International Journal of Theoretical & Applied Sciences, 10(1): 34-39(2018)

ISSN No. (Print): 0975-1718
ISSN No. (Online): 2249-3247

Data: The Core of GIS

Dr. Richa Chaturvedi
Assistant Professor, Department of Geography,
IGNTU, Amarkantak (Madhya Pradesh), India

(Corresponding author: Dr. Richa Chaturvedi)
(Received 03 January, 2018, accepted 24 January, 2018)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Geographical Information System is a new branch of information system in which system (GIS software) containing geographic data and converting useful information. The ability to integrate and analyze data organized in multiple thematic layers is a heart of Geographical Information System. Hardware, software, procedure, data and users are different components in which data is essential and core of GIS because without data GIS cannot work and cannot display the result.

Keywords: Geographical Information System, Raster, Vector, Data.

I. INTRODUCTION

Today, Geographical Information System (GIS) has become an important field of academic study, one of the fastest growing sectors of the computer industry and most important, an essential component of the information technology (IT) infrastructure of modern society (Lo and Yeung, 2005) [4]. It is made from two words Geography and Information System. In Geography, we study the physical and cultural component of the earth and information system is system in which system contain all types of electronic records. So GIS is a computer based information system used to digitally represent and analyze the geospatial data or geographical data (Bhatta, 2008) [2]. The Advent of GIS technology has transformed spatial data handling capabilities and made it necessary for re-examining the role of government with respect to the supply and availability of geographic information. GIS is very different technology from producing maps using conventional cartographic methods. A geographic information system (GIS) is a computer system for capturing, storing, checking, and displaying data related to positions on Earth's surface. GIS can show many different kinds of data on one map. This enables people to more easily see, analyze, and understand patterns and relationships. In general, it is a computer system designed to capture, store, manipulate, analyze, manage, and present all types of geographical data. It is a systematic integration of computer hardware, software and spatial data, for capturing, storing, displaying, updating manipulating and analyzing, in order to solve complex management problems. GIS provide a facility to create database, to produce maps and to develop models.

In GIS modeling facility is very useful for decision making. GIS were introduced in the 1960 as computer application for handling volumes of information (Aronoff, 1989 [1], Peuquet, and Marble, 1990 [12]) the first known use of the term "Geographic Information System" was by Roger Tomlinson in his paper "A Geographic Information System for Regional Planning". The objectives of a GIS are the management (acquisition, storage, maintenance), analysis (statistical, spatial modeling) and display (graphic, mapping) of geographic data. Many people defined GIS in different sentence like GIS is a powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world for a particular set of purposes (Burrough,1986) [3]. GIS is a computer based system that provides four sets of capabilities to handle georeferenced data: data input, data management, manipulation and analysis and data output (Arnoff, 1991). Geographic information system is a much broader and more nebulous concept- a system for the input, storage, manipulation and output of geographically referenced data (Michael, 1994) [7].

The development of GIS may be described in three stages: 1. I Stage (1960 to 1980) 2. II Stage (1980 to mid-1990) and 3. III Stage (mid 1990 to present). First stage is formative stage, in this stage computer aided graphical data processing system was successful implemented. The 1960 and 1970 thus represented the important formative year of GIS. During these two decade hundreds of software packages for handling and analyzing geographic information were produced (Marble, 1980) [6]. II Stage is maturing stage of GIS technology. In this during, concept of topology were introducing in GIS environment.

Using the concept of topology, geographic data can be stored in a simple structure that is capable of representing their attributes, location, as well as their relationship (Lo and Yeung, 7) [4]. The year 1982 is important year in GIS development because in this year ESRI (Environmental System Research Institute released Arc info GIS software. This particular GIS software package was one of the first vector based GIS to use the geo-relational data model that employed hybrid approach to geographic data processing (Morehouse, 1989) [8]. Other GIS software packages are for Map info, Arc info, Spans etc. are also introducing during those time period. In form of technology and application, GIS has become more mature and advance. III stage is devoted for information infrastructure which was started in the mid-1990s.

Table 1: Evolution of GIS.

Stage of development	The formative Year	Maturing Technology	GI Infrastructure
Time frame	1960-1980	1980-mid1990s	Mid-1990s present
Technical environment	• Mainframe computer • Proprietary software • Proprietary Data structure • Mainly raster based	• Mainframe computer and microcomputer • Geo-relational data structure • Graphical users interface • New data acquisition technologies (GPS, redefinition of datum, remote sensing)	• Workstation and PCs • Network/Internet • Open system design • Multimedia • Data integration • Enterprise computing • Object-relational data model
Major users	• Government • Universities • Military	• Government • Universities • Utilities • Business • Military	• Government • Universities and school • Utilities • Business • Military • The general public
Major application areas	• Land and resources management • Census • Surveying and mapping	• Land and resources management • Census • Surveying and mapping • Facilities management • Market analysis	• Land and resources management • Census • Surveying and mapping • Facilities management • Market analysis • Utilities • Geographical data browsing

Sources: C.P.LO and A. K.W Yeung, 2005

The concept of information infrastructure emerged in the early 1990s when the United States government proposed the National Information Infrastructure (NII) initiative (NAE, 1994 and NAPA, 1998) [10, 11]. In 1994, President Clinton issued Executive Order 12906 supporting the implementing of a National Spatial Data Infrastructure (NSDI) that he defined as the technology, policies, standards and human resources to acquire, process, share, distribute and improve utilization of geospatial data. Remote sensing and GIS technique are the basic tools in the development of Spatial Data Infrastructure (SDI). GIS today is concerned not only branch of information technology but it concerned information infrastructure for local to global level (Table 1).

II. OBJECTIVE

The purpose of this paper is presenting an importance and types of data in Geographical information system (GIS).

III. DISCUSSION

GIS is made from different components like hardware, software, procedure, user and data. In all the component of GIS, data is most important. In fact, data is the key component of GIS. The two terms, data and information are often used indiscriminately but they both have specific meaning. In fact, Data is a raw fact or material. It can be described as different observation which are collected or stored. Every data contain three components in the form of position (spatial), attribute and time (temporal). The power of a GIS is in the data analysis. Without data, GIS can't do any work. Data development and maintenance is the most costly and labour intensive work in GIS field. In GIS, Data can be entered in different forms. In reality, database is central to the GIS. Overlaying layers of different data in a common geographic space allow you to integrate and derive new information.

In GIS, mainly two types of data are used first is spatial data or geospatial or georeference data second is attribute data. Spatial data represent a specific location on the earth surface. Geographic location refers to the fact that each feature has a location that must be specified in a unique way. Attribute refer to the properties of spatial entities. It is also called non spatial data because they do not occupy a specific location. This type of data describes characteristics of the spatial features. These characteristics can be quantitative and qualitative in nature. For geographic data to be useful to GIS they must be encoded in digital form and organized as a geographic database.

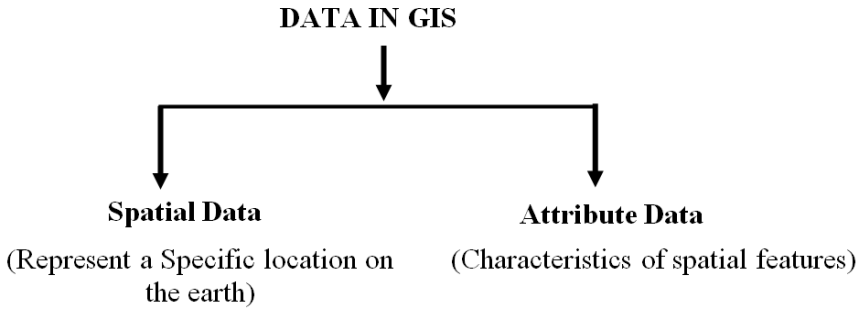

Fig. 1.

Spatial data describe the location of any feature while attribute data describe the characteristics of the spatial features. In GIS, Attribute data represent in tabular form. Once all of the desired data have been entered into a GIS system, they can be combined to produce a wide variety of individual maps, depending on which data layers are included. Often, GIS must also manipulate the data because different maps have different projections. With the help of GIS software spatial data are linked with the attribute data automatically.

When attribute data arrange in tabular form it is called database and when computer programs are used to manage and query through different procedure in a database is called database management. In table, data are arranged in row and column.

Column

Field

SEC ID	BLOCK NAME	GEOGRAPHICAL AREA (Km^2)	DENSITY	LITERACY	SEX RATIO
1.	Kaurihar	420.35	883	42.44	895
2.	Holagarh	148.46	1039	46.25	947
3.	Mauaima	150.60	992	43.12	926
4.	Soraon	134.85	1199	47.39	891

Row/ Record

Fig. 2. Database Table.

A table consists of records and each record is made from different number of fields. Each field explains a single aspect of table (Fig. 2). In GIS, no single methods are used for storing and organizing data in a database. Mainly three types of data model are used for organizing data in a database: Hierarchical, Network and Relational data model.

Hierarchical model: A hierarchical database system is one in which the DBMS supports a hierarchical structure of records organized in files at various logical levels with connections between the levels (Nag and Sengupta, 2008) [9].

In this model, data are organized into a tree like structure. Hierarchical relationship between different types of data can make it very easy to answer some question, but very difficult to answer the rest (Bhatta, 2008) [2].

Network data Model: In network data model, data are stored in network type of organization in which each record connected to several different levels records. The network model is a database model conceived as flexible way of representing objects and their relationship (Bhatta, 2008). A network structure permits rapid connection between data which physically are stored in different disk sectors. It is best suited to geographical data and frequently used in GIS application.

Relational data model: In this model, data are stored in two dimensional (2D) tables. In a relational database each records (File) has a set of attribute and a set of rows that form the table. Separate tables are linked or related to one another through a common attribute (Nag and Sengupta, 2008) [9]. A relational database contains much relationship like one to one, one to many, many to one and many to many.

There are several ways in which to bring spatial data into a GIS. Data Capture: when information Put into GIS is called data capture. Data that are already in digital form, such as images taken by satellites and most tables, can simply be uploaded into GIS. Maps must be scanned, or converted into digital information. Data capturing process is mainly based on the nature of the source data (vector/ raster), storage space and sharing of data with other system. Georeferencing is the most importing that all spatial data in GIS. Georeferencing is the process of establishing the relationship between an image (Raw, Column) coordinate system, called image space, and a map (X, Y) coordinate system, called map space. In GIS, non geo-referenced image (e.g. Satellite image, Scanned paper map) must be converted from image space to map space.

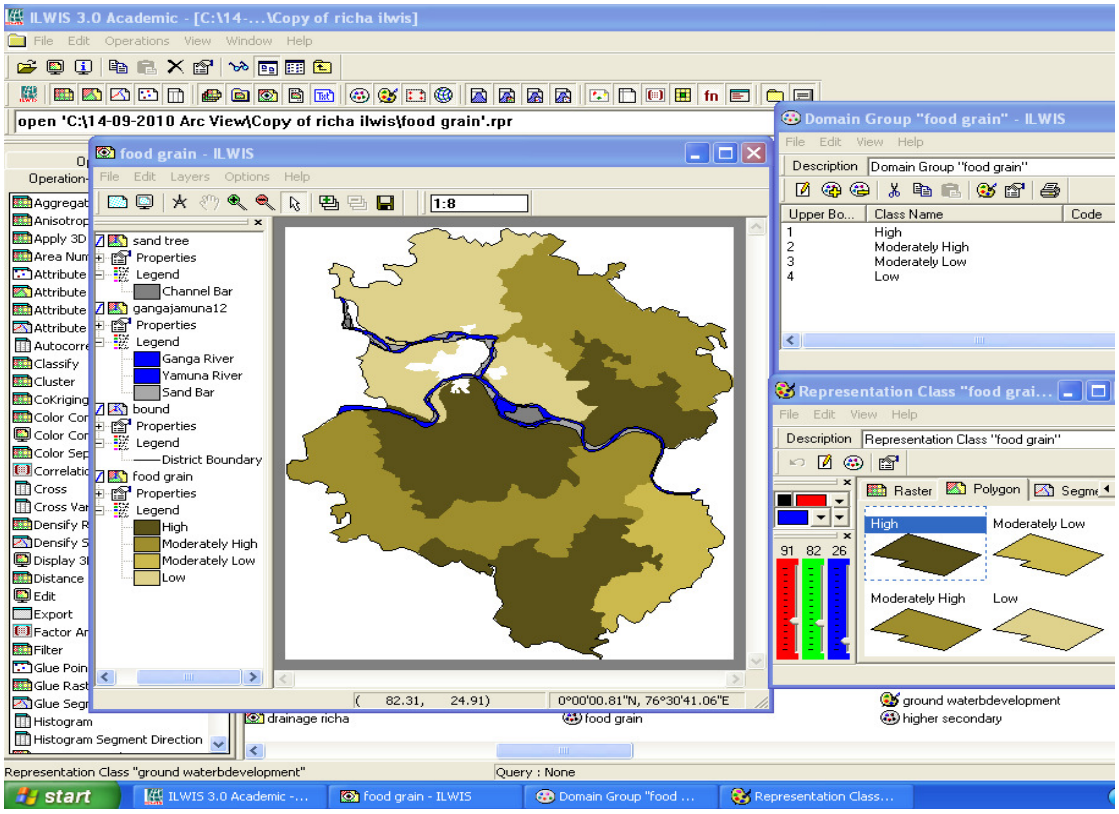

Fig. 3. Spatial feature in GIS.

Sources of Data: GIS uses the data from different sources. The primary requirement for the sources data involves knowledge of the location for the variables (Bhatta, 2008) [2]. The data sources for GIS are: conventional analog map (e.g. Topography, land use, soil, climate, Boundary, Geology etc.), Satellite remote sensing imagery (IRS, Spot, Landsat etc.), aerial remote sensing, different reports and Publication (e.g. statistical magazines, census report, town and country planning etc.), field data source and existing digital maps (Internet and World Wide Web) etc. Field data are used when existing data are not available in any formats. It is very time consuming and labour extensive work. Surveying and Global Navigation Satellite System (GNSS) are used for field data collection. GPS is a hand hold device, which are used for determine accurate position on the earth by using satellites signals.

Data input and editing: data input and data editing is the very important process for the use of GIS. Data input means inclusion of data into digital database and it is dependent on the type of data. Analog data are simply in paper form but digital data are store in computer formats and supplied by computer compatible storage device. Some possible encoding methods are: manual digitizing with the help of digitizing table, scanning, keyboard entry, text scanning, geocoding, R2V conversion (raster to vector conversion) etc. All the data always contain some types of errors and removal of error is necessary for GIS analysis. Removal of error is called editing. In fact, data editing is in response to the errors that arise during the encoding of spatial and non-spatial data. The identification of the errors is more difficult. Some GIS software removes the errors automatically but some software allows users to identify and remove or edit the errors. Some errors are: pseudo-node, undershoot, overshoot, clustered node, silver polygon, duplicate object, co-ordinate thinning, label errors and breaking crossing objects etc.

Data Structure: In GIS the data are stored in different data structure. In this process GIS convert data into different formats or GIS must be able to convert data from one structure to another. Three general features of the data within a GIS must be maintained is (a) information on the position of the features being stored, (b) topological information on the spatial relationship of the features and (c) attribute of the feature (Burrough, 1986) [3]. In GIS three types of features are represented in the form of point, linear and polygon features. Spatial data can be represented in raster and vector format. In raster format, grids are used to store the spatial features. Here, geographic space is dividing into grid cells. The linear dimensions of each cell define the spatial resolution of the data, which is determined by the size of the smallest object in the geographic space to be represented (Lo and Yeung, 2005). Some raster formats are BMP, PCX, TIFF, GIF, JPEG, PNG, GEOTIFF etc. In vector format, pair of coordinates is used to represented spatial features. Point features represented a single pair of coordinates, lines as string of coordinates with node and polygon or area are represented a closed bounded boundary of coordinates. The vector data model is an object based approach to the representation of real world features and is best used to represent discrete objects (Lo and Yeung, 2005). Some vector formats are AOI, ARC, SHP, GDF, DAT etc.

Data integration: integration of data is important aspect of GIS. In fact, it provides a facility for integration of spatial and non- spatial data (attribute data). Data integration is a method in which GIS combining of data from different sources to extract more and reliable information. The process of integrating remote sensing data into a GIS usually includes the following analytical procedures: data acquisition, data processing, data analysis, data conservation, error assessment, and final product presentation (Lunetta *et.al.,* 1991) [5]. Integration approach allows a unique identifier (ID) for spatial features and attribute database. The ID of spatial features and attribute database should be match otherwise GIS create entity attribute error. Data integration is a critical function in GIS scenario.

IV. CONCLUSION

It can be concluded that data is very important component in GIS technology. In other word, we can say that without data GIS not support any types of analysis. Spatial and non-spatial both data are very useful for solving complex problem and it provides facilities for entering to analysis of all types of data. Integration is also very important approach in GIS because without integration of spatial and attribute data GIS not became an information infrastructure.

REFERENCE

[1]. Aronoff, S., (1991). Geographical Information System: A Management perspective, WDL Publications, Ottawa, Ontario, Canada, pp. 294.

[2]. Bhatta, B., (2008). Remote Sensing and GIS, Oxford University Press, 427.

[3]. Burrough, P.A., (1986). Principles of geographical information systems for land resources assessment, Monographs on soil and Resources Survey No. 12, *Oxford University Press, New York*, 193.

[4]. Lo, C.P and Yeung, A.K.W (2005). Concept and Technique of Geographic Information Systems, Prentic-hall of India Pvt.Ltd., 80.

[5]. Lunetta, R.S., Congalton, R.G., Fenstermaker, L.K., Jensen, J.R., McGwire, K.C., and Tinney, L.R., (1991). Remote Sensing and Geographical Information System data integration: Error sources and research issues, *Photogrammetric engineering and Remote Sensing*, Vol. **57**, No.6, 687.

[6]. Marble, D.F. (1980). Computer Software for Spatial Data Handling, Ottawa, Canada: Commission on Geographic Data Sensing and Processing International Geographical Union.

[7]. Michael, F., (1994). Integrating GIS and Remote Sensing for Vegetation Analysis and Modeling: Methodological Issues, *Journal of Vegetation Science,* **5**: 615-626.

[8]. Morehouse, S. (1989). The architecture of ARC/INFO, in Auto-carto 9 Proceedings, Falls Church, VA: American Society of Photogrammetry and Remote Sensing. pp.266-277,

[9]. Nag and Sengupta, (2008). Introduction to Geographical Information System, Concept Publishing Company, New Delhi, 65.

[10]. NAE (National Academy of Engineering) (1994). Revolution in the U.S. Information Infrastructure, Washington, DC: National Academy Press.

[11]. NAPA (National Academy of Public Administration (1998). Geographic Information for the 21st Century: Building a Strategy for the Nation, Washington, DC: National Academy of Public Administration.

[12]. Peuquet, Donna J. and Marble, Duene F., (1990). What is a Geographic Information System? In Donna J. Peuquet and Marble F. Duane (eds.) introductory reading in GISs. New York.