CARTOGRAPHY: GISCIENCE AND SYSTEMS

Michael F. Goodchild, University of California, Santa Barbara

Michael F. Goodchild
Department of Geography
University of California
Santa Barbara, CA 93106-4060
USA

CARTOGRAPHY: GISCIENCES AND SYSTEMS

Keywords: layer, raster, vector, attribute table, topology, object orientation, overlay, join, buffer, map algebra, projection, geodesy, datum, geoportal, metadata

Glossary

attribute table: a table with one row allocated to each feature and one column allocated to each defining characteristic
buffer: dilation of point, line, or area features by a defined distance
datum: a mathematical reference model for global coordinates
geodesy: the science of precise Earth measurement
geoportal: a single point of entry to a distributed collection of data sets
hybrid model: a data model in which feature attributes and topological relationships are stored in a relational database, while point coordinates are stored elsewhere
join: combination of attribute tables based on common keys
map algebra: a language for manipulation of raster layers using symbolic representation
metadata: description of the properties, format, quality, and lineage of a data set
object orientation: an approach to data modeling that recognizes, among other principles, that classes of features can inherit properties from more general classes
overlay: calculation of a new topologically rich layer from two or more input layers
polygon: an area represented as a sequence of points connected by straight lines
polyline: a line represented as a sequence of points connected by straight lines
projection: a mathematical transformation of a curved surface to a plane
raster: representation of space as a collection of equal, rectangular areas ordered in a defined sequence
relational database: a data model in which all information is represented in tables linked by common keys
topology: in mathematics, properties that are invariant under distortion of the containing space; in GIS, relationships between features such as adjacency and connectivity
vector: representation of space by identification of features contained in the space, and use of coordinates to specify locations of features

Synopsis

Geographic information systems (GIS) support virtually any operation on geographic information: acquisition, editing, manipulation, analysis, modeling, visualization, publication, and storage. Geographic information science (GIScience) addresses the fundamental questions posed by GIS, and constitutes the body of knowledge exploited by GIS, and the research that will enable the next generation of GIS. The history of GIS development has been dominated by a series of applications, and today virtually any activity concerned with the Earth’s surface or near-surface has made use of GIS. GIS views the geographic domain through a particular lens, partitioning variation into a set of layers, and organizing space either as an array of cells or as a collection of point, line, or area features. Concepts of object orientation have been adopted within the past decade, allowing GIS to move significantly beyond the metaphor of the map and to include representations of phenomena that are complex and dynamic. Much effort has been expended in the past decade on exploiting the power of the Internet to support sharing of

geographic information, and the geoportal represents the most recent stage in those developments. Various efforts have been made to identify the research agenda of GIScience, and a consensus has emerged as to its most important elements.

Introduction

Geographic information can be defined as information linking places on or near the Earth's surface to the properties present at those places. In essence, any item of geographic information must possess both a specification of location, in the form of a placename, geographic coordinates, street address, or some other suitable format; and a specification of such properties as atmospheric temperature, average income of residents, or elevation above mean sea level. Thus both an informal statement such as "it is cold today in Santa Barbara" and a more formal one such as "at 12 noon local time on January 25 at latitude 34 degrees 26 minutes 41 seconds North, longitude 119 degrees 48 minutes 26 seconds West the air temperature in a Stevenson screen was 12.6 Celsius" are examples of statements meeting this definition.

Traditionally humans have handled such statements in a variety of ways: for example, as elements of the contents of maps, as observations recorded on paper, or as statements printed in text or communicated through speech. Over the decades many conventions have been adopted to codify practice, as for example in the standardization of longitude following the International Meridian Conference of 1884, or the adoption of the World Geodetic System of 1984. Increasingly, however, and in line with trends in society generally, geographic information has been converted to digital form, and has become just one of a myriad of different forms of data that pass along the Internet and are stored on the hard drives and CDs of contemporary digital technology.

The term geographic information system (GIS) was coined in the mid 1960s to describe a computer system for handling geographic information in digital form. Motivations for the adoption of GIS take many forms. One of the earliest was the need to analyze, to produce summaries of the amounts of area devoted to different uses on the Earth's surface. Digital processing appeared to offer enormous advantages in the form of precise computation, replacing manual processes that were frequently tedious, labor-intensive, and inaccurate. Another motivation arose in statistical agencies such as the Bureau of the Census, where digital processing had the potential to automate various operations, from the management of door-to-door interviews to the aggregation of statistics for multiple reporting zones. Still another arose in map-producing agencies, since digital processing offered the possibility of easy editing, rapid calculation of map projection equations, and automated plotting. By the late 1970s these applications had converged, as it became clear that the commonalities were sufficient to support a distinct type of computer application. A substantial software industry emerged, major efforts were made to standardize terms and methods, and education programs evolved in many universities and colleges.

The term geographic information science (GIScience) was coined in 1992. As with many scientific tools, including the telescope and the microscope, the development and adoption of GIS provoked a series of fundamental questions: what theory of the geographic world did GIS implement; could the traditional processes of manual cartography be formalized in computational systems; how could one measure the uncertainties inherent in any representation of the world? GIScience is variously defined as the science behind the systems; the set of fundamental questions raised by GIS; the research field that will define the next generation of GIS; the body of knowledge that GIS implements; and the use of GIS as a tool for scientific research. As a science of geographic information it represents a subset of information science; while its relationship to the discipline of geography, and to other disciplines that deal with the surface and near-surface of the Earth, is a subject of continuing debate.

Applications of GIS

A vast number of applications have been identified since GIS was first defined. Virtually all human activities deal in some sense with the surface of the Earth, and “everything that happens, happens somewhere”. Yet rates of adoption vary dramatically, and are determined by a number of factors.

The first set of adoptions occurred in activities that have relied strongly on the use of maps. Many of the first sales of GIS software occurred in the early 1980s to agencies concerned with the management of natural resources, notably the forest industry and its regulatory agencies. By the late 1980s virtually all such agencies had adopted GIS, making investments that were massive even by today’s standards. Map-making agencies have already been mentioned, and the advantages of computerized map production are so dramatic that virtually all such agencies had adopted some form of GIS by the late 1980s.

Another strong group of early adopters were corporations and agencies with large quantities of geographically distributed assets, who required GIS to maintain inventories, manage assets, and schedule maintenance. The utility companies and agencies – water, gas, electric, telecom – found the case for GIS compelling, and by the 1990s had become one of the mainstays of the GIS industry.

It has been argued that the academic GIS community has consciously or subconsciously ignored its military and intelligence roots. Many of the original technical developments were made in this domain, in remote sensing, missile guidance, positioning, and many other areas. Today digital geographic information technologies are indispensable to military operations and intelligence gathering.

Among academic disciplines, geography was understandably an early adopter of GIS. By the early 1990s any discipline concerned with physical or biological dimensions of the Earth’s surface had adopted GIS as a standard tool, but within the social sciences adoption was rather slower, due in part to the debates that still rage in these disciplines over the appropriate roles of scientific method, data, and experimentation. Adoption has been strong in the more empirically based disciplines, such as criminology and archaeology, and rather slower in the more theoretically and conceptually oriented disciplines of economics or sociology.

The 21st century has brought a new awareness of the vulnerability of modern society, and several major disasters have shown the importance of geographic information technologies in all aspects of planning, response, and recovery. GIS played a significant role in the recovery effort following the attacks of Sept 11 2001. After Hurricane Katrina it was painfully evident that anyone with a broadband connection anywhere in the world could view the events as they unfolded and at high resolution – whereas those close to the disaster had no access to the necessary power, computers, and Internet connections.

Concepts of GIS

The layer

The layer is perhaps the most conspicuous and iconic concept of GIS, having its roots very early in GIS history and capturing one of its strongest motivations. It proposes that the geographic world can be represented as a series of thematic layers, each carrying information relevant to one particular thematic domain. To mapping agencies, each layer might correspond to the information portrayed in one color of ink on a topographic map: contours of topographic elevation in brown, urban areas in pink, or wooded areas in green. To Ian McHarg, a landscape architect seeking to develop a new model for a department at the University of Pennsylvania in the late 1950s, each layer represented the perspective of one discipline; the geologic layer captured the factors believed by geologists to be important in planning; the ecological layer captured factors originating in concern for biological conservation; and various social layers captured factors related to the economy and human populations. By stacking the layers, one could combine factors in various ways. A power line, for example, might require one distinct set of weights to be applied to each discipline's layer; while a new shopping center might require a different set of weights.

This notion of partitioning geographic variation into a number of thematic layers underlies much traditional practice in cartography, where each layer might be the subject of a different printed map, or might be depicted on a single map in a distinctive color. From another perspective, it captures the ability of GIS to relate seemingly unrelated information through common geographic location. Only in a GIS, it is argued, can one combine detailed information on the ethnicity of a city's neighborhoods with detailed information on patterns of atmospheric pollution, to address the central question of environmental equity: do minorities bear a disproportionate impact from industrial pollution? Only in a GIS can one combine a map showing the distribution of schools with one showing the distribution of liquor outlets. While all of these tasks could in principle be completed manually, the effort needed to redraft onto transparent media and to correct for differences of scale and map projection is often prohibitive.

Raster and vector

Digitization is a form of coding, and in the binary representations of computers it implements a coding scheme capable of reducing the phenomena of a domain to a linear sequence of two values: 0 and 1, or "on" and "off". Most domains have by now developed standard coding schemes to do this with their particular phenomena. In music, for example, the dominant standards are MP3 and MIDI. Digital representation of geographic information similarly requires a suite of coding schemes, and in essence much of the development effort in GIS over the past few decades can be summed up as “finding the equivalent of MP3 for maps”.

Two classes of coding schemes emerged very quickly, based in part on discoveries already made in computer graphics. Raster schemes represent the map as a regular array of square or rectangular picture elements of fixed size, and record the contents of each element as a simple value in some binary code, proceeding in a standard sequence that is often row by row from the top left (northwest corner). Suppose, for example, that the first four cells of the top row contain the values 5, 5, 4, 4, denoting the presence of land cover categories 5 and 4. Then a simple coding scheme would allocate 4 bits to each cell, and record the binary equivalent of each cell’s integer value as the sequence 0101010101000100. All detail smaller than the cell size is necessarily lost in this approach, unless special rules are adopted to recognize the presence of very small features in each cell’s coded value.

Vector schemes, on the other hand, first identify each feature on the map. Features might include patches of homogeneous land cover, roads and rivers, contours, or point-like phenomena such as oil wells. Each type of feature is first categorized as point-like, line-like, or area-like, and the location of each feature is then recorded in a suitable coordinate system such as latitude/longitude. Points will be given a simple pair of coordinates; lines will be represented as sequences of points connected by straight-line segments (termed polylines), and areas will be represented as similar sequences connected in loops (polygons).

In the early days of GIS development each system was predominantly one or the other, and raster GIS and vector GIS competed for a while. Before long, however, every major GIS package developed tools for handling both, and today one expects to be able to input, store, manage, and analyze data in either form, transforming from one to the other as necessary. Figure 1 shows an example in which both raster and vector data are displayed.

[Figure 1 about here]

There are many dimensions to the choice between raster and vector, some of them more subtle than others. Simple epithets such as “raster is vaster but vector is correcter” are superficially true but largely unhelpful on closer examination. In reality the following issues are important in deciding which to use in a particular situation:

- Volume. Representing an area such as a county in vector requires the coding of a number of coordinates, and this can be done to a high precision with the use of many numerical digits. To achieve the same precision in raster would require the identification of each cell lying inside the area, and matching the precision of vector coordinates could require a vast number of cells. However many elegant methods of data compression can be applied to raster data. Moreover allocating many numerical digits to a coordinate will increase precision but will not necessarily improve accuracy, and in practice the positional accuracy of vector coordinates is often disappointingly low.

- Data source. Imagery collected from Earth-observing satellites is an increasingly important source of data for GIS. Such data invariably arrive in raster form, with a cell size determined by the sensor's designer. Raster formats also dominate data on topographic elevation.
- Processing. Although many operations can be carried out on either raster or vector representations, some are either faster or only possible on one. For example, the calculation of a viewshed, or the area visible from a point on a topographic surface, is exclusively raster, while routing tasks such as determining the shortest path through a network are exclusively vector. Although most software packages are able to handle both raster and vector, some are still primarily identified with one or the other.
- Decision context. GIS analysis is often used to support decisions of a spatial nature, such as the development of land use plans. Implementation of any plan is likely to involve existing spatial units of some kind, such as land parcels or agricultural fields. It makes sense, then, to tie the choice of raster or vector to the geometry of the decision units – for example, a plan for parcels of land would likely be developed using vector representations in which each parcel appears as a polygon.

Attribute table

In the vector world, the characteristics of points, lines, and areas are termed attributes, and represented in the form of tables. Each table stores the attributes of one class of features, all of which will have the same topological dimension (all points, all lines, or all areas), and all of which will have the same set of defined attributes. For example, a class of land parcels might have numerous associated attributes: owner, area, assessed value, identification number, etc. Each feature occupies one row of the table, and each attribute occupies one column.

Entries in attribute tables may take many forms, and the number of forms has increased as the technology for handling tables has evolved. Entries may be numeric, either whole numbers (integers) or decimal numbers (reals), or alphanumeric, including names, text descriptions, or dates. They may also include images, such as photographs of a site, or sounds, such as interview responses. Finally, entries may be keys pointing to entries in other tables – for example, a point feature denoting the home location of a patient might point to a record in another table describing the attributes of the hospital at which the patient was treated. Such keys are a central concept of the relational model, which was introduced to GIS in the 1970s and dominated database thinking until recently.

Topology

In mathematics, a property is said to be topological if it is invariant under stretching and distortion of the containing space. If a loop is drawn on a rubber sheet, for example, it is impossible to break open the loop by stretching the rubber, and thus the distinction between area, line, and point is a topological distinction. Topological properties are important in GIS, since they include such useful spatial properties as adjacency and connectivity, properties that are important in many forms of analysis. For example, a system to plan optimum routes through a road network must include representation of the connections between network links, and important properties of such connections including turn restrictions or grade separations at intersections.

The need to represent topological properties emerged very early in the history of GIS, and by the 1970s was regarded as essential to success, at least in vector systems. Figure 1 illustrates a type of map – often termed an area-class map – that is very common in land management applications. If each area on the map were digitized separately as a polygon, then each internal boundary would be represented twice, and it might be difficult to ensure that the two versions were identical. Moreover, it would be possible to edit each polygon independently of its neighbors, possibly producing gaps and overlaps that in principle cannot occur in such maps. These problems can be overcome by the simple expedient of treating the common boundary between two areas as the basic record. In the Canada Geographic Information System (CGIS), widely regarded as the first GIS, area-class maps were stored in a linear sequence on magnetic tape, each common boundary being recorded as a coded representation of a) the attributes of the area on its left, b) the attributes of the area on its right, c) the number of points needed to capture the boundary's geometry as a polyline, and d) a sequence of coordinate pairs. CGIS had no table of polygon attributes, but obtained all of its results, including measurement of polygon areas, by processing this linear sequence of boundaries.

[Figure 1 about here]

By 1980 a consensus had emerged among GIS developers on the representation of area-class maps in topologically rich form. Each internal boundary between two areas would be termed an arc, and its attributes would include pointers to the areas on its left and right, and to the boundary network junctions or nodes at its start and end. Each area would have an attribute table defining its properties, and perhaps pointers to the arcs forming its boundary. The points defining the polyline representation of each arc would be stored separately, and because the number of points per arc was variable the file containing these points would not have the normal tabular structure – for this reason this approach was termed hybrid. Simple extensions to this basic model allowed for the storage of street networks, administrative boundary maps, and many other forms of geographic information, and it survives today in the form of ESRI's coverage model.

Object orientation

By the 1990s, dramatic advances in computing power, storage capacity, and software had made many of these arguments for topology superfluous. Moreover the concept of the common boundary as a fundamental component of geographic information seemed to conflict with the intuitive ways in which humans view the world and its representations in maps. Object orientation emerged in the late 1990s as a new, more powerful and at the same time simpler way of representing the world that treated topological relationships as optional extras rather than as essential properties.

ESRI’s shapefile model provides an appropriate example. In this approach, introduced in the late 1980s, each feature is stored as a point, polyline, or polygon. Where boundaries are shared, they occur in the database twice, and powerful routines exist to compare and correct differences between them. Similarly topological properties such as connectivity can be computed by analyzing the geometry of features, and automatically ignoring small gaps or overshoots when these fall below established distance thresholds. Features can be linked during edit to ensure that moving one feature also moves any other features tied geometrically to it.

The set of relationships allowed in object-oriented databases is much richer than in the earlier relational approach. Links between features, such as between a patient and a hospital, are termed associations. The concepts of aggregation and composition allow complex features to be formed as collections of other, simpler features, matching a common property of geographic information. For example, an airport, which might be depicted at a coarse scale as a point, might be linked to its constituent runway, hangar, terminal, and other features, each of which might be depicted in fine-scale visualizations. A state might be modeled as an aggregation of counties, and a county as an aggregation of cities. None of this was possible in the earlier approach.

Finally, object orientation adds the concept of inheritance. Many classes of features can be regarded as specialized versions of other, more general classes. For example, a county is a specialized form of polygon in GIS representation; a bicycle path is a specialized form of transportation link, which is itself a specialized form of polyline. Specialized classes inherit all of the properties of more general classes, and add new properties of their own. For example, all polylines have the properties of length and number of defining points; all transportation links have these properties, plus such other properties as number of lanes, travel speed, and degree of congestion; and bicycle paths have these properties plus others more relevant to cyclists, such as whether the path is on or off road.

The adoption of object-oriented concepts is a major step in the evolution of GIS. While previous approaches were intimately tied to the map as the source of geographic information, object orientation has been applied to the successful representation of phenomena that have never been associated with maps, including events, transactions, flows, and interactions. In this sense, then, GIS has been able to move beyond the map metaphor in recent years. Nevertheless, the metaphor remains a powerful way of describing GIS, as “a computer containing maps”. But today one might also describe a GIS without using the map metaphor, as “a computer containing descriptions of features on the surface of the Earth, in which the geographic location of each feature is identified as part of its record”.

Figure 2 shows the UNETRANS model, an application of object-oriented principles to the representation of features relevant to transportation applications of GIS. It supports representation of the basic transportation network; of the physical assets associated with the network, such as traffic lights and signage; of features important to navigation on the network, such as connectivity and turn restrictions; of vehicles on the network; and of the routes and schedules of public transit.

[Figure 2 about here]

Overlays and joins

The set of possible topological relationships between features is very rich, and while topology may no longer be as important in the representation of geographic data, it plays a very important role in the kinds of analysis needed to support GIS operations. In network analysis, for example, the connections that exist between network links provide the essential basis for planning routes, determining least-cost and least-time paths, and many more sophisticated forms of analysis that fall under the broad headings of logistics and location analysis.

Some of the most notable advances in GIS research have concerned the enumeration of possible relationships between features, and the identification of these relationships is a basic function of all GIS software. Suppose that two layers exist for the same area, one depicting current land use as a collection of non-overlapping, space-exhausting area features, each showing an area of uniform land use; and a second layer showing a similar depiction of land capability for agriculture. Suppose further that the analyst wishes to determine the area of land that is both a) currently not used for agriculture and b) of high potential for agriculture. Intuitively one would think of overlaying the two maps, assuming that one was transparent, and measuring the appropriate areas of overlap (Figure 4). In vector GIS terms, the task is known as the polygon overlay problem, and results in a new layer in which each polygon is homogeneous with respect to its characteristics on both of the input layers. Operationally this requires the calculation of every intersection between the two sets of polygon boundaries, plus the assembly of new and generally shorter boundary segments into new polygons. In practice the task is often plagued by the existence of large numbers of slivers, which result when boundaries which should be identical on both maps in fact differ by small amounts. In raster the problem is formulated a little differently, since the input attributes must be combined in some way for each pixel, and the spurious polygon problem does not apply because of the raster's finite resolution.

[Figure 4 about here]

Overlays can be defined for any combination of point, line, or area features. The point in polygon problem arises when a layer of points is overlaid on a layer of polygons, in order to determine the containing polygon of each point. It is frequently applied in the analysis of point data to examine relationships between the density, count, or attributes of points and the attributes of containing areas. For example, one might use it to analyze the relationship between the number of instances of a disease in an area and the socioeconomic attributes of the area, or the area's average levels of atmospheric pollution.

A join occurs when the results of overlay are used to combine the attributes of the relevant features. For example, a point in polygon overlay might be used to add the attributes of each containing polygon to those of each point. In this way the name of the containing county might be added to each record in a set of points denoting accidents; or the name of the nearest street might be added to each record in a set of points representing tourist attractions and other landmarks.

Buffers

Another commonly employed function in GIS is the computation of a buffer. Any feature can be dilated by a specified distance to create a new feature known as a buffer. For example, a width of 100m on either side of a stream might be designated as off limits for logging, or requiring special kinds of agricultural practice, to limit erosion and contamination. In some jurisdictions it is common for buffers to be used in notifying residents of proposed new developments, or to ban certain kinds of activity within a specified distance of schools (Figure 5). In vector GIS, buffering points results in polygonal approximations to circles; buffering lines results in polygonal bands on either side of each line and around its ends; and buffering areas results in new dilated polygonal areas. Buffer widths can be varied based on the attributes of individual source features.

The raster form of the buffer operation is notably different. In its simplest form, a layer is first defined as the source, and selected cells of that layer are labeled appropriately. The buffer operation then results in a new layer in which cells are labeled if they fall within the prescribed distance of a labeled cell on the source layer. The task can be generalized so that every output cell receives a measure of distance from the nearest cell on the source layer, rather than a binary indicator of whether distance is below a specified value. It can be further generalized if a layer of travel speed or travel cost or friction is available, in which case the buffers can be made to spread at variable rates, an operation that is virtually impossible in vector GIS.

[Figure 5 about here]

Map algebra

Provided the various layers in a raster GIS are precisely co-registered, the operations of a raster GIS are in many ways simpler to organize and execute than their vector equivalents. Some of the most successful early GIS, particularly GIS designed to run on small computers, were based on raster representations, and much greater progress has been made in systematizing and standardizing raster operations. The most successful of these is map algebra, introduced in the late 1980s. Suppose that each layer in a raster GIS were addressed symbolically, just as scalars, vectors, and matrices are addressed in mathematics or in programming languages. Then the instruction “C = A + B” might be interpreted to mean: “take the values stored in each pixel of Layer A, add them cell by cell to the corresponding values in Layer B, and store the result in a new layer C.” The need to loop through every cell of the raster is implicit.

In map algebra virtually all raster GIS functions are arranged in four categories. Functions that compare layers cell by cell, such as the previous example, are termed local. Focal functions compare or convolve a cell with its neighbors, and are commonly used to smooth rasters by averaging, or to predict the direction of water flow on a raster of elevations. A zone is defined as a contiguous patch of cells of the same value, and zonal operations address such properties as area, for example by creating a new layer in which each value is the number of cells in the containing zone. Finally, global operations are performed on an entire raster layer, for example by computing the mean value of the layer or other summary statistics.

Projections and geodesy

Any analysis of the Earth's surface is inevitably complicated by the fact that the surface is not flat. Curvature matters least in the analysis of small areas, but grows in importance as the extent of the analysis approaches global dimensions. While one might argue that all computer-based analysis should respect the reality of curvature, in practice there are several strong arguments for projection, in other words for flattening the surface prior to its representation and analysis, despite the inevitable distortions that ensue. First, the paper medium is flat, and much human activity still revolves around the paper map, despite the popularity of digital media. Second, flattening is essential if one is to be able to visualize the entire Earth's surface at once. Third, it is impossible to rasterize a curved surface, so any raster must involve some form of projection. Finally, many of the operations of GIS were devised for the analysis of small areas and assume a flat surface.

To project the Earth's surface it is first necessary to find a mathematical form that approximates what is in reality a very complicated shape. Many such mathematical forms are in use in various parts of the world, though they are increasingly being replaced by a single universal form known as the World Geodetic System of 1984 (WGS84); though the latter does not provide such a good fit to any specific part of the Earth, its universality provides certain advantages. Nevertheless GIS users must confront the fact that in most areas of the Earth at least two such mathematical forms or datums are in common use; in North America WGS84 competes with the earlier North American Datum of 1927 (NAD27).

Several map projections based on these mathematical forms are commonly encountered in GIS. High-precision projections such as the Universal Transverse Mercator (UTM) divide the surface into a number of zones to minimize distortion, ensuring that distances will be distorted by only a few parts in ten thousand, and that directions and small shapes will be correctly preserved. When estimation of area is important then a projection such as Albers, which preserves area correctly while distorting distances and shapes, is commonly used. Any projection implies a planar coordinate system and a set of equations that transform to and from latitude and longitude. Finally, many GIS applications use the so-called "unprojected" or cylindrical equidistant projection, which simply maps latitude to y and longitude to x. Its exaggeration of longitude relative to latitude increases rapidly at high latitudes, it distorts shapes and areas, but its simplicity is often attractive.

Metadata

Since the advent and popularization of the Internet vast investments have been made in sharing geographic data sets, particularly data sets representing commonly used features such as roads, rivers, and political boundaries. Whenever a data set is acquired from some remote source it is essential that the prospective user understand the data's basic properties, such as its contents, quality, projection, datum, and format. Metadata are defined as data about data, and have become an indispensable part of the entire data sharing enterprise. They can be assembled into catalogs, and used to search for data meeting specific requirements. They also allow the user to determine the suitability of any data set for a specific use, to determine the data set’s format and the meaning of its various components, such as its attributes, and the conditions under which the data were compiled.

The first widely accepted standard for metadata was promulgated by the US Federal Geographic Data Committee in the early 1990s. Since then it has been refined, and adopted after suitable modification by many national and international agencies. Its description of quality is based on five components: positional accuracy, attribute accuracy, logical consistency, completeness, and lineage. Nine other components are described, including details of the projection and datum, the date of validity, and the definitions of attributes.

Geoportals

A geoportal can be defined as a single point of access to geographic data sets located on servers distributed over the Internet. By providing a single point of access, geoportals allow users to search for data that might exist on any one of thousands of servers in data warehouses and digital libraries. A geoportal includes a catalog, which is compiled either manually from information sent by data custodians, or automatically by harvesting robots that scan the Internet much as do the operators of search services such as Google. Thus a geoportal might contain metadata records describing tens of thousands of data sets, all of which are actually held on remote servers. The user is able to search the catalog, identify suitable data sets, and retrieve them from their host servers using appropriate protocols, and after addressing such access issues as licensing or use fees. Geoportals are increasingly important elements of the evolving spatial data infrastructure, or the system of institutions, standards, and protocols by which geographic data are assembled and distributed for use in many parts of the world.

The agenda of GIScience

The first attempts to define a research agenda for GIScience, and to clarify the relationship between GIScience and the technology of GIS, occurred in the late 1980s. In the UK, the Department of the Environment's Committee of Enquiry into the Handling of Geographic Information (the Chorley Committee) saw three specific stimuli: the rapidly falling costs of hardware, which had reduced the cost of entry into GIS and related activities from $500,000 at the beginning of the decade to $10,000 at the end; the advent of COTS (commercial, off-the-shelf) software to perform the basic operations of GIS; and rapid growth in the availability of spatially referenced digital data. In the U.S., the National Science Foundation announced a competition for a National Center for Geographic Information and Analysis (NCGIA), to advance the theory and methods of GIS, to promote the use of GIS across the sciences, and to increase the nation's supply of experts in GIS.

The first research agenda was published in 1988 by Rhind, who identified problems in what he termed the handling of geographic data: the volumes of data involved, the numerous types of queries that might be addressed, the prevalence of uncertainty in geographic data; the need for integration of data among organizations; and the lack of awareness of such issues as scale. He recognized that the solution of the more generic of these issues would come with time from mainstream information technology; but that issues that were more specific to the geographic case would have to be solved by an active research community focused on GIS. He saw a substantial role for knowledge-based or expert systems in the automated extraction of features from images; the integration of disparate data sets; the development of intelligent search procedures; the automation of cartographic generalization; the development of machine-based tutors; and the elicitation of knowledge from data. He also recognized the importance of research into better methods of visualization for geospatial data; the role of organizations; the legal issues of liability and intellectual property; and the costs and benefits of GIS.

The NCGIA research agenda, published in 1989, has much in common with Rhind's, but already shows signs of a search for the more fundamental issues of GIScience, in contrast to the practical issues of GIS. Five major research areas are identified:

- **Spatial analysis and spatial statistics**, the techniques used to model uncertainty in geospatial data, to mine data for patterns and anomalies, and to test theories by comparison with reality;
- **Spatial relationships and database structures**, addressing the representation of real geographic phenomena in digital form, and the interface between digital structures and human reasoning;
- **Artificial intelligence and expert systems**, reflecting Rhind's concern for the role of advanced machine intelligence in GIS operations;
- **Visualization**, and the need to advance traditional cartography to reflect the vastly greater potential of digital systems for display of geographic data; and
- **Social, economic and institutional issues**, the host of social issues surrounding GIS.

The NCGIA went on to propose 12 specific research initiatives within this general framework:

- **Accuracy of spatial databases**, focusing on error models for geographic data with strong links to the discipline of statistics, and attempting to characterize the errors inherent in both raster and vector representations;
- **Languages of spatial relations**, including principles of spatial cognition and linguistics;
- **Multiple representations**, the need to integrate different representations of the same phenomena on the Earth's surface, which might be raster or vector, or generalized at different scales;
- **Use and value of geographic information in decision making**;
- **Architecture of very large GIS databases**;
- **Spatial decision support systems**, the design of systems to support decision-making by groups of stakeholders;

- **Visualization of the quality of geographic information**, through methods that explicitly display information about the uncertainty associated with data;
- **Expert systems for cartographic design**, using intelligent systems to augment the skill of cartographers*;*
- **Institutions sharing geographic information**, including research on the impediments to sharing between agencies;
- **Temporal relations in GIS**, the extension of GIS data models to include time;
- **Space-time statistical models in GIS**, the extension of spatial analysis to include time; and
- **Remote sensing and GIS**, researching the issues involved in the integration of data acquired by remote sensing with data from other sources.

Eventually, NCGIA sponsored a total of 21 research initiatives between 1988 and 1996.

Very substantial progress was made on most of these topics in the years following their publication. In addition, four factors contributed to the evolution of these research agendas in the 1990s: first, the continued arrival of new technologies, including most notably the World Wide Web, the Global Positioning System, object-orientation, and mobile computing; second, the broadening of the research community, to include active participation by new disciplines, including linguistics and cognitive science; third, the trend away from technical issues of systems to fundamental issues of science; and fourth, the recognition that certain topics were in effect dead ends. This last perhaps accounts for the virtual disappearance of expert systems, despite their prominence in Rhind's 1988 agenda.

In 1996 the recently formed University Consortium for Geographic Information Science published the first edition of its research agenda, the result of a successful consensus-building exercise amongst the thirty or so research institutions that were then members (the number has since risen to more than 80). The agenda had ten topics:

- **Spatial data acquisition and integration**, including new sources of remote sensing, ground-based sensor networks, and fusion and conflation of data from different sources;
- **Distributed computing**, and the issues of integrating data and software over large heterogeneous networks;
- **Extensions to geographic representations**, addressing particularly the third spatial dimension and time;
- **Cognition of geographic information**, including studies of the processes by which people learn and reason with geographic data, and interact with GIS;
- **Interoperability of geographic information**, including research to overcome the difficulties of different formats and lack of shared understanding of meaning;
- **Scale**, and the complex issues surrounding representations at different levels of detail;
- **Spatial analysis in a GIS environment**, advancing the analytic capabilities of GIS;
- **The future of the spatial information infrastructure** and the institutional arrangements that provide the context for GIS;

- **Uncertainty in geographic data and GIS-based analysis**, advancing understanding of the nature of differences between raster and vector databases and the real phenomena they attempt to represent, and including the modeling and visualization of data quality; and
- **GIS and society**, the study of the impacts of GIS on society, and the societal context in which the technology is used; these issues are increasingly recognized under the more general heading of critical GIS.

UCGIS later added four emerging themes to the list:

- **Geospatial data mining and knowledge discovery**, the development of methods for extracting patterns and knowledge from very large data sources;
- **Ontological foundations of geographic information science**, addressing the fundamental components on which our knowledge of the Earth's surface is based;
- **Geographic visualization**; and
- **Remotely acquired data and information in GIScience**.

A somewhat different approach to framing the research agenda was taken by NCGIA's Project Varenius, a research effort begun in 1996 to advance the fundamentals of geographic information science. In this strikingly simple model, GIScience was anchored by three concepts -- the individual, the computer, and society – represented by a triangle, with GIScience at the core. Research about the individual would be dominated by cognitive science, and its concern for understanding of spatial concepts, learning and reasoning about geographic data, and interaction with the computer. Research about the computer would be dominated by issues of representation, the adaptation of new technologies, computation, and visualization. Finally, research about society would address issues of impacts and societal context. Many research issues would involve the interaction between the three corners of the triangle.

While all of these various attempts to define the research agenda of GIScience have intellectual merit and show a definite emergence of consensus, they all lack the compelling appeal of such scientific mega-projects as Martian exploration or the mapping of the human genome. But the concept of Digital Earth perhaps has the ability to capture popular imagination. The term was coined by then Senator Al Gore in 1992 and elaborated in a much-quoted 1998 speech:

"Imagine, for example, a young child going to a Digital Earth exhibit at a local museum. After donning a head-mounted display, she sees Earth as it appears from space. Using a data glove, she zooms in, using higher and higher levels of resolution, to see continents, then regions, countries, cities, and finally individual houses, trees, and other natural and man-made objects. Having found an area of the planet she is interested in exploring, she takes the equivalent of a 'magic carpet ride' through a 3-D visualization of the terrain. Of course, terrain is only one of the numerous kinds of data with which she can interact. Using the system's voice recognition capabilities, she is able to request information on land cover, distribution of plant and animal species, real-time weather, roads, political boundaries, and population. She can also visualize the environmental information that she and other students all over the world have collected as part of the GLOBE project. This information can be seamlessly fused with the digital map or terrain data. She can get more information on many of the objects she sees by using her data glove to click on a hyperlink. To prepare for her family's vacation to Yellowstone National Park, for example, she plans the perfect hike to the geysers, bison, and bighorn sheep that she has just read about. In fact, she can follow the trail visually from start to finish before she ever leaves the museum in her hometown. She is not limited to moving through space, but can also travel through time. After taking a virtual field-trip to Paris to visit the Louvre, she moves backward in time to learn about French history, perusing digitized maps overlaid on the surface of the Digital Earth, newsreel footage, oral history, newspapers and other primary sources. She sends some of this information to her personal e-mail address to study later. The time-line, which stretches off in the distance, can be set for days, years, centuries, or even geological epochs, for those occasions when she wants to learn more about dinosaurs."

Almost ten years later it is possible to recognize many but by no means all of the elements of this vision in such services as Google Earth, and research is under way both to realize the remaining elements of the concept, and to examine its deeper social meanings.

Google Earth has had a massive impact since its launch in 2005, in introducing many of these concepts of GIS to tens of millions of users. Google Maps, Microsoft Virtual Earth, and many other Web services have only increased the impact. Today, the term neogeography is being used to describe a renewed interest in geography on the part of users of these services. Moreover, numerous Web sites such as Wikimapia, Flickr, and OpenStreetMap are allowing millions of individuals to create and integrate their own geographic information in large repositories, potentially bypassing the traditional means of geographic information production, which has been dominated by the national mapping agencies.

Other technologies, such as remote sensing, GPS, and RFID (radio-frequency identification) are challenging traditional notions of locational privacy, as it becomes increasingly possible to track the geographic positions of individuals and vehicles, and to obtain information from them in real time. The term spatial web is being used to describe a future in which many features in the geographic world will be responsive in the way that aircraft are responsive to air traffic controllers: able to report their positions and identify themselves automatically. This future already exists in many retail stores where items for sale carry RFID tags, in many areas of livestock farming, where every animal is tagged, and in modern construction, where the components of a building are able to identify themselves.

Despite the longstanding nature of many of the concepts reviewed in earlier sections, the world of GIS is evolving more rapidly than ever. The number of people directly engaged in some aspect of its use has increased by many orders of magnitude since the early days of GIS in the 1980s, and seems set to continue increasing as services become easier to use, and as people recognize the importance of information about location in their daily lives.

Further reading

Arctur, D. and Zeiler, M. (2004). *Designing geodatabases: case studies in GIS data modeling*. Redlands, CA: ESRI Press.

Bugayevskiy, L. M. and Snyder, J. P. (1995). *Map projections: a reference manual*. London: Taylor and Francis.

Cho, G. (1998). *Geographic information systems and the law: mapping the legal frontiers*. Chichester, UK: John Wiley & Sons.

Department of the Environment (1987). *Handling geographic information. the report of the Committee of Enquiry chaired by Lord Chorley*. London: Her Majesty's Stationery Office.

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

Goodchild, M. F. (1992). Geographical information science. *International Journal of Geographical Information Systems* 6(1): 31-45.

Goodchild, M. F., Egenhofer, M. J., Kemp, K. K., Mark, D. M., and Sheppard, E. S., (1999). Introduction to the Varenius project. *International Journal of Geographical Information Science* 13(8): 731-746.

Gore, A. (1992). *Earth in the balance: ecology and the human spirit*. Boston: Houghton Mifflin.

Knowles, A. K. (ed.) (2002). *Past time, past place: GIS for history*. Redlands, CA: ESRI Press.

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

Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W. (2005). *Geographic information systems and science*. Second Edition. New York: Wiley.

National Research Council (2003). *IT roadmap to a geospatial future*. Washington, DC: National Academy Press.

National Science Foundation (1999). *Geographic information science: critical issues in an emerging cross-disciplinary research domain*. Washington, DC: National Science Foundation. http://www.geog.buffalo.edu/ncgia/workshopreport.html

Pickles, J. (ed.) (1995). *Ground truth: the social implications of geographic information systems*. New York: Guilford Press.

Rhind, D. W. (1988). A GIS research agenda. *International Journal of Geographical Information Systems* 2(1): 23-28.

University Consortium for Geographic Information Science (1996). Research priorities for geographic information science. *Cartography and Geographic Information Science* 23(3).

Zeiler, M. (1999). *Modeling Our World: The ESRI Guide to Geodatabase Design*. Redlands, CA: ESRI Press.

Figure captions

1. A GIS display of land cover in an area north of Santa Barbara, California. The map covers approximately 21km from left to right. The California vegetation map of 1977 provides the coarse vector polygons (note the straight edges) superimposed on the raster National Land Cover Dataset with a spatial resolution of approximately 30m. The polygons are colored by dominant species, while the raster cells are colored according to a land-cover classification.

2. A screen shot of ESRI's ArcMap showing census data by tract for the city of Milwaukee. The screen simultaneously displays part of the tract attribute table and a map in which each tract is classified by its percent black population.

3. The UNETRANS model, an object-oriented data model for transportation applications of GIS. Each box represents one class of features, and the links between the boxes represent various types of relationship.

4. A screen shot of ESRI's ArcMap showing census tracts (red) overlaid on postal ZIP code boundaries (red) for an area of central California. Polygon overlay might be used in such a case to estimate missing attributes of ZIP codes, such as population, based on areas of overlap with tracts.

5. A screen shot of ESRI's ArcMap showing half-mile exclusion buffers drawn around every school in South-Central Los Angeles.