Geographic Information System, or GIS, a computer system for input, storage, manipulation, analysis, and display of geographic data. Broadly defined, geographic data consists of knowledge and facts about specified locations on the surface of the Earth, such as are found on maps or in atlases.

A GIS includes (1) an input subsystem, capable of converting information from maps, satellite images of the Earth, text, or tables into digital form, with essential facilities for editing and error correction; (2) a database designed to store geographic features, such as points, lines, or areas, together with their defining attributes and important relationships between them; (3) a set of processing routines to retrieve information from the database and carry out important functions, such as calculation of the areas of selected features, determination of watershed boundaries, calculation of driving time between places, or search for areas meeting suitability criteria; (4) a set of display and output routines to present data, or the results of analysis, to the user in understandable form as computer displays or paper maps; and (5) a user interface to allow interaction with the system in a straightforward, user-friendly fashion. GIS software is available for a wide range of computing environments from personal computer to mainframe, depending on the volume of data, performance requirements, and number of users.

Many GIS databases are constructed as sets of layers, each layer containing information on one type of geographic data. One layer might include information on streets, represented as a set of lines, with attributes such as travel direction, traffic volume, street name, or pavement quality; another might represent the characteristics of soils for the same region, as a patchwork of differently-shaped areas, each area identifying a patch of some consistent type of soil; another might represent the elevation of the land surface. A typical GIS database might contain from ten to one hundred such layers, although some systems avoid the use of layers entirely by combining all information together. The ability to compare and combine layers for a region is one of the most powerful features of a GIS, and is often presented as its most important defining characteristic.

GISs fall into two broad classes, vector and raster, depending on how they choose to store the information found in each layer. Vector GISs store each layer's contents as collections of points, lines, and areas, with their attributes and interrelationships, and have similarities with computer assisted design (CAD) systems. Examples of relationships between features are adjacency (one area adjoins another), containment (a point lies inside an area), intersection (two lines join), or relative proximity (one point is nearest to another). Raster GISs store each layer by dividing it into a regular array of cells, and storing the contents of each cell, using the same regular array for each layer in the database. Broadly speaking, raster GISs lack the ability to store relationships between features, or complex sets of feature attributes, and have similarities with image processing systems. Increasingly, GIS software designers are combining raster and vector approaches, although full integration is still some way off.

Roughly 100,000 GIS are currently installed. Their uses range from the management of the complex networks of wires, transformers, poles, and switches owned by electrical utilities, to maintenance of information on land ownership by county governments, to scientific research on global environmental change, to tracking of timber resources on forest land, to dispatch of ambulances, fire, or police vehicles in response to emergencies. Although detailed functions will vary with applications, each of these uses of GIS requires a system with the database and capabilities outlined earlier. Much of the money spent developing a GIS goes into building its database, and ensuring its quality.

The first GIS, the Canada Geographic Information System, was built in the 1960s by the Government of Canada to process and analyze the data collected by the Canada Land Inventory. Much of the development work was carried out by IBM Canada Ltd., under the direction of Roger Tomlinson. Computerization was considered essential, despite the primitive systems available at that time, because of the large volumes of data, the need to compare several layers of information, and the massive numbers of calculations needed for the project's reports. GIS remained an exotic specialty of governments, and research institutions such as the Harvard University Laboratory for Computer Graphics and Spatial Analysis, until the late 1970s, when sufficiently powerful hardware and software became available at reasonable cost. GIS sales blossomed in the 1980s, and the GIS software industry continues to grow rapidly. Leading software vendors include Intergraph Corporation (Huntsville, AL) and Environmental Systems Research Institute (Redlands, CA). A range of textbooks and magazines are available on GIS, and an active international research community is dedicated to improving the technology's power and usefulness.

Image caption

The GIS-generated image shows two layers of a database covering Los Angeles County, and illustrates the power of a GIS to combine different kinds of information for a geographic area. The points, stored in one layer, represent locations of industries releasing toxic chemicals, from the Toxic Release Inventory (TRI) maintained by the U.S. Environmental Protection Agency. The underlying areas are the census tracts of the U.S. Bureau of the Census, colored according to the largest racial or ethnic group living there in 1990. The image is part of a larger study carried out by Lauretta Burke at the University of California, Santa Barbara, to examine relationships between pollution, race, and residential patterns.