GEOGRAPHIC INFORMATION SYSTEMS

A *geographic information system* (GIS) is capable of performing virtually any conceivable operation on geographic information, from editing and compilation through analysis, mining, and summary to visualization and display. *Geographic information* is a particularly well-defined type of information, since it refers specifically to the surface and near-surface of the Earth, linking observations and measurements to specific locations (to all intents and purposes *geospatial* is synonymous with *geographic*). Maps are the most familiar traditional form of geographic information, so a GIS can be considered simplistically as a computerized collection of maps, but a far wider assortment of types of information can be included in a GIS, including customer records that are tagged with geographic locations such as street addresses, or images of the Earth's surface from remote sensing satellites, or information gathered using the Global Positioning System (GPS).

Today GIS is a major application of computing technology, with an annual market for software, data, and services totalling on the order of $10 billion. The general public is likely to encounter GIS through Web-based services such as MapQuest that offer maps and driving directions computed from digital maps. Most municipal governments will use GIS to track, manage, and plan their assets and activities, as will utility and telecommunication companies, resource management agencies, package delivery companies, and departments of transportation. GIS is extensively used in the military, for targeting missile systems, planning battlefield tactics, and gathering intelligence.

GIS representations

Two major forms of representation are used in GIS: *raster* and *vector*. In raster form, an area is represented as an array of rectangular cells, and variation of some phenomenon of interest over the area is expressed through values assigned to the cells. This form is used for remotely sensed images from satellites, and today GIS users have easy access to a wide range of such images, from government sources such as NASA and the US Geological Survey, to commercial sources such as IKONOS and Quickbird. Images of interest to GIS users will have spatial resolutions (cell sizes) ranging down to 1m or less. Raster data is also the preferred format for digital elevation models (DEMs), which represent the Earth's topographic surface through measurements at regular intervals. DEM data is available for most of the US at 30m resolution, and for the world at 1km resolution.

In vector representation, phenomena are represented as collections of points, lines, or areas, with associated attributes. Vector representation is widely used to disseminate data from the Census, for example, providing summary statistics for states, counties, cities, or census tracts, and representing each reporting zone as an area. Lines and areas are most often represented as sequences of straight-line segments connecting points, and as such are termed *polylines* and *polygons* respectively. Vector representation is also used for the *street centerline* databases that describe the locations of streets, roads, and highways, and are widely used to support wayfinding.

A GIS database makes use of both raster and vector formats, and typically will contain several distinct *layers*, or representations of different phenomena over the same geographic area. For example, layers might include representations of maps of topography, soils, roads, rivers and lakes, and bedrock geology. By including all of these layers in a single database, it is possible to use the GIS to explore relationships and correlations, for example between soils and bedrock geology, and to combine layers into measures of suitability for various types of land use, or vulnerability to pollution.

A vector GIS is also capable of representing relationships between objects, for example between points representing incidents of crime, and the neighborhoods in which the crimes occurred. Relationships allow places of work to be linked to workers' home locations, or connections to be made between bus routes. Because relationships are in general unaffected by stretching or distortion of the geographic space, they are generally termed *topological* data, to distinguish them from geometric data about object positions and shapes.

GIS functions

The most important parts of a GIS are those that support its basic functions, allowing users to compile, edit, store, and display the various forms of geographic information. Defining location on the Earth's surface can be a complex task, given the large number of alternative coordinate systems and projections available to mapmakers. A GIS thus needs the ability not only to convert between raster and vector representations, but also to overcome differences between latitude-longitude and other coordinate systems; differences between map projections such as the Mercator and the Lambert Conformal Conic; and differences between the mathematical figures used to approximate the shape of the Earth. In the US, for example, geographic data may use either of two such mathematical figures: the North American Datum of 1927, based on the Clarke ellipsoid of 1886; or the newer North American Datum of 1983, based on a unified global geodetic system. Many distinct coordinate systems are in use, ranging from the high-accuracy State-Plane Coordinate systems defined by each state, to the lower-accuracy Universal Transverse Mercator system originally devised for military applications.

Once the foundation of such housekeeping operations has been built, a GIS developer can quickly add a vast array of functions and capabilities. These might include sophisticated algorithms for designing and printing hard-copy maps; algorithms to identify optimum routes for vehicles through street networks, or optimum locations for new retail outlets; methods for computing correlations between data in different layers, or for combining layers into measures of suitability; and methods for evaluating potential land use decisions. All of these methods are termed *spatial analysis,* and when applied to extremely large data sets in an exploratory mode they are termed *data mining*.

The list of supported forms of spatial analysis is huge, and an industrial-strength GIS will offer literally thousands. In addition, there is an active market in *extensions* to the basic GIS products, offered by third parties as scripts, or as applications written in such languages as Visual Basic and designed to be compatible with a specific vendor's base product.

GIS as human-computer interaction

The original motivation for the development of GIS in the 1960s came from the need to automate certain basic operations. One was editing, which is very difficult and time-consuming if performed by hand, given the technical issues involved in moving or deleting hand-drawn lines on paper maps. The other was the measurement of the area of arbitrarily shaped zones on maps, as required for example in the task of inventorying land use, or the planning of new subdivisions. This view of GIS as a personal assistant, performing tasks on geographic data that the user found too tedious, expensive, inaccurate, or time-consuming to perform by hand, drove almost all of the first thirty years of GIS development. As such, it required modes of human-computer interaction (HCI) that were suited to the task, providing a comparatively skilled user with easy access to information and the results of analysis.

More recently, however, a number of other requirements have come to dominate developments in HCI. A GIS that is used in a vehicle to provide the driver with wayfinding instructions must address the issues that arise when distracting visual displays conflict with the driving task. Many of the in-vehicle navigation systems that are now being installed in cars, either as original equipment or as optional accessories, provide for auditory instructions as well as visual output, and may include voice-recognition functions for input as well. HCI issues also arise in the use of GIS by people with visual impairment, and there have been several interesting developments along these lines in the past ten years.

Advanced GIS use requires a high level of skill and training on the part of its user. Courses in GIS often include advanced work in map projections and datums, and in spatial statistics. Thus another set of HCI issues arise in GIS applications that are designed for use by children, or by other groups whose knowledge of advanced GIS concepts is limited. Microsoft's Encarta, for example, offers a number of functions associated with geographic information, including simple map-making, and retrieval of information using maps as organizing frameworks. A child cannot be expected to understand map projections, so it is common for such systems to display information on a three-dimensional globe, rather than on a flattened or projected Earth. A child cannot be expected to understand the cartographer's concept of scale, or representative fraction, so such systems resort to clever metaphors as the basis for specifying level of detail, for example by allowing the user to "raise" or "lower" the viewpoint relative to the Earth's surface, revealing less and more detail respectively.

Virtual reality and uncertainty

A GIS contains a representation of selected aspects of the Earth's surface, combining raster and vector formats to achieve a representation using the binary alphabet of digital systems. When used at the office desk, as is typically of the vast majority of GIS applications, the representation in effect replaces the real world, limiting its user's perception of reality to the information contained in the database. The real geographic world is infinitely complex, revealing more detail the closer one looks apparently *ad infinitum*, so it follows that any database representation must be at best a generalization, abstraction, approximation, or sampling of the real thing that it purports to represent. Only in certain very limited circumstances, such as the representation of objects that are truly mathematical, including the straight lines of land surveys, is it possible to achieve close to perfect representation.

This fundamental principle of GIS representations has led to great interest in the topic of *uncertainty*, which can be defined as the difference between what the database tells the user about the real world, and what the real world would reveal to the user if visited directly. In some cases uncertainties can be resolved from the user's own memory, particularly if the user has visited the area that is represented in the database. GIS use is thus always most successful if combined with personal knowledge.

There are many sources of uncertainty, including measurement error (it is impossible to measure location on the Earth's surface exactly), generalization (the omission of local detail from representations in the interests of simplicity or limiting data volume), vagueness in the definitions of key terms (*e.g.*, soil class or land cover class definitions), and confusion on the part of the user about what the data are intended to represent. This last type of uncertainty commonly arises when data are shared, and the user misunderstands the intent of the creator of the data, perhaps because of poor documentation.

Uncertainty has been studied within several theoretical frameworks, including *geostatistics*, *spatial statistics*, and *fuzzy set theory*. Each has its benefits, and each is suited to particular settings. Statistical approaches are most appropriate when uncertainty arises because of measurement error, or when it can be characterized using probabilistic models. Fuzzy sets, on the other hand, appear to be more appropriate when dealing with imperfect definitions, or when experts are uncomfortable making precise classifications of geographic phenomena.

Augmented reality and mobile GIS

In recent years the development of wireless networks and miniaturized devices has raised the possibility of a fully mobile GIS, no longer confined to the desktop. Laptop computers now have virtually the same computational power and storage capacity as desktop workstations, and wireless networks can provide bandwidths approaching those available via Ethernet and other local-area networks. Laptops are relatively cumbersome, however, with heavy battery consumption, and personal data assistants (PDAs) offer better mobility with some sacrifice in computational power and storage capacity.

Wearable computers are increasingly practical, packing CPU and storage devices into a cigar-box-sized package to be worn on the belt, and providing visual output through devices clipped to eyeglasses. In summary, then, we are approaching the point where it will be possible to do GIS anywhere, at any time. This clearly has the most significant implications when GIS is done in the subject area, allowing the user to be in direct sensory contact with the phenomena being studied and analyzed.

Mobile GIS is already in use in many applications. Utility company workers, surveyors, and emergency incident managers already routinely have access to GIS capabilities through suitably configured PDAs, although these devices are more likely to be used for data input than for analysis. Information technologies are routinely used by scientific workers in the field to record observations, and GPS transponders are used to track animals to develop models of habitat use.

These uses of mobile GIS are distinct in the sense that the user is in contact both with the database and with the reality represented by the database. The term *augmented reality* (AR) is often used, since sensory reality is being extended through information technology, allowing the user to see things that are for one reason or another beyond the senses. AR can be used to "see" under the surface of the street when digging to install new pipes, allowing the construction crew to avoid accidentally damaging existing facilities. AR can be used to address the inability of visually impaired people to see their surroundings, and exciting developments have occurred recently in the development of systems to aid such personal wayfinding. AR can be used to superimpose historic views on the field of view, creating interesting opportunities in tourism. The ability of a cellphone user to "see" the locations of nearby businesses displayed in map form on the cellphone screen is also a form of AR.

The long-term implications of AR are profound, since they give people the ability to sense aspects of their surroundings that are beyond their senses -- "around the corner", "under the street", in the past, or in the future.

AR raises interesting issues of HCI. The displays provided by laptop computers and PDAs are adversely affected by strong light conditions typical of the outdoors. Displays that clip on eyeglasses offer comparatively high resolution (similar to a PDA), but make it very difficult to implement "point-and-click" interaction. Displays on cellphones are too small for many GIS applications, which tend to require large display areas for suitable resolution (it is difficult, for example, to annotate street maps with names on a cellphone screen). Reference has already been made to the problems of visual display for drivers. Input devices, such as one-handed keyboards, are also difficult to use. Finally, heads-up display, in which information from the GIS is superimposed directly on the field of view, requires head-mounted displays that are much more cumbersome and difficult to use than the more-versatile display clipped to the eyeglasses.

GIS and data sharing

Traditional paper maps are very efficient repositories of geographic information, and as such represent major investments. The typical national topographic map sheet of the US Geological Survey, for example, covering an area approximately 15km on a side at a scale of 1:24,000, costs on the order of $100,000 to create, and must be regularly updated with new information if it is to remain current. It takes some 50,000 such sheets to cover the 48 contiguous states, and if the entire series were to be recreated today the total investment would be in excess of $5 billion. Remote sensing satellite programs require investments in the hundreds of millions; a 1993 study by the US Office of Management and Budget found total annual investment in geographic information by federal agencies to exceed $4 billion. Not surprisingly, then, society has traditionally relied on national governments to make these kinds of investments, through national mapping agencies and national space programs. Only national governments have been able to afford the cost of the complex systems needed to create maps.

Today, this picture is changing radically. Anyone with $100 can purchase a GPS receiver capable of determining location to better than 5m, and can use it to make digital maps of local streets or property boundaries. Mapping software is available for the average PC, with the result that today virtually anyone can be a cartographer, making and publishing maps on the Internet.

Moreover, national governments find it increasingly difficult to justify the kinds of annual expenditures needed to maintain mapping programs. In 1993 the US National Research Council began to advocate the concept of a National Spatial Data Infrastructure (NSDI), a set of institutional arrangements and standards that would coordinate a new form of decentralized production of geographic information. NSDI is intended to support a *patchwork* approach, replacing uniform, government-produced series of maps with coverage at varying scales produced as appropriate by local, state, or federal agencies. NSDI provides the format standards to ensure sufficient uniformity in these data sets, and the *metadata* standards to allow parts of the patchwork to be described effectively.

Since 1993 the US has made enormous investments in information-technology infrastructure for data sharing. In accordance with US law, the vast majority of geographic information produced by the federal government is in the public domain, free of copyright restrictions, and available for no more than the cost of reproduction. Today the amount of such data available from Web sites is on the order of petabytes, and growing rapidly. This free resource has in turn stimulated the development of a wide range of applications, and an industry dedicated to adding value to data by making it easier to use, more current, or more accurate.

The term *geolibrary* has been coined to describe the Web sites that provide geographic information. By definition a geolibrary is a library that can be searched based on geographic location – that is capable of answering queries of the form "What have you got about *there*?" The National Research Council has explored the concept and status of geolibraries in one of a series of reports on NSDI.

Geolibraries present interesting issues of user interface design. All allow users to display a world map, and to zoom in to an area of interest, refining the search criteria with additional requirements. But the area of interest for many users is defined not by a location on a map, or by coordinates, but by a placename; and many users will not be able to locate that placename on a map. This issue is solved through the use of a *gazetteer*, an index that converts placenames to coordinates. But most gazetteer entries provide only a point reference, which is problematic for extended features of complex shape, such as the Missouri River.

The US National Geospatial Data Clearinghouse is an example of a geolibrary that allows search and retrieval across a distributed archive – in effect allowing its users to visit and search several libraries simultaneously, and with minimal effort. Such capabilities are made possible by metadata, the information that describes the contents of data sets in standard form, allowing vast catalogs of metadata to be searched quickly and easily. The dominant metadata standard for geographic information is the Federal Geographic Data Committee's Content Standard for Digital Geospatial Metadata. It provides some hundreds of potential fields for the description of the contents, lineage, quality, and production details of a data set.

Web-based GIS

Early efforts to build geolibraries, beginning in the mid 1990s, focused on the need to distribute data sets, by analogy to the traditional library whose responsibility ends when the book is placed in the hands of the reader. Under this model each user was required to maintain a full GIS capability, since all transformation and analysis occurred at the client. Since the advent of the Web and the widespread popularity of Web browsers, more and more services have been developed by servers, allowing the use of thinner and thinner clients. Today, a user of a standard Web browser can access services for many basic GIS operations. The task of geocoding, for example, which consists of converting street mailing addresses to coordinates, is now available from a number of sites, including MapQuest. Similarly it is possible to access remote services for converting placenames to coordinates, and it is expected that more and more GIS services will be available in this form in the coming years.

To be automatic and transparent, such Web services require adherence to standards, and many such standards have been developed in recent years by such organizations as the Open GIS Consortium. They allow a user's client GIS to request data from a geolibrary, or a geocoding service from a provider, taking care of such issues as coordinate system transformation, or clipping of data to match a user's study area. A vast range of mapping, geolibrary, and other services are now available and fully interoperable with popular GIS software. For example, a user of ESRI's ArcGIS might determine that a needed data set is not available on the desktop computer's hard drive, and might search ESRI's Geography Network Web site for suitable data. The search would be initiated over a distributed archive of registered contributors, and would return suitable "hits". The user would then be able to use a chosen data set – but rather than copying it to the client, the data set would be accessed transparently over the Internet.

Most GIS vendors now offer software to support the development of GIS services and Web-based mapping. Some services are available free, and others on a subscription basis. But it remains to be seen whether the provision of services is capable of providing sufficient cash flow to a company, and whether Web-based GIS is a viable long-term commercial proposition.

The IT mainstream

The history of GIS has been one of specialized application of information technology. The development of the first GIS in the 1960s required many original developments and inventions, including the first map scanner, the first topological data structure, and the first algorithm for map overlay. Today, however, the majority of the software in a GIS is industry-standard, implementing mainstream solutions for operating systems, application development, object-oriented database design, and graphic interfaces. Undoubtedly GIS has become closer over the years to the IT mainstream, and today it is common for records in large database solutions to be tagged with geographic location. For example, the locations of credit-card transactions are routinely tagged with location in space and time to support mining for evidence of misuse, through the detection of anomalous behavior. Mainstream solutions are attractive to software developers, of course, because they allow massive economies of scale through the adoption of standard technologies that can serve many disparate applications.

On the other hand it is clear that GIS applications will always be to some degree distinct from the mainstream. Cellphone mapping applications, for example, push the limits of available screen area and resolution. GIS database applications raise difficulties when the phenomena to be represented are fundamentally continuous rather than discrete: for example, roads and rivers are continuous features, not easily broken into the discrete chunks of database records. Topography is a continuous surface, and not easily broken into squares or triangles for discrete representation. In all of these cases the need to discretize causes downstream issues for applications.

Moreover, GIS is about the representation of an infinitely complex real world, and its effective use will always require an understanding of the nature of that world, and the consequences of the inevitable generalization, approximation, and sampling that occurs in digital representation.

Michael F. Goodchild

FURTHER READING

Chrisman, N. R. (1997). *Exploring geographic information systems*. New York: Wiley.

Clarke, K. C. (1999). *Getting started with geographic information systems*. 2nd Edition. Upper Saddle River, NJ: Prentice Hall.

DeMers, M. N. (2000). *Fundamentals of geographic information systems*. 2nd Edition. New York: Wiley.

Duckham, M., Goodchild, M. F., & Worboys, M. F. (2003). *Fundamentals of geographic information science*. New York: Taylor and Francis.

Hearnshaw, H. M., & Unwin, D. J., editors (1994). *Visualization in geographical information systems*. New York: Wiley.

Kennedy, M. (1996). *The Global Positioning System and GIS: An introduction*. Chelsea, MI: Ann Arbor Press.

Leick, A. (1995). *GPS satellite surveying*. New York: Wiley.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W., editors (1999). *Geographical information systems: Principles, techniques, management and applications*. New York: Wiley.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2001). *Geographic information systems and science*. New York: Wiley.

MacEachren, A. M. (1995). *How maps work: Representation, visualization, and design*. New York: Guilford Press.

Medyckyj-Scott, D., & Hearnshaw, H. M., editors (1993). *Human factors in geographical information systems*. London: Belhaven Press.

National Research Council (1993). *Toward a coordinated spatial data infrastructure for the nation*. Washington, DC: National Academy Press.

National Research Council (1999). *Distributed geolibraries: Spatial information resources*. Washington, DC: National Academy Press.

O'Sullivan, D., & Unwin, D. J. (2002). *Geographic information analysis*. Hoboken, NJ: Wiley.

Peng, Z. R., & Tsou, M. H. (2003). *Internet GIS: Distributed geographic information services for the Internet and wireless networks*. Hoboken, NJ: Wiley.

Peuquet, D. J. (2002). *Representations of space and time*. New York: Guilford.

Raper, J. (2000). *Multidimensional geographic information science*. New York: Taylor and Francis.

Snyder, J. P. (1997). *Flattening the Earth: Two thousand years of map projections*. Chicago: University of Chicago Press.

Worboys, M. F. (1995). *GIS: A computing perspective*. New York: Taylor and Francis.

Zhang, J. X., & Goodchild, M. F. (2002). *Uncertainty in geographical information*. New York: Taylor and Francis.