# 6

# Access to Spatial Data

## 6.1 FINDING SPATIAL DATA

When the users of cartographic data are municipalities, utilities, or state and local governments, there is often a significant amount of geocoding to be done before a mapping project can begin. This data capture can be extremely costly, however, and increasingly these mapping organizations are looking toward the use of digital data that already exist, even if that data acts only as a base upon which to expand. If the cartographer is seeking data for a map that are likely to already exist in digital form, then the problem is not to geocode, but to find and access the data.

As we have seen in Chapter 5, there is no shortage of existing data suitable for analytical and computer cartography. This situation means that new data need to be geocoded only in those circumstances where new layers or themes are needed, where new data have to be integrated with existing maps, or when digitally unmapped areas are to be used. Rather than actually dealing with the specifics of data storage formats, data models, and data structures, the computer cartographer is often faced with simply finding *where* data sets reside, whether they are *accessible*, and through which storage media distribution is possible.

As more and more data are demanded by computer mapping and information management, the more the advantages of digital data become apparent. Digital data can flow easily across computer networks, making networking the second computer revolution in cartography and making the problem of distribution and storage vastly simpler. This chapter discusses some of the means to find and access spatial data, a task that is becoming more and more common. We examine the role of map libraries, the role of the national spatial data infrastructure, and the power of the Internet, and we conclude with a discussion of spatial data transfer standards.

### 6.1.1 Using a Map Library

The search for paper maps is often conducted in a library. Libraries most likely to carry maps and support cartographic research are the research libraries in the largest cities and those attached to major universities. Map librarians make use of computer networks to share information and conduct searches, and they are increasingly making census and other digital maps available both in libraries and via networks.

Map librarians have faced difficulties in preparing for the digital cartographic revolution due to a text-only library tradition, tight budgets, and an inability to educate sufficient numbers of map librarians in digital mapping techniques. Nevertheless, lower costs for hardware and software, new technology such as CD-ROM, extensive demand for Census Bureau TIGER data, and a new awareness of the need for digital map libraries have led to significant changes. Some pioneering libraries began the transition in the mid-1980s, while now about half of the nation's libraries either retrieve and store digital map information directly or have used network links to establish a new working relationship between map libraries and their users. This transition has been helped in some cases by educational programs, such as the American Research Libraries initiative to educate and equip libraries in the provision of spatially referenced data in all formats. This program has led over 70 libraries into the digital era, providing training, hardware, and software.

The map librarian now plays a role as a broker of spatial information, linking the right data with the user and perhaps even providing the first software-based display of the data and a hard-copy computer-generated map. The network link has also changed the view of the map librarian, from that of an acquirer of a copy of a map for a user to that of a custodian of a section of unique data on a network for access by users of all libraries. Network access to library catalog information has stressed this distributed database model.

In addition, commercial companies often sell cartographic data and will conduct searches. Landsat imagery, for example, from EOSAT, can be searched for and browsed using an on-line database. Major U.S. government programs such as the Global Land Information System and NASA's Earth Observation Systems have evolved similar concepts, including image browsing by network.

### 6.1.2 The National Spatial Data Infrastructure

The U. S. Geological Survey makes its data available through the Earth Science Information Centers, where questions about digital map data availability within the federal government are answered. Some states, such as South Carolina and Wisconsin, have data clearinghouses that make data available to users. There is also a large variety of journals, newsletters, and information sources to suggest data sources.

A prevailing attitude in the United States has been the concept of digital map data, at least raw "base" map information, as a public good. The growth of geographic information science, with its powerful suite of tools in geographic information systems technology, has taken place largely because of the abundance and low cost of data. This is particularly obvious when the system in the United States is compared with full or partial cost-recovery systems, such as that used in the United Kingdom by the Ordinance Survey (Rhind, 1992). In many applications of computer mapping and GIS, data production or conversion often accounts for over 80% of the cost. With free or low-cost data, the cost of projects obviously drops significantly.

At the national level in the United States, an initiative is now under way to develop the NSDI, or National Spatial Data Infrastructure (National Research Council, 1993). This system would make digital spatial data broadly available to the public, probably over computer networks such as the Internet or through “published” CD-ROMs. In addition, the system would coordinate the mapping activities of government and other mapping agencies, to reduce redundancy and to make map data sets match each other across scales and at their boundaries.

The ability to search and retrieve data quickly will give rise to new cartographic applications, such as the mapping of natural disasters as they occur, the use of computer cartography in search-and-rescue, the use of in-vehicle navigation systems, and the development of many other possible uses, especially when coupled with the position-locating capabilities of the Global Positioning System. Both analytical and computer cartography have a considerable amount to gain from the National Spatial Data Infrastructure, as it has been proposed.

Early systems that had some of these characteristics are NASA’s Land Pilot Data System, USGS’s Global Land Information System, and NASA’s EOSDIS Version Zero. What has changed significantly since the origins of these systems is the emergence of the Internet as a highly effective mechanism for both data searching and data distribution.

## 6.2 USING THE INTERNET AND FTP

There is little doubt that the vehicle of choice for the distribution and searching of spatial data is the Internet, a network of computer networks available worldwide (Krol, 1993). The Internet is a network of computer networks and is accessible to all users through a computer that is attached to the system. In addition to e-mail (electronic mail) worldwide, the Internet allows the users of computer mapping systems access to the existing National Spatial Data Infrastructure, which is assembling itself spontaneously at a remarkable rate of speed. The Internet contains several information and data sources of great value for cartographers, primarily network conference groups, file transfer mechanisms, and network search capabilities.

### 6.2.1 Network Conference Groups

In terms of information, several network conference groups exist for cartographers, and others focus on specific software products. In each case, a list is subscribed to by sending a mail item to the Internet address in the following way (which uses `Unix Mail`):

```
mail LISTSERV@address
Subject:
subscribe LIST-L Your Name
<control>-D
```

where the address is the Internet address, the e-mail system query for a subject is left blank, the list name and your name follow the subscribe command, and the message consists only of these items. Alternatively, the news reader system can be used, which “screens” your mail and allows it to be read selectively. Other systems work similarly.

Most lists have an archive of FAQs (frequently asked questions), which should be located and read *before* you begin using the list. It is a good idea to monitor a list as a passive reader for a few days before becoming active. A common mistake for new users is to send subscription requests to the list name instead of the list server. This error is extremely frustrating to newsgroup users, who are then bombarded with every single start and stop request for subscriptions. **This error is to be avoided at all costs. It is a severe breach of network etiquette.**

MAPS-L at uga.cc.uga.edu is a medium-volume moderated newsgroup for map librarians and often has interesting locational queries. A Canadian equivalent is CARTA, at sask.usask.ca. GIS-L at ubvm.cc.buffalo.edu is a high-volume newsgroup for GIS users. It is often highly technical, but is frequently lively and contains much of interest to analytical and computer cartography.

This list is cross listed at comp.infosystems.gis on Usenet, where e-mail tools make monitoring messages and screening by subject easier. As do many lists, this list has an FAQ (frequently asked questions) list. To get the FAQ list one can anonymously ftp (see Section 6.2.3) from abraxas.adelphi.edu (file /pub/gis/FAQ) or send a message to gis-faq-request@abraxas.adelphi.edu.

Other GIS lists include TGIS (temporal issues in GIS), UIGIS (user interface issues), and CPGIS (Chinese Professionals in GIS). INGRAFX is a very low volume list at psu-vm.psu.edu. It is a list designed for those interested in issues and research in mapping and visualization. IMAGRS-L at csearn.bitnet is a list devoted to remote sensing and has some spillover into cartography.

Support lists for specific software packages can be used for help in implementing software. These include grassu-request@zorro.cecer.army.mil for the U.S. Army Corps of Engineers GRASS GIS and image processing package (both software and data are available), and idrisi-l@toe.towson.edu (send subscribe message to MAILSERV, not LISTSERV. The NSDI discussion list is NSDI-L, with the subscription message to listproc@grouse.umesve.maine.edu. Some commercial companies also use lists for technical support.

Finally, a note of caution. Network lists are somewhat ephemeral. Lists evolve, move, and sometimes die. The use of a network browsing tool such as mosaic to locate active lists is recommended.

### 6.2.2 Network Searching

It is also possible to conduct searches over the network. Searches are often conducted using network browsing tools such as ARCHIE, GOPHER and WAIS. An overview tool, mosaic (Ritter, 1994), allows access to each of the other searching systems, acts as a gateway into systems, and allows menu and window-based queries across the network (Figure 6.1). These tools allow a user to search publicly available files located on servers.

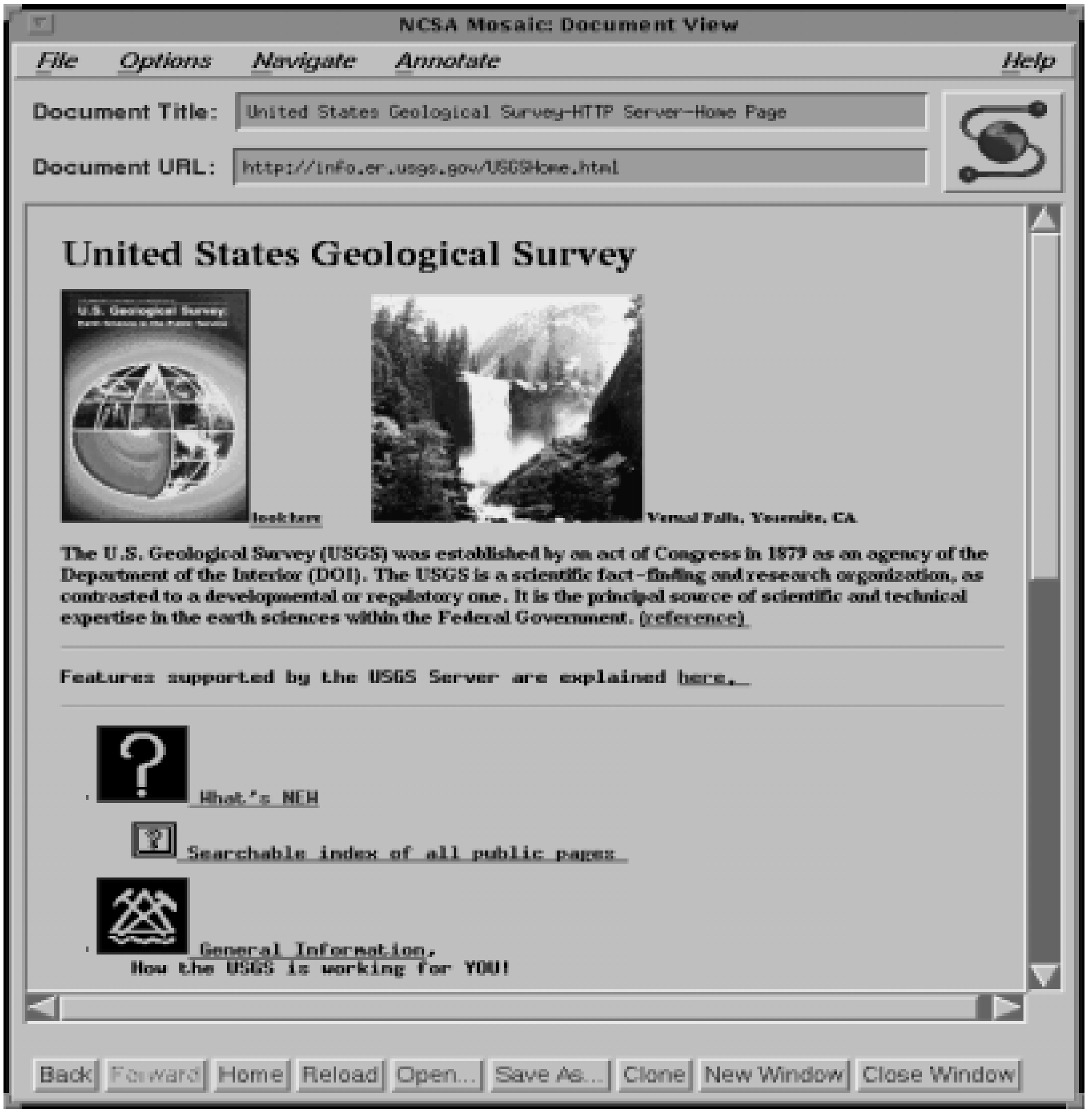


**Figure 6.1** The mosaic gopher Internet gateway to the USGS.

*Servers* are file archives located on the network that are accessible over the network either by remote login or by the file transfer protocol FTP. The user can also search directly, if a specific Internet address is known.

NASA maintains a Master Directory of all its data sets that is searchable by latitude and longitude. It can be used by first setting your workstation or terminal in vt100 emulation mode (in Unix, `setenv TERM vt100`). The system is comprehensive, and can be used to access several other systems and many individual sites. To use the system, `telnet` to `nssdca.gsfc.nasa.gov (128.183.36.22)`. At the prompt, enter

NSSDC, for username, enter NSSDC, and then select option 1 from the top-level menu.

The Master Directory allows browsing and queries to many other databases, including those of the EROS Data Center and the Pilot Land Data System. It is likely that this mechanism will also contain EOSDIS, NASA's Earth Observing System Data and Information System, which will be used to archive and distribute data from many satellites suitable for mapping, including LANDSAT 7. Another search directory of use is the USGS GLIS database (Global Land Information System), which includes index references to many USGS map products including DLGs and DEMs and satellite imagery such as AVHRR (glis.cr.usgs.gov or 152.61.192.54).

### 6.2.3 Network File Servers

In many cases, digital map data are available from individual file servers. The means by which access is made is to use the *File Transfer Protocol* (FTP). Using this protocol allows one to login to a guest account on the file server, usually using the login name of anonymous. By courtesy, it is common to use your own Internet address as the password, thus leaving a trail for the owner to trace who is accessing the databases. The individual servers can be located by using the network search tools, or they are announced in network news groups.

Many of the cartographic data sets in the formats described in the previous chapter can be found by searching and then transferring via FTP. Table 6.1 contains a set of such internet addresses for servers. From the USGS server, for example, the 3-arc second DEMs, the GIRAS, the ETOPO5 , and several other data sets are available. The availability of data is continuously increasing. News about new data listings and postings is usually available at a top level menu. Many of the data sets are compressed to make them easier for network transfer, usually with the UNIX compress facility, which results in a file with a ".Z" filename suffix. These files should be FTP transferred using the binary format, rather than as ASCII. The FTP command binary selects binary file transfer.

It is likely that the Internet will remain a primary means of data distribution and will expand considerably in the future. A significant amount of data and in many cases software as shareware can be downloaded for microcomputers and workstations from the file servers attached to the network. As a network user, you have an obligation to return useful or interesting software or data of use to other users via the network.

## 6.3 CARTOGRAPHIC DATA TRANSFER STANDARDS

Moving map or any other data over a network requires that the recipient be able to use the data when it arrives. In Chapter 5, we saw that two types of data formats have developed for cartographic data: those that are the result of data production by an agency with a need to distribute the information, often along the lines of an existing paper map product, and the industry standard formats, which have evolved out of casual transfer of image and other data. Broad-scale transfer of data over networks is time-consuming and can rather quickly fill a disk. Proliferation of formats for exchange have resulted in costly and duplicative efforts to write file format converters and translators.

**Table 6.1 Network File Servers with Cartographic Data**

| Address | Login and Directory | Contents | Comment |
|---|---|---|---|
| sseop.jsc.nasa.gov | login anonymous | NASA space shuttle images | |
| ames.arc.nasa.gov | login anonymous cd pub/GIF | NASA space shuttle images | |
| vmd.cso.uiuc.edu | login anonymous, cd wx | GOES weather images | Available every hour |
| inidata.ucar.edu | login anonymous cd images | GOES weather images | Available every hour |
| aurelie.soest.hawaii.edu | login anonymous, cd /pub/avhrr/images | GOES weather images | Available every hour |
| hanauma.stanford.edu | login anonymous, cd pub/World_Map | The CIA World Data Bank I and II | Includes software for drawing the files, start with the README file |
| martini.eecs.umich.edu | telnet | Latitude, longitude and elevation for U.S. cities | |
| glis.cr.usgs.gov | telnet | Master directory and descriptions of most types of data | |
| alum.wr.usgs.gov | anonymous, cd pub/maps | Geology maps | Includes display software |

Although many different organizations, both private and public, have developed their own formats for digital cartographic data, the disparities between formats have become disadvantageous only since the ability to share and distribute these data has become possible, especially across Internet but also on local area networks. Distributed computing has evolved to a level where databases normally reside at host locations and are made network accessible to the actual applications,that is, the mapping programs, which actually use them. Thus a map data set could remain on disk in one location, but be mapped and displayed at another.

Designing a system to accept data from two formats needs either a format translator or two translators into a third common or “standard” format. A system to convert data between 10 formats, however, must have either 10 times 9 translators or 10 forward and 10 reverse translators to and from a single, universal standard. In the long run, clearly a commonly accepted standard is most desirable. There have been some pioneering efforts to standardize digital cartographic data formats. Among the standards are several developed for data exchange within government agencies. These agencies include the Committee on Exchange of Digital Data (International Hydrographic Organization), the Federal Geographic Exchange Format of the Federal Interagency Coordinating Committee on Digital Cartography, the Geographic Data Interchange Language of the Jet Propulsion Laboratory (NASA), the Standard Digital Data Exchange Format of the National Ocean Service, plus those discussed previously (Langran and Buttenfield, 1987).

Several standards for digital map data have emerged as a result of these efforts. The Spatial Data Transfer Standard, approved as a Federal Information Processing Standard (FIPS) in 1992, will be considered in the following section. Parallel efforts have resulted in DIGEST from NATO and DX90 from the International Hydrographic Organization. Some other countries, such as Australia and Germany with ATKIS, have initiated standards efforts. In addition, DLG-E, mentioned in Chapter 5 as a revision of the USGS’s Digital Line Graph system, can be considered a standard, although it clearly had an influence on SDTS (Guptill, 1991).

DIGEST (Digital Geographic Exchange Standard) grew out of the military’s and the Defense Mapping Agency’s standardization efforts, such as the Standard Linear Format (see Section 5.4.3). After several versions, the standard received Military Standard status in 1992. DIGEST has several versions and parts. DIGEST-A is feature based, while DIGEST-C uses a relational data model (see Chapter 9). DIGEST-C is also known both as VRF (Vector Relational Format) and as the Vector Product Format (VPF). The Digital Chart of the World is stored in VPF on CD-ROM. DIGEST-A uses a telecommunications standard (ISO 8824) in the place of the static medium file storage standard ISO8211, which is also used by SDTS.

DX90 is a digital exchange format sponsored by the Committee for the Exchange of Digital Data and it is accepted by the International Hydrographic Organization. It is intended to support the use of digital map data for charting, safety, and navigation at sea. DX90 is the popular name for IHO special publication number 57, the Transfer Standard for Spatial Data. This standard is under revision and is undergoing a learning process as the various nations of the IHO begin to use the standard to exchange data.

Out of these standardization efforts grew the National Committee on Digital Cartographic Data Standards, more commonly called the Moellering Committee, which was formed in 1982 under the auspices of the American Congress on Surveying and Mapping as a result of a request from the USGS. In 1987 this committee completed the first draft for a new set of common data standards for digital cartographic data. These standards have now been approved by the National Institute for Standards and Technology as a FIPS.

The standards have substantially influenced computer cartography. This has been due to their 10-year evolution, and the substantial involvement of the user communities in government, academia, industry and other interested parties. The Spatial Data Transfer Standard, as the final standard is known, provides a uniform and consistent terminology, a set of definitions, and a set of formats for information exchange that now form a mandatory framework for federal data transfers.

The scope of the standard, however, goes well beyond federal needs, and the standard has already influenced software producers, state and local governments, and the activities of all cartographers. As such, the SDTS forms the remainder of the discussion for this chapter, concluding with a set of C language programming data structures that are consistent with the terminology of the standard.

## 6.4 THE SPATIAL DATA TRANSFER STANDARD

The Spatial Data Transfer Standard, issued August 28,, 1992, as FIPS 173 (Department of Commerce, 1992) culminated a 10- year standardization effort that began as an attempt to standardize digital cartographic data and ended with a broad standard with wide scope for application to all spatial data and spatial data transfers. The purpose of the standard is to promote and facilitate the transfer of digital spatial data between dissimilar computer systems. The standard provides a common mechanism for the transfer of data, provides a set of clearly defined spatial objects and relationships to represent real-world spatial entities, and provides a transfer model to facilitate the translation of user-defined objects into the standard.

The standard consist of two parts. Part 1 contains the logical specifications, including concepts, object definitions, data quality, and transfer module specifications. Part 2 consists of a dictionary of entities, that is, definitions of the cartographic features to be encoded in the standard. The spatial data transfer component of the standards attempts to meet the requirements for moving digital cartographic data between systems. The USGS serves as the maintenance authority for this standard, which lists a series of steps to be taken to attain conformance.

The definitions and references within SDTS are a systematic attempt to define a set of cartographic primitives as zero-, one-, and two-dimensional objects, with which digital cartographic feature representations, that is, symbols and maps, can be built. It should be noted that the standard definitions are restricted to these object definitions and do not propose any particular method or type of symbolization. The language of the geographic entity and object is used—concepts that are used throughout this book.

The section on data quality proposes that digital cartographic data include a quality report, either as a paper document or as part of the data set. Elements of the report include lineage of the data (source and modification history, for example), positional accuracy, attribute accuracy, logical consistency, and completeness of the data. These requirements seem critical to the long-term survival of data sets, especially within GISs, and also are of critical importance for the provision of an assessment of reliability for a particular map product made from digital cartographic data.

The final section of the standard relates to cartographic features. Central to this section is a complete set of cartographic entity and attribute definitions. This set of features in intended to be as complete as possible, though provisions for update are made. This means, for example, that when a digital data set references a "water tower," there is complete agreement on what constitutes a water tower. For example, the cartographic entity CHURCH is an included term number for the entity BUILDING, defined as "a permanent walled and roofed construction."

Two other important sections of the proposed standard document are of direct interest. First, under the cartographic objects section, a glossary is included. Second, , the standard also includes a bibliography, to which the reader is referred for additional information.

Implementation of the SDTS has proceeded using the mechanism of the profile. A profile is a subset of the standard, reflecting the demands of a particular type of geographic data. The first profile, now added as part of the FIPS standard, is the Topological Vector Profile (TVP). Specification of the TVP has allowed many U. S. federal data sets to be restructured into SDTS, including SDTS-DLG and SDTS-TIGER. The maintenance authority for the standard, the USGS, has released public domain software designed to read and write data into the SDTS-TVP file structures. Many software vendors are now incorporating SDTS-TVP readers into their programs. Information about the profiles and the standard in general are available over the Internet from `sdts@usgs.gov` and can be found using Mosaic, WAIS, or Gopher.

The second stage for SDTS is to implement the Raster Profile of SDTS. The raster profile will result in an extension to the FIPS standard in the near future. Therefore, the implications for the distribution of government image data, such as GOES and AVHRR data, as well as Landsat and future EOSDIS data, are significant.

## 6.5 REFERENCES

Department of Commerce (1992). *Spatial Data Transfer Standard.* FIPS PUB 173, Federal Information Processing Standards Publication, Computer Systems Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.

Guptill, S. C. (1991). "Spatial Data Exchange and Standardization." Chapter 34 in D. J. Maguire, M. F. Goodchild, and D. W. Rhind, *Geographical Information Systems: Principles and Applications, Volume 1: Principles*. London and New York: Longman Scientific and Technical /Wiley.

Kottman, C. A. (1992). *Some Questions and Answers about Digital Geographic Information Exchange Standards*. Reston, VA: Intergraph Corp.

Krol, E. (1992). *The Whole Internet: Users Guide and Catalog*. Sebastapol, CA: O'Reilly and Associates, Langran, G. and B. P. Buttenfield (1987). “Formatting Geographic Data to Enhance Manipulability.” *Proceedings, AUTOCARTO 8,* Eighth International Symposium on Computer-Assisted Cartography, Baltimore, March 29–April 3, pp. 201–210.

National Committee for Digital Cartographic Data Standards (1988). “The Proposed Standard for Digital Cartographic Data.” *American Cartographer,* vol. 15, no. 1, pp. 9–142.

National Research Council (1993). *Toward a Coordinated Spatial Data Infrastructure*. NRC Mapping Sciences Committee, Commission of Geosciences, Environment and Resources. Washington, DC: National Academy Press.

Rhind, D. (1992). “Data Access, Charging and Copyright and Their Implications for Geographical Information Systems.” *International Journal of Geographical Information Systems,* vol. 6, no. 1, pp. 13–30.

Ritter, M. E. (1994). “Mosaic and the World Wide Web.” *Cartographic Perspectives,* no. 19, Fall 1994, pp. 44–47.

Tosta, N. (1993). “The Data Wars: Part I.” *Geo Info Systems,* January, pp. 25–27.