NEW AGE

GIS BASICS

SHAHAB FAZAL

NEW AGE INTERNATIONAL PUBLISHERS

Geographic Information Systems: A Generic Definition

GIS is a special-purpose digital database in which a common spatial coordinate system is the primary means of reference. Comprehensive GIS require a means of:

1. Data input, from maps, aerial photos, satellites, surveys, and other sources.
2. Data storage, retrieval, and query.
3. Data transformation, analysis, and modelling, including spatial statistics.
4. Data reporting, such as maps, reports, and plans.

Three observations should be made about this definition

First, GIS are related to other database applications, but with an important difference. All information in a GIS is linked to a spatial reference. Other databases may contain locational information (such as street addresses, or zip codes), but a GIS database uses geo-references as the primary means of storing and accessing information.

Second, GIS integrates technology. Whereas other technologies might be used only to analyze aerial photographs and satellite images, to create statistical models, or to draft maps, these capabilities are all offered together within a comprehensive GIS.

Third, GIS, with its array of functions, should be viewed as a process rather than as merely software or hardware. GIS are for making decisions. The way in which data is entered, stored, and analyzed within a GIS must mirror the way information will be used for a specific research or decision – making task. To see GIS as merely a software or hardware system is to miss the crucial role it can play in a comprehensive decision-making process.

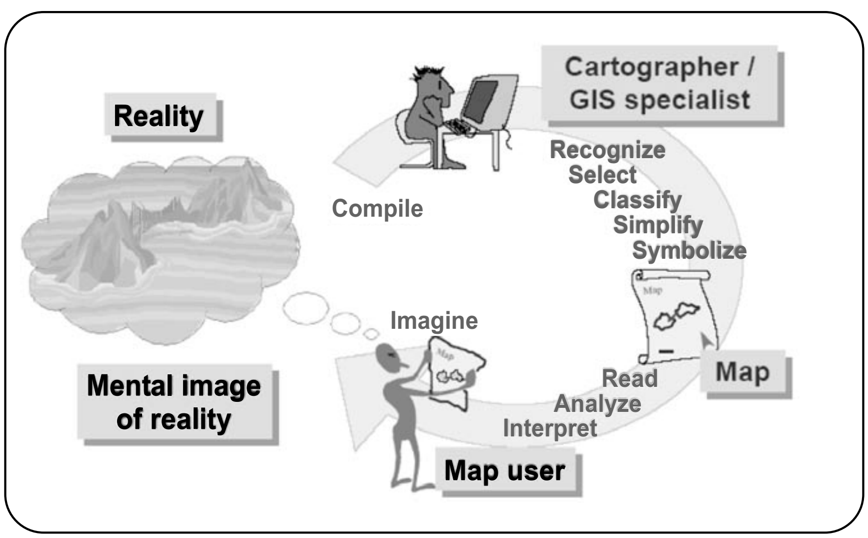

Figure 1.3: Different stages of information transfer in GIS.

What Actually GIS is?

GIS is expressed in individual letters G – I – S and not at pronunciation GIS. It stands for geographic or geographical information systems. Geographic Information Science is a new interdisciplinary field. It is built upon knowledge from geography, cartography, computer science, mathematics etc.

GIS can be defined as ***'A system for Capturing, storing, checking, integrating, manipulating, analysing and displaying data which are spatially referenced to the Earth. This is normally considered to involve a spatially referenced computer database and appropriate applications software'.***

GIS needs spatial data, this makes it unique. Here spatial means – related to the space – the real world location. That is why GIS is based on basic geographic concepts.

A Geographic Information System is an integration of computer hardware and software which can create manipulate, and analyze a geographically referenced data base to produce new maps and tabular data GIS includes the capabilities of Computer Aided Design (CAD) and Data Base Management Systems (DBMS), but is more than just a combination of those systems. In a GIS, a relationship between the graphic map data and the tabular data base is maintained so that changes to the map are reflected in the data base GIS allows automatic determination of the relationships between maps, and can create new maps of those relationships.

Geographic Information System (GIS) can also be defined as:

The organized activity by which people

- Measure aspects of geographic phenomena and processes;
- Represent these measurements, usually in the form of a computer database, to emphasize spatial themes, entities, and relationships;
- Operate upon these representations to produce more measurements and to discover new relationships by integrating disparate sources; and
- Transform these representations to conform to other frameworks of entities and relationships.

These activities reflect the larger context (institutions and cultures) in which these people carry out their work. In turn, the GIS may influence these structures.

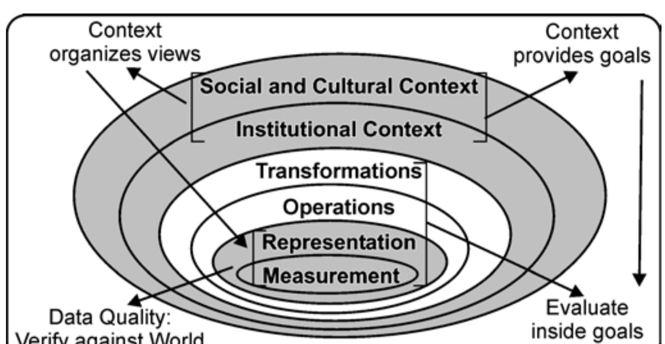

Figure 1.4: GIS framework.

Other definitions

Many people offer definitions of GIS. In the range of definitions presented below, different emphases are placed on various aspects of GIS. Some miss the true power of GIS, its ability to integrate information and to help in making decisions, but all include the essential features of spatial references and data analysis.

A definition quoted in William Huxhold's Introduction to Urban Geographic Information Systems:

> *'. . . The purpose of a traditional GIS is first and foremost spatial analysis. Therefore, capabilities may have limited data capture and cartographic output. Capabilities of analyses typically support decision making for specific projects and/or limited geographic areas. The map data-base characteristics (accuracy, continuity, completeness, etc.) are typically appropriate for small-scale map output. Vector and raster data interfaces may be available. However, topology is usually the sole underlying data structure for spatial analyses.'*

C. Dana Tomlin's definition, from Geographic Information Systems and Cartographic Modelling:

> *'A geographic information system is a facility for preparing, presenting, and interpreting facts that pertain to the surface of the earth. This is a broad definition . . . a considerably narrower definition, however, is more often employed. In common parlance, a geographic information system or GIS is a configuration of computer hardware and software specifically designed for the acquisition, maintenance, and use of cartographic data.'*

From Jeffrey Star and John Estes, in Geographic Information Systems: An Introduction:

> *'A geographic information system (GIS) is an information system that is designed to work with data referenced by spatial or geographic coordinates. In other words, a GIS is both a database system with specific capabilities for spatially-reference data, as well [as] a set of operations for working with data . . . In a sense, a GIS may be thought of as a higher-order map.'*

The GIS View of the World

GIS provide powerful tools for addressing geographical and environmental issues. Consider the schematic diagram below. Imagine that the GIS allows us to arrange information about a given region or city as a set of maps with each map displaying information about one characteristic of the region. In the case below, a set of maps that will be helpful for urban transportation planning have been gathered. Each of these separate thematic maps is referred to as a **layer, coverage, or level.** And each layer has been carefully overlaid on the others so that every location is precisely matched to its corresponding locations on all the other maps. The bottom layer of this diagram is the most important, for it represents the grid of a locational reference system (such as latitude and longitude) to which all the maps have been precisely registered.

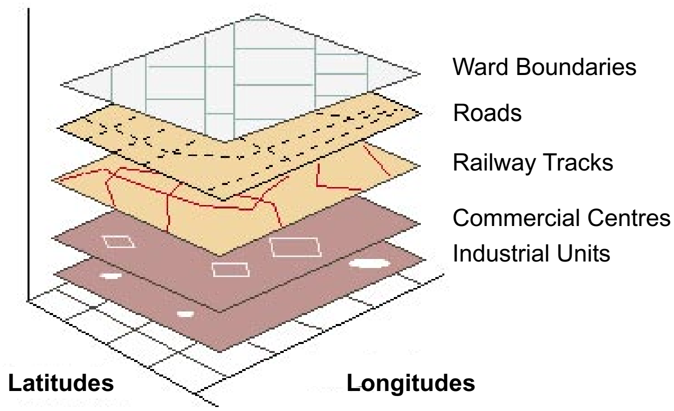

Figure 1.5: GIS: an integrating technology.

Why is GIS Important?

- 'GIS technology is to geographical analysis what the microscope, the telescope, and computers have been to other sciences.... (It) could therefore be the catalyst needed to dissolve the regional-systematic and human-physical dichotomies that have long plagued geography' and other disciplines which use spatial information.
- GIS integrates spatial and other kinds of information within a single system – it offers a consistent framework for analyzing geographical data.
- By putting maps and other kinds of spatial information into digital form, GIS allows us to manipulate and display geographical knowledge in new and exciting ways.
- GIS makes connections between activities based on geographic proximity
  - looking at data geographically can often suggest new insights, explanations.
  - these connections are often unrecognized without GIS, but can be vital to understanding and managing activities and resources.
  - *e.g.* we can link toxic waste records with school locations through geographic proximity.

Box 2: Definitions of GIS and the groups who find them useful.

A container of maps in digital form	the general public
A computerized tool for solving geographic problems	decision makers, planners
A spatial decision support system	managers, operations researchers
A mechanized inventory of geographically distributed features	utility managers, resource managers
A tool for revealing what is otherwise invisible in geographic information	scientists, investigators
A tool for performing operations on geographic data that are too tedious if performed by manual methods	resource managers, planners, GIS experts

- GIS allows access to administrative records – property ownership, tax files, utility cables and pipes – via their geographical positions.
- Maps are fascinating and so are maps in computers and there is increasing interest in geography and geographic education in recent times. GIS gives a ‘high tech’ feel to geographic information.

Contributing Disciplines

GIS is a convergence of technological fields and traditional disciplines. GIS has been called an ‘enabling technology’ because of the potential it offers for the wide variety of disciplines which must deal with spatial data. Each related field provides some of the techniques which make up GIS. Many of these related fields emphasize data collection – GIS brings them together by emphasizing integration, modelling and analysis, as the integrating field, GIS often claims to be the science of spatial information.

Geography: Geography is broadly concerned with understanding the world and man’s place in it. Geography has long tradition in spatial analysis. The discipline of geography provides techniques for conducting spatial analysis and a spatial perspective on research.

Cartography: Cartography is concerned with the display of spatial information. Currently it is the main source of input data for GIS is maps. Cartography provides long tradition in the design of maps which is an important form of output from GIS. Computer cartography (also called ‘digital cartography’, ‘automated cartography’) provides methods for digital representation and manipulation of cartographic features and methods of visualization.

Remote Sensing: This emerging technique which records images from space and the air are major source of geographical data. Remote sensing includes techniques for data acquisition and processing anywhere on the globe at low cost, consistent update potential. The main advantage of it is that interpreted data from a remote sensing system can be merged with other data layers in a GIS.

Photogrammetry: Using aerial photographs and techniques for making accurate measurements from them, photogrammetry is the source of most data on topography (ground surface elevations) used for input to GIS.

Surveying: Surveying is concerned with the measurement of locations of objects on the Earth’s surface, particularly property boundaries. Surveying provides high quality data on positions of land boundaries, buildings, etc.

Statistics: Many models built using GIS are statistical in nature, many statistical techniques used for analysis in GIS. Statistics is important in understanding issues of error and uncertainty in GIS data.

Computer Science: Computer science is one of the main engines for GIS development. Artificial intelligence (AI) uses the computer to make choices based on available data in a way that is seen to emulate human intelligence and decision-making – computer can act as an 'expert' in such functions as designing maps, generalizing map features. Computer-aided design (CAD) provides software, techniques for data input, display and visualization, representation, particularly in 3 dimensions. Advances in computer graphics provide hardware, software for handling and displaying graphic objects, techniques of visualization. Similarly, database management systems (DBMS) contribute methods for representing data in digital form, procedures for system design and handling large volumes of data, particularly access and update.

MATHEMATICS: Several branches of mathematics, especially geometry and graph theory, are used in GIS system design and analysis of spatial data.

Major Areas of Application

GIS technology, data structures and analytical techniques are gradually being incorporated into a wide range of management and decision-making operations. Numerous examples of applications of GIS are available in many different journals and are frequent topics of presentations at conferences in the natural and social sciences.

In order to understand the range of applicability of GIS it is necessary to characterize the multitude of applications in some logical way so that similarities and differences between approaches and needs can be examined. An understanding of this range of needs is critical for those who will be dealing with the procurement and management of a GIS.

FUNCTIONAL CLASSIFICATION: One way to classify GIS applications is by functional characteristics of the systems; this would include a consideration of characteristics of the data such as themes, precision required and data model. Secondly, GIS a function as which of the range of possible GIS functions does the application rely on? *e.g.* address matching, overlay? Thirdly, a product *e.g.,* does the application support queries, one-time video maps and/or hardcopy maps? A classification based on these characteristics quickly becomes fuzzy since GIS is a flexible tool whose great strength is the ability to integrate data themes, functionality and output.

GIS AS A DECISION SUPPORT TOOL: Another way to classify GIS is by the kinds of decisions that are supported by the GIS. Decision support is an excellent goal for GIS, however: decisions range from major (which areas in India are best suited for establishing SEZ with foreign aids?) to minor (which way to turn at next intersection?). Decision support is a good basis for definition of GIS, but not for differentiating between applications since individual GIS systems are generally used to make several different kinds of decisions.

GIS USERS: GIS field is a loose coalescence of groups of users, managers, academics and professionals all working with spatial information. Each group has a distinct educational and 'cultural' background with varied interests and priorities. As a result; each identifies itself with particular ways of approaching particular sets of problems. The core groups of GIS activity can be seen to be comprised of:

a. mature technologies which interact with GIS, sharing its technology and creating data for it such as surveyors and engineers, cartographers, scientists using remote sensing techniques.
b. management and decision-making groups such as resource inventors, and resource managers, urban planners, municipal officials managing land records for taxation and ownership control, facilities managers, managers involved in marketing and retail planning or vehicle routing and scheduling.
c. science and research activities at universities and government labs – these groups of GIS activity seeking to find distinctions and similarities between them.

SOME IMPORTANT AREAS WHERE GIS IS BEING USED ARE:

- *Different Streams of Planning:* Urban planning, housing, transportation planning architectural conservation, urban design, landscape planning etc.
- *Street Network Based Application:* It is an addressed matched application, vehicle routing and scheduling: location, development and site selection and disaster planning.
- *Natural Resource Based Application:* Management and environmental impact analysis of wild and scenic recreational resources, flood plain, wetlands, acquifers, forests, and wildlife.
- *View Shed Analysis:* Hazardous or toxic factories siting and ground water modelling. Wildlife habitat study and migrational route planning.
- *Land Parcel Based:* Zoning, sub-division plans review, land acquisition, environment impact analysis, nature quality management and maintenance etc.
- *Facilities Management:* Can locate underground pipes and cables for maintenance, planning, tracking energy use.

THE APPEAL AND POTENTIAL OF GIS

The great appeal of GIS stems from their ability to integrate great quantities of information about the environment and to provide a powerful repertoire of analytical tools to explore this data. Imagine the potential of a system in which dozens or hundreds of maps layers are arrayed to display information about transportation networks, hydrography, population characteristics, economic activity, political jurisdictions, and other characteristics of the natural and social environment. Such a system would be valuable in a wide range of situations – for urban planning, environmental resource management, hazards management, emergency planning, or transportation forecation, and so on. The ability to separate information in layers, and then combine it with other layers of information is the reason why GIS hold such great potential as research and decision-making tools.

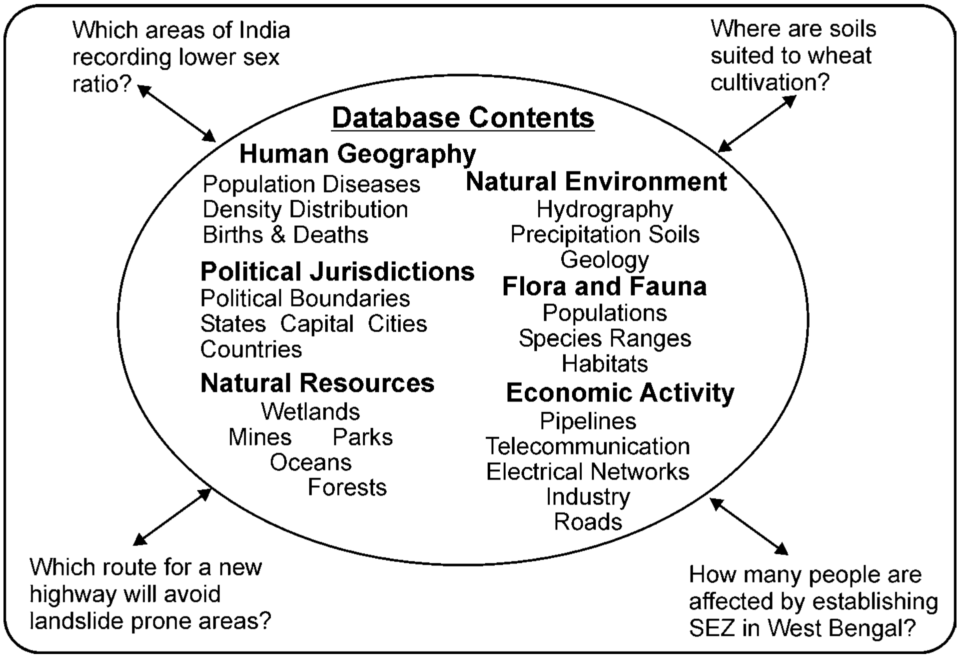

Figure 1.6: Application potential of GIS for geographical studies.

Development of GIS

Since the mid-1970s, specialized computer systems have been developed to process geographical information in various ways. These include:

- Techniques to input geographical information, converting the information to digital form.
- Techniques for storing such information in compact format on computer disks, compact disks (CDs), and other digital storage media.
- Methods for automated analysis of geographical data, to search for patterns, combine different kinds of data, make measurements, find optimum sites or routes, and a host of other tasks.
- Methods to predict the outcome of various scenarios, such as the effects of climate change on vegetation.
- Techniques for display of data in the form of maps, images, and other kinds of displays.
- Capabilities for output of results in the form of numbers and tables.

Components of GIS

HARDWARE: It consists of the computer system on which the GIS software will run. The choice of hardware system ranges from Personal Computers to multi user Super Computers. These a computers should have essentially an efficient processor to run the software and sufficient memory to store enough information (data).

SOFTWARE: GIS software provides the functions and tools needed to store, analyze, and display geographic information. The software available can be said to be application specific. All GIS software generally fit all these requirements, but their on screen appearance (user interface) may be different.

DATA: Geographic data and related tabular data are the backbone of GIS. It can be collected in-house or purchased from a commercial data provider. The digital map forms the basic data input for GIS. Tabular data related to the map objects can also be attached to the digital data. A GIS will integrate spatial data with other data resources and can even use a DBMS.

METHOD: A successful GIS operates according to a well-designed plan, which are the models and operating practices unique to each task. There are various techniques used for map creation and further usage for any project. The map creation can either be automated raster to vector creator or it can be manually vectorized using the scanned images. The source of these digital maps can be either map prepared by any survey agency or satellite imagery.

PEOPLE: GIS users range from technical specialists who design and maintain the system to those who use it to help them perform their everyday work. GIS operators solve real time spatial problems. They plan, implement and operate to draw conclusions for decision making.

NETWORK: With rapid development of IT, today the most fundamental of these is probably the network, without which no rapid communication or sharing of digital information could occur. GIS today relies heavily on the Internet, acquiring and sharing large geographic data sets.

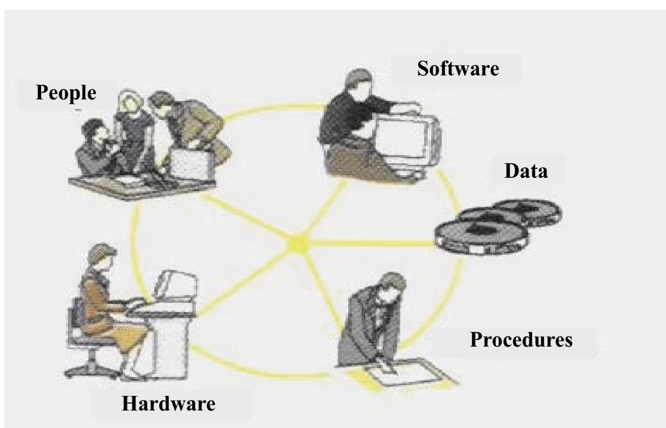

Figure 1.7: Six basic components of GIS.

Although it is very easy to purchase the constituent parts of a GIS (the computer hardware and basic software), the system functions only when the requisite expertise is available, the data are compiled, the necessary routines are organized, and the programs are modified to suit the application. A computer system can function at what may appear to be lightning speed, yet the entire time span of a GIS project can stretch to months and even years. These facets of an overall GIS are interlinked. In general, procurement of the computer hardware and software is vital but straightforward. The expertise required is often underestimated, the compilation of data is expensive and time consuming, and the organizational problems can be most vexing. These facets of an overall GIS are discussed in detail later.

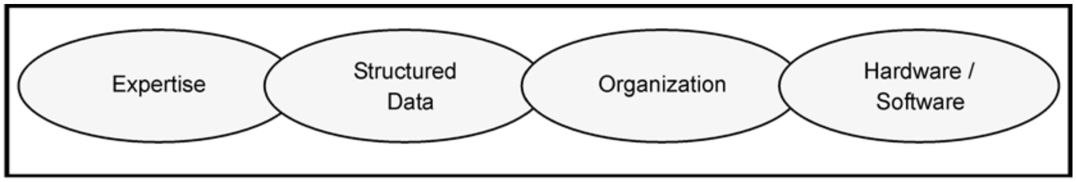

Figure 1.8: A GIS chain – Equal role of the above links in GIS organization.

Traditionally, geographical data are presented on maps using symbols, lines, and colours. Most maps have a legend in which these elements are listed and explained – a thick black line for main roads, a thin black line for other roads, and so on. Dissimilar data can be superimposed on a common coordinate system. Consequently, a map is both an effective medium for presentation and a bank for storing geographical data. But herein lies a limitation. The stored information is processed and presented in a particular way, usually for a particular purpose. Altering the presentation is seldom easy. A map provides a static picture of geography that is almost always a compromise between many differing user needs. Nevertheless, maps are a substantial public asset. Surveys conducted in Norway indicate that the benefit accrued from the use of maps is three times the total cost of their production.

Compared to maps, GIS has the inherent advantage that data storage and data presentations are separate. As a result, data may be presented and viewed in various ways. Once they are stored in a computer, we can zoom into or out of a map, display selected areas, make calculations of the distance between places, present tables showing details of features shown on the map, superimpose the map on other information, and even search for the best locations for retail stores. In effect, we can produce many useful products from a single data source.

The term geographical information system (GIS) is now used generically for any computer-based capability for the manipulation of geographical data. GIS is computer-based capability for the manipulation of geographical data. A GIS includes not only hardware and software, but also the special devices used to input maps and to create map products, together with the communication systems needed to link various elements. The hardware and software functions of a GIS include:

- Compilation
- Storage
- Updating and changing
- Management and exchange
- Manipulation
- Retrieval and presentation

- Acquisition and verification
- Analysis and combination

All of these actions and operations are applied by a GIS to the geographical data that form its database. All of the data in a GIS are georeferenced, that is, linked to a specific location on the surface of the Earth through a system of coordinates. One of the commonest coordinate systems is that of latitude and longitude; in this system location is specified relative to the equator and the line of zero longitude through Greenwich, England. But many other systems exist, and any GIS must be capable of transforming its georeferences from one system to another.

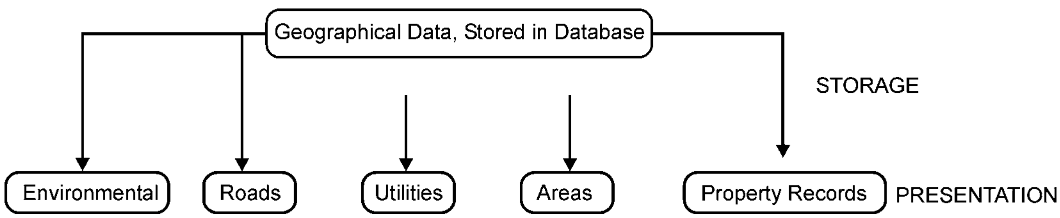

Figure 1.9: A map can be a presentation medium and a storage medium. GIS manipulates data to produce results.

Geographical information attaches a variety of qualities and characteristics to geographical locations (Figure 1.10). These qualities may be physical parameters such as ground elevation, soil moisture level, or classifications according to the type of vegetation, ownership of land, zoning, and so on. Such occurrences as accidents, floods, or landslides may also be included. We use the general term attributes to refer to the qualities or characteristics of places, and think of them as one of the two basic elements of geographical information, along with locations.

In some cases, qualities are attached to points, but in other cases they refer to more complex features, either lines or areas, located on the Earth's surface; in such cases the GIS must store the entire mapped shape of the feature rather than a simple coordinate location. Examples of commonly mapped features are lakes, cities, counties, rivers, and streets, each with its set of useful attributes. When a feature is used as a reporting zone for statistical purposes, a vast amount of information may be available to be used as attributes for the zone in GIS. In market research, for example, it is common for postal codes to be used as the basis for reports on demographics, purchasing habits, and housing markets.

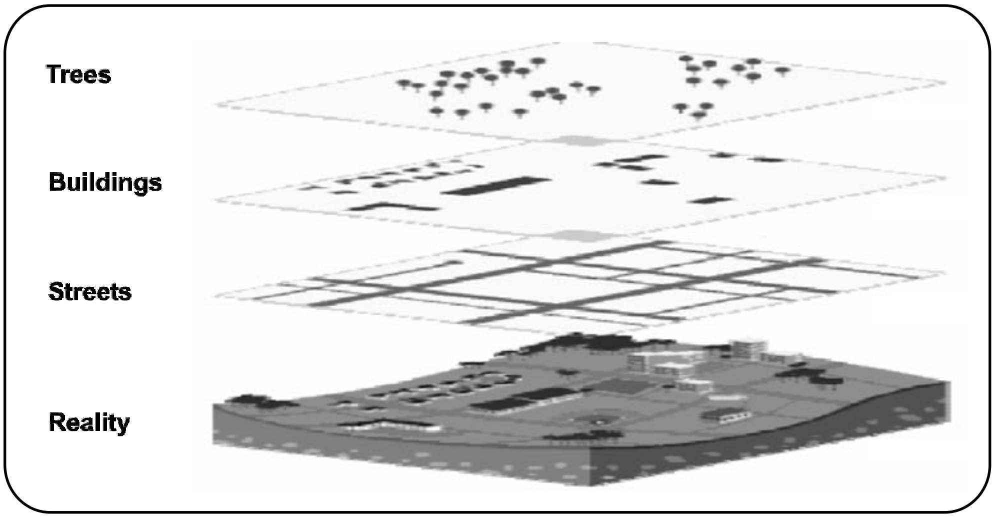

Figure 1.10: GIS stores data in different theme layers in the computer, each layer is linked to a common referencing system.

The relationships between geographical features often provide vital information. For example, the connections of a water supply pipe network may be critical for technicians, who need to know which valves to close in order to increase water pressure in certain sectors. The details of properties bordering a road are necessary if all property owners affected by roadwork are to be properly notified. Connections between streets are important in using a GIS to assist drivers in navigating around an unfamiliar city. The ability of a GIS to store relationships between features in addition to feature locations and attributes is one of the most important sources of the power and flexibility of this technology. Some GISs can even store flows and other measures of interaction between features, to support applications in transportation, demography, communication, and hydrology, among other areas.

Stored data may be processed in a GIS for presentation in the form of maps, tables, or special formats. One major GIS strength is that geographical location can be used to link information from widely scattered sources. Because the geographical location of every item of information in a GIS database is known, GIS technology makes it possible to relate the quality of groundwater at a site with the health of its inhabitants, to predict how the vegetation in an area will change as the irrigation facilities increases, or to compare development proposals with restrictions on land use. This ability to overlay gives GIS unique power in helping us to make decisions about places and to predict the outcomes of those decisions. The only requirement is that the geographical information from each source be expressed in compatible georeferencing systems.

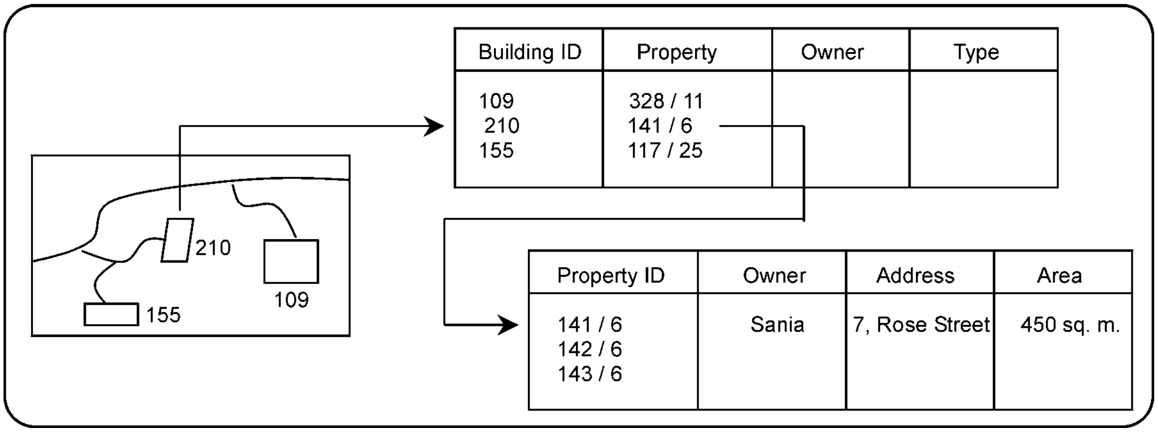

Figure 1.11: GIS functions on the interaction between digital map data and its attribute informations.

A GIS can process georeferenced data and provide answers to questions involving, *e.g.,* the particulars of a given location, the distribution of selected phenomena, the changes that have occurred since a previous analysis, the impact of a specific event, or the relationships and systematic patterns of a region. It can perform analyses of georeferenced data to determine the quickest driving route between two points and help resolve conflicts in planning by calculating the suitability of land for particular uses.

A GIS can process georeferenced data and provide answers to questions involving, *e.g.,* the particulars of a given location, the distribution of selected phenomena, the changes that have occurred since a previous analysis, the impact of a specific event, or the relationships and systematic patterns of a region. It can perform analyses of georeferenced data to determine the quickest driving route between two points and help resolve conflicts in planning by calculating the suitability of land for particular uses.

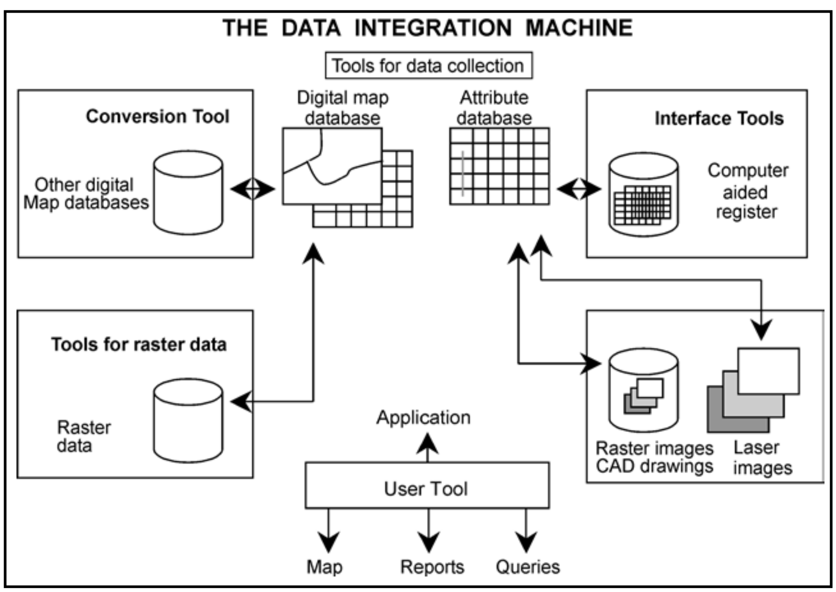

Figure 1.12: GIS is a typical data integration machine. It receives, process and transmits data.

GIS can process data from a wide range of sources, including data obtained from maps, images of the Earth obtained from space satellites, video film of the Earth taken from low-flying aircraft, statistical data from published tables, photographs, data from computer-assisted design (CAD) systems, and data obtained from archives by electronic transmission over the Internet and other networks. Data integration is one of the most valuable functions of a GIS, and the data that are integrated are more and more likely to be obtained from several distinct media-multimedia is an active area for research and development in GIS (Figure 1.12).

Technically, a GIS organizes and exploits digital geographical data stored in databases. The data include information on locations, attributes, and relationships between features. But a database can only approximate the real world, since the storage capacity of a database is minuscule in comparison with the complexity of the real world, and the cost of building a database is directly related to its complexity. The contents of a book of 100,000 words can be stored in digital form in roughly 1 million bytes (the common unit of computer storage is a byte, defined as 8 bits; 1 megabyte is slightly more than 1 million bytes). The information on a topographic map is comparatively dense, and it commonly takes 100 megabytes to capture it in digital form. A single scene from an Earth observing satellite might contain 300 megabytes, the information content of 300 books. Thus even crude approximations to the complexity of real-world geography can rapidly overtake the capacity of our digital storage devices.

Although the contents of a GIS database are equivalent to a map, there are important differences. On a map, a geographical feature such as a road or a power line is shown as a

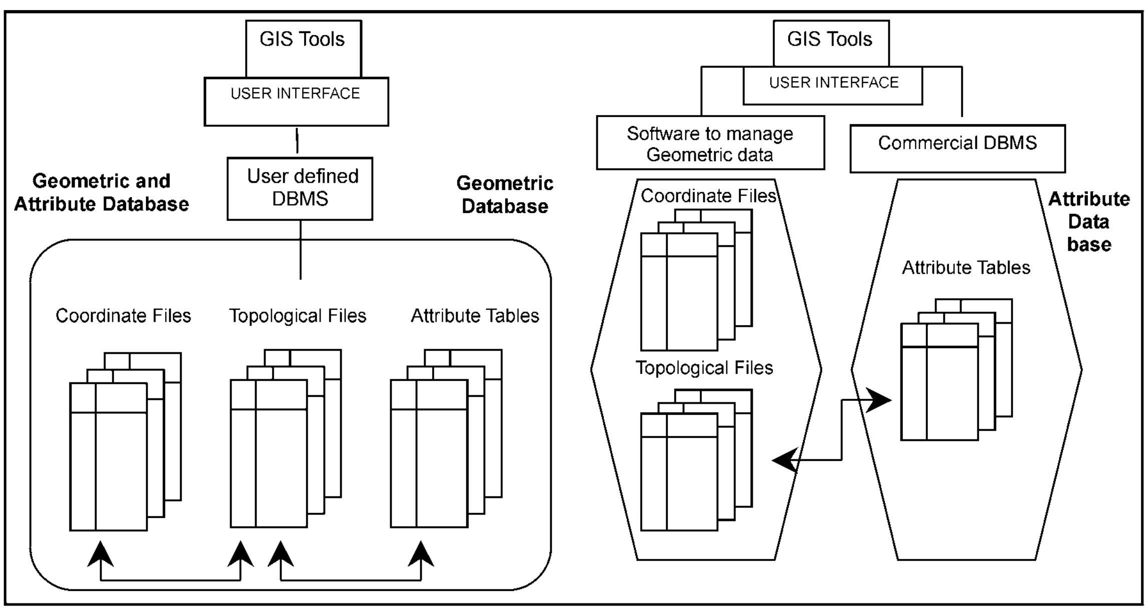

Figure 1.13: DBMS solution for GIS.

symbol using a graphic that will readily be understood by the map reader. In a geographical database a road or power line will be represented by a single sequence of points connected by straight lines, and its symbolization will be reattached when it is displayed. A tube well will be represented by a single point, with the attribute 'tube well', and will be replaced by a symbol when displayed. This approach is economical since the geometric form of the tube well symbol will be stored only once rather than repeated at each tube well location, and it also allows analysis to be more effective.

Databases are vital in all geographical information systems, since they allow us to store geographical data in a structured manner that can serve many purposes. Many GISs impose further structure by using a database management system (DBMS) to store and manage part or all of the data in a largely independent subsystem under the GIS itself. A DBMS is a general-purpose software product, and GISs that use this approach are often able to function in conjunction with a wide range of DBMS products. The database underlying a GIS achieves many objectives. It ensures that data are:

- Stored and maintained in one place
- Stored in a uniform, structured, and controlled manner than can be documented
- Accessible to many users at once, each of whom has the same understanding of the database's contents
- Easily updated with new data

This contrasts with the traditional way of organizing and storing data on paper in filing cabinets, in which data are often:

- Stored in ways that are understandable to one person only
- Easily corrupted by use, or edited in ways that are meaningful only to the editor
- Inaccessible to anyone other than the creator of the system
- Stored in formats and at scales that are so diverse that they cannot be compared or collated
- Difficult to update

GIS Diversity

Although the general definition of GIS given here is quite valid, in practice the diversity of GIS has spawned various definitions. First, users have contrived working definitions suited to their own specific uses. Thus they may vary according to whether operators are planners, water-supply and sewage engineers, support service personnel, or perhaps professional and public administrators or Earth scientists. Second, those with a more theoretical approach, such as research workers, software developers or sales and training staff may use definitions that are different from those used in practical applications. Systems can be tailor made by assembling them from available software tool kits of semi-independent modules, assorted computer hardware components, and other interoperable devices. Many applications can be addressed by acquiring a single, generic GIS product and a standard configuration of hardware. There are many views of GISs, including: