ISSN 0976 – 1411

Available online at www.internationaleJournals.com

International eJournals

International eJournal of Mathematics and Engineering 30 (2010) 293 - 303

INTERNATIONAL eJOURNAL OF MATHEMATICS AND ENGINEERING

www.InternationaleJournals.com

AN OVERVIEW OF GEOGRAPHIC INFORMATION SYSTEM

S. Sreekanth, I. V. Murali Krishna

Civil Engineering Department, JBIET, Hyderabad, A. P, India

Consultant ASCI, Hyderabad, Former Director R&D, JNTU, Hyderabad, A.P., India

ABSTRACT

We are presently positioned at the outset of twenty-first century with the fast Growing trends in computer technology, information systems and virtual world to obtain data about the physical and cultural worlds, and to use these data to do research or to solve practical problems. The current digital and analog electronic devices facilitate the inventory of resources and the rapid execution of arithmetic or logical operations. These Information Systems are undergoing much improvement and they are able to create, manipulate, store and use spatial data much faster and at rapid rate as compared to conventional methods.

1. INTRODUCTION

An Information System, a collection of data and tools for working with those data, contains data in analog form or digital form about the phenomena in the real world. Our Perception of the world through selection, generalization and synthesis give us information and the representation of this information that is, the data constitutes a model of those phenomena. So the collection of data, the data base is a physical repository of varied views of the real world representing our knowledge at one point in time. Information is derived from the individual data elements in a database, the information directly apparent i.e. information is produced from data by our thought processes, institution or what ever based on our knowledge. Therefore in a data base context the terms data, information and knowledge are differentiated. It can be summarized that the data is very important and added value as we progress from data to information, to knowledge. The data, which has many origins and forms, may be any of the following:

1. Real, for example the terrain conditions etc.
2. Captured, i.e. recorded digital data from remote sensing satellites or Arial photographs of any area.
3. Interpreted, i.e. landuse from remote sensing data.
4. Encoded i.e. recordings of rain-gauge data, depth of well data etc.
5. Structured or organized such as tables about conditions of particular watershed.

2. DEVELOPMENTS IN THE FIELD OF INFORMATION SYSTEMS

Geographic Information Management Technology encompasses many fields including Computer Science, Cartography, Information Management, Telecommunications, Geodesy, Photogrammetry and Remote Sensing and is flavored with it's applications of engineering, environmental analysis, landuse planning, natural resource development, infrastructure management, and many others. Geographic Information Management Technology has almost as many names and acronyms as uses.

One common name is Geographic Information System (GIS). Another is automated mapping/facilities management (AM/FM). Although GIS has recently became more widely accepted as a generic term for the technology, the term Geographic Information System was first published in a 1965 Northwestern University discussion paper by Michal Dacy and Dvane Marble. Key terms associated with geographic information management technology include:

- Automated Mapping (A.M.)
- Computer Assisted or Computer Aided Mapping (CAM)
- Computer Aided Drafting (CAD)
- Computer Aided Drafting and Design (CADD)
- Geographic Information System
- Automated Mapping/ Facility Management (AM/FM)
- Geoprocessing and Network Analysis
- Land information System
- Multipurpose Cadastre

All these terminology's are often used interchangeably even though they denote different capabilities and concepts (**Figure 1**).

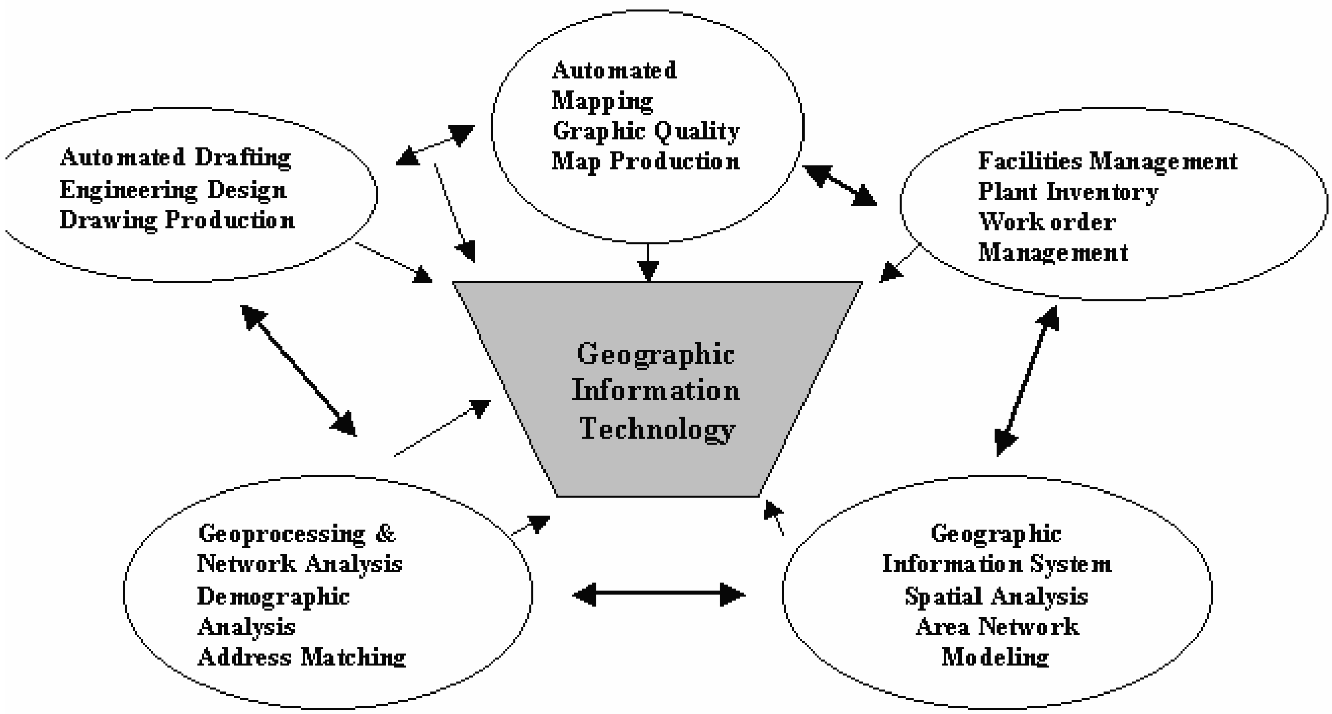

Figure 1: Geographic Information Technology

3. HISTORY of GIS

The GIS history dates back 1960 where computer based GIS have been used and their manual procedures were in life 100 years earlier or so. The initial developments originated in North America with the organizations such as US Bureau of the Census, The US Geological Survey and The Harvard Laboratory for computer graphics and Environmental Systems Research Institute (commercial). Canadian Geographic Information Systems (CGIS) in Canada, Natural Experimental Research Center (NREC), Department of Environment (DOE) in notable organizations in U.K. involved in early developments. Even the other parts of the world are noticed but they have taken place in recent past. In India the major developments have happened for the last one-decade with significant contribution coming from Department of Space emphasizing the GIS applications for Natural Resources Management. Recently the commercial organizations in India have realized the importance of GIS for many applications like infrastructure development, facility management, business/market applications etc.

4. DEFINITIONS OF GIS

The definitions of a GIS given by various authors are as follows

- “A spatial data handling system" (Marble et at, 1983).

- “A computer - assisted system for the capture, storage retrieval, analysis and display of spatial data, within a particular Organization" (Clarke, 1986).

- “A powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world" (Burrough, 1987).

- . “An internally referenced, automated, spatial information system" (Berry, 1986).

- "A system which uses a spatial data base to provide answers to queries of a geographical nature" (,Goodchild, 1985).

- “A system for capturing, storing, checking, manipulating, analyzing and displaying data which are spatially referenced to the Earth” (DOE, 1987:132)

- “Any manual or computer based set of procedures used to store and manipulate geographically referenced data”(Aronoff, 1989:39)

- “An institutional entity, reflecting an organizational structure that integrates technology with a database, expertise and continuing financial support over time”(Carter, 1989:3)

- “An information technology which stores, analyses and display both spatial and non – spatial data”(Parker,1988:1547)

- “A special case of information systems where the database consists of observations on spatially distributed features, activities, or events, which are definable in space as points, lines or areas. A GIS manipulates data about these points, lines and areas to retrieve data for ad hoc queries and analyses”(Dueker,1979:106)

- “A database system in which most of the data are spatially indexed, and upon which a set of procedures operated in order to answer queries about spatial entities in the database”(Smith,1987:13)

- “An automated set of functions that provides professionals with advanced capabilities for the storage, retrieval, manipulation and display of geographically located data”(Ozemoy, Smith and Sicherman, 1981:92)

- “A powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world”(Burrough, 1986:6)

- “A decision support system involving the integration of spatially referenced data in a problem solving environment”(Cowen, 1988:1554)

- “A system with advanced geo – modeling capabilities”(Koshkariov, Tikunov and Trofimov, 1989:259)

- "A form of MIS (Management Information System) that allows map display of the general information"(Devine and Field, 1986:18)

Although the above definitions cover wide range of subjects and activities best refer to geographical information. Some times it is also termed as Spatial Information Systems as it deals with located data, for objects positioned in any space, not just geographical, a term for world space. Similarly, the term 'a spatial data' is often used as a synonym for attribute data (i. e. rainfall/ temperature/ soil chemical parameters/ population data etc are referred as attribute data).

GIS Terminology

Some of the terms used in GIS are briefly explained below:

Data Planes, are discrete sets of data. For example, an imagery, a thematic map, a topographic sheet, a page of survey data each constitutes a data plane.

Themes, are maps containing different types of information. For example, a toposheet contains contours, roads, railways, boundaries of forests etc. Each of these constitutes a theme.

Registration, the themes in a given data plane are spatially related to each other. We say that the data is registered. Data in different planes may not be so readily related. Scales may vary, there may be transitional and rotational errors. This process of correction and development of an invariant spatial relationship between different data planes is called registration.

Database, the spatially registered set of data constitutes a spatial database. In addition, each spatial object has an associated attribute. This could be a name, a number, a range of values etc. For example, a contour has a number, a road has a name. Such attributes also form a part of the database. Further, there may be other data sets associated demographic data.

Spatial objects, all spatial objects can be represented by points, lines and polygons. A city is a point, a road is a line and a forest area is a polygon. The manner in which these fundamental units are represented are defined by the spatial data model. For example, we can have a chain as a set of line segments, a closed chain forms a polygon, an open chain is a line, a line segment of zero length is a point.

Scale, this is the relationship between distances on the ground and distances on a map. Scale always apply to linear measures, never to area or elevations.

Resolutions, this is the smallest element which can be distinguished in a data set. In case of imagery, this usually is the pixel size or a multiple to the pixel size. However, in a map this term can be confusing. This may be taken to mean the smallest mappable feature. However, some features are mapped by symbols even if their size is small.

5. GIS OBJECTIVES :

1. Maximise the efficiency of planning and decision making
2. Provide efficient means for data distribution and handling
3. Elimination of redundant data base - minimise duplication
4. Capacity to integrate information from many sources
5. Complex analysis/query involving geographical referenced data to generate new information.

For any application there are five generic questions a GIS can answer :

- Location - What exists at a particular location ?
- Condition - Identify locations where certain conditions exists.
- Trends - What has changed since ?
- Patterns - What spatial pattern exists ?
- Modelling - What if ………. ?

6. ELEMENTS OF A GIS :

The GIS has been divided into four elements. They are hardware, software, data, liveware. The following table gives complete details of different elements:-

S. No.	Elements of GIS	Details
1.	Hardware	<u>Type of Computer Platforms</u> Modest Personnel Computers High performance workstations Minicomputers Mainframe computers <u>Input Devices</u> Scanners Digitisers Tape drivers CD Keyboard Graphic Monitor <u>Output Devices</u> Plotters Printers
2.	Software	Input Modules Editing MRP Manipulation/ Analysis Modules Modeling Capability
3.	Data	Attribute Data Spatial Data Remote Sensing Data Global Database
4.	Live ware	♦ People responsible for digitising, implementing using GIS ♦ Trained

7. GIS DATA :

Geographical data deals primarily with two types of data : Spatial and Non-spatial. Spatial data is that which has physical dimensions and geographic locations on the surface of earth. Some examples are a river, a state boundary, a lake, a state capital etc. Non-spatial data is that qualifies the data. Attribute or Non-spatial data describes some aspects of spatial data, not specified by its geometry alone.

8. REPRESENTATION OF SPATIAL INFORMATION :

Geographical features are depicted on map by POINT, LINE & POLYGON

Point feature - A discrete location depicted by a special symbol or label. A single x,y coordinates

LINE feature - Represents a linear feature. A set of ordered x,y coordinates

POLYGON feature - An area feature where boundary encloses a homogeneous area

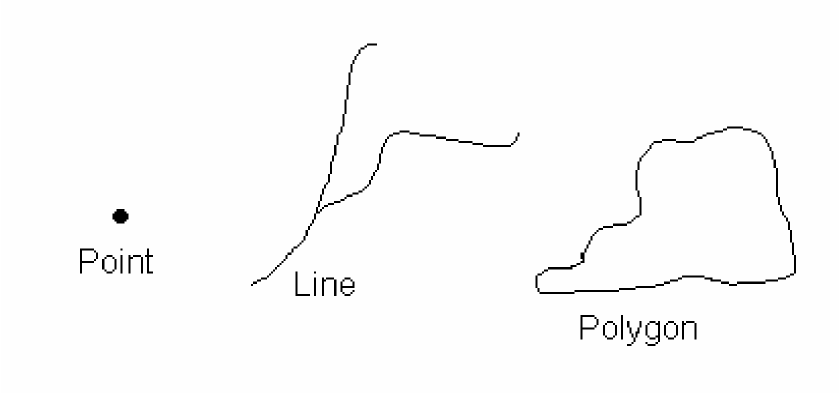

Figure 2. Elements of a map

9. REPRESENTATION OF NON-SPATIAL (ATTRIBUTE) INFORMATION

It consists of textural description on the properties associated with geographical entities. Attributes are stored as a set of numbers and characters in the form of a table. Many attribute data files can be linked together through the use of a common identifier code.

10. TOPOLOGY

Geographic data describes objects in terms of location, their attributes and spatial relationship with each other. Topology is a mathematical procedure that determines the spatial relationship of features.

Some of the advantages of topology are :

Polygon network is fully integrated, Optimal storage and reduction in redundant information neighbours are identified and polygon in polygon can be represented.

11. DATA MODELS

Conversion of real world geographical variation into discrete objects is done through data models. It represents the linkage between the real world domain of geographic data and computer representation of these features. Data models discussed here are for representing the spatial information.

Data models are of two types : Raster and Vector. In raster type of representation of the geographical data, a set of cells located by coordinate is used, each cell is independently addressed with the value of an attribute. Each cell contains a single value and every location corresponds to a cell. One set of cell and associated value is a LAYER. Raster models are simple with which spatial analysis is easier and faster. Raster data models require a huge volume of data to be stored, fitness of data is limited by cell size & output is less beautiful.

Vector data model uses line segments or points represented by their explicit x,y coordinates to identify locations. Discrete objects are formed by connecting line segments which area is defined by set of line segments. Vector data models require less storage space, outputs are appreciable, Estimation of area/perimeter is accurate and editing is faster and convenient. Spatial analysis is difficult with respect to writing the software program.

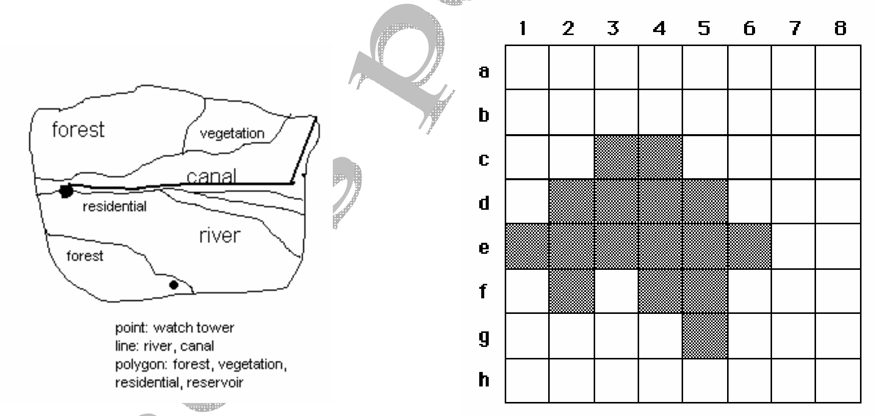

Figure 3. Vector and raster data examples

12. DATA STRUCTURES

There are number of different ways to organize the data inside the information system. The choice of data structure affects both; Data storage volume and processing efficiency. Many GIS have specialized capabilities for storing and manipulating attribute data in addition to spatial information. Three basic data structures are – Relational, Hierarchical and Network.

Relational data structure organizes the data in terms of two-dimensional tables where each table is a separate file. Each row in the table is a record and each record has a set of attributes. Each column in the table is an attribute. Different tables are related through the use of a common identifier called KEY. The information is extracted by relation which are defined by query.

Hierarchical data structure stores the data in a way that a hierarchy is maintained among the data items. Each node can be divided into one or more additional node. Stored data gets more and more detailed as one branches further out on the tree.

Network data structure is similar to hierarchy structure with the exception that in this structure a node may have more than one parent. Each node can be divided into one or more additional nodes. Nodes can have many parent. The network data structure has the limitation that the pointers must be updated every time a change is made to database causing considerable overhead.

13. ERRORS IN GIS

(i)Errors in GIS environment can be classified into following major groups :

- Age of data - Reliability decreases with age
- Map scale - Non-availability of data are proper scale or Use of data at different scales
- Density of observation - Sparsely dense data set is less reliable
- Relevance of data - Use of surrogate data leads to errors
- Data inaccuracy - Positional, elevation, minimum mapable unit etc.
- Inaccuracy of contents - Attributes are erroneously attached

(ii)Errors associated with processing

- Map digitization errors - due to boundary location problems on maps and Errors associated with digital representation of features
- Rasterisation errors - due to topological mismatch arising during approximation by grid
- Spatial Integration errors - due to map integration resulting in spurious polygons
- Generalization errors - due to aggregation process when features are abstracted to lower scale
- Attribute mismatch errors
- Misuse of logic

14. SPATIAL ANALYSIS

GIS is used to perform a variety of Spatial analysis, including overlaying combinations of features and recording resultant conditions, analyzing flows or other characteristics of networks and defining districts in terms of spatial criteria. Its uses in various fields are: facility management, planning, environmental monitoring, population census analysis, insurance assessment, health service provision, hazard mapping and many other applications. Although GIS and AM/FM Systems have similar capabilities, GIS traditionally has refer to systems that emphasize Spatial Analysis and modeling while AM/FM systems management of geographically distributed facilities. A complete GIS or Spatial Information System consists hardware, Software, humanware(i.e. trained experts in GIS).

A GIS can acquire and store data by import from external sources or by capture from maps and reports. Once in storage the data must be kept backed up, and updated when new information becomes available. Since more than 70% of the cost in GIS Project lies in data capture; the database is the primary asset of a GIS. Spatial data is collected from a variety of sources. Remotely sensed data from satellite is a primary data source. The other information coming from modern survey instruments is also a primary data source as it can be read directly into GIS similar to remote sensing data. Where as the secondary data capture involves processing information which has already been compiled but requires converting into a computer readable format by manual or automatic digitization.

15. COST IN GIS

A complete GIS needs hardware, software and expertise. Initially when the GIS hardware and Software entered the market are with limited capabilities and priced at higher cost. However due to competition the rates are reducing and it will follow for certain years. The following figure shows the relative cost with time. Only the cost of spatial data is likely to remain high.

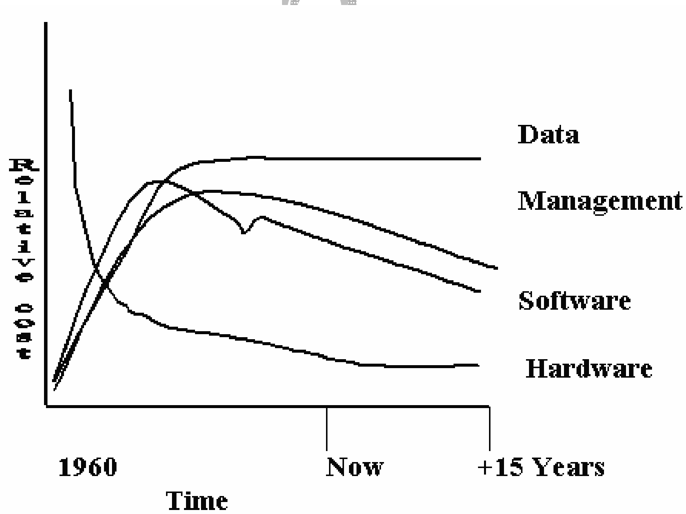

Figure 4. Cost in GIS

16. APPLICATIONS OF GIS AND REMOTE SENSING

1. Agricultural development
2. Land evaluation analysis
3. Change detection of vegetated areas
4. Analysis of deforestation and associated environmental hazards
5. Monitoring vegetation health
6. Mapping percentage vegetation cover for the management of land Degradation
7. Crop acreage and production estimation
8. Wasteland mapping
9. Soil resources mapping
10. Groundwater potential mapping
11. Geological and mineral exploration
12. Snow-melt run-off forecasting
13. Monitoring forest fire
14. Monitoring ocean productivity etc.

17. TRENDS IN GIS

1. Natural Resources Management
2. Telecom GIS
3. Automated mapping and facility management
4. Virtual 3-D GIS
5. Internet GIS
6. Spatial Multi-media etc.

18. REFERENCES

1. “Principles of Geographical Information Systems for Land Resource Assessment” by P.A.Burrough.
2. “Geographical Information Systems –Principles” Vol.1 by Goodchild.
3. International Journal of Geographical Information Science.
4. International Journal of Remote Sensing
5. Geocarto International