

# Gis basics: spatial data structure and module



**ASSIGN  
BUSTER**

## GIS Basics: Spatial Data Structure and Module

### Introduction

The fourth chapter of the book, *GIS Basics*, deals with spatial data structures and models. The author organizes this chapter in a way that gives a breakdown of different elements that comprise the topic and creates the relationship between them, thus forming a background with which to understand the differences between data structures and models as well as their application in geography. The essence behind such elaborate explanations is to allow for entry and application of various data types and information into computer applications and programs that allow the utilization of the same in the form of useful information. Spatial data comprises data mostly applicable in the field of geography concerning physical elements and features from the earth and human interaction and relation to such features and structures.

### Data and information

There exist several differences between data and information. The main difference between the two is that data serves as a source of information, but information does not necessarily entail data. The presumption in this statement is that data is an ingredient of information. Data undergoes processing to create a transformation that results in a form with more meaning to the recipient, especially in terms of understanding the various aspects that prove important in making a decision. The usefulness of data in creating useful information depends on the application of such information after the conversion process. In establishing the usefulness of information,

<https://assignbuster.com/gis-basics-spatial-data-structure-and-module/>

several principles apply. These principles include relevance, reliability, timeliness, intelligibility, consistency, completeness, and convenience among others. The relevance of useful information depends on the intention of such information and the appropriate level of detail. Reliability means that the user of the information has to ensure that it is accurate and it emanates from a verifiable source, which is often acquirable via independent means. The principle of timeliness requires information to remain useful depending on the purpose for the conversion of the data. The principle of consistency incorporates the need to check with other sources while convenience means that information should be easy to handle for the user and obtain protection from malware and unsupervised access.

An information system changes data into information through various processes. The first process, viz. conversion, involves the transformation of data from one format, unit of measurement, or feature of classification to another in order to match the usage. Organization of data forms the second process, which often involves arrangement of data according to database management rules and procedure for easy access and use. Structuring means that data has to undergo formatting or reformatting so that it is acceptable to a certain software application. On the other hand, modeling involves the inclusion of spatial analysis and visualization of data so that it is useful to the user in terms of understanding and decision-making.

Organization and structuring are elements of crucial importance to the proper functioning of information systems as their absence makes turning data to information impossible.

Information organization

<https://assignbuster.com/gis-basics-spatial-data-structure-and-module/>

## 1. The data perspective of information organization

People understand information organization from four main perspectives, viz. data, relationship, operating system, and application architecture. In the data perspective, people consider the organization of data in terms of their descriptive and graphical elements. Therefore, the two elements possess distinctive features necessitating different storage requirements as well as storage options. A person thus needs to understand the correct sequence in which entities occur and build up until they eventually form a data file. A data item that falls under descriptive data is one of the most basic elements in the organization of information. It is the smallest unit of storage in a database and it goes by the term 'stored field' in the database terminology. It may appear in the form of a number, date, an expression, or character string. A group of related data items forms a record and often appears in the form of different characteristics pertaining to the same entity. A set of related record forms a data file.

The element of relation often occurs in terms of different occurrences of the same type or class of entities, regardless of whether the said entities are people, things, events, or phenomena. A collection of data items of the same type and size goes by the term 'array' and it can occur either in one dimension or two. When the organization of data takes the form or arrangement of entries in rows and columns, the final product is a table, which often applies to relational databases. A list, on the other hand, is a finite sequence of data items and it may follow a specific arrangement or lack any sort of order. A tree constitutes yet another form of data arrangement that falls under relational data in which each data item has an

attachment to one or more data items and often takes the shape of an inverted tree.

The concept of a database is one that has developed due to the introduction of computers as media for data storage. Essentially, a database and a data file contain very similar information with slight differences. The main differences that set the two apart are the type of information and medium of storage they demand. A data file contains records with the same data type and format description. A database, on the other hand, contains a group of related records organized in one or more data files with similar or different data types or formats. The type of storage for a data file is flexible enough to be manual or digital while that of the database relies strictly on computers.

These differences occur due to the capacity of a computer to process more information at a time than a person does, the ability to process different data files, create a relationship between them, and store the data files within the shortest time possible. The creation of data files often occurs manually, thus limiting the amount of processing that is applicable to a particular data type or format description at any one time. Secondly, the aim for data file processing usually touches on the creation of a particular solution and often stops after the establishment of the solution. Database processing often aims at a myriad of solutions for the different data files, the creation of relations between such data files and sometimes the formulation of predictable variables that aid organizations in the decision-making process. Thirdly, a database often complies with the central control of data in order to ease the redistribution of the same within different departments in an organization. Through computer networking, this characteristic ensures that

<https://assignbuster.com/gis-basics-spatial-data-structure-and-module/>

different departments within an organization receive the same information, depending on the need for such information. Databases are classifiable into relational - table like, network - have pointers linking them to associated files, hierarchical data - tree like relationship, and object-oriented data, which are associated with specific objects.

Graphical data, which is the second organization of information in the data perspective, has its most basic element known as basic graphical element. There exist three basic graphical elements, viz. point, line, and polygon or area. These elements can be employed to represent geographical features as single entities or collectively to form complex geographical features. The use of these basic graphical elements to represent geographical data yields vector data. The vector data is conventionally organized into layers of related themes, which yield entities such as base maps, vegetation, soil, and political boundaries among many others. Several themes of vector data about a specific geographical region constitute the spatial component of a geographical database. This method of representation is based on the object view of the real world.

Graphical data yielded by imaging devices gives another form of graphical data known as raster data. This form of data comes from the representation of geographical data in the form of picture elements (pixels). Thus, raster pixels capture a generalized representation of a given area. This form of data can also be arranged into themes, which eventually give information such as vegetation cover and land use among others. This method of representation is based on the field view of the real world.

## 2. The relationship perspective of information organization

Relationships are important in information organization and they can be either categorical or spatial based on what they describe. Categorical relationships are concerned with how individual features in a classification system are linked. Classification follows the concept of scales of measurement of which there are four distinct types, viz. the nominal scale (qualitative, non-ranking, non-numerical), ordinal scale (nominal, with ranking), interval scale (ordinal, with ranking, numerical values based on arbitrary data), and ratio scale (interval scale with numerical values based on absolute data).

Categorical relationships that use measurement scales, which involve ranking, have their data sorted into varying levels of detail. At the highest level of classification, data is broadly classified, but this aspect changes down the classification hierarchy. Descriptive data follows this system of classification. On the other hand, spatial relationships are concerned with how different features in space are linked to one another. In graphical data, one can effortlessly make out spatial relationships, but transferring these graphical spatial relationships into a database remains a challenge. Implicitly capturing spatial relationships into databases is characterized by the need for large storage and slow data computation. Yet spatial relationships are very important in geographical data handling. Thus, the aim of information organization and data structure in this context is to establish ways of handling spatial relationships with the least possible storage or computation thresholds.

## 3. Operating system perspective of information organization

<https://assignbuster.com/gis-basics-spatial-data-structure-and-module/>

In this perspective, information is arranged in the form of directories, which are special computer files that arrange other files into a hierarchy. With reference to systems that employ graphical user interfaces, directories are also known as folders. Directories fall into different levels such as root directories (top most), sub-directory (under another), and parent directory (above another). Usually, files of similar characteristics are placed in one directory such that the path that leads to a file comprises the directory name and the file name. Geographical information systems borrow the same concept, but they refer to it as the workspace. This aspect implies that in geographical information system terms, a workspace is a directory that contains files relating to a given project.

#### 4. The application architecture perspective of information organization

Today, computer software replicates a client/server system in their architecture. This system denotes a relationship among computers on telecommunication network, or several processes within a single computer. A client thus denotes a process that seeks services from one or many servers simultaneously. A server, on the other hand, is a process that provides the requested services to one or many computers at once. Information systems have many ways by which they can replicate the client/server. However, there are five commonly used ways, viz. database, file, web, groupware, and transaction servers. The aim of information organization from this perspective is to come up with means of easing the transfer of resources between clients and servers. This goal is achievable by ensuring that data is strategically placed at the appropriate location alongside similar data to ease access to the data.



## Data - Fundamental concepts

Data conventionally refers to facts. Some are meaningful to the users while others are not. The data that users consider as important is protected in arrangements known as databases. Data can be spatial or non-spatial. Spatial data is concerned with location, orientation, size, and shape. The relationship between these elements leads to spatial relationships, which is typical of spatial data. Non-spatial data, on the other hand, is conventionally linear and autonomous. The difference between spatial and non-spatial data is so pronounced that their storage and management differs.

The complex nature of spatial data and its numerous relationships necessitated the development of databases. Databases underscore the information itself, not the storage medium that holds the information. GIS is in a position to be developed and managed due to databases for they form the building blocks for GIS. This aspect is made possible by the concept of database management systems (DBMS).

A larger system of information organization and management is the repository. A repository is an arrangement developed with the aim of storing and protecting data. It could consist of several databases, which possibly contain related information or sometimes the databases can be completely unrelated. A repository is developed such that it supports the addition, retrieval, and deletion of the information contained therein. Some allow the changing or updating of data. Repositories are comparable to bank vaults since their primary purpose is to protect their content from theft or destruction. Repositories are known for two key features, viz. security and

robustness. Mostly, there is a need for a password in order to access the contents of a repository. The robustness feature also ensures that accidental destruction of data in a repository is minimized. This goal is achieved through the transactional mechanism, whereby a series of database manipulations are designed such that in case of any interruption, the database restores itself to the pre-transactional state.

### Database management systems (DBMS)

This system is a type of repository, which allows for the manipulation of a database and whose user interface allows for the administration of the database. A phonebook is the best example of a DBMS. While a repository was likened to a bank vault, a DBMS can be likened to a full-fledged bank with all its services. Thus, they provide comprehensive database manipulation functionalities.

### Discussion Points

The distinction between data and information evades many people. They often find themselves using these two terms interchangeably, that is, one in place of the other. However, it is apparent that the two terms denote very distinct concepts such that using one instead of the other is incorrect and misleading. In the light of this observation, what are the fundamental elements of information that clearly set it apart from data? In highlighting these elements, it is necessary to outline the relationship between the two concepts as well.

The advent of computers has revolutionized every field of study including geography. It is now easier to manage data, files and databases because of the improved functionality provided by computer applications that have been developed to enhance these functionalities. In the field of geography, this improvement can be seen in the development of Geographical Information Systems (GIS). With this development in mind, what are the key additions that computers have brought to the field of geography, without which, they would be considered inconsequential to this field?

In the current age, information access, sharing and transfer has become easy due technological advancement. This has led to this age being termed as the information explosion age. Thus, the development of information organization systems can be seen as an attempt at making meaningful use of the information at the disposal of humanity. The three information organization perspectives discussed in this chapter all have some relevance to geography. In your assessment, is there a particular information organization perspective that can be considered more appropriate to the field of geography? What evidence supports your answer?