Library in Stockholm has been making use of GIS, especially as since 1994 Sweden has witnessed legal deposit for electronic documents in hand-held form, i.e. not databases. An example is ‘ArcSverige,’ an administrative and statistical dataset in ArcView format. Map Librarian Göran Bäärnhielm has also downloaded the Swedish “Digital chart of the World” data in ArcView export format that has been made available from the Pennsylvania State University website.

In addition to these specific examples, some European countries are legislating in a manner which appears to necessitate the incorporation of GIS into the map library, with both Estonia and Finland now including datasets under the terms of their respective legal deposit arrangements.

It seems that GIS activity in the map libraries of Europe is lagging some way behind our North American counterparts, and it is only reasonable to reflect that while Jan Smits’ comments flatter British map librarianship, there is much learning required before GIS finally breaks through in the next millennium.
Toward a Catalog for the Millennium: Digital Geospatial Metadata and Data in the Alexandria Digital Library

Mary Lynette Larsgaard
Map & Imagery Laboratory, Davidson Library
University of California, Santa Barbara
Mary@atlas.sdc.ucsb.edu

Introduction

This article began as a review of both GIS (geographic information systems) in libraries and of the progress of the Alexandria Digital Library (ADL) during the four years (1994-1998) that it was funded as a National Science Foundation/Advanced Research Projects Agency/National Aeronautics and Space Administration Digital Library Initiative (DLI). It has ended - as those readers with high expectations, considerable patience, or both, will discover - as a brief overview of digital geospatial data and its presence, with ADL as a logical part of developments in the field, followed by a more detailed statement of ADL’s cataloging/metadata and how it has changed over the years of the grant and thereafter.

The Beginnings

In the late 1970s and the early 1980s, it was becoming increasingly obvious to map librarians that cartographic materials in digital form were slouching closer and closer, as computer memories grew larger and software capable of dealing with digital data in a more sophisticated fashion (Auto-Carto). More tellingly, the U.S. Geological Survey (USGS) was becoming increasingly involved in the automation of mapping processes, as were the topographic surveys of many other countries (USGS 1983). Given the substantial percentage of almost any map collection in the United States that is composed of USGS maps, this was fair warning indeed of what was to come. Also at about this time, ESRI (Environmental Systems Research Institute) - which had begun modestly in an historic home in Redlands, California, in 1969 - had its first user conference, in 1981, during the same year that it launched Arc/INFO, its flagship GIS software (ESRI).

Matters stayed relatively low-level, if visible at all, in most map libraries, although in the mid- and late-1980s, some map libraries did begin to offer digital geospatial data and the software and hardware to manipulate it. But computer hardware was still fairly expensive at the time; one library - the Map and Imagery Laboratory, Davidson Library, University of California at Santa Barbara - went in with a few academic departments in order to purchase an image-processing workstation that included an Eikonix scanner, for scanning hard-copy items into digital form.

The 1990s

Then the CD-ROM (computer disc - read-only memory) came into general use, and no “map” library has been the same since. The Federal government seized with alacrity upon the CD as a method of transporting, rapidly and inexpen...
sively, large amounts of data. Before long, map librarians were discovering that CDs were breeding and multiplying like coat hangers in dark closets, as the U.S. government began distributing such data as Census in digital form. The first modest rivulets of CDs that crossed our thresholds mounted to an ocean in the mid-1990s and now, with the increasing use and omnipresence of the World Wide Web, has abated somewhat. But still, most map libraries find themselves adding several new CDs every month.

Another development, not in the least coincidentally at about the same time as the flood of U.S. Census CDs hit our hard-copy shores, was that in June of 1992 the Association of Research Libraries (ARL), in partnership with members of the GIS community (most notably with ESRI), initiated the ARL GIS Literacy Project. The library director for a given library was responsible for supplying appropriate computer hardware to run GIS software - in this case, ESRI's ArcView - while ESRI supplied the software gratis to the library. While it was originally planned to be a one-year project for twenty-five research libraries, it was so successful that by early 1995 over seventy libraries were participating, and a similar program had been launched with twenty-eight Canadian research libraries (Making GIS 1995; Cline and Adler, p. 111). Each year, the number of librarians attending the annual ESRI user conference (held generally in late May or late June) has increased, as more and more map librarians find more and more users of GIS in their respective constituencies.

Also in about the mid-1990s, we began to see that the U.S. Government Printing Office recognized that manipulating digital geospatial data takes more computer horsepower than it does to manipulate statistical data in digital form, and began issuing recommendations for computer work stations for geospatial data. The most recent is a draft issued in early 1999, with the most recent non-draft edition being from mid-1998 (1998 Recommended...). And also in the mid-1990s, the Alexandria Digital Library (ADL) started up - on October 1 of 1994.

The Alexandria Digital Library

The Alexandria Digital Library (ADL) had as its expressed goal: "... [to] design, develop and evaluate a distributed, high-performance digital library of spatially-indexed information that includes collections of maps and images in digital form. The main output of the Alexandria Project will be a distributed testbed system that provides geographically dispersed users with access to a geographically dispersed set of library collections. Users will be able to access, browse and retrieve specific items from the collections of the library by means of user-friendly interfaces that integrate visually-based and text-based query languages. Librarians will be provided with facilities that enable them to extend their collections of appropriately formatted materials and to add meta-information to their electronic catalogue. The testbed is being designed to scale to a distributed library at the national-level" (Alexandria 1994 p. 2).

With a beginning date of October 1, 1994, and an ending date of September 30, 1998, that was a large amount of work to get done. As happens not only in research but also in real life, along the way the researchers found out a fair number of things that they didn't really want to
know, fortunately along with considerably more positive discoveries. Those interested in reading research findings are referred to the bibliography of research papers at http://www.alexandria.ucsb.edu/frames3.html (select “Publications” and then under “Research Papers,” “Current Bibliography,” which was 137 papers as of December of 1998). This article will focus on the side of matters in which this author, as resident map librarian and cataloger, was most involved - that is, on the building of the ADL user interface and catalog.

As a part of its written agreement with the National Science Foundation (NSF) concerning how work was to proceed, ADL promised to have a prototype system completed within six months of the start date. Two computer engineers worked many late hours in order to have that occur. The prototype was to be software, plus metadata and digital data, for about one hundred geospatial-data items. By far the largest portion of this was about sixty aerial photographs, which had first to be scanned; a few maps were scanned, or rather portions of them, since ADL’s scanner could scan a maximum size of eleven inches by seventeen inches. Also loaded were TIGER (U.S. Bureau of the Census) files, DLG (U.S. Geological Survey Digital Line Graphs), Landsat and SPOT images for the general area of Santa Barbara County and east to Los Angeles County, and a few miscellaneous other items.

While the scanning was going on, one ADL computer programmer was wrestling with building a metadata schema - using a schema-building software called ERWin - based on the U.S. Federal Geographic Data Committee’s (FGDC) standard (U.S. FGDC 1994), plus fields from USMARC for non-digital data (U.S. Library of Congress 1969- ). The FGDC standard is deeply hierarchical and therefore so was the ADL Metadata Schema. When printed out, it occupied a sheet of paper about three feet wide and seven feet long, filled to nearly overflowing with 455 fields (a drop in the bucket compared with USMARC’s about 3,000 fields) in 81 tables. At the same time that the programmer was building the schema in ERWin (http://www.platinum.com/products/appdev/erwin_ps.htm), he was also hunting down a sturdy relational-database software, preferably already on campus and license fees already paid for. Sybase (http://www.sybase.com) filled the bill.

As the programmer was building the schema and learning as much about Sybase as he could as quickly as possible, on the cataloging side an interface was being constructed in preparation for entering metadata for the prototype’s contents. Since the focus by the funding agencies was on the user interface, it was on that interface that the lion’s share of the programmers’ time was spent, with the metadata-entering interface being given extremely short shrift. Entering catalog records into the metadata interface was easily the worst cataloging experience in this author’s cataloging life. While it was lovely to look at - pleasant colors and deceptively simple-looking “fill in the blank” in nature, built with TCL/TK software - the problem was that the fields had been placed in tables that had everything to do with what the database software found convenient and almost nothing to do with how cataloging is most efficiently done. That is, there were three tables with a large number of fields in each. Note that all fields except those in the “mandatory” table are optional except when the information is readily available. Also note that the USMARC table contains many fields that no cataloger will recognize; they were placed in the USMARC table because we used field tags in 590 and 591; there were about twenty additional USMARC fields that in the main had their own

...along the way, the researchers found out a fair number of things that they didn’t really want to know, fortunately along with considerably more positive discoveries.

The prototype was to be software, plus metadata and digital data for about one hundred geospatial data items.

Entering catalog records into the metadata interface was easily the worst cataloging experience in this author's cataloging life.
...there were about twenty additional USMARC fields that in the main had their own tables:

1. Mandatory fields: abstract; purpose; w_b_coord; e_b_coord; n_b_coord; s_b_coord; access_constr; use_constr; progress; maint_a_update_freq; current_ref; md_date; md_stand_name; md_standancia; md_stand_vers.

2. Optional fields:
   - dataset_credit;
   - security_class_sys;
   - sec_class;
   - sec_hand_dex;
   - native_dataset_env_dex;
   - logical_consistency_report;
   - completeness_rep;
   - perc_cloud_cover;
   - attribute_acc_rep;
   - horiz_pos_acc_rep; v_pos_acc_rep;
   - horiz_dat_name; eleip_name;
   - semi_m_axis; denom_flat_ratio;
   - altit_dat_name; altit_dist_units;
   - altit_encod_meth; depth_dat_name;
   - depth_dist_units;
   - depth_encod_meth; md_use_constr;
   - med_time_conv; md_access_constr;
   - md_sec_class; md_sec_class_sys;
   - md_sec_hand_desc;
   - md_future_rev_date;
   - indir_spat_ref_meth;
   - direct_sp_ref_meth; local_desc;
   - local_georef_info; lat_resol;
   - long_resol; geog_coord_units;
   - v_count; row_count; col_count;
   - raster_object_type; and

3. USMARC fields:
   - date_time_latest_trans;
   - other_phys_details; accomp_mat;
   - form_cont_note; project;
   - c_point_long; c_point_lat; orbit;
   - path; row; solar_azimuth;
   - solar_eleve; nadir_long; nadir_lat;
   - proc_enh; cloud_cover_desc;
   - spectral_bands; bytes_p_pixel;
   - multi_level_desc_ident; spatial_res;
   - main_entry_pers_name;
   - main_entry_ass_dates_w_name;
   - main_entry_corp_name;
   - main_entry_corp_name_subord_unit.

Almost all the other fields were much easier to deal with - they were in very small tables, with either only one field or less than ten fields within the table. The first screen of the template listed the tables; the cataloger then called up the tables - one at a time! - and filled in the information, table after table. It was all too easy to miss a table (and therefore its fields), and proofreading was impossible; we finally gave up and just printed out the contents of each table in the ADL Catalog, and proofread those. Even so, we missed a few typos. It was at this point that this author began more fully to appreciate the highly specialized integrated-library-system (ILS) software that we who work in libraries tend to take for granted.

In the pressure of getting the metadata ready for the prototype (for an NSF site visit), the cataloger cataloged both the hardcopy aerial photograph and the scan of a photograph on exactly the same record, instead of creating one record for each. As a temporary stopgap, it was acceptable, but not as a long-term solution; each version requires its own record for optimal ease of use to the users of the data. After the metadata was entered for the prototype, the cataloger (this author) flatly refused to use the metadata interface ever again. Fortunately, there were several previously existing metadatabases upon which ADL needed to concentrate on loading into the digital library:

1. GEODEX: ca. 335,000 sheet-level catalog records for U.S. and foreign map series, mainly topographic but some thematic, a product of the American Geographical Society Collection at the Golda Meir Library, University of Wisconsin-Milwaukee (http://leardo.lib.uwm.edu) (Baruth 1988);
2. NASA/Ames: ca. 430,000 frame-level records for flights flown by NASA/Ames from the early 1970s to the mid-1990s;
3. Miscellaneous other metadata and data collections, such as the U.S. Central Intelligence Agency (CIA) page-size maps of foreign countries that had been scanned by the Map Collection of the Perry-Castaneda Library, University of Texas-Austin (http://www.lib.

Also by this time, spring of 1995, it was obvious that the next version of ADL was going to have a Web interface, written in HTML. Since that time, the ADL interface has stayed a Web interface; about a year ago, it moved to being written in Java, and a new version is currently being written.

In addition, though, an HTML version - much simplified from the current interface, which utilizes multiple windows open at any one time while a user is searching - is being written, in preparation for ADL being available through the California Digital Library (CDL) in September of 1999.

Back to the cataloging. Over the next few years, 1995 through spring of 1998, GEODEX and NASA/Ames were loaded into the ADL Catalog, as were records for the following collections: SPOT Image coverage of California; PCBIO records (the California Biodiversity Project on campus); DRG (Digital Raster Graphics) as available from the U.S. Geological Survey and the U.S. National Park Service; SNEP (Sierra Nevada Ecosystem Project; a Federal grant program); EDAC (records for remote-sensing images held by the Earth Data Analysis Center); DOQQ (Digital Orthophoto Quarter Quads) for California; U.S. CIA maps; and about 3,000 records for Space Shuttle photographs. We had two different staff persons working consecutively on manipulating the metadata during this time period; we feel very fortunate to have had two such excellent staff, one of whom is still working with the project. At the same time that ADL's metadata staffers were working on loading metadata into ADL, this author was working at current and retrospective cataloging of the Map and Imagery Lab's maps (ca. 480,000 sheets) and CD-ROM (ca. 3,000) collection into the Davidson Library's online catalog, PEGASUS, using OCLC as a source for copy cataloging.

A major problem that ADL experienced during the project was the length of time it took to build a metadata record from all of those tables. In the worst-possible case, joins would have to be made between 81 tables in order to present to the user one metadata record - which could easily take twenty minutes just for one record. During that time period, ADL engineers tried out other database-management software, specifically Informix (at that time called Illustra; being used as of this writing) and Oracle, to try to find ways to serve out data and metadata as quickly as possible. The lead engineer spent considerable time "pre-cooking" the metadata.

Finally, in early 1998, it became obvious that using the deeply hierarchical metadata schema was condemning ADL to convoluted, byzantine and time-consuming ingest of metadata into the Catalog, and to poor response time when the Catalog was searched. The lead engineer therefore constructed a much simpler schema, composed in the main of "bucket" (search category; e.g. Originator; type; Format; Free-Text; etc.) tables. But, now for the bad news, this meant that the ADL Catalog, which was based on the hierarchical metadata schema, would have to have all of its content switched to another catalog, which we decided to call ADL Catalog'99. We began with GEODEX, and the system seemed to work well enough, although it needed some tweaking -
A few more changes were made in what we were now calling the Bucket Schema, and we are now about to reload into Catalog'99, plus add several new collections.

Toward the Millenium

Let's take another look at the summary statement of what ADL had set out to do, and break it into parts:

A. "...[to] design, develop and evaluate a distributed, high-performance digital library of spatially-indexed information that includes collections of maps and images in digital form." The design of the library was carried out, but the library was never able to do all that we wished for. It seems difficult to believe in these days of such books as Brandon Plewe’s 1997 publication, "GIS Online: Information Retrieval, Mapping and the Internet," but it is still not possible to run anything other than fairly simple GIS software for persons that are on the Internet as compared with an intranet - that is, a net that serves one office in a geographically very limited area. ADL serves up geospatial data and metadata; it has no GIS software available for users. Beyond prototype work, with loading a mirror site and digital data at the San Diego Supercomputer Center (SDSC), it is not a distributed library, although it does have maps and images in digital form. The next step on this road is to have ADL be served out through the California Digital Library effective September of this year - as previously noted - and to aim toward having the SDSC serve out terabytes of digital geospatial data.

B. "The main output of the Alexandria Project will be a distributed testbed system that provides geographically dispersed users with access to a geographically dispersed set of library collections."

This did happen in a pilot test; users from anywhere on the Internet who signed up as test users could indeed access data that was stored at various places on the UCSB campus and at one point at SDSC.

C. "Users will be able to access, browse and retrieve specific items from the collections of the library by means of user-friendly interfaces that integrate visually-based and text-based query languages." Here again, this did happen. The ADL general-user interface does indeed have a graphic map browser, plus a selected number of text searches (including Type (genre), Form (format), Assigned Text (subject headings), and Free Text (searching all fields), plus a Date search.

D. "Librarians will be provided with facilities that enable them to extend their collections of appropriately formatted materials and to add meta-information to their electronic catalogue." Although in a few cases, metalogers at non-UCSB locations (e.g., the Sierra Nevada Ecosystem Project/SNEP; the U.S. National Park Service office in Denver) did put together and send metadata to ADL in ADL's database form, this proved to be of less importance than making sure the interface worked well, and getting as much metadata and data loaded as quickly as possible.

E. "The testbed is being designed to scale to a distributed library at the national-level." We will see what happens with the CDL work this fall (Alexandria 1994 p. 2).

The ADL project did accomplish more in research terms than it had originally promised. As a part of the serendipity of research, problems that had not been thought of at the time the proposal was written in late 1993 and early 1994 soon appeared and had to be addressed. For reports on this, see the various ADL major reports on the Website; go to http://www.alexandria.ucsb.edu/frames3.html and select any of the project documents. From a standard library's
By the fall of this year, the new HTML interface will have been completed, and we hope also to have on board an ADL help-desk staffer, in preparation for ADL becoming available through the California Digital Library. ADL digital data will probably be available only to University of California IP addresses, because of various legal agreements made with sellers of data. During the year 2000, we plan on adding more metadata and data sets; one may be the map records on the Davidson Library's online catalog, PEGASUS - almost none of which (the exception being the USCIA maps) is currently available through ADL.

Bibliography

ESRI. 199-. About ESRI: the history of ESRI.
Redlands, Calif.: ESRI. http://www.esri.com/company/about/history.html


- Deckelbaum, David. GIS in libraries: an overview of concepts and concerns.
- Hensel, Barbara; and Moss, Melissa; Harry, Steven. Identifying issues and concerns: the University of Connecticut's MAGIC - a case study.
- Mangano, Michael J. The making of a standard.
- Rand, Roberta Y. Assisted search for knowledge (ASK): a navigational tool set to global change data and information.


Spatial data infrastructures in the 1990s. 1998. Cartography and geographic information systems 25(3) (special issue), July.

- Buttenfield, Barbara P. Looking forward; geographic information services and libraries in the future.
- Coleman, David J.; and Nebert, Douglas. Building a North American spatial data structure.
- Lamont, Melissa; and Marley, Carol. Spatial data and the digital library.

Special issue of geographic information systems (GISs) and academic libraries. 1995. Journal of academic librarianship 21(4), July.

- Hensel, Barbara. Editorial: GIS in academic libraries, the best kept non-secret, or the genie is now out of the bottle.
- Adler, Prudence S. Introduction.
- Chevron, Joan F. Getting started: ready, set, ... get organized!
- Cobb, David A. Developing GIS relationships.
- Cox, Allan B. An overview to geographic information systems.
- Hunt, L. and Joselyn, Mark. Maximizing accessibility to spatially referenced digital data.
- Longstreth, Karl. GIS collection development, staffing, and training.
- Strasser, Theresa C. Desktop GIS in libraries, technology and costs: a view from New York State.


The University of Illinois at Urbana-Champaign's Map and Geography Library owns the largest library-held collection of aerial photographs, approximately 170,000 frames, in Illinois. The black and white photography was flown between 1936 and 1993 by private firms under contract to federal and state agencies. Before they were donated to the library, the photographs produced prior to 1988 were used by the United States Department of Agriculture Soil Conservation Service offices to support fieldwork. These are county-based sets at an approximate scale of either 1:20,000, prior to the 1970s, or 1:40,000. The sets are not complete nor does a frame-by-frame inventory exist. The post-1988 1:40,000-scale sets, 1988, 1993 and soon-to-be-received 1998, were produced as statewide sets and purchased from the Illinois Department of Transportation. Although these state-wide sets are intended to exhaustively cover the entire state, they do not because either the photographs were not taken or were not accepted from the contracting company by the Department of Transportation or because the photographs were loaned and never returned. Again, there is no comprehensive inventory in place.

The aerial photographs are filed on edge in metal four-drawer cabinets. Some of the older sets have been slipped into polyester film sleeves but the vast majority are unprotected. All of the pre-1988 photographs show signs of field and survey use; lead pencils, ball point pens, water-based markers, wax pencils, and pin holes had all been used to mark locations and boundaries and inscribe descriptions. Filing the photographs on edge has allowed many to curve with the emulsion/image side inside the curve. Nearly all of the photographs are losing image at the edges because of being carried in insufficiently sized folders and careless filing which allows the drawer above to riffle edges protruding too high. Sets from the mid-1970s were mounted on cloth which is acting as an abrasive on neighboring photographs.

The climate in which the photographs are stored is less than optimal. Temperature and humidity vary seasonally, and occasionally in the space of a single day. High levels of humidity were demonstrated unwittingly during summer 1997 when a photograph was left on top of a case in the office for a week; it rolled completely into a cylinder with a three-inch diameter without human assistance.

Regardless of completeness or physical condition, the aerial photography collection has been in heavy demand particularly by researchers and firms interested in land use or land cover changes.
Potential aerial photograph users are required to inspect set indexes in order to select appropriate frames. They complete a photograph request form listing county, year, flight line, and photo number and submit the form to Map and Geography Library staff who pull the requested photographs. The photograph storage area is in an office that is locked in the evening and during weekends; photographs can be retrieved only during standard business hours. Unless an aerial photograph user tells library staff that he is in Urbana from a distant location for only that day, users are told to expect that their photographs will be available for use the next business day. Once the photographs are pulled, they are kept underneath the unit’s circulation desk and are available for use any time that the library is open.

Aerial photographs may not be borrowed, except for making copies on a self-serve photocopier in a nearby library, so the Map and Geography Library assists patrons in preparing photoreproduction orders and hand-carries photographs to the university’s graphic services division for duplication.

Aerial photograph users must determine themselves which photographs may be useful. A service of filling aerial photograph requests received by fax, which required Map and Geography Library staff to determine appropriate photographic coverage based on maps marked and faxed by distant patrons, ceased in December 1997 because the unit was not adequately staffed to support the level of requests being received. Nor could a mechanism be put in place easily to recover costs. Distant patrons unable to travel to Urbana, or without a local proxy already arranged, are referred to Library Researchers, an organization within the Graduate School of Library and Information Science which provides for-fee library research. Interestingly, since this policy has been enacted, the number of requests received in particular from distant engineering firms has decreased substantially!

The deteriorating physical condition of the collection and difficulty in providing access to both local and distant patrons has spurred the library into developing an aerial photography pilot web site with search, display, and eventually download capabilities. It is hoped that making scanned images available through a web site will reduce the amount of handling the paper photographs receive while increasing their availability to a wider audience.

Scanning

The Illinois Historic Aerial Photography Imagebase project (URL http://images.grainger.uiuc.edu/aerial_photos/) began in early 1997. A proposal was submitted to the University Librarian by Beth Sandore, the library’s Digital Imaging Coordinator, and Jenny Marie Johnson, the Map and Geography Librarian, describing a potential pilot project to address both the preservation and access needs of the aerial photography collection. The proposal included developing a Web access mechanism, going a step beyond “simple” scanning of aerial photographs and writing files to CD-ROM to providing an interface of searchable feature terms. Scanning activities would focus on photography flown between 1935 and 1955, an estimated 65,000 photos. Shortly after the proposal was submitted to the University Librarian, Douglas Johnston, Director of the university’s Geographic Modeling Systems Laboratory, was invited to join Sandore and Johnson to form the core of the project team. Johnston’s geographic information systems, remote sensing, and interface design expertise was a strong complement to Sandore’s strength in imaging and Johnson’s in use of cartographic materials and information.

Photographs covering the greater...
Joliet area in Will County from 1936 and 1954 were selected as the initial pilot group for a number of reasons. The region includes rural and urban areas, four different methods of transportation, a human-modified water body, and unquestionable change in land use and land cover during the eighteen years between flights. The area south of Joliet was of particular interest because of the development of oil refinery activities and a federal munition production area. The project team also hoped that Joliet would be close enough to Chicago to possibly generate some funding interest there while not being so closely aligned with the Chicago metropolitan area as to alienate possible funding sources elsewhere in the state.

In April 1997, Scantech Color Systems, Inc. donated the scanning of 270 aerial photographs. The scanning was supposed to follow benchmarks set in 1994. A group composed of concerned state geologists, library employees, and Scantech Color Systems personnel had experimented with some of the earliest photographs to determine appropriate sampling levels for scanning using a drum scanner taking into account photograph scale, size, and print quality. They determined that the ideal sampling rate fell between 31 and 42 micrometers per pixel, resulting in ground resolutions of 2.8 to 3.8 feet per pixel and file sizes between 55 and 32 megabytes. They compared the scans with the original prints and concluded that scanning at 31 micrometers per pixel did not increase the level of feature details captured; instead it seemed that some features became more blurred. The group ultimately decided that having a greater ground resolution was not sufficient enough to increase the amount of storage space by nearly 60 percent (Luman et al 1995).

Six photographs were attached to the drum of a Crosfield drum scanner at one time that spun for two hours while scanning. Nine gigabytes of data, TIF and JPG files, were delivered on Jaz drives to the project team. The TIF files, averaging thirty-two megabytes in size, were 8-bit gray scale, 600 dots per inch scans. The project team calculated that the images had an approximate one-meter resolution.

The project was awarded an Illinois State Library “Educate and Automate: Digitizing Illinois’ Collections” grant in 1998 to support further development of the database. Project activities during summer 1998 occurred in three separate locations, scanning in the Grainger Engineering Library where the Digital Imaging Initiative is located, feature indexing and metadata compilation in the Map and Geography Library, and geographic registration in the Geographic Modeling Systems Laboratory. With sufficient student employees and images to work with, activities can occur concurrently in all locations.

Summer 1998 target areas were chosen, like the Joliet sets, intending to interest potential funding sources, to expand the kinds of land cover being examined and indexed, and to extend the geographic dispersion of coverage. The photographs selected were from the oldest sets for southern Cook County, central Champaign County, and the Illinois River between Peoria and Beardstown. The project could afford to work with only one year of photographs for each area selected so the Will County groups remain the demonstration sets for land use change studies.

After a target region was defined, the appropriate photographs were identified using set indexes and listed on a workform that guided a student in retrieving the photographs from the files. The workform also was the basis for initial data entry in a tracking matrix that was developed during summer 1998 to control the locations and status of 750 photographs. The matrix was updated each time a photograph was

The area south of Joliet was of particular interest because of the development of oil refinery activities and a federal munition production area.

The group ultimately decided that having a greater ground resolution was not sufficient enough to increase the amount of storage space by nearly 60 percent.

Summer 1998 target areas were chosen, like the Joliet sets, intending to interest potential funding sources, to expand the kinds of land cover being examined and indexed, and to extend the geographic dispersion of coverage.
...although the erasers removed some emulsion, they seemed to be the simplest and most benign cleaning choice.

The watermark is not intended to control use but rather to enable the university library to identify its files.

Some problems with faulty scans that looked correct during visual inspection also have been discovered through the registration process.

carried from one library to another and any time that a process was completed for a particular image. When photographs were pulled from the files, a notice was left behind indicating the range of flight lines and photo numbers removed and when they were taken. Occasionally, photographs were required for the project that had been pulled for patron use. In these cases, a notice listing each individual frame was filled out for the patron so that the missing photographs could be requested and retrieved specifically and quickly.

Pencil, both lead and wax, marks were removed as completely as possible using soft kneaded erasers before scanning occurred. The project was not located in an area appropriately vented for solvent or any kind of liquid use, and although the erasers removed some emulsion, they seemed to be the simplest and most benign cleaning choice.

Two Epson Expression 836xl scanners were purchased, and Adobe Photo Shop was selected as the imaging software package. Student employees were able to scan four photographs per hour, producing 12-bit gray scale, 720 dots per inch, forty megabyte files of 750 photographs for four Illinois counties and in the process meeting standards set at Cornell University for digital archiving.

The students also imbedded a “watermark” in each file using Digimarc software. The watermark is not intended to control use but rather to enable the university library to identify its files. The project team and summer student staff compared images marked with a “strong” watermark with the same images marked with a “medium” strength watermark. All could see artifacts left by the watermarking software when it was set to “strong” when images were enlarged but almost no artifacts were observed at the “medium” strength. “Medium” strength was decided to be sufficient because there was not intention of control, only identification.

Metadata Compilation
Metadata describing the scan files were compiled using a subset of the Federal Geographic Data Committee’s “Contents Standards for Digital Geospatial Metadata” (FGDC 1994). The subset was selected by project members using the “Content Standards for Digital Geospatial Metadata Workbook,” version 1.0 (FGDC 1995). The workbook includes both definitions of metadata elements and graphics to enable interpretation and decision making. Each section of the metadata standard is depicted as a diagram with elements shaded to indicate if they are mandatory, mandatory if applicable, or optional. Elements were evaluated in regard to their applicability, usability, and the ability of the project to encode appropriate or meaningful data. Figure 1 shows the metadata elements selected for each record level.

In some cases, mandatory if applicable became defined as mandatory if available. For instance, Section 2, “Data Quality Information,” is mandatory if applicable; three of its six sub-sections are mandatory, two are mandatory if applicable, and one is optional. But for two of the three mandatory subsections, Logical Consistency and Completeness, the project team had no formal way to evaluate the data being produced or compose appropriate descriptions beyond opening the files after scanning and visually checking for anomalous patterns. Some problems with faulty scans that looked correct during visual inspection also have been discovered through the registration process. The third mandatory sub-section, Lineage, was compiled based on paper records describing the original photography and documentation prepared by the project team describing equipment and processes.

Much of the standard metadata record is intended to describe the
**Metadata Elements by Record Level**

<table>
<thead>
<tr>
<th>PHOTO LEVEL</th>
<th>SET LEVEL</th>
<th>BASE LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>County ID</td>
<td>Metadata Date</td>
<td>Abstract and Purpose</td>
</tr>
<tr>
<td>Flight Line and Photo Number</td>
<td>Progress and Maintenance Frequency</td>
<td>Thesaurus Information</td>
</tr>
<tr>
<td>Flight Date</td>
<td>Set Bounding Coordinates</td>
<td>Access and Use Constraints</td>
</tr>
<tr>
<td>Photo Bounding Coordinates</td>
<td>Place Keyword (County Name)</td>
<td>Description of Source Material</td>
</tr>
<tr>
<td>Browse Graphic File Name and Description</td>
<td>Source Scale Denominator</td>
<td>Process Description</td>
</tr>
<tr>
<td>Latitude and Longitude Resolutions</td>
<td>Responsible Agencies</td>
<td>Spatial Reference Method</td>
</tr>
<tr>
<td>Column and Row Counts</td>
<td>Set Title</td>
<td>Raster Object Type</td>
</tr>
<tr>
<td>Network Resource Name</td>
<td>Publication Information</td>
<td>Geographic Unit Coordinates</td>
</tr>
<tr>
<td>File Size</td>
<td>Set Beginning and Ending Dates</td>
<td>Distribution Liability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Usage Fees</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Metadata Standard Information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Contact Information</td>
</tr>
</tbody>
</table>

**FIGURE 1** The metadata elements, listed in the table above, varied by record level.

Data so that potential users can determine if the set is appropriate for their projected needs. The project team also felt strongly that the metadata should include terms describing features included on each photograph. An indexing work form was completed for each photograph indicating, besides basic set, flight line and photo number, date of flight, county(ies) covered, quadrangle(s) included, transportation infrastructure (highways, railroads, airports, water ways), natural water bodies, towns, other administrative units, and other structures without proper names that seemed to have an impact on the landscape such as oil wells, quarries, and railroad round houses. During the first summer of the project, a graduate student in geography was hired to develop an indexing process and workflow. The second summer, the same students who had been hired to scan and register the photographs also indexed their contents, using content analysis forms like the one shown in Figure 2.

Aside from the time spent actually indexing, pre-indexing set up was the most time and labor intensive. The paper indexes, either photocomposite or graphic, for the aerial photograph sets were laid on a large table in the Map and Geography Library and covered with a large sheet of polyester film. Based on lists generated when the photographs were initially pulled to be

The project team also felt strongly that the metadata should include terms describing features included on each photograph.
cleared and scanned the boundaries of the target area were marked over the indexes on the film. Maps of the area were consulted to identify and mark major landmarks such as towns, rivers, and major highways. Then the photographs were arranged in flightline order; some students found it useful to clip the photographs together creating a panoramic view. The students worked through an entire flight line in photographic sequence identifying as many named or landscape-impacting features as possible. Doing this work in a sequence promoted an economy of effort because the students would be identifying the same features as they repeated sequentially.

Appropriate source materials to generate names from were inconsistent. One of the most useful tools was a county highway map produced at approximately the same time as the photographs were flown. These maps, copied from originals at the Illinois State Library, enabled the students to index highways and railroads with names appropriate to the situation shown in the photographs. This is important because some highways no longer exist, Route 66 is an ideal example, and railroad lines have changed ownership more than once in the past sixty years. Railroad identification proved to be the most problematic early in the project because names and the available information sources often conflicted. United States Geological Survey 7.5- and 15-minute topographic quadrangles were used to index water bodies, airports, towns, other administrative areas, and other features. Often the highway map was used to double-check names to be sure that they were consistent with the era of the photograph.

The project team is aware that the depth or completeness of feature indexing has differed between students and even between indexing sessions. The students were instructed to not include roadways lower in the transportation hierarchy than state highways but beyond that firm guidelines generally were not established. Examples were given along with instructions on how to complete data for each kind of feature. The students were trusted to use "cataloger's judgment" in combination with the expertise they were building in interpreting photographs and the cultural landscapes of specific regions. The manual feature indexing performed during 1997 and 1998 was checked against the United States Geological Survey's Geographic Names Information System (GNIS) for completeness and very few feature names were found in GNIS that did not appear in the manual indexing.

Early in the project when metadata elements were being selected, the project team had an opportunity to evaluate the metadata data entry tool created by the Alexandria Digital Library but decided that it was more complex than the team felt comfortable in assigning student employees to do data entry with. A volunteer developed data entry tool using Visual Basic that fed into an Access database for the Illinois project. The structure of the database and data entry tool reflected the team's determination of the purpose of each selected metadata element: did the element describe an individual frame, a set of photographs, or information that could apply to many sets of photographs. The database, and the data entry tool, has three levels, photo, set and base. In order for a specific photograph to be completely described, information from all three levels must be incorporated, but information from only set and base are needed to describe a set of photographs. The information included at each level does not duplicate that of the other levels.

The photo-level records include the data gathered on the feature indexing workform and during the second summer of the project were entered by student employees. This
information is augmented by data generated through the registration process such as coordinates of the east, west, north, and south bounds of the photograph along with number of pixels in rows and columns, and pixel size.

The set-level records are entered by professional staff and include data necessary to describe an entire set of photographs: agencies responsible, set title, scale, year of flight, and coordinate bounds for the entire set. Controlled language for the agencies from Library of Congress authority records is being used. Because project staff were initially much more familiar with MARC record format, as interpreted by OCLC, the set-level records were first compiled in the MARC format and then "translated" into the data entry form.

Professional staff also enter the base-level records, perhaps better named "umbrella records." These broadest of all deal with library policies regarding data use and contact information.

Multilevel description is not unusual for cartographic materials collections; it is almost instinctive for map librarians to describe entire sets such as the 1:1,000,000 map of the world produced by the United States Defense Mapping Agency or the United States Geological Survey's 7.5-minute topographic quadrangles for the state of Illinois as a single unit in a catalog and then supplement the set record with a graphic index describing individual sheets. Each sheet is a bibliographic individual that has individual characteristics of name and date of situation, but for complete description information, the set is also referenced.

**Registration**

Each photograph’s file was registered to geographic coordinates using ERDAS Imagine 4.0 at the Geographic Modeling Systems Laboratory. It was not possible to perform a full rectification of the image files because critical information about the camera used to take the original photographs is not available. A first-order conversion to correct scale and rotation was performed using digital raster graphics files of United States Geological Survey 7.5-minute topographic quadrangles. Experienced student employees averaged registering three image files per hour achieving a root mean squared error of less than four meters. The aim of the project’s registration is locational accuracy,

**FIGURE 2** Sample of the workform used to do feature indexing and metadata compilation for individual aerial photographs.

<table>
<thead>
<tr>
<th>Air Photo Content Analysis</th>
<th>initials/date</th>
</tr>
</thead>
<tbody>
<tr>
<td>County Will</td>
<td>5.30.54</td>
</tr>
<tr>
<td>Photo ID (Flight Line &amp; Photo Number) BXK - 2N - 7W</td>
<td></td>
</tr>
<tr>
<td>Quad Name Joliet</td>
<td></td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
</tr>
<tr>
<td>• Highways 30; 4 (Maple St); Alt 65; S2; 4A</td>
<td></td>
</tr>
<tr>
<td>• Railroads Chicago Rock Island Pacific; NY Central; Elgin Joliet Eastern; Atchison Topeka and Santa Fe</td>
<td></td>
</tr>
<tr>
<td>• Airports</td>
<td></td>
</tr>
<tr>
<td>• Water Chicago Sanitary &amp; Ship Canal; Ille Mts Canal</td>
<td></td>
</tr>
<tr>
<td>Water Bodies De Plaine River; Hickory Creek; Sprig Creek</td>
<td></td>
</tr>
<tr>
<td>Towns Ingalls Park; Ridgewood; Forest Park; Joliet; Ladd; Raynor Park</td>
<td></td>
</tr>
<tr>
<td>Other Administrative Units Pitcher Park; Rivals Park</td>
<td></td>
</tr>
<tr>
<td>Other RR yards, quarries</td>
<td></td>
</tr>
</tbody>
</table>
The interface by which the images were accessed needed to be an improvement beyond the traditional paper indexes.

It was envisioned that users would narrow their search area by drawing a bounding box on a map or by selecting a county from an alphabetical list.

A number of users were observed "driving" or "walking" roads with their fingers on the indexes to their final photographic destination.

The images by which the road phase seemed in able to accurately registering a with aerial photograph. "TIF files also the photo composite on place by aerial photograph part attention to details, centering a grid second a displacement. fall on multiple quadrangles had to resort to manually locating the photographs the photogmphs th~ use points could not images registration or depicted registering coordinates, and Geo-referenced made that ecit project does not quad­ Illinois' consistent landscape shows no more than project team needed are images, inter­ use different standard look the on top the in a experience, points was the reliable diffi­ complex and difficult to differ­ center wu e.nvisioned lholt After created or posi­ discilrded of points one-half being the patrons. because coordinates could be available on the topographic quad­ overlaying of the proj­ scanned. or their the student be done to for the templates that can their the student calculated as users select all of the points same! difficulty with this polygon photoglyph-sized a be available on the width segments, A second automation team discussed extent predetermined size group of the project, to resolu­ inspection, first each square-mile number of rows and image. overlaying of the proj­ scanned. or their the problems of geo-locating photographs on identically featured areas. Illinois' consistent landscape of flat fields and country roads intersecting in a square-mile grid makes appropriate points for registration difficult to locate or positively identify; all of the road intersections look the same!

Interface Design
Early in the project, the decision was made that the Illinois Historical Aerial Photography Imagebase project would have two emphases, conservation/preservation and access. The project team hoped that making photographs available through the Web would lead to fewer photographs being handled while allowing greater numbers of
potential users to view the images. The interface by which the images were accessed needed to be an improvement beyond the traditional paper indexes. It also needed to be much more sophisticated than a mere list of flightline and frame numbers; aerial photographs rarely have distinctive titles. The interface developed by the Alexandria Digital Library (http://alexandria.sdc.ucsb.edu/) at the University of California, Santa Barbara was examined, as was the interface for the Environmental Planning and Geographic Information Systems (http://elib.cs.berkeley.edu/) digital library project at the University of California, Berkeley. Alexandria was too complex for the narrow geographic scope of the Illinois aerial photograph project, and the Berkeley project largely replicated the process of using paper indexes through the identification of flightlines and thumbnail sketches.

Aerial photograph users were casually observed while they used set indexes to select appropriate photographs. Anecdotally, it seemed that aerial photograph users came to the indexes with descriptive verbal terms which they were forced by the graphic nature of the indexes to either completely abandon or somehow translate into a graphic representation. Typical verbal descriptions are “a couple miles northwest of Champaign,” “halfway between Chicago and Joliet on the Illinois River,” and “the southwest quarter of Section 18, Township 5 North, Range 3 East.” Inexperienced photography users were often dismayed at the lack of verbal clues on the aerial photograph indexes, particularly on the photocomposite indexes. Users had to find a way to “pattern match” between the index and either a mental map or a paper map such as a county highway map or topographic quadrangle. A number of users were observed “driving” or “walking” roads with their fingers on the indexes to their final photographic destination.

The project team decided that a Web access interface had to break the tie to graphic indexing and retrieval as much as possible. Assumptions could not be made about either the cartographic literacy of distant and anonymous users or the reference materials they had at hand to assist in using a graphic index. Instead, the terms collected by feature indexing would be the foundation for Boolean searches of the metadata to retrieve appropriate JPG files to view and, ultimately, TIF files to download (along with their connected metadata).

The team “storyboarded” a sequence of steps for the website to lead a potential air photo user through. It was envisioned that users would narrow their search area by drawing a bounding box on a map or by selecting a county from an alphabetical list. Once an appropriate area is selected, its coordinates would be submitted to an Access database of metadata describing the photographs, and feature names that fall within the coordinate bounds would display in scrollable pick lists. The user then would select the names of the features he is interested in, a Boolean search performs against the Access database, and lists of appropriate photographs are displayed.

Regardless of how the search is performed, by narrowing first or last, when a photograph is selected from the list, the user will view a JPG image of the file. The current version of the interface allows users to choose a “small” or “large” image. Mr. SID is not a viable option because of hardware incompatibilities.

Regardless of image transfer methods, users will receive both the image and the appropriate metadata.
The project team is aware that the depth or completeness of feature indexing has differed between students and even between indexing sessions.

The primary web interface (shown in Figure 3) and connections between the Access metadata database and the display were constructed using Allaire's Cold Fusion application development environment. Java applets were written to manipulate the zoomable map of Illinois and to pass bounding coordinates between the map and the Access metadata database. The map is served using Internet Map Café which has been notoriously unstable.

Mr. SID from LizardTech was investigated as a possible browse-image display option; the project team is very excited about users being able to zoom and pan within an image displayed through the website. Unfortunately, at this time, Mr. SID is not a viable option because of hardware incompatibilities; the software is written for UNIX, and the aerial photography website's server is an NT machine.

**Next Steps**

Stability of the pilot aerial photography site is of primary concern to the project team. The site had been served from a machine intended for development work, and for a variety of reasons was not always available; it was moved to a "production" machine during fall 1998 but still is not stable because of incompatibilities with Microsoft's new 4.0 version of Internet Information Server. In early 1999, the Access database was moved to a machine that supports SQL-based queries. The demands being placed on the metadata database were beyond what could be done by the Access database quickly, and response time had been slowing as the database grew and queries became more complex. At this time the project team does not know when a fuller commitment by the library or through grant funding will be available to acquire enough disk space so that the TIF images can be made available for downloads. Once TIF images are available for downloads, buttons will be placed on the interface to allow for single-image transfer. The team has fleetingly discussed the need to serve individuals and agencies requiring multiple images, including cost recovery and fee-based services, but no solution has been agreed upon. Regardless of image transfer method, users will receive both the image and the appropriate metadata.

In contrast to the registration process, feature indexing may be able to be partially automated by using data for Illinois from GNIS. Users would still perform initial geographic limiting using a bounding box or county designation, but instead of searching against names attached to photo records, the search that generates the pick lists would work with a single list of names and their coordinates derived from GNIS. When users select feature names from the pick lists, the search would again go to the list of names to generate coordinates to be searched for in the photo descriptions. Before this method can be put in place, GNIS must be tested to determine exactly how coordinates are assigned for linear features such as rivers and large areas such as cities and forest preserves. The project team suspects that photographs will not be retrieved because they do not contain appropriate coordinate locations such as the headwaters or mouths of rivers and center points of cities, the coordinates recorded in GNIS. If data from GNIS proves to be acceptable, the pick lists will be streamlined and reorganized to reflect GNIS data categories and how the project elects to collapse the many categories into six or fewer. Unfortunately, GNIS does not include information about
highways or railroads; these will continue to be indexed manually. Movement between images has not yet been implemented. The project team would like to enable users to view a browse image and then move to an adjoining image. Individual images are not topologically linked to their neighbors so a search algorithm will need to be written to compare the coordinate bounds of the initial photograph against the coordinate bounds of all photographs to determine the next photograph in the sequence. The search may be further complicated because the photographs do not merely adjoin but actually overlap.

The Illinois Historic Aerial Photography Imagebase project will continue as a library-wide and inter-college project. No single unit has the expertise or staff available to direct all of the project’s multiple facets. Each of the project participants brings a different understanding and knowledge that has broadened the project’s goals and strengthened the possibilities of meeting those goals. Together, the Map and Geography Library, the Digital Imaging Initiative, and the Geographic Modeling Systems Laboratory will find solutions for creating, indexing, storing, and serving data representing an estimated 65,000 aerial photographs of Illinois, over 3 terabytes of data.

Literature Cited


Each of the project participants brings a different understanding and knowledge that has broadened the project’s goals and strengthened the possibilities of meeting those goals.
Book Review

Map Librarianship, 3rd ed.

Mary Larsgaard


Reviewed By:
Charley Seavey,
Zia Library Services,
Santa Fe, NM

Mary Larsgaard has been actively dealing with cartographic formats for almost thirty years. To my knowledge only the editor of this journal has been at it longer. Few, if any, have applied the amount of thought to the topic that Larsgaard has. Writing a book on a subject is a marvelous way of focusing the mind, and Larsgaard's is well focused indeed.

The first edition of Map Librarianship appeared in 1978, the second in 1987. Surely three editions of a work covering twenty years of change in a rapidly changing field have a lot to tell us about where we have been, and where we might be going? We cannot claim that the twenty years from the first to third editions did not move us from plane tables and spirit levels to the computerization of everything. On the other hand plane tables were a fairly recent memory in 1978, and computerization of everything, while far from complete, will probably happen within the careers of some of our younger readers. Larsgaard's books represent, among other things, the cusp of change unequaled since Gutenberg started thinking about the printing press in the 15th century.

While the field has been changing over twenty years, the third edition of Map Librarianship remains organizationally similar to the first. Larsgaard's approach organizes the book along the life-cycle of information in a library: arrival, intellectual and physical organization, and provision of access to the objects, whatever they may be, within the library.

Starting with the first editions, the specifics of this general structure work out thusly:
Chapter 1: Selection and Acquisition-- all three editions.
Chapter 2: Classification-- all three editions.
Chapter 3: Cataloging and Computer Applications-- first two editions, becomes Simply Cataloging in the third Ed.
Chapter 4: Care and Storage-- all three editions,
Chapter 5: Public Relations and Reference-- becomes simply Reference in the second and third editions.
Chapter 6: Administration-- becomes PR and Marketing in the second and third editions.
Chapter 7: Map Librarianship: An Overview-- becomes Education in the second and third editions.

The content of the original Chapter 3 clearly refers to computerized cataloging, not any other sort of computer applications. There is also a considerable amount of space in the first edition devoted to "form cards" -- which were a kind of pre-AACRII method of keeping track of maps if one did not want to follow the then extant cataloging codes. By the third edition
the amount of space devoted to form cards has decreased dramatically, although they are apparently still in use for certain items at the University of California-Santa Barbara where Larsgaard works.

Some other indications of change in the field over time: the index term “aerial photography” in the first edition changes to “Remote sensing imagery” in the second. The term “spatial data” is not introduced until the third.

There is a heavy emphasis on classification and cataloging in all three editions. In part this is because Larsgaard is a cataloger. In part the emphasis, particularly in the first edition, is a reflection of the lack of uniform practice in map collections across the country. In those long gone days of twenty years ago, there had been no cataloging code that was both intellectually suited to the cartographic format, generally accepted by non-map librarians, or easy to use. Many map librarians simply invented their own systems rather than invest huge amounts of time in employing AACRI and its predecessors.

Classification was similarly a problem. The true elegance of the LC G Schedule did not emerge until the fourth edition of 1976. While today the answer to map classification (the G Schedule) is obvious (except to those who persist in assigning Superintendent of Documents numbers simply because they are there), as of the first edition of Map Librarianship the answer was by no means clear. The amount of space devoted to classification in the first and third editions is about the same, but the third devotes a lot more space to non-traditional methods such as geocoding than does the first. For devotees of classification theory the chapter is wonderful reading, although Larsgaard makes plain her preference for the G schedule in the third edition.

One of the valuable things about Map Librarianship is, and has been, the numerous appendices. The first edition had 15 appendices—ranging from a sample map acquisition policy to the data used to prepare the end paper map, which involved maps per 100 persons, by state, in the U.S.A. in 1969. Lists of journals, lists of libraries issuing map acquisition lists, lists of manufacturers of map storage and handling equipment.... By the third edition the book is down to six appendices: A sampling of digital data on CD-ROM and diskette (385-410); Specialized bibliographies (411-443); US publishers and distributors of globes and three dimensional maps (445-446); A sample request form for free maps (from the University of Northern South Dakota at Popple, Broken Asp, CA, no less. Deconstructing the Larsgaard sense of humor is worthy of an article all to itself.); A draft of proposed rules for cataloging digital spatial data (449-455); and the equipment manufacturer/supplier list (457).

Finally, the bibliography. Larsgaard knows the literature. If it isn’t in one of her bibliographies, it probably isn’t worth reading if investigating the topic at hand. That probably overstates the case, and one must always remember that book manuscripts are in press for a good while, so searches for more recent material are always necessary.

Larsgaard has, mercifully, elected to not give us a list of map-related URLs with which to set forth surfing the World Wide Web. URLs are notoriously unstable, and many are bound to be out of date the instant a book is submitted for publication. Until PURLS become standard, URLs will change every time somebody re-organizes their hard disk, or gets a new computer. In a book length treatise, such efforts are very close to a waste of time and book bandwidth.

Is the book recommended? Of course. This is a book that belongs on the desk of anybody even tangentially involved in the acquiring, organizing, storing, and providing access to collections of cartographic materials. Any collection ranging from a few thousand topographic maps to those well into the hundreds of thousands of sheets of traditional formats, let alone all the digital material coming along, will benefit from having this book available. In slang terms, a purchase decision is a real “no brainer.”
Circulars

No. 1 Cartographic Citations: A Style Guide

No. 2 Index to the Library of Congress "G" Schedule: A Map and Atlas Classification Aid

Occasional Paper Series

No. 1 Exploration and Mapping of the American West, Selected Essays

No. 2 A Guide to Historical Map Resources for Greater New York

No. 3 Mapping the TransMississippi West, 1540-1861: An Index to the Cartobibliography

No. 4 The Mapping of National Parks

Available from: Jim Coombs
Maps Library
Southwest Missouri State University
901 S. National, Box 175
Springfield, MO 65804-0095
(417) 836-4534