Crossroads – Bridging the Digital Divide

by DAVID COBB

INTRODUCTION

It's so good to be in Cambridge, it feels almost like home. Let me start by stating that in the past year ITMB, a successful map publisher in British Columbia, Canada published more paper map titles than at any time in their history. Similarly, the U.S. Geological Survey (USGS) recently announced that they have ceased producing paper from their aerial photography archive and will only produce digital copies. I believe that both of these facts speak to the future of maps and digital data. It means there will be paper maps well into the future and there will be an increasing array of digital data - some of it reformatted, as in the USGS case, and most of it will be born digital. When asked to speak about GIS and its role in libraries I often find myself in a conundrum - am I here to slay the dragon, or to pet the dragon. The role of technology in libraries is not one that has been embraced by everyone, and often the technology itself seems to have been force-fed upon us. The library profession is not one that has historically been a proponent of change and the very nature of GIS is change. In one sense, we have been given the choice of becoming paper museums or, at the very least, making GIS technology available in our collections.

Today, I would like to review the many ways that GIS is, or will, affect our collections. I will divide the presentation into a general overview of GIS in libraries, how it affects our acquisitions or collection development policies, its effect on cataloging, on reference services, staffing, and our web services. Then I will shift the focus a little and discuss the current situation at the Harvard Map Collection, the future role of legacy collections, and a look to the future.

GIS IN LIBRARIES

My own experience with mapping goes back to my early cartography days - drawing block diagrams in pen and ink - and I still remember the frustration of creating large piles of computer cards, turning them into the computer center late in the evening, and hoping that I would have a meaningful printout the next morning. My experience with GIS began in the late 1980's when I was working with the USGS as a member of their Earth Science Information Center network and as a member of the Cartographic User's Advisory Council. This council is an advisory group representing different library organizations and meets with US federal agencies. At that time USGS was still distributing various kinds of information via microfiche and the idea of using the Internet to communicate and distribute information was still considered experimental. USGS was experimenting with computer mapping, as were several academic centers with geography departments. I remember one of the their first projects that was so proudly unveiled at a meeting I attended: it was a river mapping project in Montana that showed flood prone areas, land use, forest types, and animal habitats. My first reaction was: these are really fascinating maps, but I'm really not interested in this 10 acres of riverbed outside of Missoula, Montana! I saw GIS as a potentially valuable tool, but only if it was going to get to my 10 acres or when I could map similar characteristics for the nation or my state.

Two developments were to have profound impacts on the role of libraries and their experimentation with GIS - at least in the United States. The first was the 1990 decennial census of the United States. Many of you may be familiar with the old GE-50 and GE-70 series of maps produced by the U.S. Census Bureau that showed various census characteristics. They would map, for example, black population, native American population, and the so-called 'nighttime' population map using white on black to show urban and rural population densities. It was announced that with this 1990 census these series would cease and very few maps were going to be printed at all. In addition, the tract maps (those showing streets and boundaries for census enumeration areas) would be printed on-demand only and users would be charged for their production. As many of you are aware, we in the US are not used to being charged for maps produced by our government and this was only the beginning. We also soon

learned that many, and actually most, of the detailed census numeric tables were going to be distributed in digital format only. This meant that not only we had to learn the Census Bureau's software to extract data but we also had to make our own maps - a fact that was very frustrating at the time but now one that we appreciate more and more. The simple result is that we have far more control over the numbers of maps that we can produce - hundreds more than if the Census Bureau had printed them - and we can control the scale, size, colors, etc.

The second development was an agreement between the Association of Research Libraries (ARL) and ESRI's President Jack Dangermond in 1992 to introduce, educate, and equip librarians with GIS skills. It was at this time that the US federal government was distributing an increasing amount of information in digital format only and with no accompanying software. Librarians were ill-equipped to handle this influx of data and the use of GIS to handle large datasets was still problematic. It was important that libraries address the situation or we would be pushed aside and university data centers would become the providers of digital data and services. This affected not only map collections, but also government document collections, which were also receiving large amounts of digital data. Since the majority of this data was geo-referenced, it became apparent that a better knowledge of GIS technology would significantly improve our ability to take advantage of the growing number of CD-ROM data being distributed. Therefore, the ARL GIS Literacy Project (as it was formally called) defined its goals as follows:

- introduce GIS to a number of libraries to address diverse user information needs with an initial focus on the US Census Bureau information;
- develop a core GIS expertise in the research library community willing to share
- their time, applications expertise, user training, and education related to GIS;
- stimulate and encourage connections between federal, state, local, and other library GIS users and information;
- promote research, education, and the public right to know through improved access to government information using GIS;
- initiate library projects to explore new applications of spatially referenced data and evaluate the introduction of these services in research libraries.

Initially, 17 libraries chose to participate in Phase I of this project but it quickly increased to over 30 libraries and then expanded to Canada. I would conclude that the results of this project were mixed. There are some participating libraries who offer little, if any, GIS services and there are others who have a wide array of services and designated staff for GIS. I suspect that those who have been successful are those that would have been successful with or without such a project. However, the project accomplished several unstated goals: (1) it raised an awareness of GIS among a variety of librarians and has resulted in GIS at least being introduced in several library education programs, and (2) it raised the need for GIS technology among library administrators and that was required for funding. I have to smile when I look back at my initial computer request - because it seems so small now but was considered the biggest and fastest then - but it was important to invest in new equipment and maintaining our computer inventory remains a challenge. I'll come back to that later.

ACQUISITIONS

Let's turn our attention to GIS's effect on our collection development practices. Essentially, if we are to take full advantage of this technology's power, we move from being just the user's of maps to being also the producer's of maps. Now that is cartographic empowerment! Being a map producer has implications for staff qualifications and expertise that will be addressed later. As we collect more and more digital data we are acquiring the potential power to produce hundreds and thousands of maps on demand. I believe that it is important that we think of GIS not as 'geographic information systems,' but rather as 'geographic information services' which seems more appropriate for our library setting. We have always been providing 'geographic information services' and this allows us to place GIS in its proper perspective as one of our services.

One of the unstated goals of the Harvard Map Collection (which means it is my goal) is to provide 1:50,000-scale coverage for as many areas of the world as possible. Realistically, this is an impossible goal, but it does drive our collection development ideals. One of our greatest concerns is to have a professor tell his class: "Go to

the Map Collection and research your hometown." As with many of you, this means that students will be asking for maps of India, Nigeria, Japan, and the U.K., as well as maps of Chicago, San Francisco, and New York. Being surrounded by an international student body, and an internationally educated faculty, places certain demands on your collection development policies. Not only are our geographic needs diverse but so are the language requirements. This was enough of a challenge when we were looking for paper maps, but it is especially challenging today as we have multiple formats being offered.

Today, we have many choices: paper, digital data, and digital images. They all represent a myriad of geographies and the formats present us with numerous conundrums. As a colleague once remarked: "Paper is democratic." Essentially, paper is a format that we are familiar with, a format that we know how to store, and a format that, given proper care, is known to last for centuries. Digital data has the potential, combined with powerful GIS software, to produce hundreds, if not thousands of maps again and again. Digital data represents the thematic palate. We are offered climatic data, toxic release data, road data, and census data. In other words, it mirrors those topics represented by paper maps and all contained on a convenient CD, or downloadable from a website. One of the major requirements of GIS is the need for detailed boundary data. Boundary layers are the first requirement for those who have data and both government and commercial sources provide it. EDINA provides significant geographical data for the UK but, unfortunately, is limited to UK institutions. However, Bartholomew provides high quality data for both the UK and Ireland and both are used in our collection. ESRI, a leading GIS software developer, also offers ESRI Data & Maps, a variety of map data in their shape file format at many scales of geography. They also offer a continuously updated GIS data website on their Geography Network. One of the more complete set of international boundary files is offered by GfK Macon, a German GIS data provider. Map libraries must redefine and reprioritize their budgets in order to acquire these datasets that will significantly enhance their GIS services. While we may be expected to acquire this information, it is unusual to learn of budgets specifically created for digital collection development.

Digital image products range from reformatting of historical maps or atlases to new publications produced digitally or both digitally, and in paper, such as the Nationalatlas Bundesrepublik Deutschland. How comfortable do we feel purchasing a digital format that we are unable to predict will operate in 10 years, much less than100 years. GIS information is more closely related to our paper counterparts than we may think. Like paper, we will not be able to save everything. These products are appearing almost every day, and are too numerous to mention, but offer our users high quality images of materials that are often only available as black and white images in the research literature. The obvious issue becomes one of access vs. ownership. Do we need a paper copy in our library if the web provides us with a satisfactory substitute and allows us to consider discarding our paper copies and decrease storage requirements? As libraries choose electronic online formats and cancel print options, it places libraries in a vulnerable position, and assumes that issues related to preservation of digital files will be resolved. And yet, in both the fields of GIS data delivery and in the history of cartography, websites are being created almost daily. I must say that I am pleasantly surprised that in this world, after 2001, that I have seen few restrictions on the availability of GIS data. Of course, there are some and we have encountered it in our access to large-scale information for New York City. On the whole, restrictions on access to GIS data are really no different than restrictions on access to paper maps. In fact, the Rand Corporation's recent book entitled Mapping the Risks (Baker, 2004), basically concludes that most data considered to be restricted is available so widely that it would be impossible to do so.

Digital data has had significant impacts on a variety of libraries collection development guidelines. Specifically, a library in the US that owns Delorme's Map Expert (for a mere $50) has all of the street maps for the entire United States. Why then should I collect paper street maps? It provides my library with geospatial coverage for areas we would most likely never collect paper maps for, such as Tolono, Illinois or Whitewater, Wisconsin. Digital web services have also impacted libraries. The mapping services offered by MapQuest and Yahoo! have, almost, placed libraries on the periphery when it comes to route planning or address location.

Similarly, I believe that we are moving away from dependence on federal or national sources and are experiencing a greater need for more local and commercial resources. Part of this transition is that our national resources are offering more of their small and medium-scale coverages via the web. While this may be more of a US phenomenon, large-scale mapping has been left to local governments while the federal government concentrated on base geographic data. As a part of our national geospatial infrastructure, local governments have been responsible for tax assessment data showing cadastral information, building footprints, large-scale aerial photography and it has seldom been shared with libraries. As these local governments move into the GIS sphere

they use the federal information as their regional base and concentrate their efforts with the larger-scale information that they know far better.

One of my colleagues at Harvard has a motto: “Free the Data!” He is, of course, correct. Why was the data collected - to be stored in a file drawer, on a computer hard drive, or purposely kept from public use? It is gratifying to witness more and more sources that are willing to share their data - especially to libraries for academic research. We have several large GIS datasets for surrounding communities, including our fair City of Cambridge. The sharing of geospatial information, spurred on by the open GIS forums, is a benefit to all: users and producers. Such a supportive policy framework is essential as map compilers themselves often are limited by their own restrictions.

CATALOGING

Even cataloging, that essential component of bibliographic control is not untouched by GIS services. Unfortunately, it must be said that most national GIS metadata standards exceed library-cataloging rules. Here again, libraries are not institutions embracing change although better access should be welcomed. One of my oft-used examples is the 1994 Atlas of Florida. This atlas happens to be a CD-ROM, not available in print, and contains in excess of 150 data layers and is given a subject heading of “Florida-Maps.” We have no idea that this title actually includes maps of different fish species, manatee breeding grounds, alligator populations as well as land use and plant species, among many other subjects. The cataloging challenge is to review a worldwide dataset, realize that it includes detailed administrative boundaries for Nigeria, and not describe it as “World-Maps?”

Like all human endeavors, it’s all about time and really good cataloging requires a lot of time. Digital collections, because of their nebulous and virtual identity, require more attention. The GIS metadata standards allow for it and library cataloging should work to bridge the gap. Our digital environment allows for sheet level cataloging and access (something that sends shivers through catalogers when I talk about it). Computer cataloging allows us to easily duplicate large portions of a record so that individual sheets in an atlas, or individual sheets in a large series of maps, could be accessible.

An important new international project will soon allow seamless access to all of the sheets from the 19th-century Austro-Hungarian 1:75,000 series, and some of the project participants are in this audience. Not only will the maps be available for study, but they will be significantly enhanced with the creation of a searchable gazetteer that will include current and historical names. This project will no doubt be a success and will hopefully lead to the inclusion of other important map series in the future.

REFERENCE

Let us now turn to reference, the quintessential service that is performed by librarians daily. Today, the library that limits its services to paper maps, is one that is not only restricting its services, but limiting the reference information it can offer for its users. Using the web will allow a collection to offer a diversity of information, but acquiring data and offering GIS expertise will obviously enhance any geography reference service. When someone walks into the map collection, we can no longer assume that a paper map will be the answer to their question - but it might be. The map library today has changed in the 21st century and more changes are to come. As our services expand, more users will walk through the door expecting to use GIS services. Our users are more sophisticated today. Many users will have searched the Internet precluding their visit and they have not found what they wanted. We also must adjust to the role of less person-to-person reference, more email, and the ever more virtual ‘snatch the data and run’ from our websites. Like books in a library, we will be offering more data (i.e. information), but will be unaware of how it is being used.

In spite of the technology, and its complexities, the library’s major advantage is that it represents a neutral site. It remains a location, technically, but one without agenda’s, without departmental affiliations (usually), and yet continues to display a priority for acquiring, cataloging, and storing. Therefore, it is important to integrate GIS as a regular option for answering reference questions and - vice versa - that paper maps be used to complement GIS.

STAFFING

Staffing is another area that is seriously affected by GIS. There are currently four open positions in the US for map librarians and all four require a knowledge of GIS. In the past, map library staff would usually have solid geography or history backgrounds. Most libraries also required a library science degree that familiarizes an individual with the basics of library organization and management. Today, these remain valuable qualities for librarians managing a variety of GIS services. It is accepted now that librarians will be web literate and aware of sources of virtual information that they can draw upon. In addition, geospatial librarians will need to have database management skills beyond their digital mapping skills. It will be critical to lobby library administrators that GIS services require a specialized technical staff. It will be more and more difficult to expect general librarians to be familiar with map compilation, cartographic design, address matching, geo-referencing, and image processing to name a few. This will likely result in staff members who may have no library training. Quality GIS professionals are more akin to our IT staff whose knowledge of computers and databases is more important than their knowledge of librarianship.

GIS AND WEB SERVICES

Is there any research library that does not have its own website? Web addresses are an important signature on the web that allows a library to communicate essential information regarding its holdings, hours, services, facilities description, etc. Here, you can see a recent months use of our website with over 8,000 hits daily. The web allows us to 'communicate' in a way with users that we would never be able to achieve in person. A website is also an avenue used to distribute geospatial information and to access geographic information. Websites vary, from the well known raster image sites at the Library of Congress and David Rumsey's site, to government data distribution sites such as those provided by the USGS and EDINA in the UK as mentioned earlier.

However, beyond the well-know sites, bookmarked by most of us, can we find the information we want on the web? Oddens' Bookmarks, known to us all, had over 22,000 links earlier this summer and a useful search engine. On the other hand, a search of Google last week for the term "maps" resulted in an excess of 132 million hits; Yahoo exceeded 122 million. I then revised my search for "Cambridge Maps". Google returned in excess of 3.8 million and Yahoo 1.4 million - we are truly not alone! Usually our searches are more precise and we often find what we are looking for but, inevitably, we will miss some potential sources because of the sheer magnitude.

GIS AT HARVARD

As many of you are aware there is no geography department at Harvard today although Harvard had a rich tradition of geography until the department was disbanded after WWII - but that is a story for another time. Spatial studies are concentrated in the Harvard Graduate School of Design today which includes departments for architecture, landscape architecture, and urban and regional planning. In 1992, with the retirement of Harvard's map curator, the Library made a decision to introduce GIS based on the ARL project mentioned earlier. I arrived that year to a library that did not have any computers, not even one that linked to the online library system - we were going to build from the ground up. As a member of the ARL project, we were able to acquire our first computer and began to offer very basic GIS services using one of the early versions of ArcView. It literally became one of those phenomena's that "If you build it, they will come" - and they did.

It quickly became apparent that we needed someone who would be able to concentrate on developing GIS services. We started with a part-time graduate student and, within one year, we were fortunate to hire someone full-time who had worked with a suite of ESRI products during graduate studies. It soon became apparent that we had a conflict between being a cartographic specialist and a GIS specialist - we were already becoming specialized, and the demand for someone to assist with mapping projects versus higher end data delivery and analysis was already upon us. In 1998, six years after we had begun to offer GIS services we hired our first professional GIS specialist. Three years later, in 2001, we hired a Digital Cartography specialist, whose major focus was to work with faculty and students to improve their use of GIS services and produce maps for their

research. Our GIS specialist then turned their attention to the development of the Harvard Geospatial Library (HGL). Three years later, and just last month, we have hired a GIS instructor and a GIS data technician. Our GIS specialist was spending a significant amount of their time working on the development of HGL and processing data for input while the Digital Cartography specialist was making appointments 2-3 weeks into the future to work with individuals. Our instruction was almost always limited to one on one meetings and proving to be time consuming and inefficient. Similarly, development for HGL and processing data usually meant that less and less data was getting processed.

Now, for the first time, we have more of our staff devoted to digital data (4) than to paper reference, cataloging, and preservation (3.5). One might assume from this that we will be turning our budget over to digital acquisitions, and that is a national trend in the US Between 1994 and 2002, the Association of Research Libraries reported that expenditures on electronic resources for the typical university library increased by almost 400%, while library materials budgets only increased by 61%. While I do not have exact figures, I believe that our current budget used for paper is ca. 70% and our digital resources are ca. 30%. Our equipment expenses certainly reflect a similar trend. We have purchased very few new map cases in the last few years but have purchased a new wide format scanner, a new wide format printer, DVD burners, upgraded four public computer workstations and acquired numerous software upgrades.

One of the major challenges for map collections revolves around the area of copying and reproduction rights. Like so many other collections, we have struggled with the demands of users versus the needs for collection preservation. Several years ago we made arrangements with a local copying firm, located in Harvard Square, to pick up and deliver materials that they would then copy on an oversize black and white copier. This worked for current materials, but the demands for historical materials, color, and digital copies were not being met. We did have the choice of sending materials to our Photo Lab but the associated costs were usually too high, especially for our students. Large flat-bed scanners, while perhaps representing the ideal solution, were, and remain, very expensive solutions. Scanners themselves are not stand-alone operations. They require an accompanying computer setup and an output printer. We also knew that many of our users would not be interested in print. Often, our users prefer a digital file that they then take back to their office to create maps or illustrations for their research. We chose a flat roll-through scanner produced by Vidar in consultation with our preservation office. Our particular model allows for materials of different thickness to be scanned without undue pressure on the item and we encapsulate our materials so that the original artifact does not come in contact with the soft rubber rollers. Users have the option of choosing a digital file or having a print made on our recently upgraded printer, an Epson 9600. To illustrate this equipment's role with GIS, and certainly to geographic information services, the following statistics will be useful. In an 11 month period ending in May 2004 we processed 227 orders, scanned 1,041 items, printed 411 images, and created 630 digital files. Therefore, 61% of the items scanned preferred to have digital copies.

Two projects will illustrate how GIS has changed our services to our community. First, an MIT student approached our staff and asked to look at large-scale USGS topographic maps for various US cities. Although this student seemed to know exactly what she was looking for, a sizeable pile of maps was accruing over several days. A follow-up interview revealed she was looking at the wrong maps and really needed the smaller scale 15-minute maps (similar to the OS 1 inch series). She was using these maps to create a link between railroad density, and future abandonment, and economic development for many cities. She then took advantage of our scanning facilities, produced digital scans, and worked with our staff to learn image tiling and geo-referencing. This thesis project has now evolved from a traditional paper thesis to one incorporating GIS raster imagery and spatial analysis.

Similarly, a Harvard faculty member was interested in having a map made for a publication that would show early Celtic place names. He had been researching the place names of an early Irish saga - translated it is the 'Cattle raid of Cooley.' He knew nothing of GIS and had compiled his place names from various manuscript sources and was comparing them with maps in our collection. The more complicated his research became, and as the size of his list of place names grew, it became all too obvious that 8.5 x 11 maps were not going to capture his results. It was then that we suggested him to use some basic aspects of computer mapping to capture and archive his research. We acquired the Bartholomew Ireland boundary data specifically for this project and, combined with current place names from Irish OS maps, he is preparing a series of maps to illustrate the many names mentioned in the saga. It is not unusual to find him in the Map Collection weekly working on one of our computers with several printed maps on the table right behind him, and he moves back and forth easily. A typical

day will find all four of our workstations being used at various times and often paper maps or atlases are close at hand. For me, that is a successful map collection entering the 21st century.

LEGACY COLLECTIONS

I believe there are certain collections, especially those large collections represented by LIBER in Europe, and ARL in the US, that may be described as legacy collections. We might refer to them as super research collections, or even collections of last resort. In most instances, these collections contain many thousands of maps representing a diversity of time and they have collected their materials over many decades. "Whither the Paper Map" is not an option for these libraries, in my opinion. We may embrace the newer technologies, but I believe there is an inherent responsibility to maintain our traditional collections and to continue to develop them. As stated earlier, it will not be possible to save everything. However, we still must prioritize, understand the role and value of our collections, and establish goals to ensure that these collections are preserved, cared for, and accessible. I realize that I am preaching here to a converted group, but we need to realize that we are keepers of cultural artifacts that reflect the evolution of mapmaking: its techniques, its different formats, and the mapmakers themselves.

THE FUTURE

I am both confident in the future for map collections and worried. I am confident that we have a new range of services and tools to supplement what we have provided for many years. I am worried that, as libraries become enamored with the new technologies, that we move more and more of our historic collections to the back room. There are those who would suggest that we scan the originals and send them to storage. In some cases, this will be a very appropriate scenario, but we must resist policies that would restrict access for our collections to computer screens. I worry if we can maintain professional librarians that understand the importance of balancing budgets between digital information and traditional collections. I hope that services in a map collection incorporate technology and, while it is certainly useful onto itself, it should also be used to enhance access to our historical collections. I worry that many map collections, including Harvard's, has very little expansion room and that portions of our collections will be stored in offsite storage facilities. I worry that I will not be able to access many of the electronic resources I have already acquired because the software will no longer be readable in the near future.

In conclusion, we have the technology. Only time will tell if we have the intelligence to use it wisely. It would be pure folly if libraries were to ignore traditional collections that have been built over centuries. The successful map collections of the 21st cetury will be those that maintain respect for their legacy collections and adapt the new technologies to develop new services and enhance better access to their collections as a whole.

REFERENCES

Baker, John C. [et al]: Mapping the Risks : Assessing the Homeland Security Implications of Publicly Available Geospatial Information. Santa Monica, CA : Rand Corporation, 2004.

WEB SITES REFERRED TO IN THE TEXT

ArcView. http://www.esri.com/software/arcview/

ARL - Association of Research Libraries. http://www.arl.org/

Bartholomew. http://www.bartholomewmaps.com

EDINA - Edinburgh Data and Information Access. http://edina.ac.uk/maps

ESRI. http://www.esri.com/

Epson. http://www.epson.com

Geography Network. http://www.geographynetwork.com

GfK Macon. http://www.globalmaps.com/

GIS – Geographic Information Systems. http://www.gis.com/index.html

Harvard Geospatial Library. http://peters.hul.harvard.edu:8080/HGL/jsp/HGL.jsp

Harvard Graduate School of Design. http://www.gsd.harvard.edu/

Harvard Map Collection. http://hcl.harvard.edu/maps/

ITMB – International Travel Maps & Books. http://www.itmb.com/

LIBER - Ligue des Bibliothèques Européennes de Recherche. http://www.kb.dk/liber/

MapQuest. http://www.mapquest.com/

Nationalatlas Bundesrepublik Deutschland. http://www.ifl-leipzig.com/daten/deutsch/nationalatlas/demo/index2.htm

Oddens' Bookmarks. http://oddens.geog.uu.nl/index.php

Vidar. http://www.vidar.com/

U.S. Census Bureau. http://www.census.gov/

USGS – U.S. Geological Survey: http://www.usgs.gov/

Yahoo! – Maps. http://dir.yahoo.com/Science/Geography/Cartography/Maps/