Nature and Source of Data lec 2

Dr. Supriya

Assistant Professor (Guest)

Department of Geography

Patna University, Patna

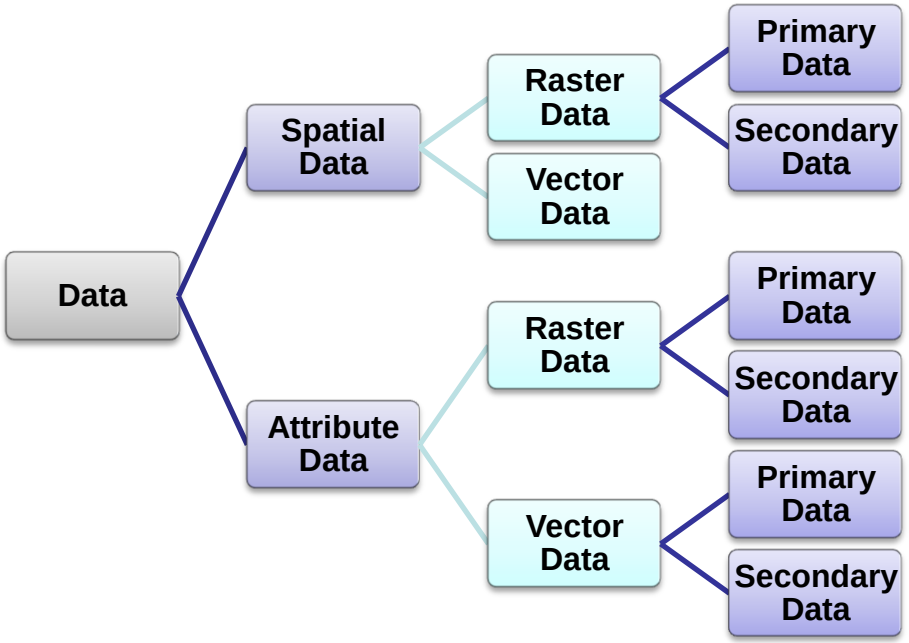
CONTENTS

- Introduction
- General classification of geographic data & its sources
- Geographic data capture from primary sources
- Geographic data capture from secondary sources
- Capturing attribute data
- Conclusion
- References

©Dr.Supriya

INTRODUCTION DATA & ITS SOURCES IN GIS

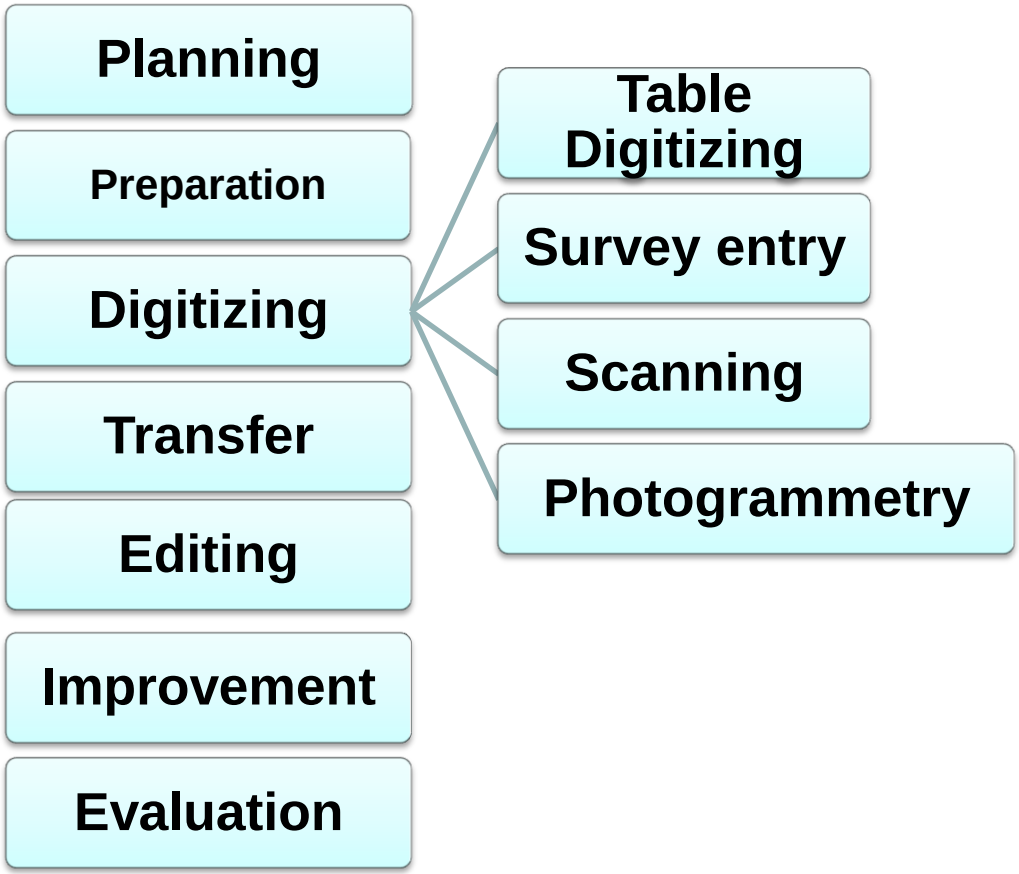
- GIS can contain a wide variety of geographic data types originating from many diverse sources. From the perspective of creating geographic databases, it is convenient to classify raster and vector geographic data as primary and secondary.



Data in GIS

©Dr.Supriya

Data collection involve a series of sequential stages:



- The workflow commences with planning, prepration, digitizing (here taken to mean a range of techniques such as table digitizing, survey entry, scanning, and photogrammetry) or transfer, editing and improvement, finally, evaluation of Data & Data Collection Methods.

©Dr.Supriya

General Classification of Geographic Data & Its Sources

- **Primary data sources** are those collected specifically for use in GIS. Typical examples of primary GIS sources include raster SPOT and IKONOS Earth satellite images, and vector building survey measurements captured using a total survey station.
- **Secondary data sources** are those that were originally captured for another purpose and need to be converted into a form suitable for use in a GIS project. Typical secondary sources include raster scanned colour aerial photographs of urban areas, and paper maps that can be scanned and vectorized.

©Dr.Supriya

Geographic Data Capture from Primary Sources

- Primary geographic data capture involves the **direct measurement** of objects.
- It can be in both raster and vector data capture methods.
- **RASTER DATA CAPTURE:**
- The most popular form of primary raster data capture is **remote sensing**. Information is derived from measurements of the amount of electromagnetic radiation, reflected, emitted, or scattered from objects.
- There are three basic aspects: **Resolution:** spatial, spectral, and temporal. All sensors need to trade off spatial, spectral, and temporal properties because of storage, processing, and bandwidth considerations.
- From the GIS perspective, **resolution** is the key physical characteristic of remote sensing systems.

©Dr.Supriya

Geographic Data Capture from Primary Sources

Resolution in GIS

- **Spatial resolution:** It refers to the size of object that can be resolved and the most usual measure is the pixel size. Satellite remote sensing systems typically provide data with pixel sizes in the range 1 meter – 1 km.
- **Spectral resolution:** refers to the parts of the electromagnetic spectrum that are measured. Since different objects emit and reflect different types and amounts of radiation, selecting which part of the electromagnetic spectrum to measure is critical for each application area.
- **Temporal resolution:** or repeat cycle, describes the frequency with which images are collected for same area.

©Dr.Supriya

Geographic Data Capture from Primary Sources

- **VECTOR DATA CAPTURE:**
- Primary vector data capture is a major source of geographic data.
- The two main branches of vector data capture are ground surveying and GPS.
- **Surveying:** Ground surveying is based on the principle that the 3D location of any point can be determined by measuring angles and distances from other known points. Surveys begin from a benchmark point. If the coordinate system of this point is known, all subsequent points can be collected in this coordinate system. If it is unknown then the survey will use a local or relative coordinate system.

©Dr.Supriya

Geographic Data Capture from Primary Sources

- **GPS:** The Global Position System (GPS) is a collection of 27 NAVSTAR satellites orbiting the Earth at a height of 12,500 miles, five monitoring stations, and individual receivers.
- The **GPS** was originally funded by the US Department of Defence, and for many years military users had access to only the most accurately data.
- Fortunately this selective availability was removed in May 2000, so that now civilian and military users can fix the x, y, z location of objects relatively easily to an accuracy of better than 10 m with standard equipment. **21 satellites** with three operational spares, **6 orbital planes, 55 degree inclinations**, **20,200 kilometer, 12 hour orbit**.

©Dr.Supriya

Geographic Data Capture from Secondary Sources

- Geographic data capture from secondary sources is the process of creating raster and vector files and databases from maps and other hardcopy documents.
- **Scanning** is used to capture raster data. Table digitizing, heads-up digitizing, stereo- photogrammetry, and COGO data entry are used for vector data.
- Data input by a scanner There are three different types of scanner generally used for data entry:

1. **Flat-bed scanner** – A common PC peripheral, it is small and inaccurate.
2. **Rotating drum scanner** – It is expensive and slow but accurate.
3. **Large-format feed scanner** – most suitable for capturing data in GIS. It is quicker, cheaper and accurate.

©Dr.Supriya

Precautions for map scanning in GIS:

- **Output Quality:** The output quality of map is very crucial in GIS, it needs to be sharp and clear.
- **Resolution:** This is the density of the raster image produced by the scanning process. The resolution of scanners is usually measured in dots per inch (dpi) as a linear measurement along the scan line.
- **Accuracy:** The accuracy of the scanned image is important if the image needs to be used in GIS.
- **Georeferencing:** The output of a map from scanner needs to be correctly referenced according to the coordinate system used in GIS.
- **Vectorization:** The output from scanned maps are often used to generate vector data. This involves, automatic or user controlled raster to vector conversion.

©Dr.Supriya

Secondary Raster & Vector data capture

- **Raster data capture using scanners**: A scanner is a device that converts hardcopy analog media into digital images by scanning successive lines across a map or document and recording the amount of light reflected from a local data source.
- **Vector Data Capture:** Secondary vector data capture involves digitizing vector objects from maps and other geographic data sources. The most popular methods are:
  1. Manual digitizing
  2. Heads-up digitizing and vectorization
  3. Photogrammetry.
  4. Coordinate Geometry data entry (COGO).

©Dr.Supriya

OBTAINING DATA FROM EXTERNAL SOURCES (DATA TRANSFER)

- One major task that needs to be faced at the start of a GIS is related to build or buy a database.
- GIS uses imported or transferred data captured by other sources. Some of these data are freely available, but many of them are sold as a commodity or agencies.
- Some examples of geographic data formats **Geographic Data Formats** One of the biggest problems with data obtained from external sources is that they can be encoded in many different formats.
- Many GIS software systems are now able to read directly **Auto CAD DWG** and **DXF**, **Microstation**, **DGN**, and **Shapefile, VPF**, and many image formats.

©Dr.Supriya

Capturing Attribute Data

- **Capturing Attribute Data:** All geographic objects have attributes of one type or another. Although attributes can be collected at the same time as vector geometry, it is usually more cost-effective to capture attributes separately.
- **Metadata** are a special type of non-geometric data that are increasingly being collected. Some metadata are derived automatically by the GIS software system (for example, length and area, extent of data layer, and count of features), but some must be explicitly collected (for example, owner name, quality estimate, and original source).
- Explicitly collected metadata can be entered in the same way as other attributes as described above.

©Dr.Supriya

Problems in creating Database

- There are several unique issues faced in capturing and managing GIS Database.
- **Cost & Time:** Capturing high quality data quickly is possible, but it is very expensive. Lower quality data can be captured over a longer of time.
- **Data Editing:** The process of data encoding is so complex that an error free data input is next to impossible. Data may have errors derived from the original source data or may be during encoding process. The process is known as data editing or 'cleaning' Includes:
  - Detection and correction of errors.
  - Re-projection.
  - Transformation and generalization.
  - Edge matching and rubber sheeting.

©Dr.Supriya

Examples of spatial errors

- **Data Conversion:** While manipulating and analyzing data, the same format should be used for all data.
- When different layers are to be used simultaneously, they should all be in vector or all in raster format.
- Usually the conversion is from vector to raster, because the biggest part of the analysis is done in the raster domain.
- Vector data are transformed to raster data by overlaying a grid with a user- defined cell size.
- Sometimes the data in the raster format are converted into vector format. Data reduction because the data storage needed for raster data is much larger than for vector data.

©Dr.Supriya

Examples of spatial errors

- **Geographic Data –** Linkages And Matching Linkages: A GIS typically links different sets.
  - **Exact Matching:** Exact matching means when we have information in one computer file about many geographic features and additional information in another file about the same set of features. Hierarchical Matching: Some types of information, however, are collected in more detail and less frequently than other types of information. For example, land use data covering a large area are collected quite frequently.
  - **Fuzzy Matching**: On many occasions, the boundaries of the smaller areas do not match those of the larger ones. This occurs often while dealing with environmental data. A GIS can carry out all these operations because it uses geography, as a common key between the data sets. Information is linked only if it relates to the same geographical area.

©Dr.Supriya

CONCLUSION

- GIS deals with geographic data ; Spatial and attribute data. Both data have derived or captured from primary and Secondary sources.
- Satellite images, and vector layers captured from total survey station are obtained from primary sources. Scanned colour maps, aerial photographs, GPS location and attribute data from different agencies obtained from secondary sources.
- Crating GIS database takes high cost and long duration of time. But reproduction and updation of these high quality data retrieve very quickly from GIS platforms.

©Dr.Supriya

Suggested Readings

- Albert , C.P.Lo. & Yeung. K.W., (2007) Concept and techniques of Geographic information system ; PHI Learning PVT Ltd. New Delhi
- Burrough,P.A. and McDonnell, R.A., (1998) : Principles of Geographic Information Systems, Oxford University Press, Oxford. De Mers, Michael N., (1999) : Fundamentals of Geographic Information Systems, John Wiley & Sons, NewYork.
- Fraser Taylor, D. R. (1991): Geographical Information System, London.
- DevidattChauniyal, Sudoor Samvedanevam Bhaugolik Soochna Pranali.
- Heywood, I. et al. (2004) : An Introduction to Geographic Information Systems, Pearson Education.
- Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W., (2001), Geographic Information Systems and Science, Wiley, Chichester.
- Siddiqui, An Introduction to Geographical Information System.

Contact me:
Mob/ WA: 9006640841
Emai:supriyavatsa52256@gmail.com
Webpage: http://bhuvankosh.com