Spatial Data Visualization and Analysis Support in the Library

Bruce Reid and Francis W. Derby

The Pennsylvania State University
P. O. Box PSU
Lehman, PA 18627

Abstract

In many institutions of higher education, the primary mission of the library is to be a resource for students, faculty and researchers. In meeting this objective, the library has served as a repository for print materials. With regard to spatial data, a collection of monographic materials, maps and other materials of cartographic nature, have often been the sole source of support for geographically referenced data.

In today's information age, the need for current and up-to-date data has become essential for decision-making and resource management. The utility of spatially referenced materials in hard copy format has become greatly curtailed due to the difficulty and cost for keeping them current. The need for spatially referenced materials in digital or electronic format to support new technologies and applications has become increasingly apparent in many institutions.

One major application of spatially referenced data is in geographic information systems (GIS) technology. A decision support and resource management technology, GIS is a tool for analyzing spatially referenced data to identify trends, and assess possible impact of options for solving spatially related problems. No longer restricted to engineering and the social sciences, the application of GIS technology extends to all forms of activities where geographically referenced data are used. In addition to academia, the technology is being used extensively by Federal, State and local agencies, public and private institutions, as well as individuals.

In order to provide data and visualization support for such activities, libraries are compelled to provide the requisite data, computer hardware, and software for patrons to perform spatial analysis and visualization. Many institutions are investing financial resources to acquire electronic data, computer hardware, and software to equip the libraries and to provide the requisite staff training. The functions of the library have thus, been extended beyond the original

mission as a repository for print materials. This paper looks at the approach which has been adopted by Penn State, Wilkes-Barre campus to extend its support services to include spatially referenced data acquisition, analysis and visualization activities.

1. Introduction

Over the last two decades, advances in information and communication technologies have influenced the manner in which businesses are conducted. The digital technology revolution has also influenced the need for complete and up-to-date data to support research, data analysis, decision making and resource management activities. Complex questions that are being addressed both nationally and internationally require that scientists and researchers have access to new technologies such as geographic information systems as well as the most complete and up-to-date electronic data. In order to continue supporting their clientele and provide required materials in a timely manner, libraries are extending their resources to include electronic materials as well as the tools to access and process them. But this change has come with additional cost in data acquisition, computer hardware and software, and staff training. This paper looks at the approach the library at the Wilkes-Barre campus of The Pennsylvania State University adopted to provide spatial data visualization and analysis support, in view of the fact that the university has geographically dispersed campuses.

2. The role of libraries in academia

Historically, the mission of most libraries in academic environments has been to act as a primary information and data resource for students, faculty and researchers. The library at the Wilkes-Barre campus of The Pennsylvania State University had operated in a traditional fashion, not unlike most of the libraries in the region. The library had been supporting research and curricular activities with an extensive array of print materials such as monographs and journals, along with microfilm and some proprietary databases. In addition, the library personnel had been conducting workshops that were subject specific as well as serving as a primary source and data repository for the campus. Spatial data was traditionally obtained through a collection of monographic and cartographic materials, and in most instances, was the only source of support for geographically referenced spatial data. The information gathered was typically one-dimensional in nature and always in hard-copy format.

In today's fast-paced information-driven world, the library's methodology concerning information retrieval must be continually reflective of the technological advances. Many of the resources for geographic information that libraries utilize do not meet today's needs for currency and accuracy, but changes are being made. For example, print materials are being supplemented and often replaced by computer generated formats. Proprietary databases and specialized software now provide information that libraries must possess to satisfy the needs of our students and faculty that demand rapid and accurate decision responses. The library must adapt to the

changing technology landscape and be actively searching for new avenues to expand its supportive role in the academic community.

3. Drivers for change

Several factors have influenced the need for libraries to adapt to change and extend support services into electronic and other data types. Primary factors include the digital technology which provides an efficient and more accurate method for capturing, processing, and disseminating spatial data; the revolutionary information and communication technology; and the need for researches, entrepreneurs, managers, decision, and policy makers to operate with the most complete and up-to-date information. Other factors such as the computer, data storage, and analysis technologies have also been catalysts for the change. However, this discussion will be restricted to the primary drivers mentioned here.

Geoscientists, environmental scientists, and researchers need to analyze large volumes of data to identify trends and patterns that exist within spatially referenced phenomena. New and evolving technologies have facilitated methods for capturing, processing, and analyzing spatial data. Examples of data capturing technologies include positioning systems such as the Global Positioning System (GPS) in the USA and the Russian GLONAS system, and digital data capture with cameras using charge coupled devices (CCD) technology, and remote sensing technologies especially high resolution satellite imagery. Data processing and storage technologies have been influenced by computers with very high processing and multitasking capabilities. Extensive data storage capabilities have made it possible to store large data sets and thereby improved the ability to analyze and disseminate large spatial databases. The technology for manipulating and analyzing large geospatial data arguably has had the greatest influence on the geospatial information environment.

Using these new and enabling technologies, scientists and researchers are harnessing new technologies such as geographic information systems, remote sensing techniques, powerful climatic and environmental simulation models to extrapolate global impacts of human activities on the environment. Application of geospatial information science in land and natural resource management has arisen from the knowledge that land, the most valuable resource of mankind, is crucial for the economic, social, and environmental advancement of all communities. Land and natural resources are neither inexhaustible nor indestructible. Therefore, accurate knowledge of the extent of such resources and record of such knowledge are essential to their rational use and conservation. Land management is a process whereby the resources inland are put to their best use toward the benefit of mankind. Environmental and earth scientists are also applying these technologies to effectively manage spatially distributed resources such as water, oil, and gas in an effort to balance the ecosystem and to ensure that the impacts of human activities are not detrimental to our own existence. Other factors such as the Internet, data communication, and networking technologies have also played major roles to influence the geospatial information technology revolution.

Over the last few decades, there has been an obvious upsurge in information and communication technologies. Among other things, information technology provides the means for analyzing large data sets to derive knowledge that is not obvious just by inspection. Just as information and communication technologies have been key to enabling growth of businesses, spatially referenced information systems have become critical resources for making informed decisions concerning business and commerce, resource management as well as socioeconomic and sustainable development. The current geospatial information revolution is rapidly becoming an essential component of the Information Technology (IT) revolution (Derby and Ofosu, 2000). The need for spatial data to support IT activities has resulted in the availability of on-line data about land, land parcels, and spatially referenced environmental data. Informed decisions based on knowledge derived from an analysis of spatial data are changing the dynamics of civil society, sustainable development, as well as the global impact of man's actions on the environment. Geospatial information systems and communication technologies are being employed in land and natural resource management, engineering applications, public and environmental policy administration, and in businesses.

4 The need for adaptation

An unofficial review of the effectiveness of the library services with respect to data availability and applicability to education and research showed that, whereas there was adequate material to support the liberal arts and humanities, the availability of digital spatial data to support engineering, environmental science, and the geosciences was lacking. Although such materials were available in other campuses within the Penn State system, their retrieval from remote campuses through communication networks could not be easily facilitated. The need for a local repository of spatially referenced materials in digital or electronic format to support new technologies became increasingly apparent. To facilitate education and research involving geospatial data, the library is obligated to provide the necessary support services. In today's IT environment, this responsibility includes the availability of up-to-date spatial data in electronic format. It also includes a need to increase the number of subject specific and proprietary databases along with specialized software for processing, analyzing, and visualizing the results of geospatial research activities. To meet these responsibilities, the library needs to acquire computer hardware and software products that are able to satisfy the demands of spatial analysis and visualization.

Geographic information systems (GIS) technology is such a tool, and it has become increasingly important in many institutions. In installing a GIS at the library, other factors need to be considered such as:

- training of library personnel to provide the requisite assistance to students and faculty,
- acquiring electronic data and digital databases,
- connecting and/or developing links with centers and repositories of proprietary spatial databases,

- providing a list of freely accessible digital spatial database centers such as the Pennsylvania Spatial Data center (PASDA), Federal Geospatial Data Center (FGDC), Bureau of Land Management (BLM), National Geodetic Surveys (NGS) and many others, and
- planning system maintenance protocols.

Most of these issues are administrative in nature. Major administrative decisions regarding the acquisition and dissemination of geospatial data had to be handled at senior levels within the library system. Other decisions such as hardware configuration, system support, and maintenance, are matters to be addressed at the local level. These matters are discussed below under administrative arrangements, training and support, and implementation.

5. Administrative arrangements

Having established the need to implement a GIS support in the library, administrative arrangements had to be made regarding availability of electronic databases, accessibility across the communications network, and collection development policy. Penn State has a collection of standard electronic proprietary data such as regional, statistical, and demographic data which are useful for GIS applications. These databases exist on CD-ROMS which are shared between campuses and departments. A collection development policy at Penn State ensures that future needs for geospatial information are addressed accordingly.

Additional spatial data sources could be made available to students and researchers by providing links to the Web sites of freely accessible spatial databases. Links to sites such as the Pennsylvania Spatial Data Center (PASDA), Bureau of Land Management (BLM), Federal Geographic Data Center (FGDC), National Geodetic Surveys (NGS) and many others are valuable resources to students and researchers. Such links would be provided at the online Library Information Access System (LIAS) at Penn State

Also, a decision had to be made about gaining access to a wider collection of proprietary spatial databases. However, before acquiring extensive spatial databases, it was necessary to review what was available under current licensing agreements with proprietary database providers. Penn State already has agreements with other libraries and several repositories which already have access to some electronic spatial data. It was necessary to investigate the availability of spatial data within existing networks of repositories and to determine the limits of existing agreements with regard to spatial data in electronic format. Where necessary, such agreements could be modified to include sharing of electronic spatial databases. Additional linkages will be explored when demand for materials increases beyond existing agreements. However, it was important to recognize that at as demand for data increases, licensing agreements with specific data providers may be required in order to support specific needs. Linkages with other institutions are currently being explored.

To facilitate access to electronic spatial databases by students and faculty beyond local campuses, another administrative decision had to be made to provide online access to such data, rather than keeping these resources at the reference section of the library. Whereas online

availability of electronic spatial data could facilitate access from other campuses, licensing agreements may prohibit such action. Again, it was necessary to investigate whether existing licensing agreements included online dissemination and to seek modification to the terms of the agreement where necessary. Through these administrative arrangements, it was determined that sufficient data were available to satisfy immediate local needs. Although there are plans to provide geospatial data online, legal issues associated with online delivery of geospatial within the Penn State university system need to be addressed first. It is anticipated that the issue will be addressed in due course.

Other administrative issues such as responsibilities for computer networking and communications support, updating and maintaining information at the website, and overall system maintenance are included in existing administrative arrangements within the Penn State system. Such matters are the responsibilities of specialized personnel who are already employed at Penn State.

6. Implementation

Although there are several varieties of GIS software available and in use by various institutions, Penn State has a university-wide software license agreement with Environmental Systems Research Institute (ESRI) for all their GIS software products. The license agreement includes the use of all ArcGIS and ArcView products as well as extensions. Through this agreement, it was possible to install ArcView software and extensions on a local computer for a fraction of the cost. The next issue to consider was a suitable configuration of computer hardware for the GIS software.

With assistance from the computer systems administration staff and a faculty member who is knowledgeable in GIS development, the appropriate hardware configuration was designed. Essential features of the configured computer consisted of a 2.80 GHz. Pentium 4 processor, 512 Megabytes of RAM, a CD-RW drive, a 250-MB Zip drive, and 128 MB of video RAM. The system included a 21-inch monitor. Other peripherals included a large format color Inkjet printer. Although this GIS hardware is connected to the communications network, it is dedicated to geospatial data processing and is located away from the reading rooms to avoid disturbing patrons who need to work in a quiet and tranquil environment.

7. Training and support

The next objective was to plan the initial training so that the library personnel could provide adequate support to users. The Head Librarian identified training and support needs. One staff member, the Head Librarian, pursued the required training in order to provide assistance to users. Training involved self-study of GIS manuals and related materials complemented by support and supervision from an experienced faculty member from the engineering department. Additional

training was obtained from personnel from the GIS section of the Maps Library at the University Park campus of Penn State University. Support provided at University Park included training on the use of available digital databases, access to ArcView and ArcGIS courses at the Virtual Campus of the Environmental Systems Research Institute, funding for the required hardware, release-time for the Head Librarian to attend training sessions, and additional funding to acquire print materials to be made available to users. As a result of these training opportunities, the Head Librarian is knowledgeable enough to complete certain basic GIS tasks. The Head Librarian will travel to the University Park campus on a regular basis for additional training.

Support for users are currently provided by the Head Librarian only. Most requests for assistance have been in the form of a need to access specific databases. But most of the current users of the system are knowledgeable in the use of the software. It is anticipated that a student work-study who is experienced in the use of Arcview and ArcGIS will be needed in the library to provide assistance to advanced users when demand increases. Fortunately, there is no immediate need to conduct workshops on GIS applications because there are GIS and related courses offered on campus which cover these topics. Until the need arises, any related instruction will be to demonstrate the types of databases that are available online, on CD-ROMs and at other locations together with instructions on how they can be accessed.

8. Conclusions

Although requests were sent out regularly for faculty members to submit lists of materials needed in the library to support teaching and research, most members limited their requests to print materials. It also became obvious that there was a decline in the use of the library simply because students, especially those in the geosciences and engineering, chose to use the Internet rather than the library. A review of the effectiveness of the library services highlighted a need for geospatial information that was not being addressed. It became clear that geospatial data visualization and analysis support was needed in the library at the Wilkes Barre campus. Students who were aware of the source from which to obtain digital spatial data did so without support or assistance from the library. In effect, the library was losing a portion of its customer base, simply because it did not have the kinds of materials that some students and faculty needed.

In addressing this shortcoming, there has been a need to redefine objectives, extend the mission of the library to include support for spatial data visualization; invest in the requisite hardware, software, electronic databases; and training necessary to support GIS and other activities. Providing access to electronic databases requires a review of existing administrative policies. This process entails redefining objectives for providing electronic databases through the Penn State system and reviewing policies regarding access and distribution of such databases in relation to existing license agreements with proprietary database providers.

In all, it has been possible to develop a GIS visualization support system at the library without incurring extensive financial expenses. Since its implementation, the library has seen an

increased use of the system. The GIS workstation that has been installed at the library has proven to be an important resource. Some faculty members are applying the technology in their courses. New courses are being developed in non-engineering disciplines for students to gain exposure to the strength and benefits of the technology. A new course in spatial data communications has been developed with the aim of exposing non-engineering students to the technology. Finally, an entirely new program is under development which will incorporate spatial information technology and data analysis to other disciplines such as business management, environmental science and transportation management.

References

Adler. P. (1995) *The ARL GIS Literacy Project*, Association of Research Libraries, Washington D.C. (ftp://www.wrl.org/info/gis/gis.descrip).

Beard, K (1995) *Digital Spatial Libraries: A context for engineering and Library Collaboration*, Information Technology and Libraries Vol. 14, pp 79-86.

Boxall, James (2003) *Geolibraries: Geographers, Librarians, and Spatial Collaboration*, Canadian Geographer, Toronto, Vol. 47, Iss. 1; pg 18.

Derby, F. W. and Ofosu, W. (2000) *Geographic Information Systems as a Complement to Information Science and Technology Program at Penn State*, Proceedings of the 2000 ASEE conference, St. Louis, MO, ON CDROM.

Frazer, S. L. (2002) *Virginia Geographic Information System (GIS) Programs and Resources for Libraries*, Virginia Libraries, September 2002, pp. 18-20.

Biographical information

BRUCE REID is the Head Librarian at Penn State University, Wilkes-Barre campus. He has a B.S. degree in Business from Farleigh Dickinson University in New Jersey, and a Library Information degree from the University of Minnesota. His subject areas are Business, GIS applications, Telecommunications, and Land Surveying.

FRANCIS DERBY, is currently Associate Professor of Surveying and Geographic Information Systems at Penn State University. He has extensive international experience in cadastre and Land Information Systems and GIS. His current interests include land tenure issues, implementation of Cadastral, Land and Geographic Information Systems.