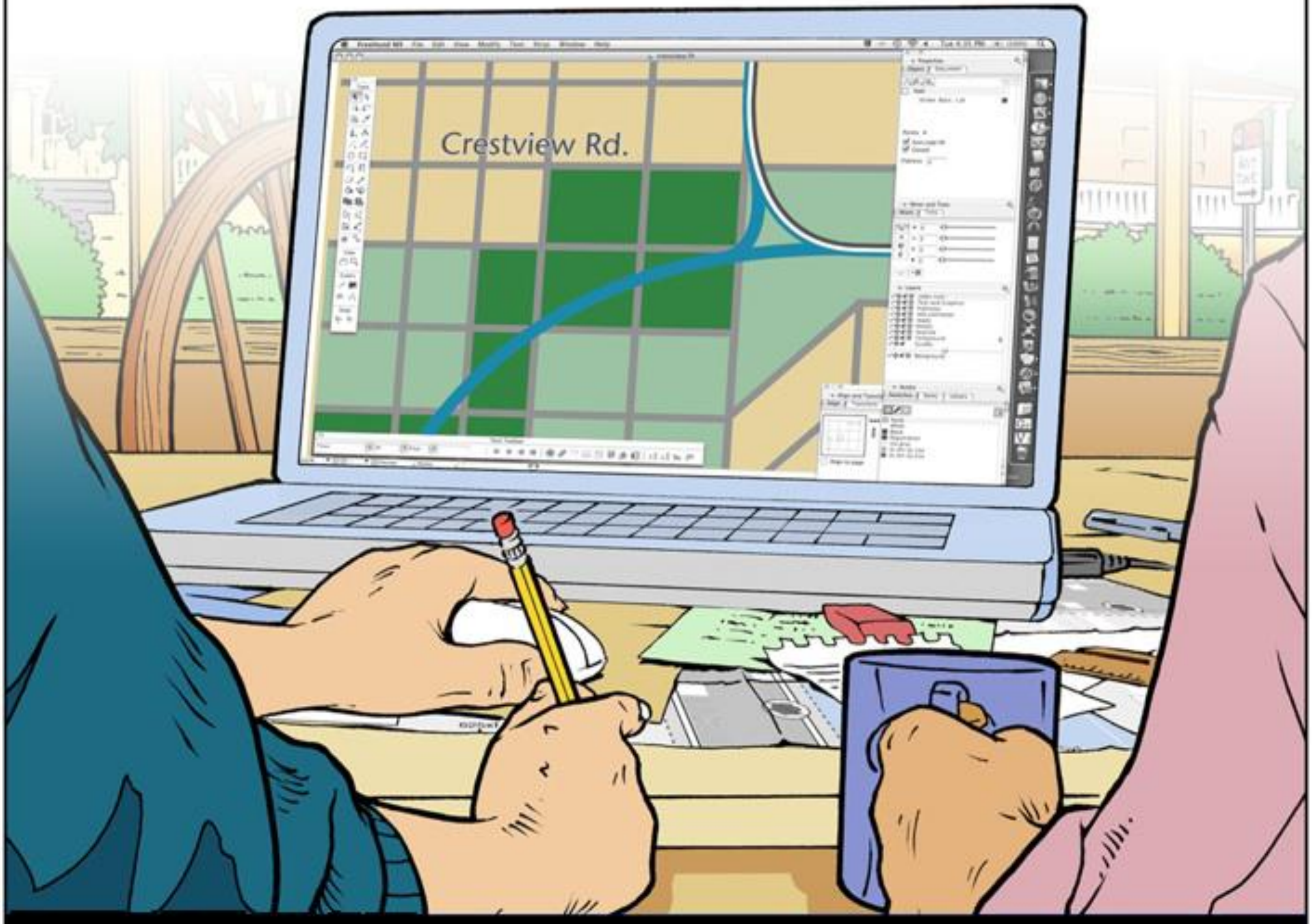
FIRST EDITION

GEOGRAPHICAL INFORMATION SYSTEM



BISHWO PRAKASH POKHAREL

SYLLABUS

Course Title: Geographical Information System
Course no.: CSC-459 **Full Marks: 60+20+20**
Credit Hours: 3 **Pass Marks: 24+8+8**
Nature of Course: Theory (3 Hrs.) + Lab (3 Hrs.)

Course Synopsis: Basic concept of Geographical Information System

Goal: The course covers about spatial data modeling and database design, capturing the real world, spatial analysis and visualization, overview of open GIS

Course Contents:
Unit 1: Introduction **6 Hrs.**
1.1 Overview, History and concept of GIS
1.2 Scope and application areas of GIS
1.3 Purpose and benefits of GIS
1.4 Functional components of GIS
1.5 Importance of GPS and remote sensing data in GIS

Unit 2: Digital Mapping Concept **3 Hrs.**
2.1 Map Concept: map elements, map layers, map scales and representation
2.2 Map projection: coordinates system and projection system

Unit 3: Spatial Data Modeling and Database Design **9 Hrs.**
3.1 Introduction to geographic phenomena and data modeling
3.2 Spatial relationship and topology
3.3 Scale and resolution
3.4 Vector, raster and digital terrain model
3.5 Spatial database design with the concept of geo-database

Unit 4: Capturing the Real World **8 Hrs.**
4.1 Different Methods of data capture
4.2 Map projection and spatial reference
4.3 Data Preparation, Conversion and Integration
4.4 Quality aspects of Spatial Data
4.5 GPS
4.6 Remote Sensing

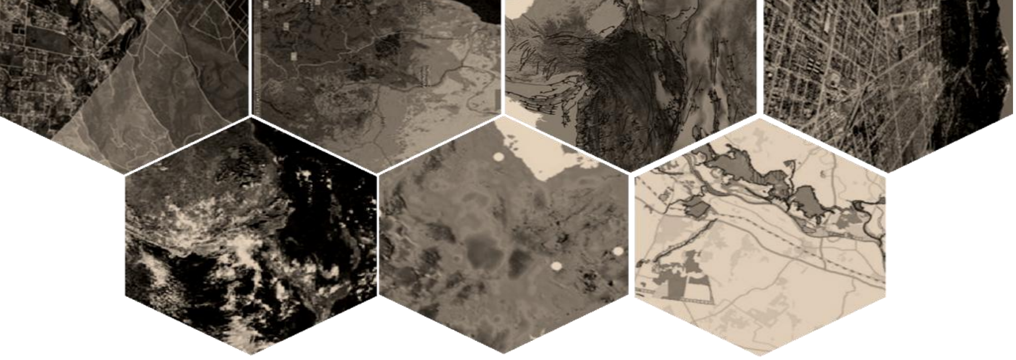
Unit 5: Spatial Analysis and Visualization **7 Hrs.**
5.1 Spatial Analysis
5.1.1 Overlay
5.1.2 Buffering
5.2 Map outputs and its basic elements

Unit 6: Introduction to Spatial Data Infrastructure **8 Hrs.**

6.1 SDI Concepts and its current trend
6.2 The concept of metadata and clearing house
6.3 Critical factors around SDIs

Unit 7: Open GIS **4 Hrs.**

7.1 Introduction of Open Concept in GIS
7.2 Open Source Software for Spatial Data Analysis
7.3 Web Based GIS System
7.4 System Analysis and Design with GIS



CHAPTER 1

Introduction

CHAPTER OUTLINE	
1.1	Overview, History and concept of GIS
1.2	Scope and application areas of GIS
1.3	Purpose and benefits of GIS
1.4	Importance of GPS and remote sensing data in GIS

Geographic information systems (GIS) (also known as Geospatial information systems) are Computer Software and hardware systems that enable users to capture, store, analyze and manage spatially referenced data. GISs have transformed the way spatial (geographic) data, relationships and patterns in the world are able to be interactively queried, processed, analyzed, mapped, modeled, visualized, and displayed for an increasingly large range of users, for a multitude of purposes.

In a general sense, the term describes any information system that integrates stores, edits, analyzes, shares, and displays geographic information. GIS applications are tools that allow users to create interactive queries (user-created searches), analyze spatial information, edit data in maps, and present the results of all these operations. Geographic information science is the science underlying geographic concepts, applications, and systems

It is a special case of information system where the database consists of observations on spatially distributed features, activities or events, which are definable in space as points, lines or area. A geographic information system manipulates data about these points, lines and areas to retrieve data for ad hoc queries and analyses.

Whether setting a new business, finding the best soil for growing bananas, or figuring out the best route for an emergency vehicle, local problems also have a geographical component GIS will give you the power to create maps, integrate information, visualize scenarios, solve complicated problems, present powerful ideas, and develop effective solutions like never before. GIS is a tool used by individuals and organizations, schools, governments, and businesses seeking innovative ways to solve their problems.

Today, GIS is a multibillion-dollar industry employing hundreds of thousands of people worldwide. GIS is taught in schools, colleges, and universities throughout the world. Professionals in every field are increasingly aware of the advantages of thinking and working geographically.

Some definitions of GIS:

- GISs are simultaneously the telescope, the microscope, the computer, and the Xerox machine of regional analysis and synthesis of spatial data. (Ron Abler, 1988).
- A Geographic Information System (GIS) is a computer based system that facilitates the phases of data entry, data analysis and data presentation especially in cases when we are dealing with geo referenced data.
- A Geographic Information System (GIS) is a computer-based mapping tool that enables geographic or spatial data capture, storage, retrieval, manipulation, analysis, modeling and presentation of the real world scenario. Basically, GIS is working on the principle of geography. Geography or GIS is now proving its potential and widely accepted by interdisciplinary experts at various levels to better manage the earth's resources.

1.1 Overview, History and concept of GIS

Overview:

A geographic information system consists of the tools and services necessary to allow one to collect, organize, manipulate, interpret, and display geographic information. A GIS is more than just the hardware and software familiar to most people; it extends to the staff that operates the system, the databases, the physical facilities, and the organizational commitment necessary to make it all work. A GIS can be defined by how it is used (e.g., a land information system, a natural resource management information system), by what it contains (spatially distinct features, activities, or events defined as points, lines, polygons, or raster grid cells), by its capabilities (a powerful set of tools for collecting, storing, retrieving, transforming, and displaying spatial data), or by its role in an organization (a map production system, a spatial analysis system, a system for assisting in making decisions regarding basic geographic questions: where is it? what is it? why is it there?).

Types of GIS:

There are a number of Geographical Information Systems (GIS) or GIS software available today. They range from high-powered analytical software to visual web applications and each of those are used for a different purpose. These can be categorized into 3 main groups of GIS:

1. **Web based GIS:**

   Web-based GIS or WebGIS are online GIS applications which in most cases are excellent data visualization tools. Their functionality is limited compared to software stored on your computer, but they are user friendly and particularly useful as they not required data download.

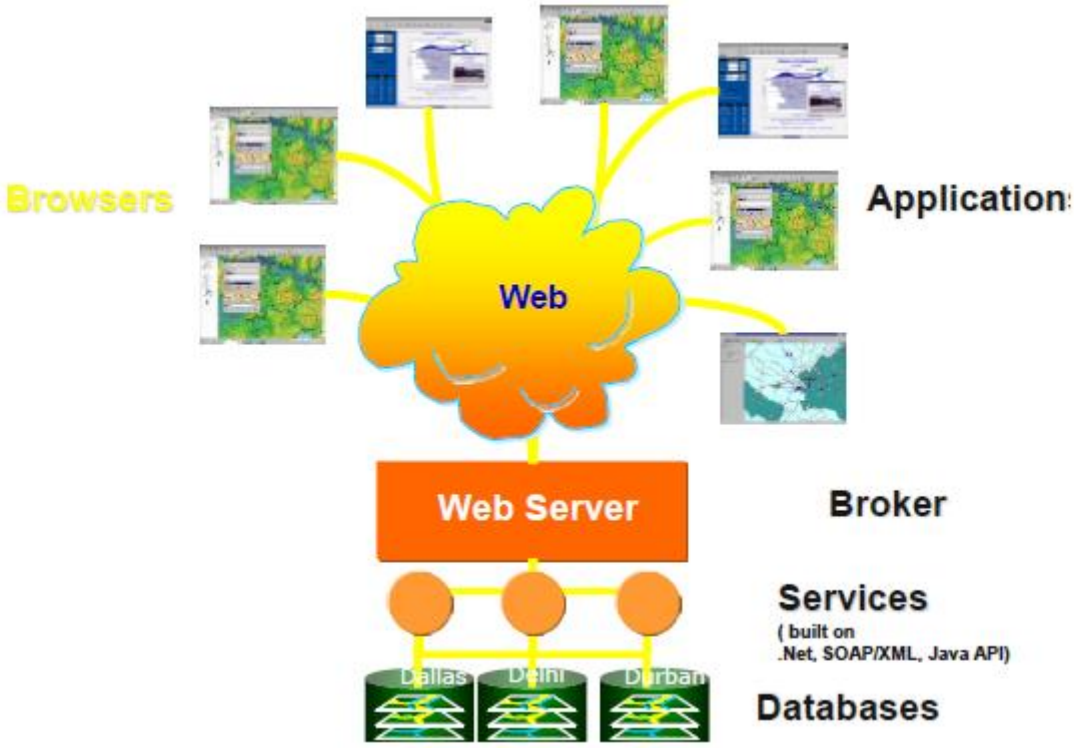


Fig: Web based GIS application architecture

2. **Geobrowser:**
   A Geobrowser is better explained with reference to a web-browser. In short, a geobrowser can be understood as an Internet Explorer for geographic information. Like the internet it allows the combination of many types of geographic data from many different sources. The biggest difference between the World Wide Web and the geographic web however is that everything within the latter is *spatially referenced.* Google-Earth is the most popular geobrowser available.

3. **Desktop GIS:**
   A GIS, or GIS software, allows us to interactively work with spatial data. A desktop GIS is a mapping software that needs to be installed onto and runs on a personal computer. ArcGIS is what ESRI refer to as a suite of products which can be tailored to our need. ArcGIS is used for a vast range of activities, covering both commercial and educational users.

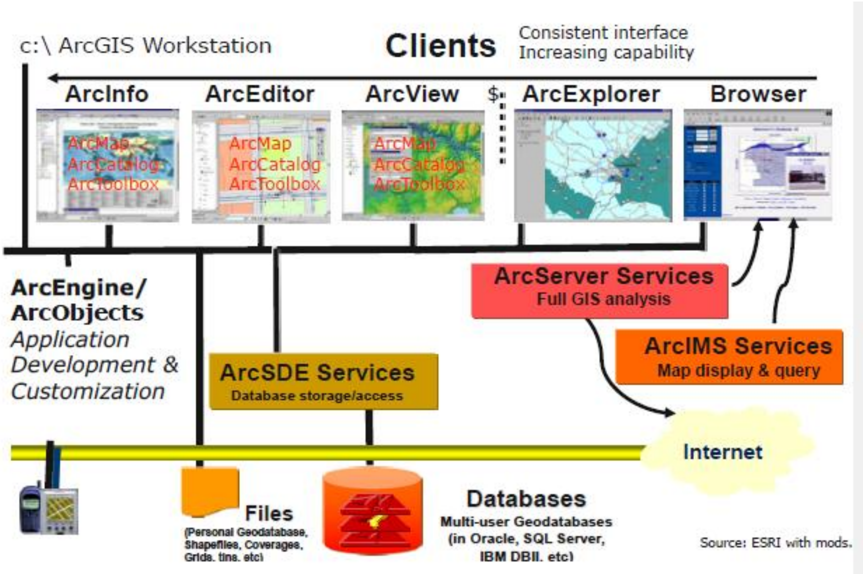


Fig: System architecture of ArcGIS

History of GIS

- GIS's origins lie in thematic cartography.
- Many planners used the method of map overlay using manual techniques.
- Manual map overlay as a method was first described comprehensively by Jacqueline Tyrwhitt in a 1950 planning textbook.
- HcHarg used blacked out transparent overlays for site selection in *Design with Nature.*
- The 1960s saw many new forms of geographic data and mapping software.
- Computer cartography developed the first basic GIS concepts during the late 1950s and 1960s.
- Linked software modules, rather than stand-alone programs, preceded GISs.
- Early influential data sets were the World Data Bank and the GBF/DIME files.
- Early systems were CGIS, MLMIS, GRID and LUNR.
- The Harvard University ODYSSEY system was influential due to its topological arc-nod (vector) data structure.
- GIS was significantly altered by the PC and the workstation.
- During the 1980s, new GIS software could better exploit more advanced hardware.
- User Interface developments led to GIS's vastly improved ease of use during the 1990s.
- During 1980s, new GIS software could better exploit more advanced hardware.

Summary:

1960's	• Canada Geographic Information System (CGIS) developed : national land inventory pioneered many aspects of GIS. • Harvard Lab for computer graphics and spatial analysis: pioneered software for data handling. • US Bureau of census developed DIME (Dual Independent Map Encoding) data format. • ESRI (Environmental Systems Research Institute) founded.

1970's	• CGIS fully operation (and still operational today) • First Lansat satellite launched (USA) • CARIS founded • USGS (U.S. Geological Survey) begins Geographical Information Retrieval and Analysis system (GIRAS) to manage and analyze large land resources databases and Digital Line Graph (DLG) data format. • ERDAS founded • ODYSSEY GIS launched (first vector GIS)
1990's	• MapInfor for windows, Intergraph, Autodesk and others. • ESRI produces ArcView and ARCGIS • $7 + billion industry

Concept of GIS

A simple five step process lets us apply GIS to any business or organizational problem that requires a geographic decision. The steps are as follows:

- **Ask:** What is the problem you are trying to solve or analyze, and where is it located?, Framing the question will help you decide what to analyze and how to present the results to your audience.
- **Acquire:** Next we need to find the data needed to complete our project. The type of data and the geographic scope of our project will help direct our methods of collecting data and conduction the analysis.
- **Examine:** We will only know for certain that our data is appropriate for study after thoroughly examining it. This includes how the data is organized, how accurate it is, and where the data came from?
- **Analyze:** Geographic analysis is the core strength of GIS. Depending on our project, there are many different analysis methods to choose from. GIS modeling tools make it relatively easy to make these changes and create new output.
- **Act:** The results of our analysis can be shared trough reports, maps, tables, and charts and delivered in printed format or digitally over a network or on the web. We need to decide on the best means for presenting our analysis, and GIS makes it easy to tailor the results for different audiences.

While creating any GIS applications we should follow the above five steps very carefully and thoroughly. Each steps are very important and have its own significance on overall application development. If we don't pay attention to any one of the above steps the overall result that we are supposed to get will be different.

1.2 Scope and application areas of GIS

- GIS technology can be used for scientific investigations, resource management, asset management, archaeology, environmental impact assessment, urban planning, cartography, criminology, geographic history, marketing, logistics, Prospectivity Mapping, and other purposes.
- For example, GIS might allow emergency planners to easily calculate emergency response times in the event of a natural disaster, GIS might be used to find

wetlands that need protection from pollution, or GIS can be used by a company to site a new business location to take advantage of a previously under-served market.

- Uses of GIS range from indigenous people, communities, research institutions, environmental scientists, health organizations, land use planners, businesses, and government agencies at all levels.
- Uses range from information storage; spatial pattern identification; visual presentation of spatial relationships; remote sensing - all sometimes made available through internet web interfaces, involving large numbers of users, data collectors, specialists and/or community participants.
- One of the primary services provided by a GIS project is the geo-referencing of various data layers for mapping projection, involving the use of satellite image data for GIS mapping including:
    - Mineral Mapping
    - Pipeline Corridor Mapping
    - Defence Mapping
    - Airport Mapping
    - Land Cover Classification
    - Urban Development
    - Pre and post 2D/3D seismic surveys
    - Environmental Impact Studies (EIS)
    - Coastal erosion studies
    - Cadastre Mapping
    - Disaster Analysis

Mineral Mapping

Satellite Imagery and aerial photography have proven to be important tools in support of mineral exploration projects. They can be used in a variety of ways. Firstly they provide geologists and field crews the location of tracks, roads, fences and inhabited areas. This is important for mapping out potential access corridors for exploration areas and considering the environmental impact of large project. These images are also useful for mapping outcrops and regolith systematic and vegetation cover across exploration blocks and over regional areas.

Pipeline Corridor Mapping

Improve Safety and Security for Pipeline and Transmission Surveys: Satellite imagery and GIS data have significant potential to reduce a number of safety and security issues for pipeline corridor planning as well as supply managers with solutions through spatial representation of data for land, lease management, exploration, production, transmissions, environmental, financial and facilities management. This information is required to make decisions that will significantly impact the operator's ability to provide the services demanded by their customers.

Defence Mapping

Defence and Security

Satellite imagery and GIS maximizes security programs which can enable local governments to better assess and understand how to develop programs to save lives, protect property and enhance the future economic stability of their communities. The current threats to a country range from incidents of terrorism and information attacks on critical infrastructure to the potential use of weapons of mass destruction and the spread of infectious diseases. Each one of these threats could cause massive casualties and disruption to a country.

Airport Mapping

Airport Mapping Database

Satellite Imaging Corporation (SIC) provides 3D airport mapping using high resolution stereo satellite imagery to support airport pre-planning and design, airport layout plans (ALPs), navigational mapping, and airport security and aviation safety operations.

3D Digital Surface Models (DSM's) and Digital Terrain Models (DTM's) can be created to support airport and aviation operations to provide details and data for the construction of airport runways, airport terminals, airport layout design, airspace analysis, obstruction surveys, facility mapping, taxiways, aprons/parking areas, 3D flight simulation for pilot training, aircraft operations, and GIS database development.

Land Cover Classification

Satellite Imagery and GIS for Land Cover and Change Detection

Satellite imagery and GIS maps for land cover, land use and its changes is a key to many diverse applications such as environment, forestry, hydrology, agriculture and geology. Natural Resource Management, Planning and Monitoring programs depend on accurate information about the land cover in a region. Methods for monitoring vegetation change range from intensive field sampling with plot inventories to extensive analysis of remotely sensed data which has proven to be more cost effective for large regions, small site assessment and analysis.

Evaluation of the static attributes of land cover (types, amount, and arrangement) and the dynamic attributes (types and rates of change) on satellite image data may allow the types of change to be regionalized and the approximate sources of change to be identified or inferred.

Urban Development

Satellite imagery for urban and land development can be used to gather strategic planning information pertaining to a district or an entire city. High resolution satellite imagery and LiDAR incorporated into a GIS (Geographic Information Systems) and CAD (Computer Aided Drafting) has gained popularity among Planners, Developers and Engineers for large scale mapping of any region for most urban and land development applications.

Information from satellite images when combined with GIS mapping is used for analysis in evaluating construction costs as well as environmental impact of alternative routes for utility and transport corridors; land cover and land use classification; identifying population groups at risk where human intervention is most needed to limit and prevent hazards during development stages.

Pre and post 2D/3D seismic surveys

No matter how remote, Satellite Imaging Corporation (SIC) can retrieve satellite images from the most difficult-to-photograph areas of the world. For heavily forested areas, we provide medium-to-high resolution "Bare Earth" DEMs. This provides weather independency, allowing us to map large areas of terrain in limited timeframes, independent of the weather and solar illumination conditions. We are also familiar with specialized retrieval methods used for satellite imagery in remote areas, highly developed areas and areas of persistent heavy cloud cover such as the tropics.

Environmental Impact Studies (EIS)

Satellite Imaging Corporation (SIC) provides satellite image data at different spatial, spectral, and temporal resolutions by using the appropriate combination of bands to bring out the geographical and manmade features that are most pertinent to your project for detecting and monitoring environmental changes.

Satellite imagery and GIS have greatly expanded opportunities for data integration, analysis, modelling, and map production for environmental monitoring and assessment. As populations grow, as countries boost their economies, as landscapes change, governments have increasingly relied on up-to-date satellite imagery and other geospatial data for applications such as environmental planning, land registration, disaster response, public health, agricultural biodiversity conservation and forestry

Coastal erosion studies

Many coastal managers are changing the way they manage coastal problems. Instead of only undertaking corrective measures, officials are moving toward prevention. Using potential models with satellite imaging technology and land cover data through GIS, managers can create scenarios for future development, as well as permitting and land use scenarios, to estimate the impacts on sensitive water bodies.
Satellite images can provide coastal management researchers and scientists with data for assessment and analysis of water temperature, salinity, phytoplankton, hydrology, shoreline changes, bathymetry, soil moisture and potential threats to our coasts.

Assessments and predictive capabilities through satellite imagery incorporated with GIS are needed to predict onset of events that may significantly affect human health, critical wetlands and ecosystems, and economic development.

Cadastre Mapping

Satellite images which forms the base for the generation of action plan maps, if used in the background of intelligent cadastral vector data, can improve the details of the thematic maps as well as action plan maps. It also helps in the monitoring of land cover changes that can be identified by detailed change detection processing procedures and implemented in the GIS cadastre mapping project.

Disaster Analysis

Satellite imagery and GIS maps can give emergency and disaster response officials a wealth of information for assessment, analysis and monitoring of natural disasters such as hurricanes, tornadoes and cyclone damage from small to large regions around the globe.

Estimates of the particular land cover classes that may be inundated by a natural disaster can enable operators and planners to better assess their region's risk and vulnerability. This information will allow for prioritizing target mitigation and preparedness activities for their area.

The use of multispectral satellite imagery is therefore critical for the separation of constituent materials within an image and for the interpretation of images of damage for pre or post-disaster assessment. View Before and after example satellite images here.

Other application areas are:

- Emergency Services (Flood Forecasting and Control)
- Environmental

- Natural Disasters (Tsunami, Flood , Plane Crash)
- Education
- Government (Socio Economic Statistics, Police,)
- Medical (Public Health Alert)
- Industry, Businesses (Tourist Information)
- Defense
- Urban management (Water Supply, Power Supply, Drainage Mgmt)
- Land Information System
- Transportation Management
- River Management
- Shipping Route Management
- Railway GIS

The following are some of those areas where GIS can be fruitfully applied:

One of the first major areas of application was in **natural resources management**, including management of

- Wildlife habitat,
- Wild and scenic rivers,
- Recreation resources,
- Floodplains,
- Wetlands,
- Agricultural lands,
- Forests.

One of the largest areas of application has been in **facilities management**. Uses for GIS in this area have included

- Locating underground pipes and cables,
- Balancing loads in electrical networks,
- Planning facility maintenance
- Local, state, and federal governments have found GIS particularly useful in **land management**.

GIS has been commonly applied in areas like:

- Zoning and subdivision planning,
- Land acquisition,
- Environmental impact policy,
- Water quality management,
- Maintenance of ownership.

More recent and innovative uses of GIS have used information based **on street-networks**. GIS has been found to be particularly useful in

- Address matching,
- Location analysis or site selection,
- Development of evacuation plans.

1.3 Purpose and benefits of GIS

What does GIS do?

- It allows users to map multiple different sources of geographic data within a single computerized environment.
- Different data sources are usually treated as layers, which may be reordered and switched on and off at will, set to varying transparencies, and manipulated through tools such as zooming, panning, and sometimes rotating.
- It allows users to employ many different & powerful tools to analyze the spatial distribution of their data.
- This spatial analysis can provide a route into discovering and unlocking previously unseen patterns in our data, shedding new light on unknown aspects of the past.
- It also allows users to produce paper and electronic maps for inclusion in their work and for the dissemination of their results to the wider archaeological, historical and public communities.
- Depending on the GIS software used, this might include animations or interactive maps delivered over the internet.
- It is easy to disguise poor quality data by entering it into a GIS, resulting in maps that convey an undue authority.
- Electronic maps output from a GIS may need tweaking in picture editing software to produce the best results for publication.
- Many of the tools provided by GIS packages can be applied to data where their use would not be appropriate.
- GIS tools can be used to support distinctly spurious ideas and to cloud what might normally be seen as unconvincing conclusions.
- Finally, the most widespread GIS packages only express a fraction of the true spatial complexity of the world around us.
- This is because the third dimension is only just starting to be properly represented and, furthermore, time is entirely absent from the majority of conventional GIS packages.
- Time is what separates geography from geometry, so current GIS software will remain incomplete until their developers begin to integrate temporality

GIS can be used to:

- Explain events
- Planning Strategies
- Integrate Information
- Solve complicated problems
- Predict outcomes
- Create “smart” maps
- Visualize scenarios
- Present powerful ideas

Who Uses GIS?

- Police and Law Enforcement Agencies
- Planning Strategies
- Foresters

- Industry
- Environmental Engineers
- Real Estate Professionals
- Telecommunications Professionals
- Emergency Response Organizations
- Local and Federal Government
- Health
- Transportation
- Geographers
- Market Developers

Benefits of GIS:

1. Cost savings resulting from greater efficiency. These are associated either with carrying out the mission (i.e., labour savings from automating or improving a workflow) or improvements in the mission itself. A good case for both of these is Sears, which implemented GIS in its logistics operations and has seen dramatic improvements. Sears considerably reduced the time it takes for dispatchers to create routes for their home delivery trucks (by about 75%). It also benefited enormously in reducing the costs of carrying out the mission (i.e., 12%-15% less drive time by optimizing routes). Sears also improved customer service, reduced the number of return visits to the same site, and scheduled appointments more efficiently.

2. Better decision making. This typically has to do with making better decisions about location. Common examples include real estate site selection, route/corridor selection, zoning, planning, conservation, natural resource extraction, etc. People are beginning to realize that making the correct decision about a location is strategic to the success of an organization.

3. Improved communication. GIS-based maps and visualizations greatly assist in understanding situations and storytelling. They are a new language that improves communication between different teams, departments, disciplines, professional fields, organizations, and the public.

4. Better geographic information recordkeeping. Many organizations have a primary responsibility of maintaining authoritative records about the status and change of geography (geographic accounting). Cultural geography examples are zoning, population census, land ownership, and administrative boundaries. Physical geography examples include forest inventories, biological inventories, environmental measurements, water flows, and a whole host of geographic accountings. GIS provides a strong framework for managing these types of systems with full transaction support and reporting tools. These systems are conceptually similar to other information systems in that they deal with data management and transactions, as well as standardized reporting (e.g., maps) of changing information. However, they are fundamentally different because of the unique data models and hundreds of specialized tools used in supporting GIS applications and workflows.

5. Managing geographically. In government and many large corporations, GIS is becoming essential to understand what is going on. Senior administrators and executives at the highest levels of government use GIS information products to communicate. These products provide a visual framework for conceptualizing, understanding, and prescribing action. Examples include briefings about various geographic patterns and relationships including land use, crime, the environment, and defence/security situations. GIS is increasingly being implemented as enterprise information systems. This goes far beyond simply spatially enabling business tables in a DBMS. Geography is emerging as a new way to organize and manage organizations. Just like enterprise-wide financial systems transformed the way organizations were managed in the '60s, '70s, and '80s, GIS is transforming the way that organizations manage their assets, serve their customers/citizens, make decisions, and communicate. Examples in the private sector include most utilities, forestry and oil companies, and most commercial/retail businesses. Their assets and resources are now being maintained as an enterprise information system to support day-to-day work management tasks and provide a broader context for assets and resource management.

Advantages of GIS

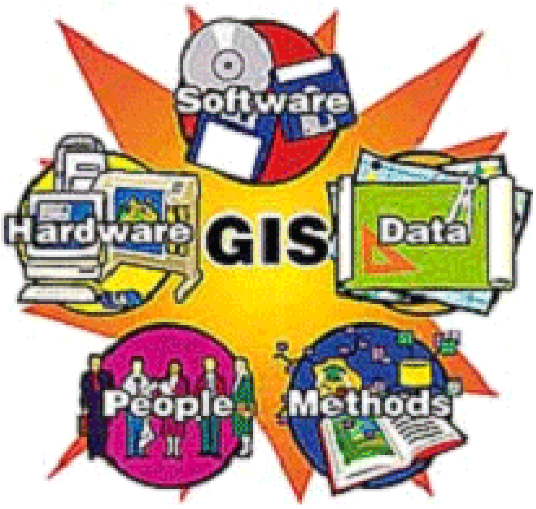
- Exploring both geographical and thematic components of data in a holistic way
- Stresses geographical aspects of a research question
- Allows handling and exploration of large volumes of data
- Allows integration of data from widely disparate sources
- Allows analysis of data to explicitly incorporate location
- Allows a wide variety of forms of visualization

Limitations of GIS

- Data are expensive
- Learning curve on GIS software can be long
- Shows spatial relationships but does not provide absolute solutions
- Origins in the Earth sciences and computer science. Solutions may not be appropriate for humanities research

Functional Components of GIS:

An operational GIS also has a series of components that combine to make the system work. These components are critical to a successful GIS.



A working GIS integrates five key components:

1. **Hardware:** Hardware is the computer system on which a GIS operates. Today, GIS software runs on a wide range of hardware types, from centralized computer servers to desktop computers used in stand-alone or networked configurations.
2. **Software:** GIS software provides the functions and tools needed to store, analyze, and display geographic information.
3. **Data:** Perhaps the most important component of a GIS is the data. Geographic data and related tabular data can be collected in house, compiled to custom specifications and requirements, or occasionally purchased from a commercial data provider. A GIS can integrate spatial data with other existing data resources, often stored in a corporate DBMS. The integration of spatial data (often proprietary to the GIS software), and tabular data stored in a DBMS is a key functionality afforded by GIS.
4. **People:** GIS technology is of limited value without the people who manage the system and develop plans for applying it to real world problems. GIS users range from technical specialists who design and maintain the system to those who use it to help them perform their everyday work. The identification of GIS specialists versus end users is often critical to the proper implementation of GIS technology.
5. **Methods:** A successful GIS operates according to a well-designed implementation plan and business rules, which are the models and operating practices unique to each organization.

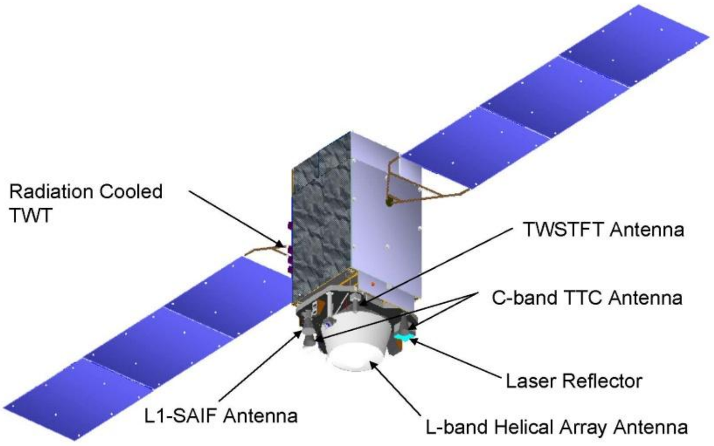
1.4 Importance of GPS and remote sensing data in GIS

There are many ways in which geographical information sciences can help with fieldwork projects, these are just a few of the possible applications:

- Logistics : Planning routes and navigation.
- Research: mapping vegetation, wildlife, urbanization, soils and geological features.
- Monitoring: Data logging of fire extents, forest loss, river channel changes
- Conservation applications: assessing biodiversity, park zonation, impact assessment
- Technology transfer: training local technical staff, donating hardware and software
- Education: maps for displays, involving school children with field work.

Global Positioning System (GPS):

- The Global Positioning System (GPS) is a satellite-based navigation system made up of a network of 24 satellites placed into orbit by the U.S. Department of Defense.
- GPS was originally intended for military applications, but in the 1980s, the government made the system available for civilian use.
- GPS works in any weather conditions, anywhere in the world, 24 hours a day.
- There are no subscription fees or setup charges to use GPS.
- Accuracy can be pinpointed to within one (1) meter with special military-approved equipment.
- GPS equipment is widely used in science and has now become sufficiently low-cost so that almost anyone can own a GPS receiver.

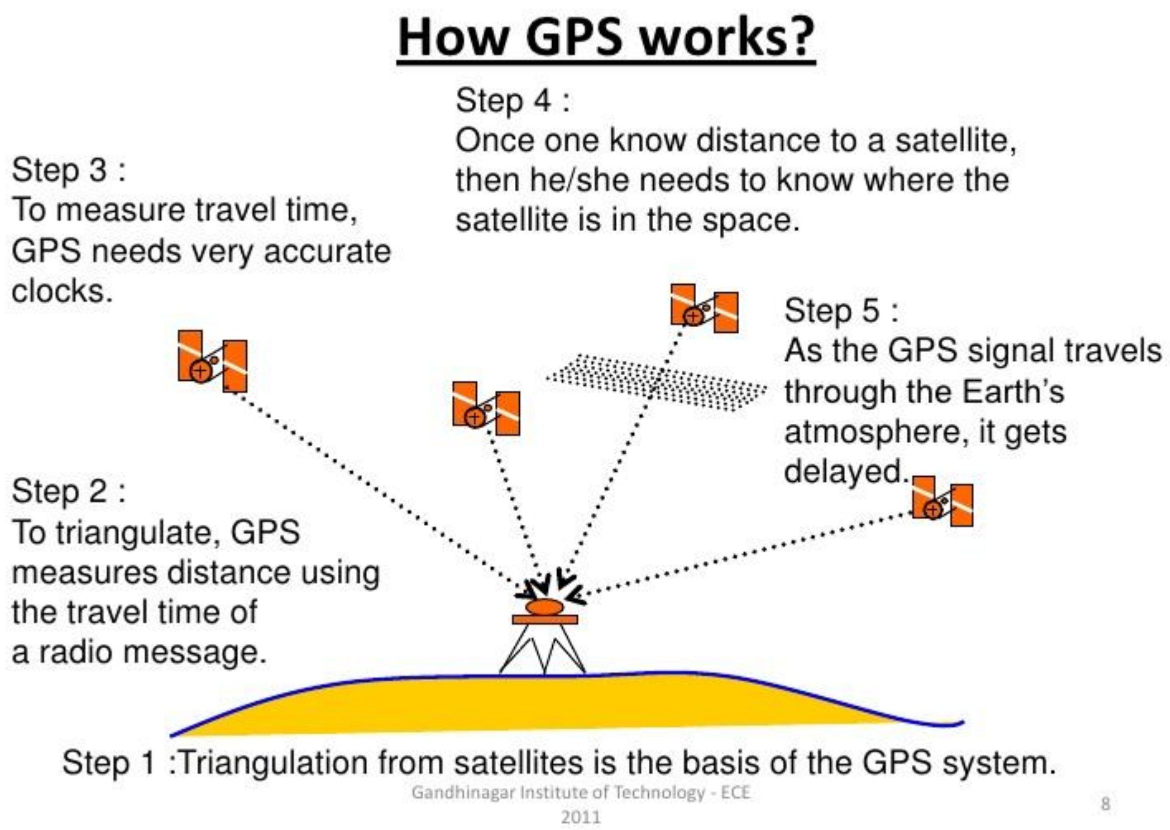


GPS Technology:

- The 24 satellites that make up the GPS space segment are orbiting the earth about 12,000 miles above us.
- They are constantly moving, making two complete orbits in less than 24 hours.
- These satellites are travelling at speeds of roughly 7,000 miles an hour.
- GPS satellites are powered by solar energy & have backup batteries. Why?
- They have backup batteries onboard to keep them running in the event of a solar eclipse, when there's no solar power.
- Small rocket boosters on each satellite keep them flying in the correct path.

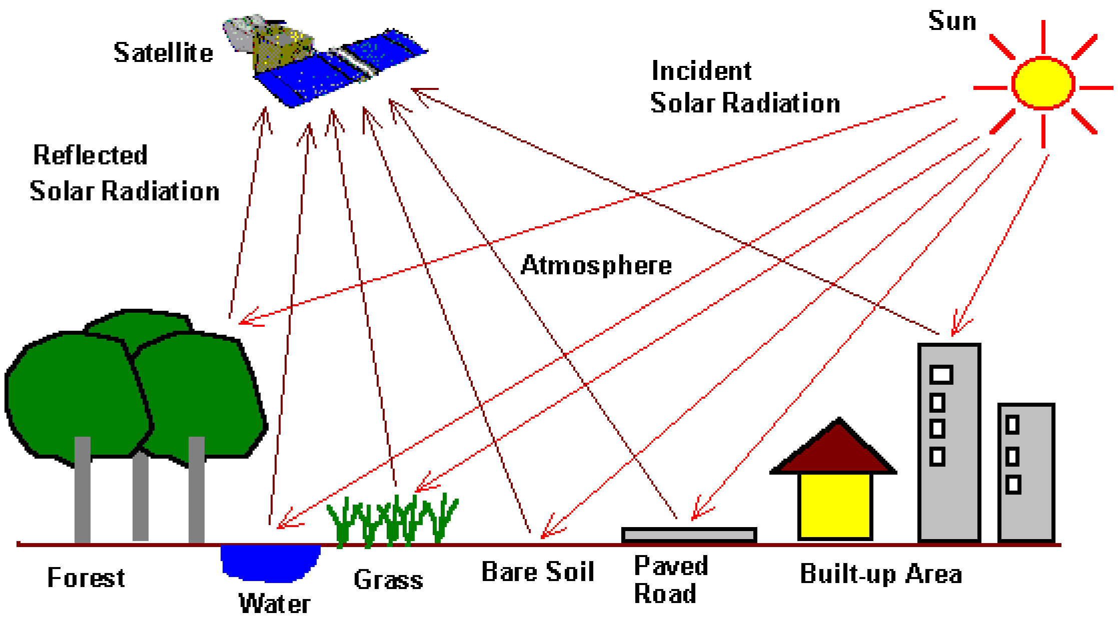
How GPS works:

- GPS satellites broadcast radio signals providing their location, status, and precise time {t_1} from on-board atomic clocks.
- The GPS radio signals travel through space at the speed of light more than 299,792 KM/Second.
- A GPS device receives the radio signals, nothing their exact time of arrival and uses these to calculate its distance from each satellite in view.
- Once a GPS device knows its distance from at least 4 satellites, it can use geometry to determine its location on Earth in three dimensions.



Remote Sensing:

- Remote Sensing is methods of observing features from a distance, such as photography or infra-red scanning
- Remote sensing technologies are used to gather information about the surface of the earth from a distant platform, usually a satellite or airborne sensor.
- Most remotely sensed data used for mapping and spatial analysis is collected as reflected electromagnetic radiation, which is processed into a digital image that can be overlaid with other spatial data.



Satellite Remote Sensing

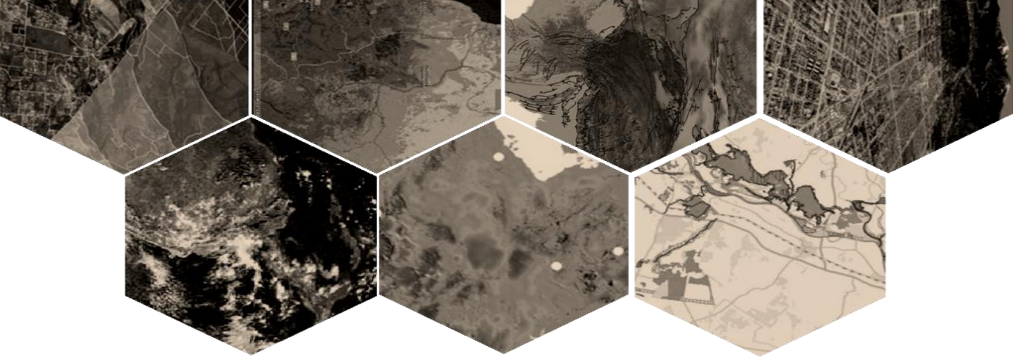
- In satellite remote sensing of the earth, the sensors are looking through a layer of atmosphere separating the sensors from the Earth's surface being observed.
- Hence, it is essential to understand the effects of atmosphere on the electromagnetic radiation travelling from the Earth to the sensor through the atmosphere.
- The atmospheric constituents cause wavelength dependent absorption and scattering of radiation.
- These effects degrade the quality of images. Some of the atmospheric effects can be corrected before the images are subjected to further analysis and interpretation.
- Remote sensing systems are often designed to operate within one or more of the atmospheric windows.
- These windows exist in the microwave region, some wavelength bands in the infrared, the entire visible region and part of the near ultraviolet regions.
- Although the atmosphere is practically transparent to x-rays and gamma rays, these radiations are not normally used in remote sensing of the earth

Optical and Infrared Remote Sensing

- In Optical Remote Sensing, optical sensors detect solar radiation reflected or scattered from the earth, forming images resembling photographs taken by a camera high up in space.
- The wavelength region usually extends from the visible and near infrared (commonly abbreviated as VNIR) to the short-wave infrared (SWIR).
- Different materials such as water, soil, vegetation, buildings and roads reflect visible and infrared light in different ways.
- They have different colours and brightness when seen under the sun.
- The interpretation of optical images require the knowledge of the spectral reflectance signatures of the various materials (natural or man-made) covering the surface of the earth.

Remote Sensing Images

- Remote sensing images are normally in the form of digital images.
- In order to extract useful information from the images, image processing techniques may be employed to enhance the image to help visual interpretation, and to correct or restore the image if the image has been subjected to geometric distortion, blurring or degradation by other factors.
- There are many image analysis techniques available and the methods used depend on the requirements of the specific problem concerned.
- In many cases, image segmentation and classification algorithms are used to delineate different areas in an image into thematic classes.
- The resulting product is a thematic map of the study area.
- This thematic map can be combined with other databases of the test area for further analysis and utilization.



CHAPTER 2

Digital Mapping Concept

CHAPTER OUTLINE	
2.1	Map concept: map elements, map layers, map scales and representation
1.2	Map projection: coordinates system and projection system

2.1 Map concept: map elements, map layers, map scales and representation

Map Concept:

Maps are the marks on a paper that stands for definable things on the earth's surface. A representation usually on a flat surface, of the whole or a part of an area. It is any concrete or abstract image of the distributions and features that occur on or near the surface of the earth or other celestial bodies.

The term 'map', however, in non-geography uses does not necessarily refer to a representation but to how things are arranged or how they relate to one other. For whatever reason, at geographic scales, 'map' means a representation of the earth and not earth's patterns themselves and it usually refers to a graphic representation, although the term 'map' can be used more broadly to refer to any representation of geographic space. To reach a graphic representation, there must be a mental conception (or representation) of the world. It determines how we map, and maps in turn influence the mental representation.

Map Resolution:

It refers to how accurately the location and shape of the map features can be depicted for a given map scale. In large-scale maps the resolution is greater because the reduction factors used to put the real-world features on a map is less. As a map scale decreases, features are simplified, smoothed or not represented at all. Features such as roads and streams must be represented as lines not areas. Millions of maps are produced and used annually throughout the world by scientists, scholars, governments, and business to meet environmental, economic, political, and social needs.

Maps gain value in three ways:

As a way of recording and storing information:

Governments, business, and society as large must store large quantities of information about the environment and the location of natural resources, capital assess, and people.

As a mean of analyzing distributions and spatial patterns:
Maps let us recognize spatial distribution and relationships and make it possible for us to visualize and hence conceptualize patterns and processes that operate through space.

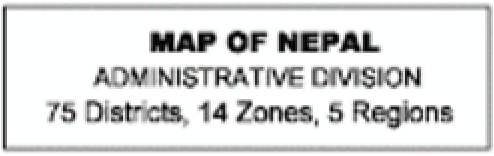
As a method of presenting information and communication findings:
Maps allow us to convey information and findings that are difficult to express verbally.

Virtual Maps vs. Real Maps
Real map: A hard copy or conventional map.
Virtual map: Information that can be converted into a real map, i.e. information on a computer screen, mental images, field information, notes, and remote sensing information.

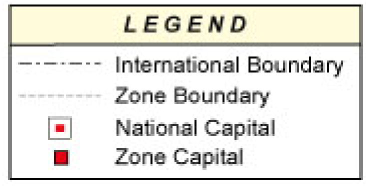
Map Elements:
Maps are the primary tools by which spatial relationships are visualized. Maps therefore become important documents. There are several key elements that should be included each time a map is created in order to aid the viewer in understanding the communications of that map and to document the source of the geographic information used.

Map usually contains the following elements:

- **Title and subtitle:** Usually draws attention by virtue of its dominant size; serves to focus attention on the primary content of the map. It should be an answer to "What, Where and When".
  Tips: Never underline a title or a subtitle and never put a colon after a title



- **Legend:** The principal reference to the map symbols; subordinated to the title. However, this is still a key element for map reading; describing all unknown or unique map symbols used.
  Tips: Only the word "legend" should be written on your map not like "map legend" , or "Nepal legend" etc.



- **Map scale:** It provides the reader with important information regarding linear relations on the map. A scale can be numerical for example 1:50000 or graphical.
  Tips: The dimension and thickness of a graphical scale has to be adapted to the map content.

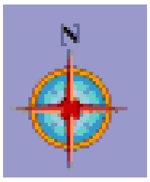
- **Credits:** Can include the map source, the author, indication of the reliability of accuracy of the map, dates, or other explanatory materials.
  Tips: Credits should always be written smallest as possible but nevertheless readable and be placed in a box without a frame.

- **Mapped areas:** Objects, land, water and other geographical features important to the purpose of the map.

- **Map symbols:** Wide variety of forms and functions; the most important element of the map, along with the geographic areas rendered.

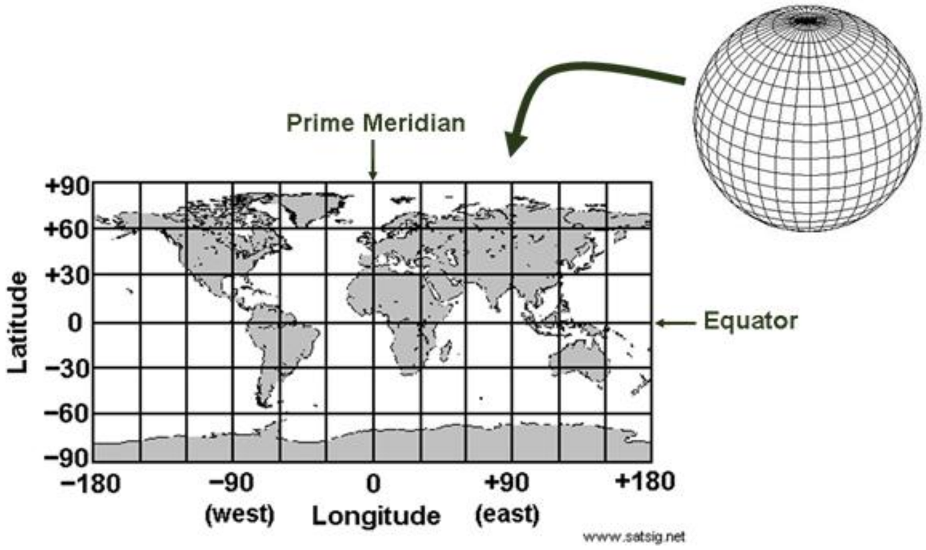
- **Place name and labeling:** The chief means of communicating with maps; serve to orient the reader on the map and provide important information regarding its purpose.
  Tips: Use the same font for the map frame, the map layout, and the map content.

- **North arrow:** According to the rules, each map should have a north arrow but if the map is north oriented, or if the geographical co-ordinate are already on the map the north arrow can be omitted.
  Tips: The north arrow must be well readable, but not be too dominant on the map.



- **Border and Neatlines:** Both optional; borders can serve to restrain eye movements. Neatlines are finer than borders, drawn inside them and often intra-parallelism, rendered as part of the graticule; used mostly for decoration.

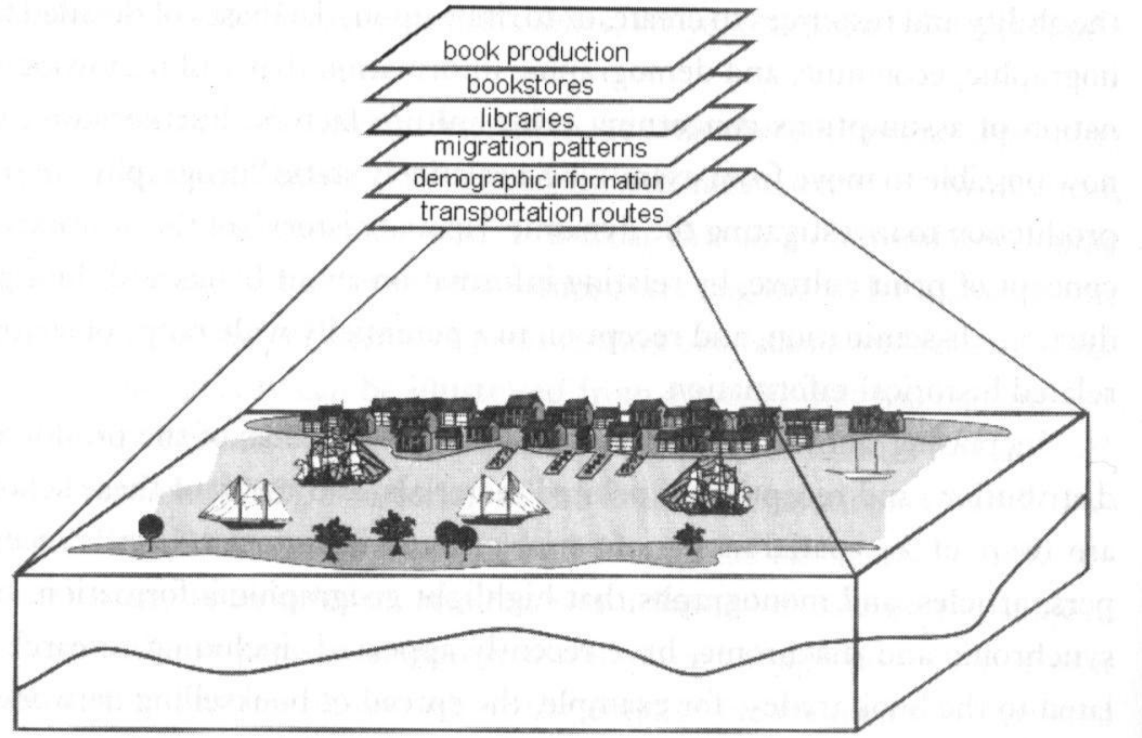
- **Graticule:** Often omitted in maps today; should be included if the location information is crucial to the map purpose. For example: into topographical maps.



Map Layers:

A map layer is GIS database containing groups of point, line or area (polygon) features representing a particular class or type of real-world entities such as customers, streets, or postal codes. A layer contains both the visual representation of each feature and a link from the feature to its database attributes. Maps in a geographic information system are made by combining multiple layers.

Fig: Combine a point layer of landmarks, a line layer of streets, and an area layer of ZIP Codes to create the GIS map shown on the right.



- Data on different themes are stored in separate layers.
- As each layer is geo-referenced, layers from different sources can easily be integrated using location.
- This can be used to build up complex models of the real world from widely disparate sources.

Map scale and representation:

Naturally it is impossible for real world features to be drawn on the map as large as their true size. Therefore in order to represent the real world, maps are made to a specific scale. Map scale is defined as the ratio of the distance between two points on the map to the corresponding distance on the ground. Maps come in a variety of scales. Large scale maps cover a small area with great detail and accuracy, while small scale maps cover a large area in less detail.

Map scale including bar, verbal and fractional scales As shown in this image, map scales can be expressed as a verbal statement, as a fraction or ratio and finally as a graphic or bar scale. Such scale expressions can be used to find the ground distance between any features from conversion of the corresponding map distance measurement.

Verbal Scale:
"1 centimeter on the map represents 500m on the ground" is a verbal scale. Clearly here a distance of 1cm on the map corresponds to 500m on the earth's surface. So if you plan a route with a total distance of 22cm on the map, that would imply that you'll be traveling (22cm x 500m) / 1cm = 11000m or 11km on the ground.

Representative Fraction (RF) - Fractional Scale - Ratio Scale:

1:50000 represents the map scale as a mathematical ratio or fraction, thus the name ratio scale or fractional scale. 1:50000 can be shown as 1/50000 as well. Here such a scale means that one unit of measurement on the map is equal to 50000 of the same unit on the ground. Such a unit can be anything such as centimeter, meter, feet, inches, your finger length, half a length of a pencil, etc. Also we can say that any distance on the map is 1/50000 of its true value on the ground.

Therefore 1cm on the map is equal to 50000cm on the ground,
i.e. 1cm on the map is equal to (50000cm x 1m) /100cm
= 500m or 0.5km on the ground.
Again a 22cm route on the map can be calculated to be equal 22 x 50000cm
= 1100000cm on the ground
or (1100000cm x 1m) / 100cm
= 11000m.

2.2 Map projection: coordinates system and projection system

Map projections are attempts to portray the surface of the earth or a portion of the earth on a flat surface. Some distortions of conformality, distance, direction, scale, and area always result from this process. Some projections minimize distortions in some of these properties at the expense of maximizing errors in others. Some projection are attempts to only moderately distort all of these properties. We need to choose a projection that will MINIMIZE distortion in our area and be best suited for our application.

1. Maps are flat but they represent curved surfaces. Transforming 3D space onto a 2D map is called a projection.
2. Projections are mathematical expressions which convert data from a geographic location (latitude, longitude) on a sphere or spheroid to a representative location on a flat surface.
3. Projection always causes distortion in one or more ways: shape, area, distance, direction. Therefore, one must choose which characteristic to be accurate at the expense of the others.

Conformality: When the scale of a map at any point on the map is the same in any direction, the projection is conformal. meridians (lines of longitude) and parallels (lines of latitude) intersect at right angles. Shape is preserved locally on conformal maps.

Distance: A map is equidistant when it portrays distances from the center of the projection to any other place on the map.

Direction: A map preserves direction when azimuths (angles from a point on a line to another point) are portrayed correctly in all directions.

Scale: Scale is the relationship between a distance portrayed on a map and the same distance on the Earth.

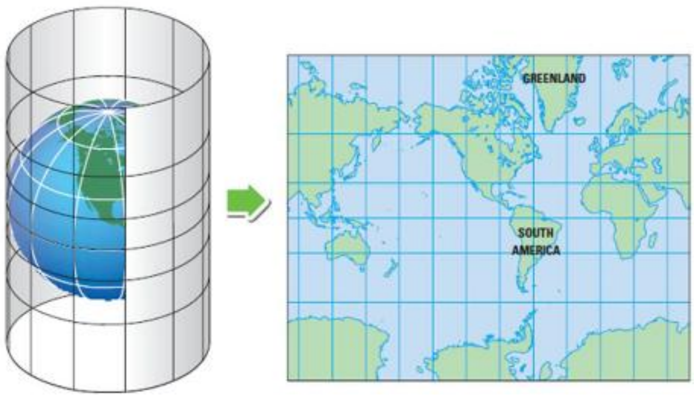
Area: When a map portrays areas over the entire map so that all mapped areas have the same proportional relationship to the areas on the Earth that they represent, the map is an equal-area map.

Classification of map projections:

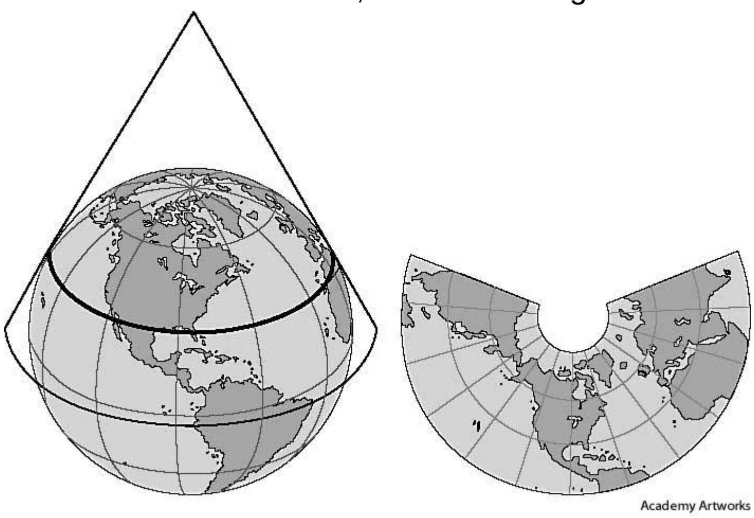
The map projection can be onto a flat surface or a surface that can be made flat by cutting, such as a cylinder or a cone. If the globe, after scaling, cuts the surface, the projection is called sec ant. Lines where the cuts take place or where the surface touches the globe have no projection distortion.

Map projections fall into three general classes:

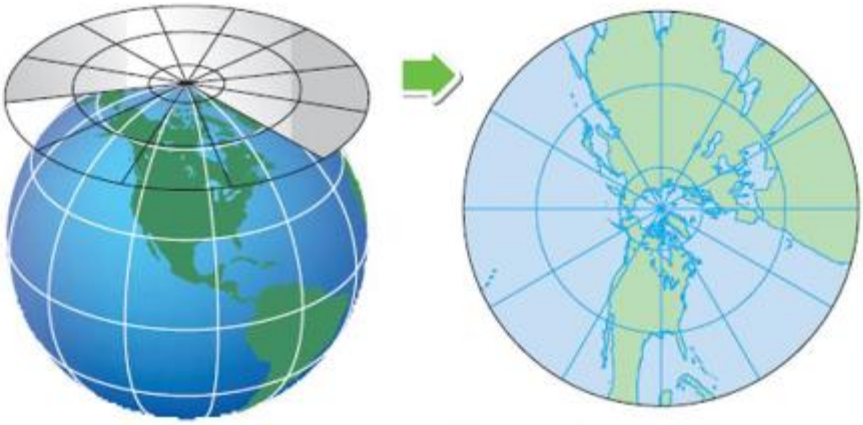
1. **Cylindrical Projection:** It is assumed to circumscribe a transparent globe (marked with meridians and parallels) so that the cylinder touches the equator throughout its circumference. Assuming that a light bulb is placed at the center of the globe, the graticule of the globe is projected on to the cylinder. By cutting open the cylinder along a meridian and unfolding it, a rectangles shaped cylindrical projection is obtained. Cylindrical are true at the equator and distortion increases toward the poles.



2. **Conical projection:** A cone is placed over the globe in such a way that the apex of the cone is exactly over the polar axis. A cone must touch the globe along a parallel of latitude, known as the standard parallel, which can be selected by the cartographer. Along this standard parallel, scale is correct and distortion is the least. When the cone is cut open along a meridian and laid flat, a fan shaped map is produced, with meridians as straight lines radiating from the vertex at equals angles, while parallels are arcs of circles, all drawn using the vertex as the center.



3. **Planar of Azimuthal projection:** A plane is placed so that it touches the globe at the north or South Pole. This can be conceived as the cone becoming increasingly flattened until its vertex reaches the limit of 180^0. The projection resulting is better known as the polar Azimuthal projection. It is circular in shape with meridians projected as straight line radiating from the center of the circle, which is the pole.



Coordinate System:

A coordinate system is a standardized method for assigning codes to locations so that locations can be found using the codes alone. Standardized coordinate systems use absolute locations. A map captured in the units of the paper sheet on which it is printed is based on relative locations or map millimeters.

Some standard coordinate systems used are:

- Geographic coordinates
    - Lat-long, geodetic lat long, Earth Centered Earth Fixed XYZ
- Universal Transverse Mercator (UTM) system
- Military grid
- State plane coordinate system

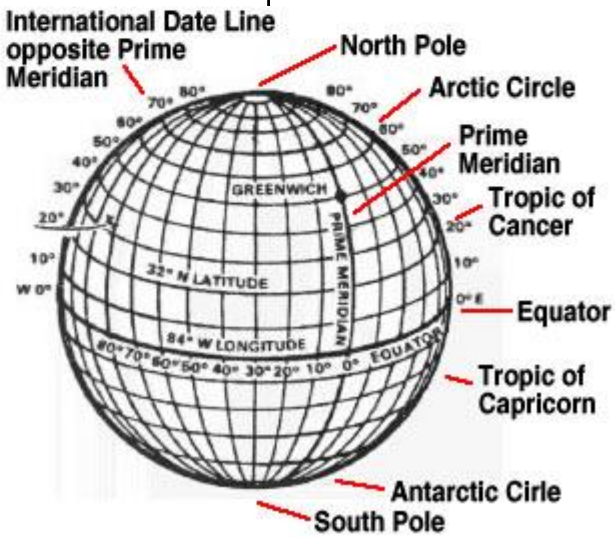
To compare or edge-match maps in a GIS, both maps must be in the same coordinate system else, the edges do not match and it gives us false information.

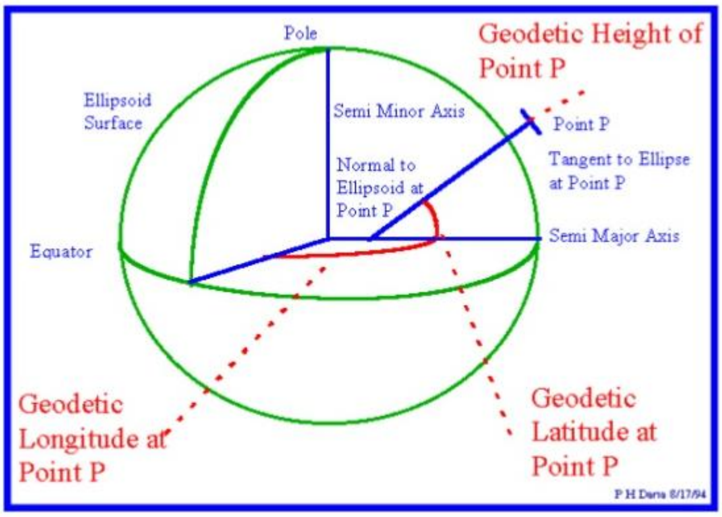
Latitude, Longitude, Height:

The most commonly used coordinate system today is the latitude, longitude and height system. The Prime Meridian and the Equator are the reference planes used to define latitude and longitude. Geographic coordinates are the earth's latitude and longitude system, ranging from 90 degrees south to 90 degrees north in latitude and 80 degrees west to 180 degrees east in longitude.

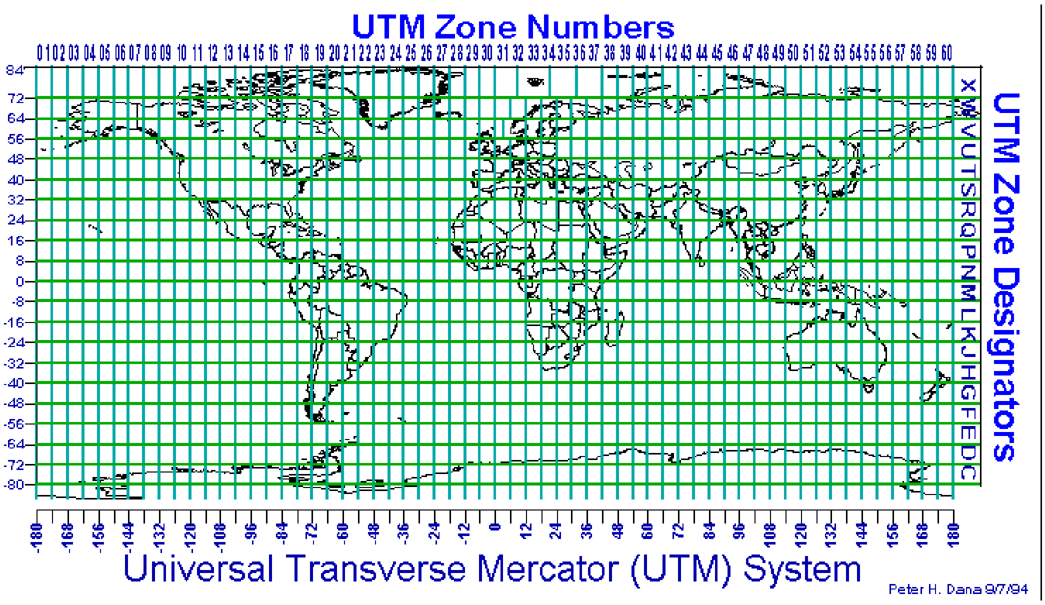
- A line with a constant latitude running east to west is called a parallel.
- A line with constant longitude running from the north pole to the south pole is called a meridian.
- The zero-longitude meridian is called the prime meridian and passes through Greenwich, England.
- A grid of parallels and meridians shown as lines on a map is called a graticule.
- The geodetic latitude of a point is the angle from the equatorial plane to the vertical direction of a line normal to the reference ellipsoid.
- The geodetic longitude of a point is the angle between a reference plane and a plane passing through the point, both planes being perpendicular to the equatorial plane.

- The geodetic height at a point is the distance from the reference ellipsoid to the point in a direction normal to the ellipsoid.

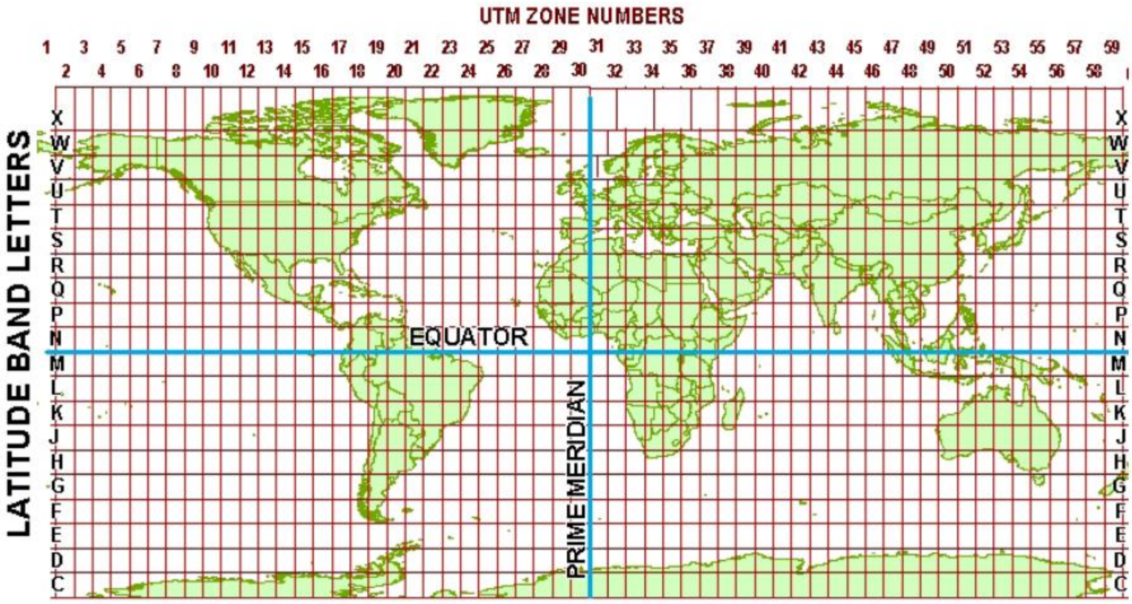




Universal Transverse Mercator (UTM): UTM is the most prevalent system used for mapping and other work. UTM zone numbers designate 6 degree longitudinal strips (60 vertical zones) extending from 80 degree South latitude to 84 degrees North latitude. Zone numbers start from the 180th meridian in an eastward direction.

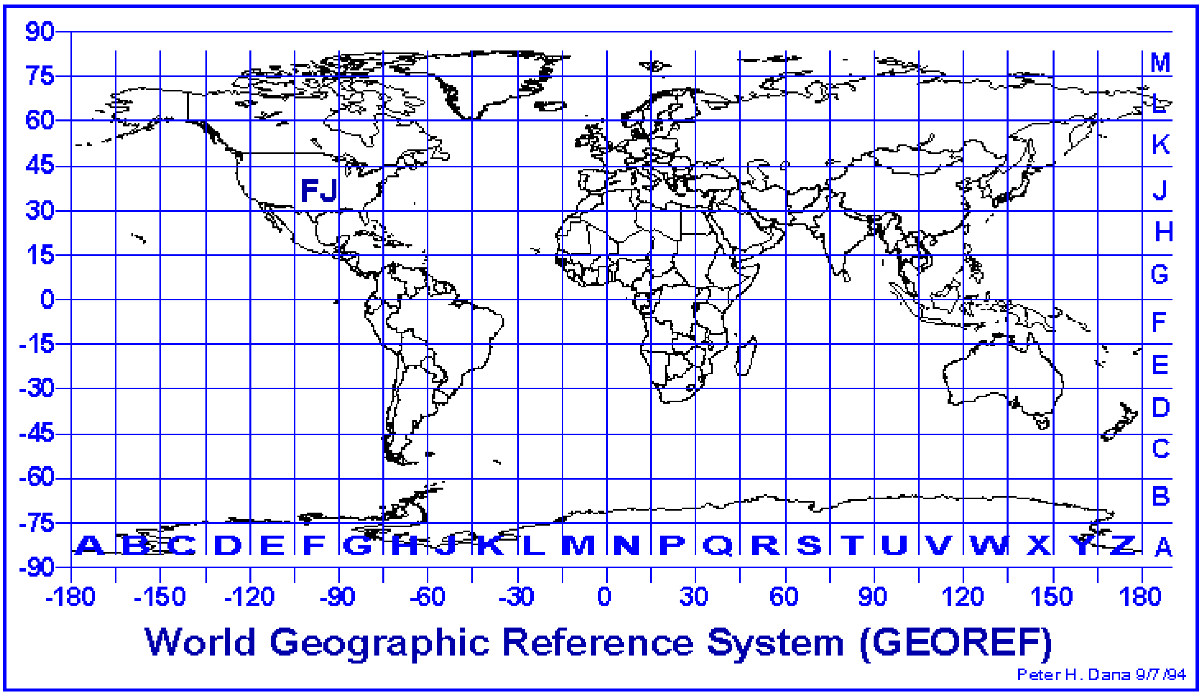


Military Grid Reference System (MGRS): MGRS is an extension of the UTM system. UTM zone number and zone character are used to identify an area 6 degree in east-west extent and 8 degrees in north-south extent. UTM zone number and designator are followed by 100km square easting and northing identifiers. The system uses a set of alphabetic characters for the 100km grid squares starting at the 180 degree meridian the characters A to Z (Omitting I and O) are used for 18 degrees before starting over. From the equator north the character A to V (Omitting I and O) are used for 100km squares, repeating every 2000km. The reverse sequence (from V to A) is used for southern hemisphere.



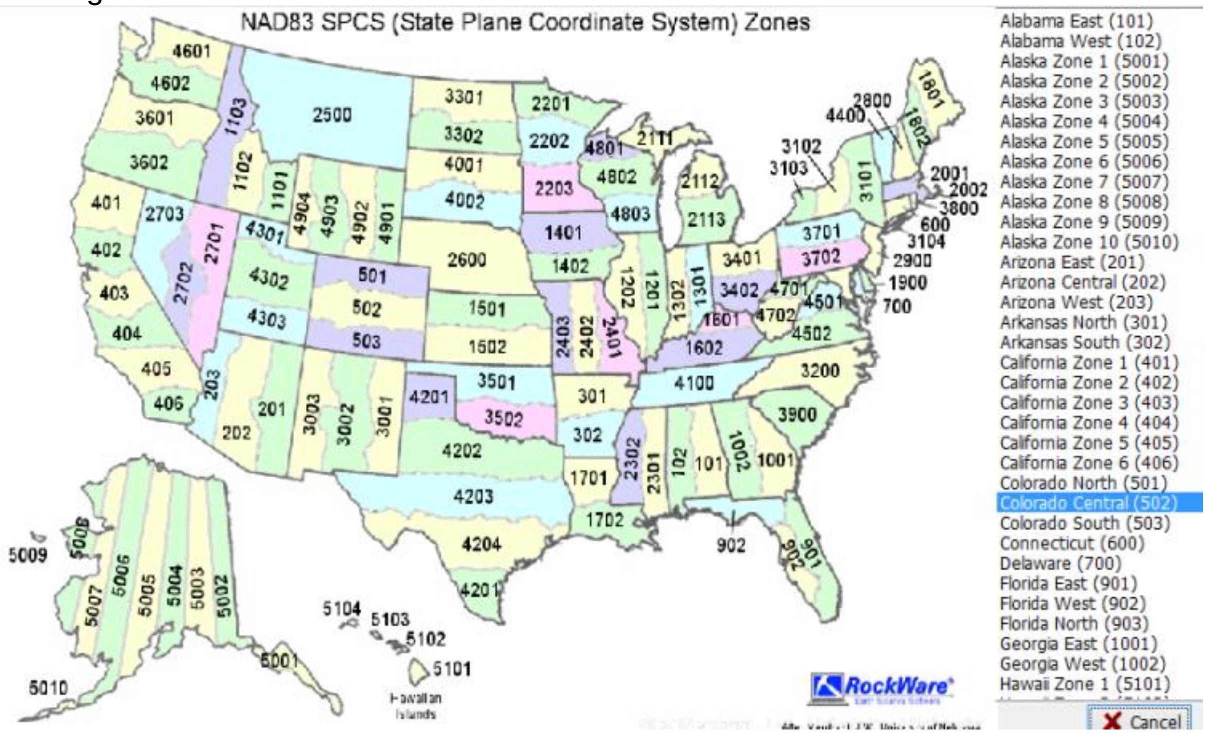
MGRS GRID ZONES

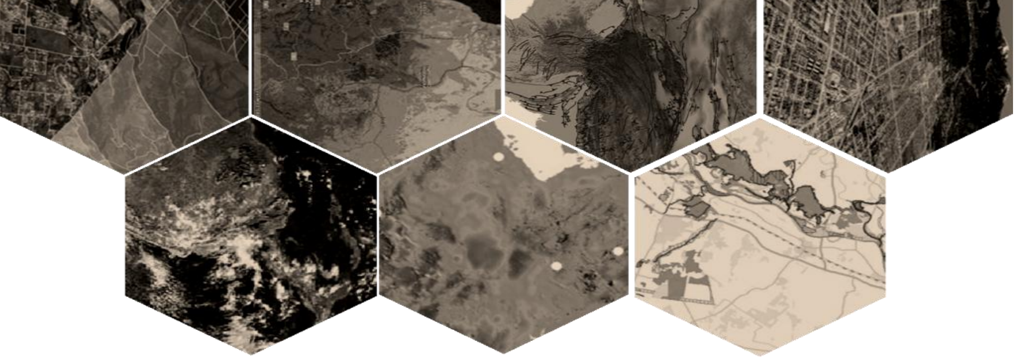
World Geographic Reference System (GEOREF): The World Geographic Reference System is used for aircraft navigation. GEOREF is based on latitude and longitude. The globe is divided into twelve bands of latitude and twenty four zones of longitude, each 15 degrees in extent.



World Geographic Reference System (GEOREF)

State Plane Coordinate System (SPCS): In the US, the State Plane System was developed in the 1930s and was based on the North American Datum 1927. State plane systems were developed in order to provide local reference systems that were tied to a national datum. Some smaller states use a single state plane zone. Larger states are divided into several zones. State plane zone boundaries often follow country boundaries. Lambert conformal conic projections are used for rectangular zones with a larger east-west than north-south extent. Transverse Mercator projections are used to define zones with a larger north-south extent.





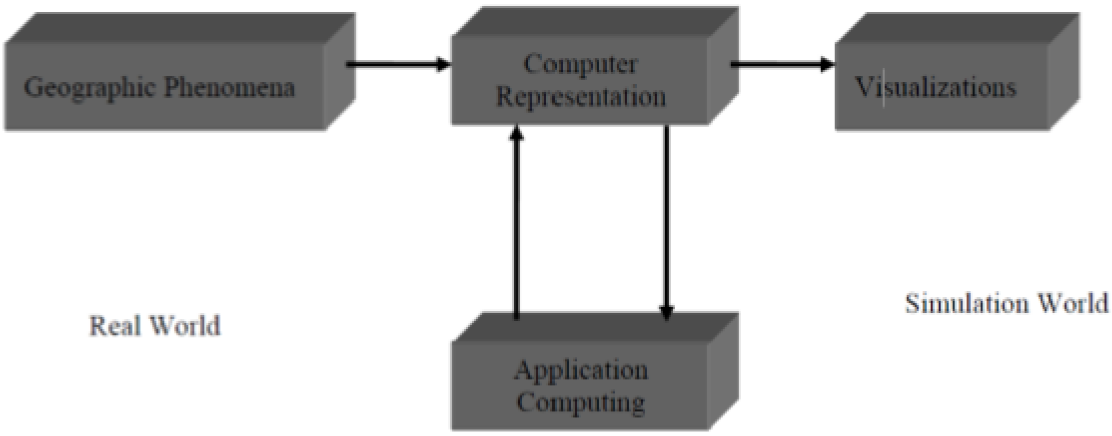
CHAPTER 3

Spatial Data Modeling and Database Design

CHAPTER OUTLINE

3.1 Introduction to geographic phenomena and data modeling

Geographic Phenomenon:



Geographic phenomena exist in the real world. In using GIS software, we first obtain some computer representation of these phenomena-stored in memory, in bits and bytes-as faithfully as possible. This is where we speak of spatial data.

Geographic phenomenon is as something of interest that

- Can be named or described
- Can be georeferenced, and
- Can be assigned a time (interval) at which it is/was present

What are relevant phenomena are for one’s current use of GIS depends entirely on the objectives that one has. For instance, in water management, the objects of study can be river basins, agro-ecologic units, measurements of actual evapotranspiration, ground water levels, irrigation levels etc…observe that all of these can be named/described, georeferenced and provided with a time interval at which each exists. In multipurpose cadastral administration, the objects of study are different houses, barns, parcels, streets of various types, land use, sewage canals and other form of urban infrastructure may all play a role. Again these can be named/described, georeferenced and assigned a time interval of existence. We do not claim that all relevant phenomena come as triplets (description, georeferenced, time interval), though many do. If the georeference is missing, we seem to have something of interest that is not position in space.

Types of geographic phenomena:

Geographic phenomena come in different flavors. To this end, first make the observation of a phenomenon in a GIS requires us to state what it is, and where it is. We must provide a description-or at least a name-on the one hand and a georeference on the other hand. There is another issue to time dependent data which is not provide much automatic support by the current GIS and it must be considered as issue of advanced GIS use. A second fundamental observation is that some phenomenon manifests themselves essentially everywhere in the study area, while others only occur in certain localities. If we define our study area as the equatorial ocean, for instance we can say that sea surface temperature can be measured anywhere in the study area. Therefore, it is the example of a geographic field.

A ***Geographic field*** is a geographic phenomenon for which, for every point in the study area, a value can be determined.

The usual examples of geographic fields are temperature, barometric pressure and elevation. These fields are actually continuous in nature. Examples of discrete fields are land use and soil classification.

Many other phenomena do not manifest themselves everywhere in the study area, but only in certain localities which are called geographic objects.

Geographic objects populate the study area, and are usually well distinguishable, discrete, bounded entities. The space between them is potentially empty.

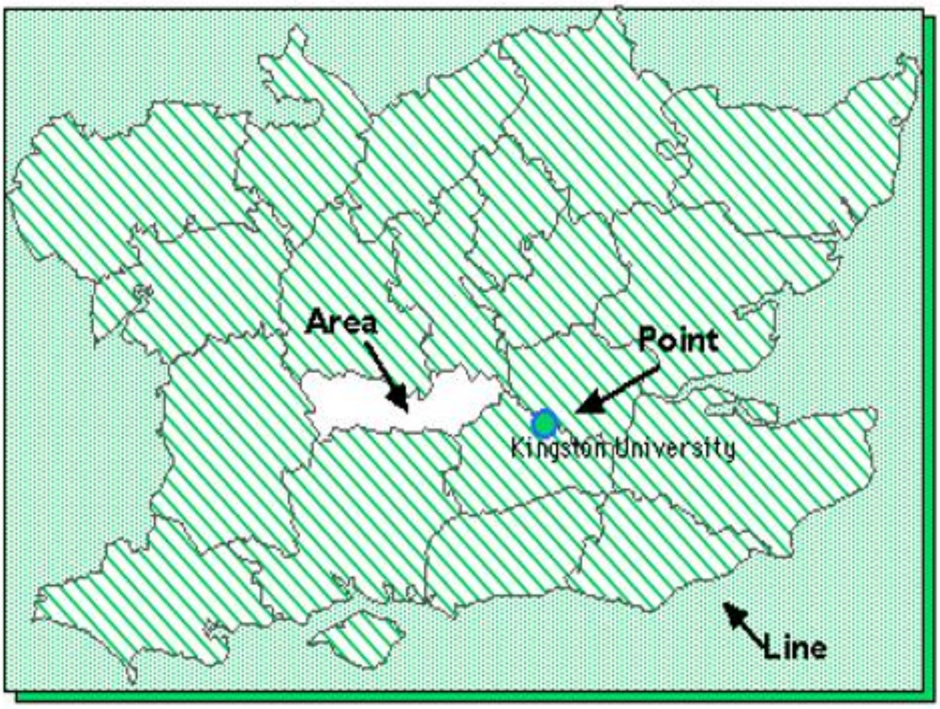
A general rule-of-thumb is that natural geographic phenomena are more often fields, and man-made phenomena are more often objects.

Modeling:

Reality is too complex for even the most sophisticated GIS software, so in order to represent reality in a spatial database, a simplification of reality is created. This simplification is known as a data model. Modeling is the process of producing an abstraction of the 'real world' so that some part of this can be more easily handled. In a data model, reality is simplified into just three spatial entities, or elements, which can be used to represent the real world. These three spatial entities are:

- The Point

- The Line
- The Area/surfaces



3.2 Spatial relationship and topology

General spatial topology:

Topological deals with spatial properties that do not change under certain transformation. A simple example will illustrate what we mean. Assume you have some features that are drawn on the sheet of rubber (as in figure). Now take the sheet and pull on its edges, but do not tear or break it. The features will change in shape and size. But some properties, however, do not change.

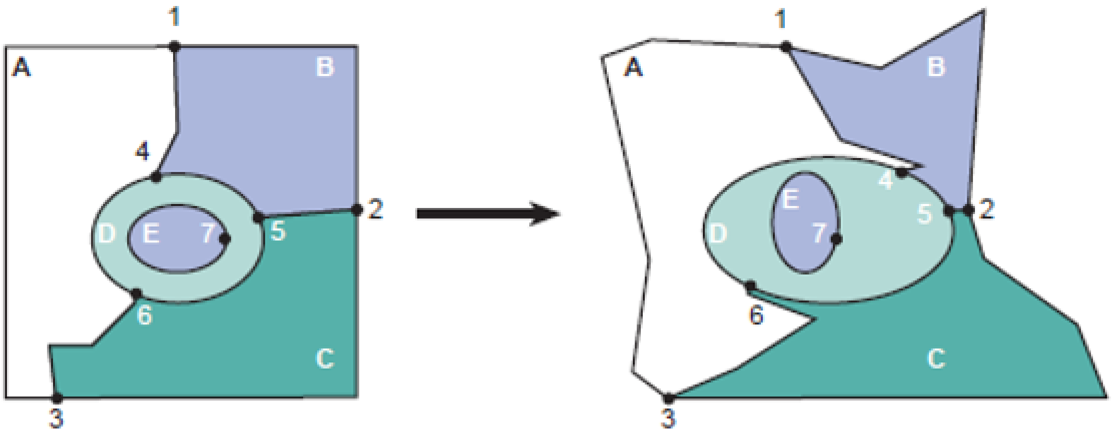


Fig: Rubber sheet transformation: the transformed, yet many relationships between the constituents remain unchanged.

- Area E is still inside area D,
- The neighbourhood relationships between A, B, C, D, and E stay intact and their boundaries have the same start and end nodes, and
- The areas are still bounded by the same boundaries, only the shapes and lengths of their perimeters have changed.

Topology refers to the spatial relationships between geographical elements in a data set that do not change under a continuous transformation.

These relationships are invariant under continuous transformations. Such properties are called topological properties, and the transformation is called a topological mapping. The mathematical properties of the geometric space used for spatial data can be described as follows:

- The space as a three dimensional Euclidean space where for every point we can determine its three dimensional coordinates as a triple (X , Y, Z) of real numbers. In this space we can define features like points , lines , polygons and volume as geometric primitives of the respective dimension. A point is zero dimensional a line one dimensional, a polygon two dimensional, and a volume is a three dimensional primitive.
- The space is a metric space, which means that we can always compute the distance between two points according to a given distance functions. Such a function is also known as a metric.
- Interior and boundary are properties of spatial features that remain invariant under topological mapping. This means that under any topological mapping the interior and the boundary of a feature remain unbroken and intact.

3.3 Scale and resolution

Scale:

Scale: Reduction of area to show portion of Earth's surface on map.

Map scale: Extend of reduction expressed as a ratio

1" = 2,000'

1' = 24,000 inches, 1:24,000

1:24,000 is example of representative fraction (RF) because amounts on either side of colon are equivalent.

Map scale indicates how much a given area is reduced. On a same-size map (or piece of paper), features on a small-scale map (1:250,000) will be smaller than on a large-scale map (1:1,200)

Small scale – can show large amount of area without much detail.

Large scale – can only show small area but lots of detail.

Resolution:

Resolution refers to how accurately the location and shape of map features can be depicted at a given scale.

Large scale maps has better resolution because the reduction is less. As scale becomes smaller, more and more features become too small to display.

Lead pencil (0.5 mm) line = 39.4 feet on 1:24,000

In the practice of spatial data handling, one often comes across questions like "what is the resolution on the data?" or "at what scale is your date set?" Now that we have moved firmly into the digital age, these questions sometimes defy an easy answer.

Map scale can be defined as the ratio between the distance on a paper map and the distance of the same stretch in the terrain. A 1:50,000 scale map means that 1cm on the map represents 50,000 cm, i.e., 500 m, in the terrain. 'Large scale' means that the ratio is large, so typically it means there is much detail, as in a 1:1,000 paper map. 'Small scale' in contrast means a small ratio, hence less detail, as in a 1:2,500,000 paper map. When applied to spatial data, the term *resolution* is commonly associated with the cell width of the tessellation applied.

When digital spatial data sets have been collected with a specific map-making purpose in mind, and these maps were designed to be of a single map scale, like 1:25,000, we might suppose that the data carries the characteristics of a 1:25,000 digital data set.

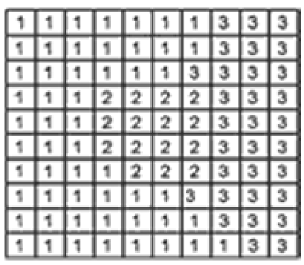
3.4 Vector, raster and digital terrain model

Vector and raster data:

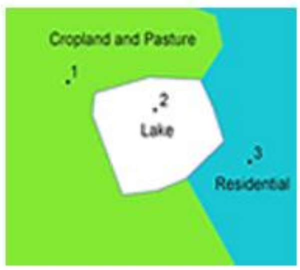
The graphical representation of spatial and attribute data in GIS software takes the form of either raster or vector graphics. The differences between raster and vector graphics, as detailed below, effect the level of detail, visual appeal, speed of manipulating graphics and data storage space required.

Aerial photographs and satellite images are generally in a raster format and are used in GIS to view a detailed map at a given extent or for the purpose of digitizing. Raster graphics are predominantly used to display spatial data and use a grid-type architecture in terms of storing spatial and graphic value data. Vector graphics are commonly used to represent features like roads, rivers, housing, and the like using points, lines and polygons. Based on scalable vector graphics, vector graphics provide a linear and detailed approach to manipulating attribute data. Raster and Vector graphics are frequently used together.

The key difference between Raster and Vector graphics is how they are structured. Raster graphics use pixels ("dots") whereby a graphic is made up of a large number of pixels, each pixel having a location & colour value in a grid-like format. A vector graphic is rendered by a mathematical manipulation referenced by co-ordinates. Given the different structure of these graphic types, the following differences arise as a result:



Raster Structure



Vector Structure

1. Storage Space: Raster graphics require more storage space than vector graphics, as they store a location & colour value per pixel.
2. Detail: Raster images are more detailed within a given extent ("zoom"), however raster images become pixelated if too tight a zoom is applied. Vector images are less detailed, but maintain their original aesthetics regardless of extent or zoom.
3. Responsiveness: performance & responsiveness when manipulating vector image is faster than raster images, as the data structured used to render vectors is

mathematically based whereas rasters requires the retrieval of individual pixel values and a manipulation of each pixel

Advantages of vector data:

- Data can be represented at its original resolution and form without generalization.
- Graphic output is usually more aesthetically pleasing (traditional cartographic representation);
- Since most data, e.g. hard copy maps, is in vector form no data conversion is required.
- Accurate geographic location of data is maintained.
- Allows for efficient encoding of topology, and as a result more efficient operations that require topological information, e.g. proximity, network analysis.

Disadvantages of vector data:

- The location of each vertex needs to be stored explicitly.
- For effective analysis, vector data must be converted into a topological structure. This is often processing intensive and usually requires extensive data cleaning. As well, topology is static, and any updating or editing of the vector data requires rebuilding of the topology.
- Algorithms for manipulative and analysis functions are complex and may be processing intensive. Often, this inherently limits the functionality for large data sets, e.g. a large number of features.
- Continuous data, such as elevation data, is not effectively represented in vector form. Usually substantial data generalization or interpolation is required for these data layers.
- Spatial analysis and filtering within polygons is impossible.

Advantages of Raster data:

- The geographic location of each cell is implied by its position in the cell matrix. Accordingly, other than an origin point, e.g. bottom left corner, no geographic coordinates are stored.
- Due to the nature of the data storage technique data analysis is usually easy to program and quick to perform.
- The inherent nature of raster maps, e.g. one attribute maps, is ideally suited for mathematical modeling and quantitative analysis.
- Discrete data, e.g. forestry stands, is accommodated equally well as continuous data, e.g. elevation data, and facilitates the integrating of the two data types.
- Grid-cell systems are very compatible with raster-based output devices, e.g. electrostatic plotters, graphic terminals.

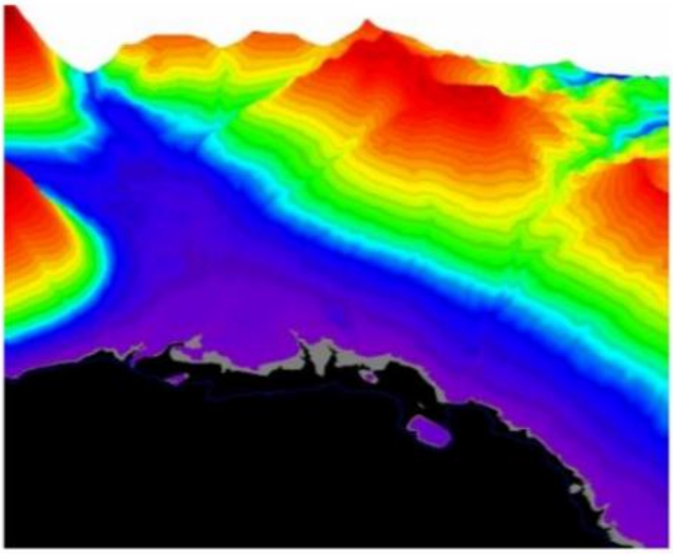
Disadvantages of raster data:

- The cell size determines the resolution at which the data is represented.;
- It is especially difficult to adequately represent linear features depending on the cell resolution. Accordingly, network linkages are difficult to establish.
- Processing of associated attribute data may be cumbersome if large amounts of data exists. Raster maps inherently reflect only one attribute or characteristic for an area.

- Since most input data is in vector form, data must undergo vector-to-raster conversion. Besides increased processing requirements this may introduce data integrity concerns due to generalization and choice of inappropriate cell size.
- Most output maps from grid-cell systems do not conform to high-quality cartographic needs.

Digital Terrain Model (DTM):

A digital terrain model is a topographic model of the bare earth – terrain relief - that can be manipulated by computer programs. The data files contain the spatial elevation data of the terrain in a digital format which usually presented as a rectangular grid. Vegetation, buildings and other man-made (artificial) features are removed digitally - leaving just the underlying terrain (on the other hand, Digital Surface Model (DSM) is usually the main product produced from photogrammetry, where it does contain all the features mentioned above, while a filtered DSM results in a DTM).



DTM model is mostly related as raster data type (opposed to vector data type), stored usually as a rectangular equal-spaced grid, with space (resolution) of between 50 and 500 meters mostly presented in cartesian coordinate system – i.e. x, y, z (there are DTMs presented in geographic coordinate system – i.e. angular coordinates of latitude and longitude). For several applications a higher resolution is required (as high as 1 meter spacing). A DTM can be used to guide automatic machinery in the construction of a physical model or even in computer games. Modeling terrain relief via DTM is a powerful tool in GIS (Geographic Information System) analysis and visualization. DTM can be stored in a GIS databases in several ways:

a. a set of contour vectors (left);
b. a rectangular grid of equal-spaced corner/point heights (middle); or,
c. an irregularly spaced set of points connected as triangles (TIN - Triangular Irregular Network) (right).


Figure 1

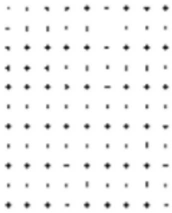


Figure 2


Figure 3

Application of DTM:

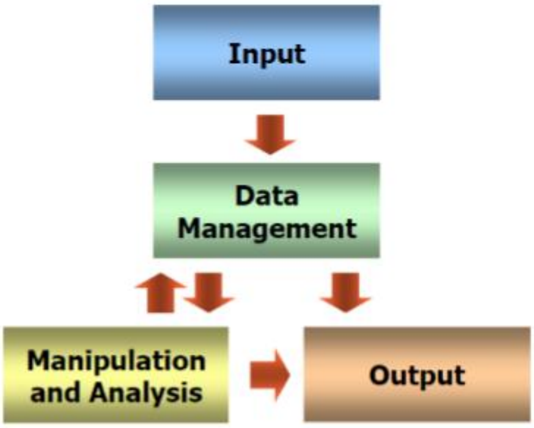
The DTM data sets are extremely useful for the generation of 3D renderings of any location in the area described. 3D models rendered from DTM data can be extremely useful and versatile for a variety of applications. DTMs are used especially in civil engineering, geodesy & surveying, geophysics, and geography. The main applications are:

- Visualization of the terrain
- Reduction (terrain correction) of gravity measurements (gravimetry, physical geodesy)
- Terrain analyses in Cartography and Morphology
- Rectification of airborne or satellite photos
- Extraction of terrain parameters, model water flow or mass movement


TIN Representation of a human skull from a low resolution laser scan


SSX3 video game featuring complex terrain relief (represented by *DTM*)

Functional components of GIS:



According to the definition, a GIS always consists of modules for input, storage analysis, display and output of spatial data. Figure above shows a diagram of these modules. For a particular GIS, each of these modules may provide many or only few functions. However, if one of these functions would be completely missing, the system should not be called a geographic information system. Beside data input (data capture), storage and maintenance, analysis and output, geoinformation processes involve also dissemination, transfer and exchange as well as organizational issues. The latter defines the context and rules according to which geoinformation is acquired and processed.

- Data input: bringing data in the GIS environment.
- Data manipulation: allowing alteration of primary data.
- Data output: moving data (or analysis results) out of the GIS.
- Data management: controlling access to data and ensuring data integrity and storage efficiency.
- Data retrieval: calling data from a stored format into use.
- Data display: visualizing primary or derived data.
- Data analysis and modeling: gathering insights into relationships in the data, and modelling spatial phenomena

3.5 Spatial database design with the concept of geo-database

Spatial DBMS:

A spatial database system may be defined as a database system that offers spatial data types in its data model and query language, and supports spatial data types in its implementation, providing at least spatial indexing and spatial join methods.

Spatial database systems offer the underlying database technology for geographic information systems and other applications. We survey data modeling, querying, data structures and algorithms, and system architecture for such systems. The emphasis is on describing known technology in a coherent manner, rather than listing open problems.

In various fields there is a need to manage geometric, geographic, or spatial data, which means data related to space. The space of interest can be, for example, the two-dimensional abstraction of (parts of) the surface of the earth or a 3d-space representing a digital terrain model. At least since the advent of relational database systems there have been attempts to manage such data in database systems.

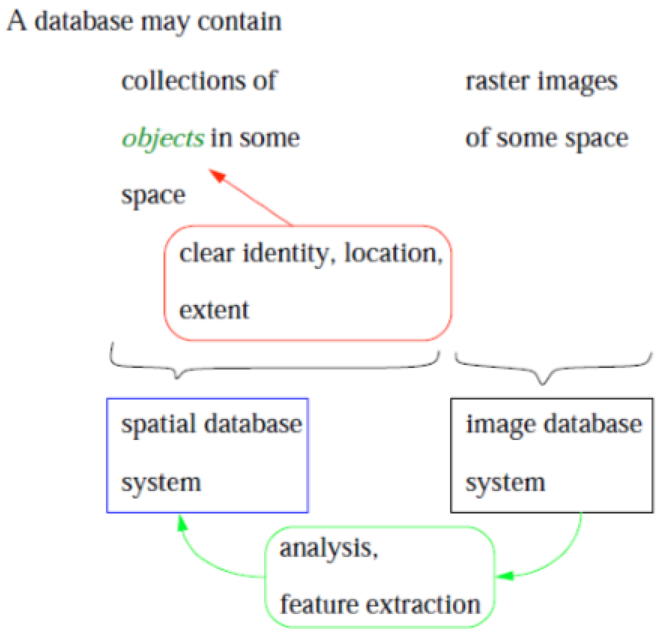
Characteristic for the technology emerging to address these needs is the capability to deal with large collections of relatively simple geometric objects, for example, a set of 100 000 polygons. Several terms have been used for database systems offering such support like pictorial, image, geometric, geographic, or spatial database system. The terms “pictorial” and “image” database system arise from the fact that the data to be managed are often initially captured in the form of digital raster images (e.g. remote sensing by satellites, or computer tomography in medical applications).

The term “spatial database system” has become popular during the last few years, and is associated with a view of a database as containing sets of objects in space rather than images or pictures of a space. Indeed, the requirements and techniques for dealing with objects in space that have identity and well-defined extents, locations, and relationships are rather different from those for dealing with raster images.

A spatial database therefore has the following characteristics:

i. A spatial database system is a database system.
ii. It offers spatial data types (SDTs) in its data model and query language.
iii. It supports spatial data types in its implementation, providing at least spatial indexing and efficient algorithms for spatial join.

Nobody cares about a special purpose system that is not able to handle all the standard data modeling and querying tasks. Hence a spatial database system is a full-fledged database system with additional capabilities for handling spatial data. Therefore spatial indexing is mandatory. It should also support connecting objects from different classes through some spatial relationship.



A spatial database includes collections of information about the spatial location, relationship and shape of topological geographic features and the data in the form of attributes. The design of the spatial database is the formal process of analyzing facts about the real world into a structured model. Database design is characterized by the following phases: requirement analysis, logical design and physical design. In other words, you basically need a plan, a design layout and then the data to complete the process.

Having a solid well designed spatial database is the key to performing good Spatial Analysis. The database can be complex and designed with expensive sophisticated software or can be merely a simple well organized collection of data that can be utilized in a geographic form.

Three main categories of spatial modeling functions that can be applied to geographic features within a GIS are:

(1) geometric models, such as calculating the Euclidean distance between features, generating buffers, calculating areas and perimeters, and so on;
(2) coincidence models, such as topological overlay; and
(3) adjacency models (path finding, redistricting, and allocation).

All three model categories support operations on spatial data such as points, lines, polygons, tins, and grids. Functions are organized in a sequence of steps to derive the desired information for analysis.

Almost all entities of geographic reality have at least a 3-dimensional spatial character, but not all dimensions may be needed. E.g. a highway pavement actually has a depth which might be important, but is not as important as the width, which is not as important as the length. Representation should be based on the types of manipulations that might be undertaken. Map-scale of the source document is important in constraining the level of detail represented in a database. E.g. on a 1:100,000 map individual houses or fields are not visible.

Spatial database management:

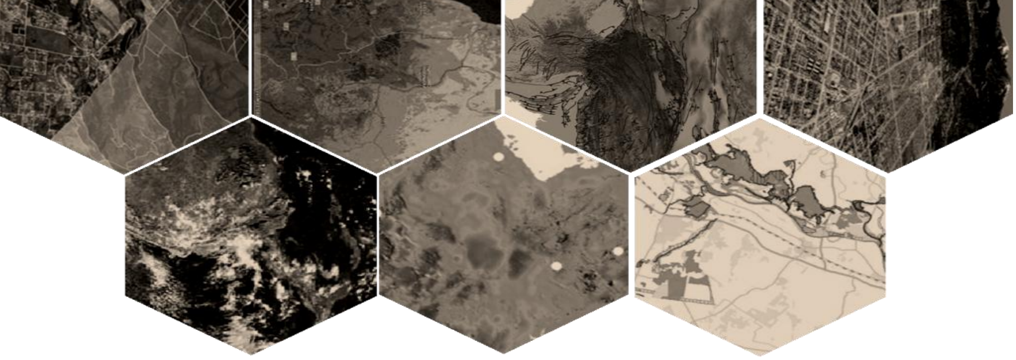
Many factors influence a successful Geographic Information System (GIS) implementation. None however are more fundamental than having the right management strategies and software to implement these. The spatial database is the foundation by which all data is uniformly created and converted. But maintaining the integrity and currency of the data is of fundamental importance. A classic mistake made by many organizations is thinking that a generic spatial database design will be sufficient for their needs. That is simply not the case. The spatial database is the end result of a series of processes that determine the specific functional requirements for the user and the key applications. Interoperability of data is also a critical area of concern in the development of spatial data information systems. As we move from newly created data to assimilation of all existing data, a properly designed spatial database is insurance for end user success. A good spatial database management software package should be able to:

1. Scale and rotate coordinate values for "best fit" projection overlays and changes.
2. Convert (interchange) between polygon and grid formats.
3. Permit rapid updating, allowing data changes with relative ease.
4. Allow for multiple users and multiple interactions between compatible data bases.
5. Retrieve, transform, and combine data elements efficiently.
6. Search, identify, and route a variety of different data items and score these values with assigned weighted values, to facilitate proximity and routing analysis.
7. Perform statistical analysis, such as multivariate regression, correlations, etc.
8. Overlay one file variable onto another, i.e., map superpositioning.
9. Measure area, distance, and association between points and fields.
10. Model and simulate, and formulate predictive scenarios, in a fashion that allows for direct interactions between the user group and the computer program.

Geo database:

The physical store of geographic information, primarily using a database management system (DBMS) or file system is called a geodatabase.

- A geo database is a collection of feature classes and tables. Feature classes can be organized into feature datasets.
- The geo database data model is an object-oriented data model that lets you make the features in your GIS datasets smarter by endowing them with natural behaviours.
- The objects in a geo database can be related to each other (it stores relationships).
- The geo database lets you implement custom behaviours by implementing domains, validations rule or writing software code.



CHAPTER 4

Capturing the real world

CHAPTER OUTLINE

4.1 Different methods of data capture

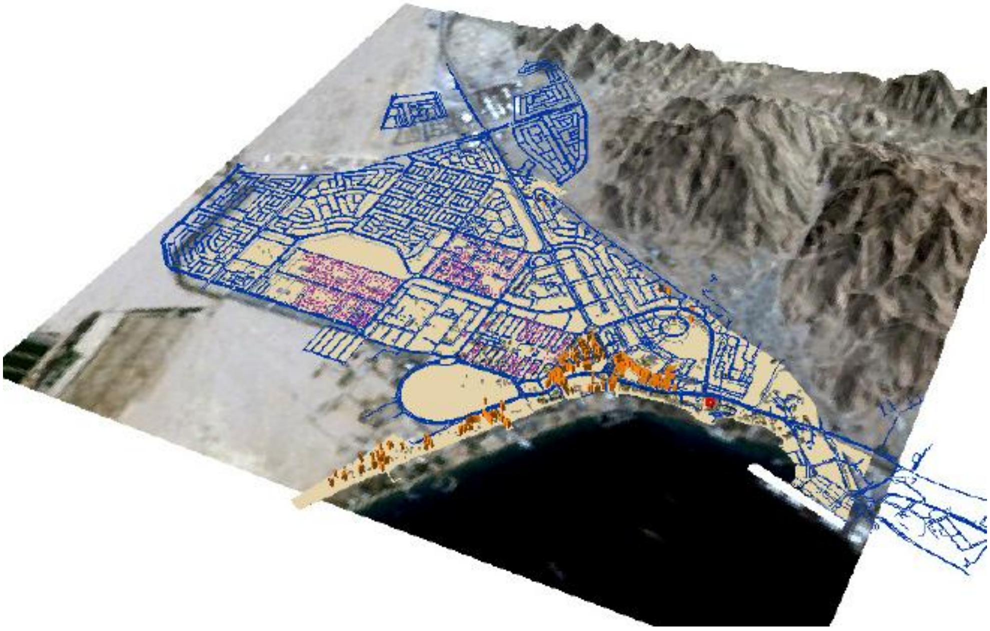
The different data capturing methods from various sources commonly used in a GIS are briefly discussed below:

1. **Photogrammetric Compilation:** The primary source used in the process of photogrammetric compilation is aerial photography. Generally, the process involves using specialized equipment (a stereoplotter) to project overlapping aerial photos so that a viewer can see a three-dimensional picture of the terrain, known as a photogrammetric model. The current technological trend in photogrammetry is toward a greater use of digital procedures for map compilation.

Fig: Aerial photo of Narayani bridge

2. **Digitizing:** A digitizing workstation with a digitizing tablet and cursor is typically used to trace digitize. Both the tablet and cursor are connected to a computer that controls their functions. Most digitizing tablets come in standard sizes that relate to engineering drawing sizes ("A" through "E," and larger). Digitizing involves tracing

features on a source map, taped to the digitizing tablet, with a precise cross hair in the digitizing cursor and instructing the computer to accept the location and type of feature. The person performing the digitizing may separate features into map layers, or attach an attribute to identify the feature.



3. **Map scanning:** Optical scanning systems automatically capture map features, text, and symbols as individual cells, or pixels, and produce an automated product in raster format as described earlier . Scanning outputs files in raster form, usually in one of several compressed formats saves storage space (e.g., TIFF 4, JPEG). Most scanning systems provide software to convert raster data to a vector format differentiating point, line, and area features. Scanning systems and software is becoming more sophisticated with some abilities to interpret symbols and text, and store this information in databases. Creating an intelligent GIS database from a scanned map will require vectorizing the raster data and manual time for entering attribute data from a scanned annotation.

Fig: E-sized map scanner

4. **Satellite data:** Earth Resources Satellites have become a source of huge amount of data for GIS applications. The data obtained from the Satellites are in digital form, which can be directly imported to GIS. There are numerous satellite data

sources such as LANDSAT or SPOT. A new generation of high-resolution satellite data that will increase opportunities and options for GIS database development is becoming available from private sources and national governments. These satellite systems will provide panchromatic (black and white) or multi-spectral data in the 1- to 3-meter ranges as compared to the 10- to 30-meter range available from traditional remote sensing satellites.

Fig: Satellite data of Bharatpur Airport

5. **Field data collection:** Advances in hardware and software have greatly increased opportunities for capture of GIS data in the field (e.g., sign of utility inventory, property surveys, land use inventories). In particular, electronic survey systems and the global positional system (GPS) have revolutionized surveying and field data collection. Electronic distance measurement services allow for survey data to be gathered quickly in automated form for uploading to a GIS. Sophisticated GPS collection units have provided a quick means of capturing the coordinates and attributes of features in the field.

Fig: Field data collection

6. **Tabular data entry:** Some of the tabular attribute data that is normally in a GIS database exists on maps as annotation and or can be found in paper files. Information from these sources will be required for GIS applications and will have to

be converted to digital form through keyboard entry. This kind of data entry is commonplace and relatively easy to accomplish.

7. **Document Scanning:** Smaller -format scanners can also be used to create raster files of documents such as permit forms, service cards, site photographs, etc. These documents can be indexed in a relational database by number, type, date, engineering drawings, etc., and queried and displayed by users. GIS applications can be built which allows users to point to and retrieve for display a scanned document (e.g., tax parcel) interactivley.
8. **Translation of existing digital data:** Existing automated data may be available from existing tabular files maintained by outside sources. Many programs are available that perform this translation and, in fact, many GIS packages can be acquired with programs that translate data to and from several "standard" formats which are accepted widely by the mapping industry and have been used as intermediate "exchange" formats for moving data between platforms (e.g., Intergraph SIF, TIGER, Shapefile and AutoCAD DXF).

4.2 Spatial reference

Overview of spatial references:

Geographic data for any particular area is stored in separate layers. For example, roads store in one layer, parcels in another, and buildings in a third. To enable the data in each layer to integrate when displayed and queried, each layer must reference locations on the earth's surface in a common way. Coordinate systems provide this framework. They also provide the framework needed for data in different regions to be referenced in different ways. Each layer in the geodatabase has a coordinate system that defines how its locations are georeferenced.

In the geodatabase, the coordinate system and other related spatial properties are defined as part of the spatial reference for each dataset. A spatial reference is the coordinate system used to store each feature class and raster dataset as well as other coordinate properties such as the coordinate resolution for x,y coordinates and optional z- and m-(Measure) coordinates. If required, you can define a vertical coordinate system for datasets with z-coordinates that represent surface elevation.

Co-ordinate system:

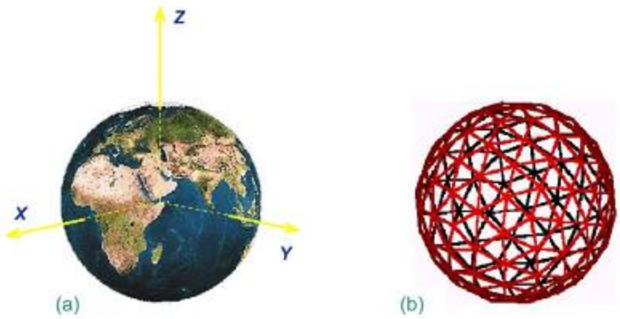
X,y coordinates are georeferenced with a geographic or projected coordinate system. A geographic coordinate system (GCS) is defined by a datum, an angular unit of measure (usually degrees), and a prime meridian. A projected coordinate system (PCS) consists of a linear unit of measure (usually meters or feet), a map projection, the specific parameters used by the map projection, and a geographic coordinate system.

A projected or geographic coordinate system can have a vertical coordinate system as an optional property. A vertical coordinate system (VCS) georeferences z-values, most commonly used to denote elevation. A vertical coordinate system includes a geodetic or vertical datum, a linear unit of measure, an axis direction, and a vertical shift.

M, or measure, values do not have a coordinate system.
For a spatial reference that includes an unknown coordinate system (UCS) you specify a tolerance only. It is not possible to georeference a feature associated with a UCS. You cannot define a vertical coordinate system if the x,y coordinate system is unknown. If at all possible, you should not use an unknown coordinate system. Because the valid area of use and unit of measure are not known, the resolution and tolerance values may not be appropriate for the data.

Spatial reference system and frames:
The geometry and motion of objects in 3D Euclidean space are described in a reference coordinate system. A reference coordinate system is a coordinate system with well-defined origin and orientation of the three orthogonal, coordinate axes. We shall refer to such a system as a spatial reference system (SRS). A spatial reference system is a mathematical abstraction. It is realized by means of spatial reference frame (SRF). We may visualize an SRF as a catalogue of coordinates of specific, identifiable point objects, which implicitly materialize the coordinate axes of SRS.

Several spatial reference systems are used in the earth sciences. The most important one for the GIS community is the International Terrestrial Reference System (ITRS). The ITRS has its origin in the center of mass of the earth. The Z-axis points towards a mean earth north pole. The X-axis is oriented towards a mean Greenwich meridian and is orthogonal to the Z-axis. The Y-axis completes the right handed reference coordinate system.



- (a) International Terrestrial Reference System: ITRS
- (b) International Terrestrial Reference Frame: ITRF

ITRS and ITRF:
Co-ordinates in an International Terrestrial Reference System (ITRS) are computed at different epochs and the solutions are called ITRF. Due to plate tectonics and tidal deformation, the co-ordinates changes for a certain point between the different ITRF. The latest version of ITRF is ITRF 2005. In simple terms the ITRF is a realisation of the ITRS.

Benefits of a geodetic datum based on ITRF:
Adopting an ITRF based geodetic datum allows for a single standard for collecting, storing and using geographic or survey related data. This will ensure compatibility across various geographic, land and survey systems at the local, regional, national and global level. This is the main reason that the ITRF based CORS networks should form the basis for Spatial Data Infrastructure (SDI) which is the enabling infrastructure to manage a country's key

spatial data sets ie it underpins or is the reference layer for the cadastre, transit / road networks, infrastructure corridors like gas, water, power, communications etc. An ITRF based geocentric datum or CORS network will also:

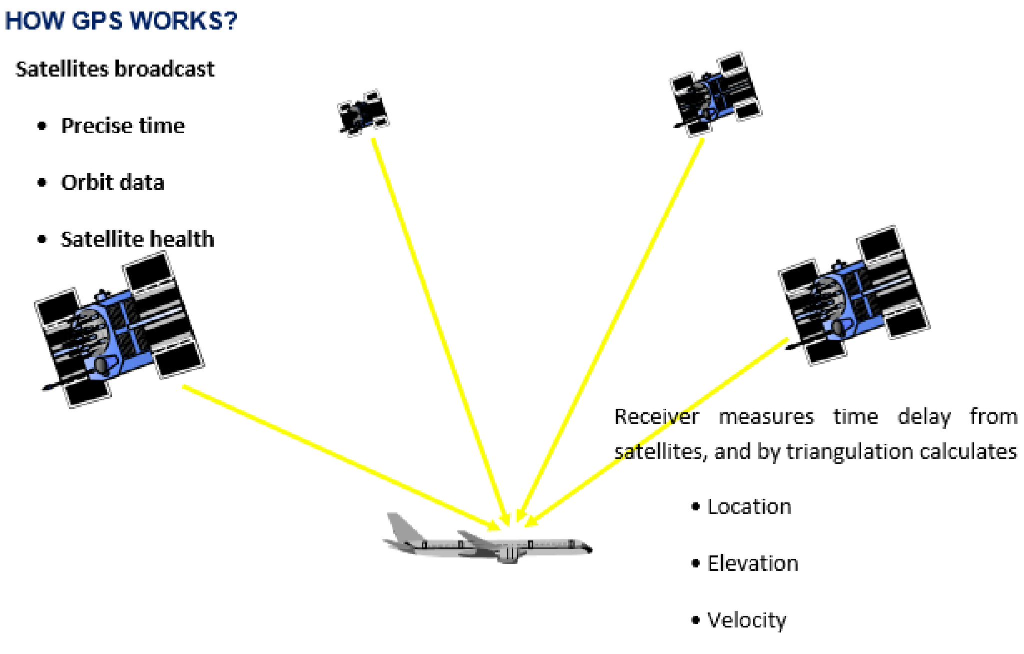
- provide direct compatibility with GNSS measurements and mapping or geographic information system (GIS) which are also normally based on an ITRF based geodetic datum;
- minimise the need for casual users to understand datum transformations;
- allow more efficient use of an organisations' spatial data resources by reducing need for duplication and unnecessary translations;
- help promote wider use of spatial data through one user friendly data environment;
- reduce the risk of confusion as GNSS, GIS and navigation systems become more widely used and integrated into business and recreational activities.

4.3 GPS and remote sensing

Global positioning system (GPS): a system of earth-orbiting satellites which can provide precise (100 meter to sub-cm.) location on the earth's surface (in lat/long coordinates or equiv.)

GPS is a worldwide radio-navigation system formed from a constellation of 24 satellites and their ground stations. It uses these "man-made stars" as reference points to calculate positions accurate to a matter of meters. GPS receivers have become very economical, making the technology accessible to virtually everyone. GPS provides continuous three-dimensional positioning 24 hours a day to the military and civilian users throughout the world. These days GPS is finding its way into cars, boats, planes, construction equipment, farm machinery, even laptop computers. It has a tremendous amount of applications in GIS data collection, surveying, and mapping. GPS is increasingly used for precise positioning of geospatial data and the collection of data in the field.

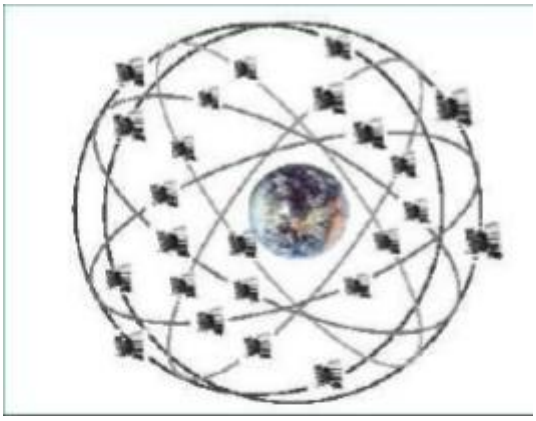
How GPS works:



- The GPS receiver compares the time a signal was transmitted by a satellite with the time it was received. The time difference tells the GPS receiver how far away the satellite is.
- With four or more satellites in view, the receiver can determine the user's 3D position (latitude, longitude and altitude).

Components of GPS:

1. **Space segment:** The nominal GPS Operational Constellation consists of 24 satellites that orbit the earth in 12 hours. There are often more than 24 operational satellites as new ones are launched to replace older satellites. The satellite orbits repeat almost the same ground track (as the earth turns beneath them) once each day. The orbit altitude is such that the satellites repeat the same track and configuration over any point approximately each 24 hours (4 minutes earlier each day). There are six orbital planes, with nominally four SVs (Satellite Vehicles) in each, equally spaced (60 degrees apart), and inclined at about fifty-five degrees with respect to the equatorial plane. This constellation provides the user with between five and eight SVs visible from any point on the earth.



2. **Control segment:** The Master Control facility is located at Schriever Air Force Base (formerly Falcon AFB) in Colorado. These monitor stations measure signals from the SVs which are incorporated into orbital models for each satellites. The models compute precise orbital data (ephemeris) and SV clock corrections for each satellite. The Master Control station uploads ephemeris and clock data to the SVs. The SVs then send subsets of the orbital ephemeris data to GPS receivers over radio signals.

Global Positioning System (GPS) Master Control and Monitor Station Network

3. **User segment:** The User Segment consists of all earth-based GPS receivers (figure 6.3). Receivers vary greatly in size and complexity, though the basic design is rather simple. The typical receiver is composed of an antenna and preamplifier, radio signal microprocessor, control and display device, data recording unit, and power supply (figure 6.3). The GPS receiver decodes the timing signals from the 'visible' satellites (four or more) and, having calculated their distances, computes its own latitude, longitude, elevation, and time. This is a continuous process and generally the position is updated on a second-by-second basis, output to the receiver display device and, if the receiver provides data capture capabilities, stored by the receiver logging unit.

GPS errors:

1. Satellite errors: Slight inaccuracies in time keeping by the satellites can cause errors in calculating our positions. Similarly, the satellite's position in space is equally important as it is the starting point of the calculations. Although the GPS satellites are at very high orbits and are relatively free from the perturbing effects of atmosphere, they still drift slightly from their predicted orbits which contributes to our errors.
2. The atmosphere: The GPS signals have to travel through charged particles and water vapour in the atmosphere which delays its transmission. Since the atmosphere varies at different places and at different times, it is not possible to accurately compensate for the delays that occur.
3. Multipath error: As the GPS signal finally arrives at the earth's surface, it may be reflected by local obstructions before it gets to the receiver's antenna. This is called multipath error as the signal is reaching the antenna by multiple paths.
4. Receiver error: Since the receivers are also not perfect, they can introduce their own errors which usually occur from their clocks or internal noise.
5. Selective availability: Selective availability (SA) was the intentional error introduced by DoD to make sure that no hostile forces used the accuracy of GPS against the US or its allies. It introduced some noise into the GPS satellite clocks which reduced their accuracy. The satellites were also given some erroneous orbital data which was transmitted as a part of each satellite's status message. These two factors significantly reduced the accuracy of GPS in civilian uses.

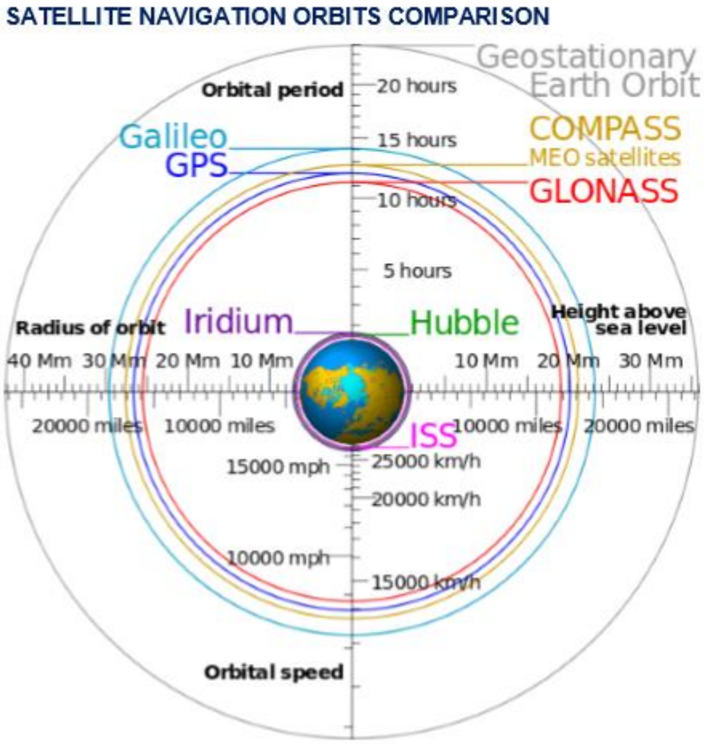
On May 1st, 2000, the White House announced a decision to discontinue the intentional degradation of the GPS signals to the public. Civilian users of GPS will be able to pinpoint locations up to ten times more accurately. The decision to discontinue SA is the latest measure in an on-going effort to make GPS more responsive to civil and commercial users worldwide.

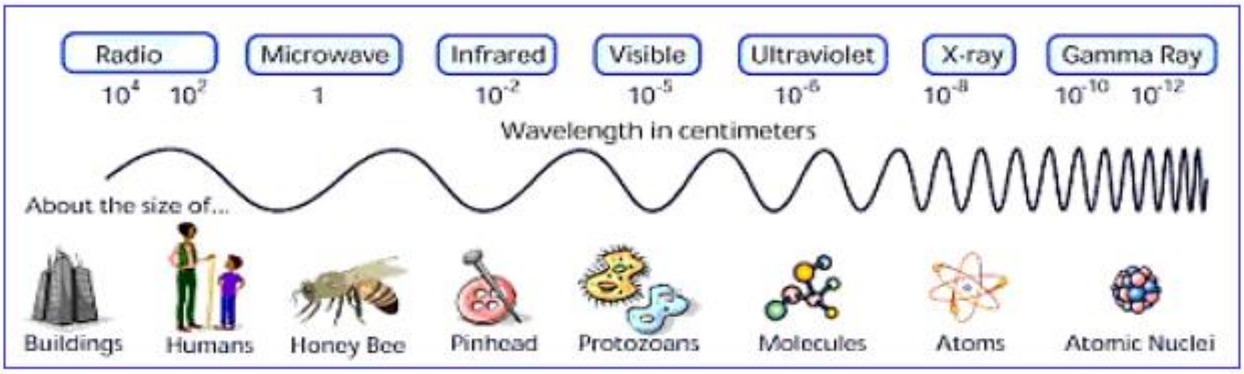
Application of GIS:

1. Precision agriculture
    a. Precise plowing, seeding, watering, spraying
    b. Localized identification and treatment of distressed crops reduces chemical use
    c. Precise leveling of fields prevents fluid runoff
    d. Machinery, asset, and personnel management
    e. Automated tractor control
2. Open pit mining
    a. Enhanced management of assets, equipment
    b. Work progress tracked in real-time, remotely
    c. Improved machine control saves time, lowers maintenance and fuel consumption, prevents accidents
    d. Rapid surveying for drilling blast holes
    e. Smaller, more empowered workforce
3. Timing applications
    a. Communications network synchronization and management
    b. Power grid management and fault location
    c. Financial transactions
    d. E-commerce signatures
4. Other civilian applications
    a. Public Safety
    b. Scientific Research
    c. Environmental Management
    d. Telematics

Remote sensing:

- Remote Sensing is the science and art of acquiring information (spectral, spatial, temporal) about material objects, area, or phenomenon, without coming into physical contact with the objects, or area, or phenomenon under investigation.
- In remote sensing, information transfer is accomplished by use of electromagnetic radiation (EMR).
- EMR is a form of energy that reveals its presence by the observable effects it produces when it strikes the matter.
- So, one is looking at the physical nature of spatially distributed features.





When you listen to the radio, or cook use microwave oven, you are using EM waves. When you take a photo, you are actually doing remote sensing.

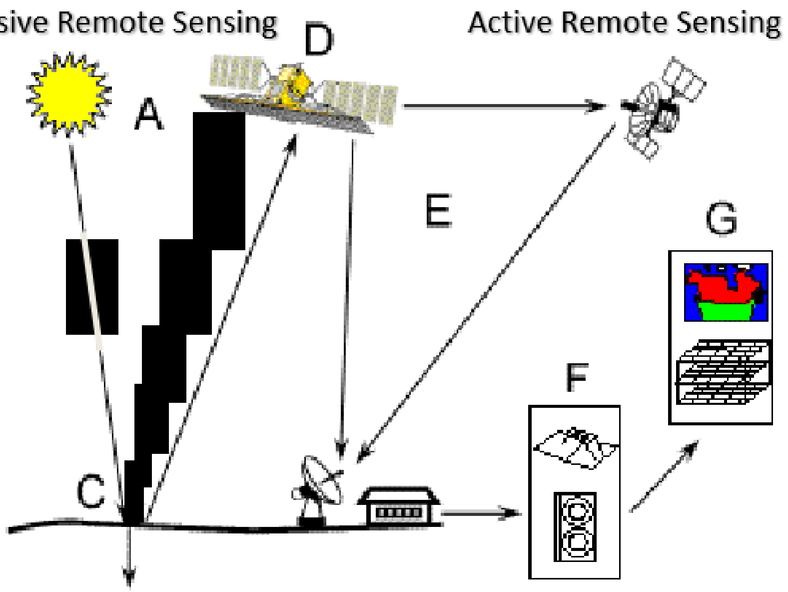
Types of remote sensing:

In respect to the type of Energy Resources:

- Passive Remote Sensing: Makes use of sensors that detect the reflected or emitted electromagnetic radiation from natural sources.
- Active remote Sensing: Makes use of sensors that detect reflected responses from objects that are irradiated from artificially-generated energy sources, such as radar.

In respect to Wavelength Regions: Remote Sensing is classified into three types in respect to the wavelength regions

- Visible and Reflective Infrared Remote Sensing.
- Thermal Infrared Remote Sensing.
- Microwave Remote Sensing.



A. the Sun: energy source
C. target
D. sensor: receiving and/or energy source
E. transmission, reception, and pre-processing
F. processing, interpretation and analysis
G. analysis and application

Remote sensing images:

Remote sensing images are normally in the form of digital images (figure). In order to extract useful information from the images, image processing techniques are applied to enhance the image to help visual interpretation, and to correct or restore the image if the image has been subjected to geometric distortion, blurring or degradation by other factors. There are many image analysis techniques available and the methods used depending upon the requirements of the specific problem concerned.

Use of remote sensing in GIS:

Remote sensing data after can be integrated with various other geographic data. There has been an increasing trend in integration of remote sensing data into GIS for analytical purpose. There many ways we could use remote sensing data and some examples are illustrated as below: Land cover maps or vegetation maps classified from remote sensing data can be overlaid onto other geographic data, which enables analysis for environmental monitoring and its change. Image data are sometimes also used as image maps, with an overlay of political boundaries, roads, rivers etc. Such an image map can be successfully used for visual interpretation.

Importance of remote sensing:

- Large amounts of data needed, and Remote Sensing can provide it
- Reduces manual field work dramatically
- Allows retrieval of data for regions difficult or impossible to reach:
    - Open ocean
    - Hazardous terrain (high mountains, extreme weather areas, etc.)
    - Ocean depths
    - Atmosphere
- Allows for the collection of much more data in a shorter amount of time
    - Leads to increased land coverage AND
    - Increase ground resolution of a GIS
- Digital Imagery greatly enhances a GIS
    - DIRECTLY: Imagery can serve as a visual aid
    - INDIRECTLY: Can serves as a source to derive information such as…
        - Land use/land cover
        - Atmospheric emissions
        - Vegetation
        - Water bodies
        - Cloud cover
        - Change detection (including sea ice, coastlines, sea levels, etc.)

Data digitizing process:

- MANUAL
    - Map is fixed to digitizer table
    - Control Points are digitized
    - Feature Boundaries are digitized in stream or point mode
    - The layer is proofed and edited
    - The layer is transformed/registered to a known system
- AUTOMATED SCANNERS
    - Digitizing done automatically by a scanner
    - There is a range of scanner qualities
    - Most utilize the reflection/transmission of light to record data
    - "Thresholding" allows for the determination of both line and point features from a hardcopy map
    - Editing still required
- DIRECT DATA ENTRY
    - Coordinate Geometry is used, with GPS playing a vital role
    - This involves directly entering in coordinates measured in the field
    - These coordinates can then be tagged with attribute data

- This data this then downloaded to a computer and incorporated into a GIS

4.4 Data preparation, conversion and integration

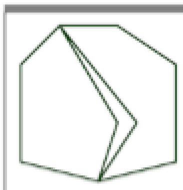
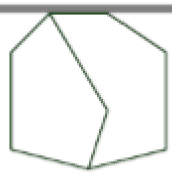
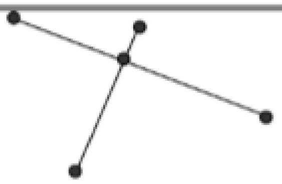
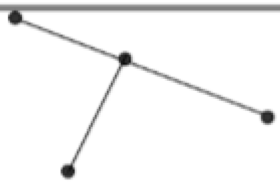
Data preparation:

Spatial data preparation aims to make the acquired spatial data fit for use. Images may require enhancements and corrections of the classification scheme of the data. Vector data also may require editing, such as the trimming of overshoots of lines at interactions, deleting duplicate lines, closing gaps in lines, and generating polygons. Data may need to be converted to either vector format or raster format to match other data sets. Additionally, the process includes associating attribute data with the spatial data through either manual input or reading digital attribute files into the GIS/DBMS.

The intended use of the acquired spatial data, furthermore, may require thinning the data set and retaining only the features needed. The reason may be that not all features are relevant for subsequent analysis or subsequent map production. In this case, data and/or cartographic generalization must be performed to restrict the original data set.

Data precision, error and repair:

Acquired data sets must be checked for consistency and completeness. This requirement applies to the geometric and topological quality as well as the semantic quality of the data. There are different approaches to clean up data. Errors can be identified automatically, after which manual editing methods can be applied to correct the errors. Alternatively, a system may identify and automatically correct many errors. Alternatively, a system may identify and automatically correct many errors. Clean-up operations are often performed in a standard sequence. For example, crossing lines are split before dangling lines are erased, and nodes are created at intersections before polygons are generated.

Before cleanup	After cleanup	Description
		Erase duplicates or silver lines
		Erase dangling objects or overshoots

Precision refers to the level of measurement and exactness of description in a GIS database. Precise location data may measure position to a fraction of a unit. Precise attribute information may specify the characteristics of features in great detail. It is important to realize, however, that precise data--no matter how carefully measured-may be inaccurate. Surveyors may make mistakes or data may be entered into the database incorrectly.

- The level of precision required for particular applications varies greatly. Engineering projects such as road and utility construction require very precise information measured to the millimeter or tenth of an inch.
- Highly precise data can be very difficult and costly to collect. Carefully surveyed locations needed by utility companies to record the locations of pumps, wires, pipes and transformers cost $5-20 per point to collect

Error encompasses both the imprecision of data and its inaccuracies

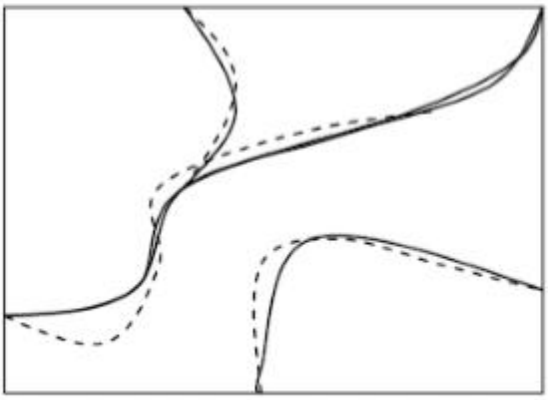
Multiple data sources

A GIS project usually involves multiple data sets, so a next step addresses the issue of how these multiple sets relate to each other. There are three fundamental cases to be considered if we compare data sets pair wise:

- They may be about the same area, but differ in accuracy,
- They may be about the same area, but differ in choice of representation, and
- They may be about adjacent areas, and have to be merged into a single data set.

Differences in accuracy

Images come at a certain resolution, and paper maps at certain scale. This typically results in differences of resolution of acquired data sets, all the more since map features are sometimes intentionally displaced to improve the map. For instance, the course of a river will only be approximated roughly on a small scale map, and a village on its northern bank should be depicted north of the river, even if this means it has to be displaced on the map a little bit. The small scale causes an accuracy error. If we want to combine a digitized version of that map, with a digitized version of a large-scale map, we must be aware that features may not be where they seem to be. Analogous examples can be given for images at different resolutions.

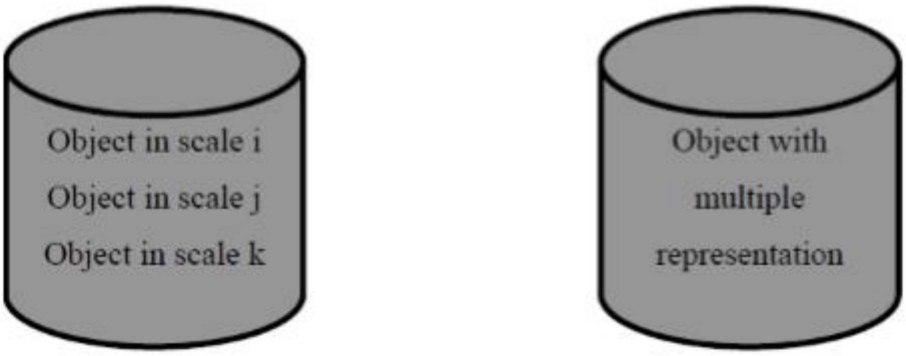


The integration of two vector data sets may lead to silver

In the figure above, the polygons of two digitized maps at different scales are overlaid. Due to scale differences in the sources, the resulting polygons do not perfectly coincide, and polygon boundaries cross each other. This causes small, artifact polygons in the overlay known as silver polygon.

Differences in representation

There exist more advanced GIS applications that require the possibility of representing the same geographic phenomenon in different ways. Map production at various map scale is again an example but there are numerous others. The commonality is that phenomena must sometimes be viewed as points, and at other times as polygons, for instance. The complexity that this requirement entails is that the GIS or the DBMS must keep track of links between different representations for the same phenomenon and must also provide support for decisions as to which representations to use in which situation.



Multi-scale and multi-representation systems compared; the main difference is that multi-representation systems have a built in understanding that different representations belong together.

For example, a small-scale national road network analysis may represent villages as point objects, but a nation wide urban population density study should regard all municipalities as represented by polygons. The links between various representations for the same things maintained by the system allows interactive traversal, and many fancy applications of their use seem possible. The systems that support this type of data traversal are called multi representation systems.

Data Transformation

In virtually all mapping applications it becomes necessary to convert from one cartographic data structure to another. The ability to perform these object-to-object transformations often is the single most critical determinant of a mapping system's flexibility.

Format Change: Raster to vector and vector to raster conversion within the same GIS system. May also include raster to vector and vector to raster data.

Issues to consider:

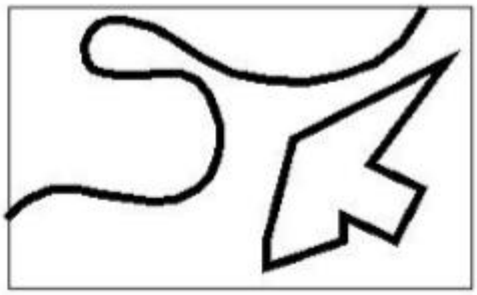
- Loss of detail: especially at features edges, generally vector data more accurately represents a feature
- Loss of attribute data: some raster formats do not allow for multiple attributes per cell

Vector and raster formats store similar GIS data in very different ways.

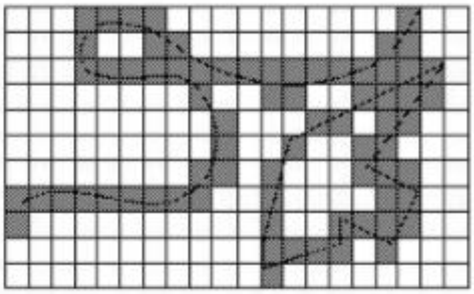
A particular GIS will adopt one of two strategies for dealing with two types of data. Some systems use only one format exclusively and provide utilities or import options to bring in the data and convert it to the needed format.

Other GIS software supports the native format of each type of data and requires the GIS operator to change the formats explicitly when operation requires commonality of formats. The computer program in both cases performs raster-to-vector and vector-to-raster conversion. Most often when converting from vector to raster the results are visually satisfactory, but the conversion techniques can produce results that are not satisfactory to the attributes each grid cell represents. It is particularly true along the edges of areas, where the user seldom knows the decision rules concerning how the partial cells are handled.

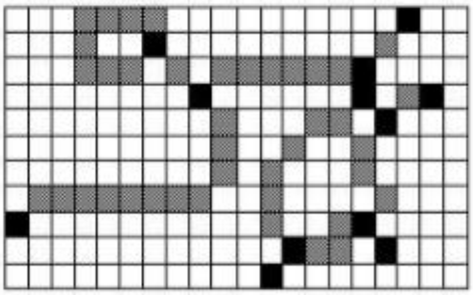
Alternatively, by converting from raster to vector, you may preserve the vast majority of the attribute data, but the visual results will often reflect the blocky, step-like form. The size of the grid cells from which conversion proceeds is an important factor controlling the "blockiness" of the resulting vector. Different mathematical smoothing algorithms can minimize this effect.



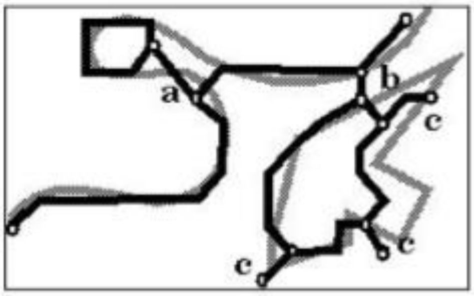
Vector lines



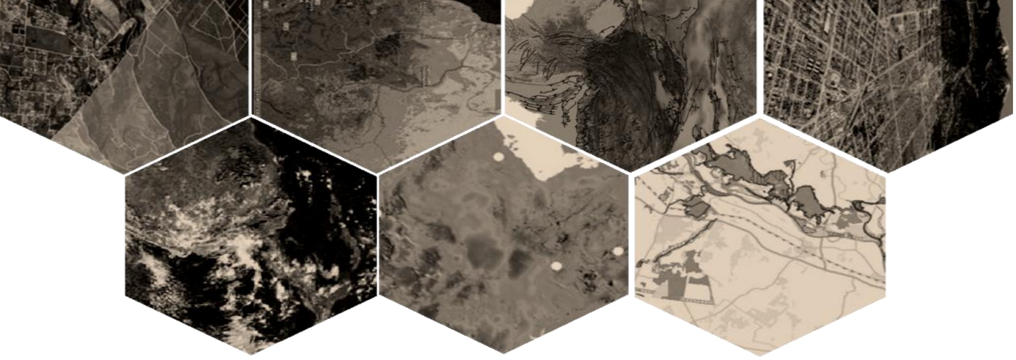
Raster scan of lines



Thinned Image: Endpoints located



Conversion to topological vectors



CHAPTER 5

Spatial Analysis & Visualization

CHAPTER OUTLINE	
5.1	Spatial analysis
5.2	Map outputs and its basic elements

5.1 Spatial analysis

Analyzing data normally comprises three principle phases:

i. **Stating the problem**
ii. **Choice of data**
iii. **Analyses of the data chosen**

All GIS provide functions for analyses of data chosen and for storing results of such analysis. Data may be analyzed at various levels:

i) **Data in attribute table:** Are stored for presentation in reports or for use in other computer system.

ii) **Operation:** Are performed on geometric data, either in search mode or for computational purpose.

iii) **Operations:** Like arithmetic, Boolean, statistical are performed in attribute tables.

iv) **Geometry and attribute table:** Are used jointly to:
   a. Compile new set of data based on original and derived attributes.
   b. Compile new sets of data base on geographic relationships.

Spatial analysis is a process for looking at geographic patterns in your data and relationships between features. The actual methods you use can be very simple – sometimes, just by making a map of the theme you are analysing, or more complex, involving models that mimic the real world by combining many data layers.

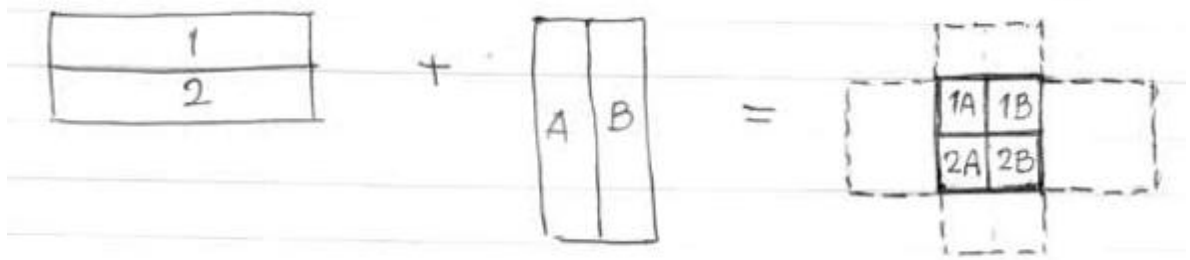
Spatial analysis allows us to study real-world processes. It gives the information about the real world including the present situation of specific areas and features, the change in situation or the trends. For instance – 'where and how much the forest areas are decreasing or increasing?', 'where the urban areas are growing up in the Kathmandu valley?' and so on.

Overlay:

Overlay is the core part of GIS analysis operation. It combines several spatial features to generate new spatial elements. In other word, overlay can be defined as a spatial operation, which combines different geographic layers to generate new information. Overlay is done using Arithmetic, Boolean, and Relational operators, and is performed in both vector and raster domain.

1. **Vector overlay:**

   This overlay operation combines the geometries and attribute of two feature layers to create the output. The geometric of output represent the geometric intersection of features from the input layers.



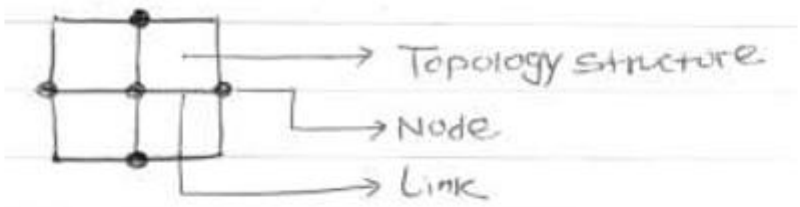
Above figure shows overlay combines the geometrics and attributes from two layers into single layers. The dashed lines are for illustration only and not included in the output. Attributes of above figure is shown below:

ID	Parcel NO	Area
101	1	12 m^2
102	2	12 m^2

ID	Parcel NO	Area
201	A	12 m^2
202	B	12 m^2

ID	Parcel NO	Area
301	1A	1 m^2
302	1B	1 m^2
303	2A	1 m^2
304	2B	1 m^2

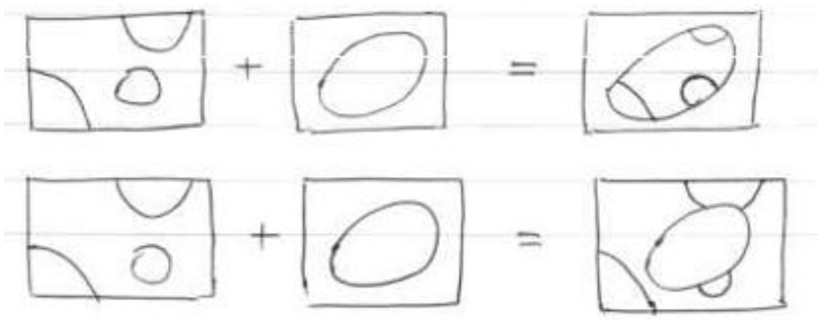
Each feature on overlay contains a combination of attributes from input layers and this combination differs from its neighbours i.e. new intersection are identified as nodes and line between the nodes as links. The new nodes and links then constitute new topological structure.



Feature layers to be overlaid must be spatially registered and based on same coordinate system. In UTM (Universal Transverse Mercator) and SPC (State Plane Coordinate) system must be in same zone with same datum.

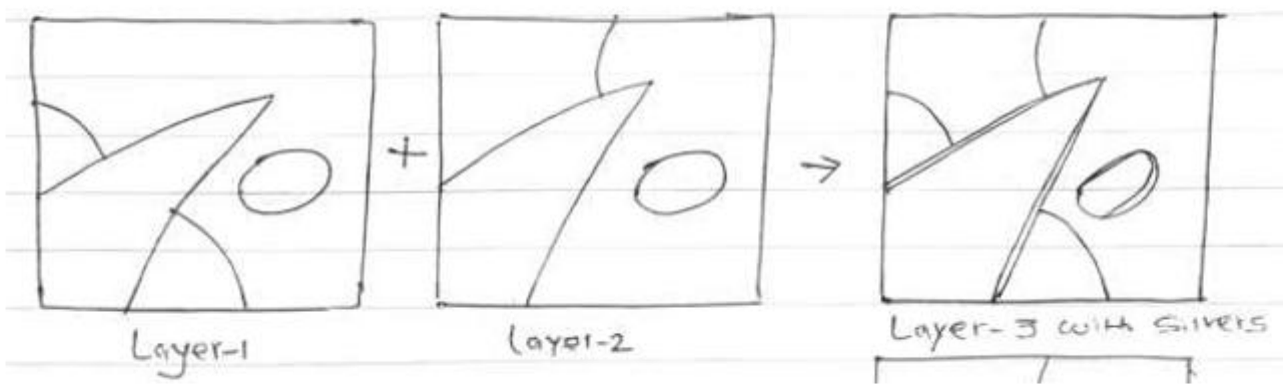
A common error from overlaying polygon layers is silvers, a very small polygon along correlated or shared boundary lines of the input layers.

Computing intersection of large number of polygons can be very time-consuming.
For example: Township border in one thematic layer must be used to clip all other thematic layers in order to produce a collection of data for that township only.



Silvers:
Each new polygon is a new object that is represented by row in attributes table. Each object has new attribute, which is represented by column in attribute table.

Superimposing and comparing two geometric data sets of differing origin and accuracy often give rise to large number of small polygons called silver.

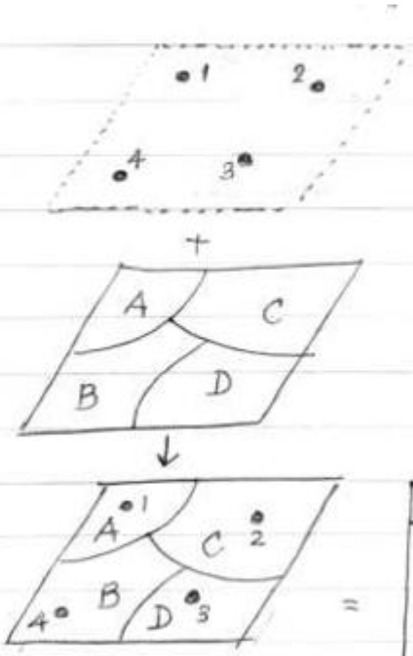


Two polygon representing land area may have slightly differing geometric border on a lake, yet on a piecemeal basis the border may concide. Superimposing the two polygon can then produce an unduly large number of smaller polygon called ‘silver’. These small polygon may be counteraced automatically by laying small zone (fuzzy tolerance).

These silver, an error in the imput can propagate to the analysis output.

Vector overlay and its feature type:
First consideration for overlay is 'feature type'. Overlay operation group uses two polygon layers (polygon – on – polygon) and uses one polygon layers on one point/line (point – in – polygon), (line – in – polygon).

1) Point – in – polygon overlay:



ID	Point No.	Land Mark
101	1	Temple
102	2	Church
103	3	Tap
104	4	Chautara

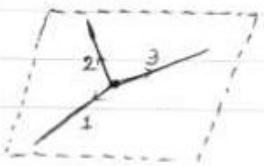
ID	Polygon No	Parcel Type
201	A	Goverment
202	B	private
203	C	Institutional
204	D	Coperative

ID	Point No	Polygon No.	Land Mark	Parcel Type
301	1	A	Temple	Government
302	2	C	Church	Institutional
303	3	D	Tap	Coperative
304	4	B	Chautara	Private

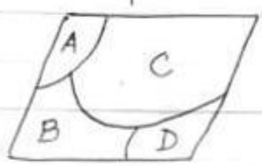
- Points superimposed on polygons.
- So many points be superimposed on polygon.
- Points are then assigned the attributes of polygons on which they are superimposed.
- Relevant geometric operation carried to associate point within polygons.
- Attribute table are updated after all points are associated with polygons.

This polygon method can find association between landmark and parcel type, building and parcel etc.

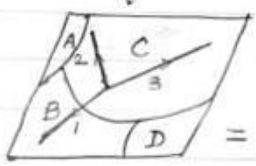
2) Line – in – polygon overlay:



ID	Line No.	Road Name
101	1	Arnico Hw
102	2	Mahendra Hw
103	3	prithivi Hw



ID	polygon No.	zone
201	A	Sagarmatha
202	B	Bagmati
203	C	Janakpur
204	D	Narayani

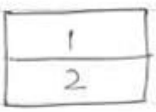


ID	Line No	Road Name	polygon No	zone
301	1	Arnico Hw	B	Bagmati
302	1	Arnico Hw	C	Janakpur
303	2	Mahendra Hw	C	Janakpur
304	3	prithivi Hw	C	Janakpur
305			A	Sagarmatha
306			D	Narayani

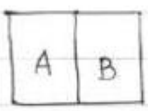
- Line superimposed on polygons.
- As a result, raw set of lines contains attributes of both original lines and plus polygon on which it falls.
- Intersection are computed, nodes and links are formed, topology is established and attribute table is updated.

For example, A line – in – polygon overlay can find soil data for proposed road. The input layer includes the proposed road. The overlay contains soil polygons and the output shows dissected proposed road each road segment having different set of soil data from its adjacent segment.

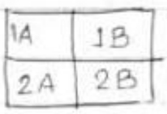
3) Polygon – on – polygon overlay:



ID	Polygon NO	Height
101	1	200 m
102	2	500 m



ID	Polygon No	Vegetation type
201	A	Maize
202	B	Potato



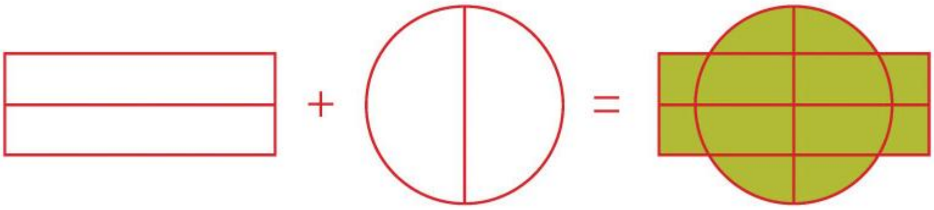
ID	Polygon No	Height	Vegetation type
301	1A	200 m	Maize
302	1B	200 m	Potato
303	2A	500 m	Maize
304	2B	500 m	Potato

- This is most common overlay.
- Output combines the polygon boundaries from the input and overlay layers to create a set of polygons.
- Each new polygon carries attributes from both layers, and these attributes differ from these of adjacent polygon.
- Example above shows the association between elevation zone and vegetation type.

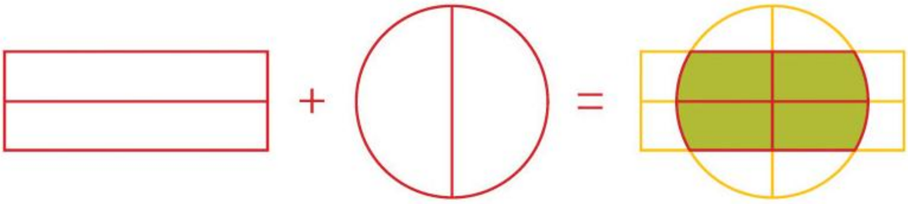
Vector overlay methods:

Overlay methods are based on Boolean connector AND, OR, XOR etc.

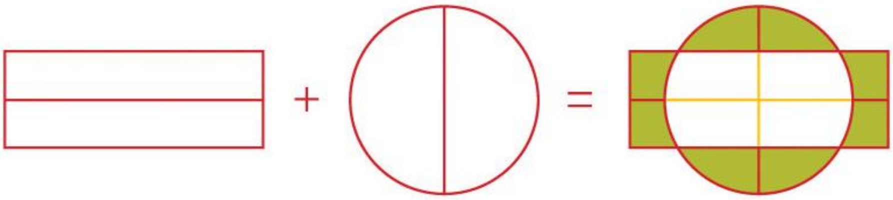
a. **Union:** Preserves all feature from the inputs. The area extent of output combines area extends of both input layers. It requires that both input layers be polygon layers.



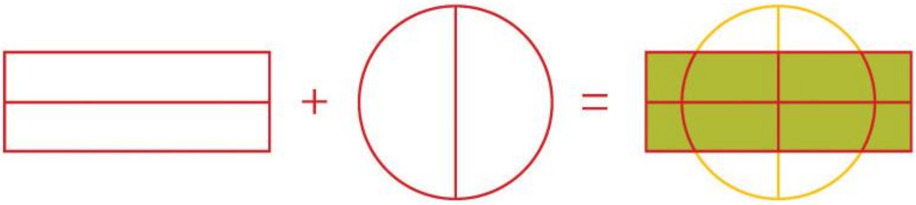
b. **Intersection:** It preserves only those features that fall within the area extent common to inputs. We preferred it because, any feature on its output has attribute data from both of its inputs.



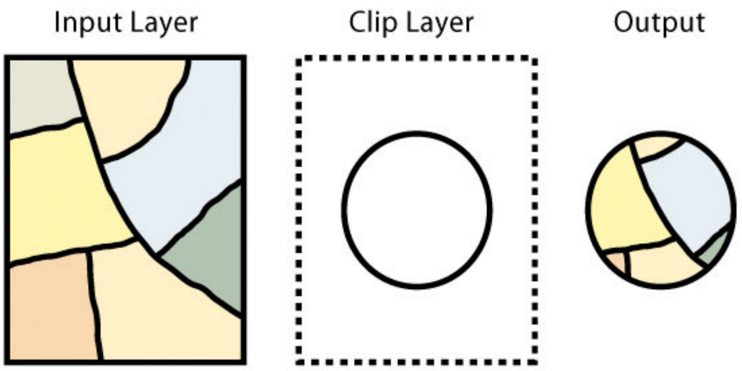
c. **Symmetrical difference:** Preserve features that fall within the area extent that is common to only one of the input. It is opposite to intersect in terms of outputs area extent. This method requires that both input layers be polygon layers.



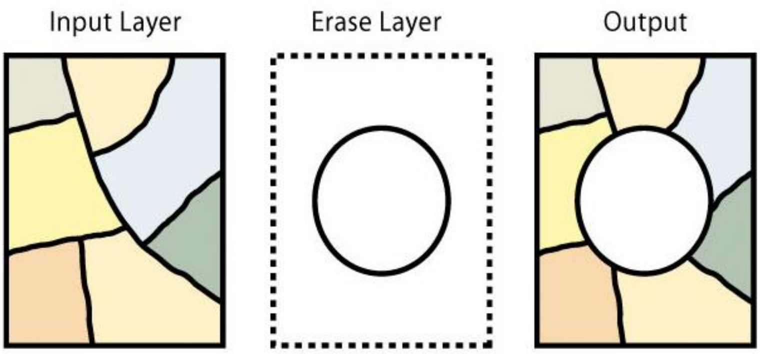
d. **Identity:** Identity preserve only features that fall within the area of layer defined as the input layer. The other layer is called identity layer. Input layer is called identity layer. Input layer may contain point, lines, polygon and identity layer is polygon layer.



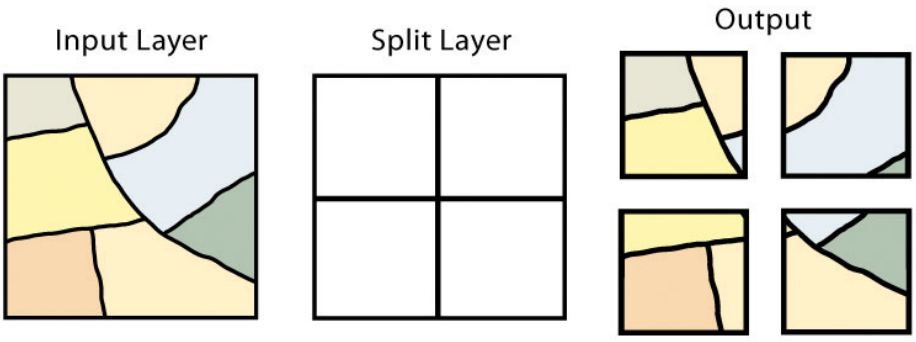
e. Clip



f. Erase

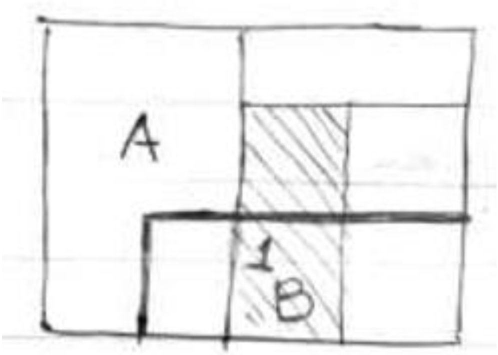


g. Split



Application of vector overlay:

- An overlay operation combines features an attributes from the input layers.
- An overlay output is useful for query and modelling purposes.
- A more specific application of overlay is to help solve a real interpolation problem that involves transferring known data from one set of polygon to another.
- For example:



- Thick lines represent census tracts and thin lines school districts. Census tract 'A' has known population of 4000 and B has 2000, overlay result shows that areal proportion of census tract A in school district 1 is 1/8 and areal proportion of census tracts B is ½.

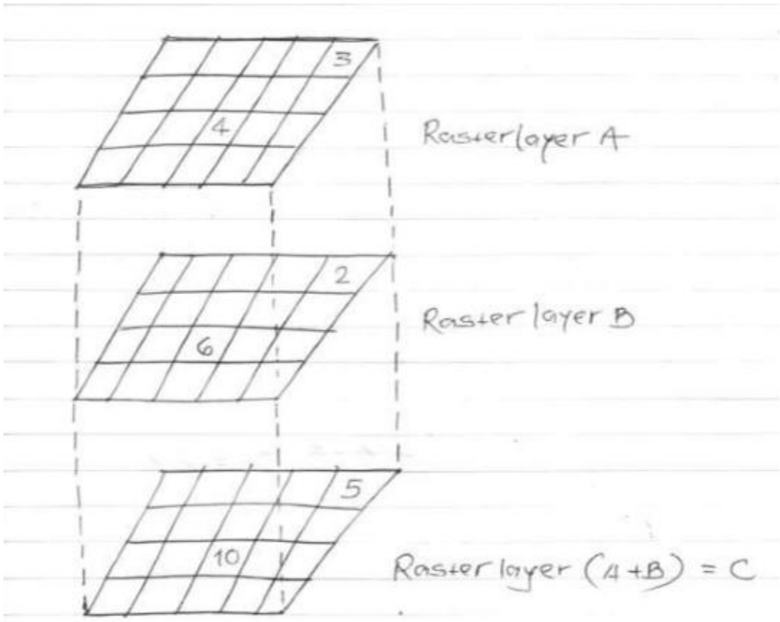
Error propagation in Vector overlay

- Silvers are examples of errors in the inputs that can propagate to the analysis output.
- Error propagation refers to the generation of errors that are due to inaccuracies of input layers.
- There are two types of error:
    - Positional: It can be caused by inaccuracies of boundaries that are due to digitizing or interpretation errors.
    - Identification: It can be caused by inaccuracies of attribute data such as incorrect coding of polygon values.
- Every overlay product tends to have some combinations of positional and identification errors.
- It depends on number of input layers and spatial distribution of errors in input layers.
- The accuracy of an overlay output decreases as number of input layers increases.
- Accuracy decreases if the likelihood of error occurring at the same location in input layers decreases.

2. Raster overlay:

- Raster overlay is simpler than vector overlay and can be carried out directly on cell values.
- It is more efficient than vector overlay, as extent of calculation is much less.
- Both raster layer must have identical geometry i.e. cell size must be same, there must be no relative rotation of transfer between grids.
- Attribute are representation of thematic layers.
- There is no need to distinguish between polygons, lines, points, because all raster data comprise cells.
- Arithmetic, logical, statistical operation may be performed directly during overlay process.
- Deviation is carried by transformation and resampling to the same cell size.

- New composite cells are composed from original cells and registered as a new thematic layer.
- There is no formation of smaller errorneous polygon.



The arithmetic operation of two thematic layers A and B produce a new layer C through operation C = A + B

Other operations are:

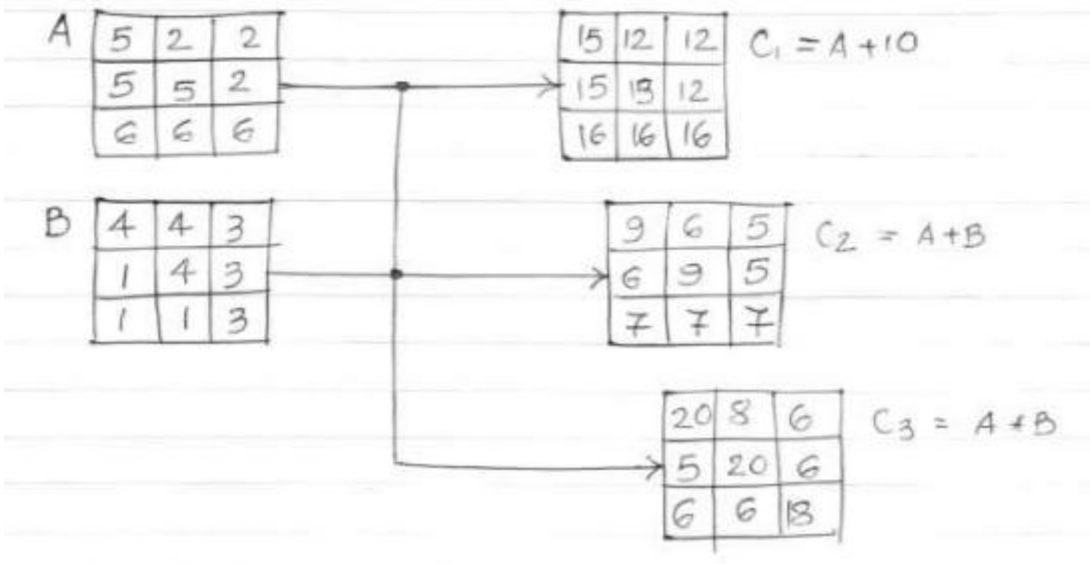
C = A – B
C = A/B
C = A * B

Logical operation is:

If A > 100, C = 10
Else, C = 0

Operators and function used in raster overlay:

a. Arithmetic operators: The operators (+, -, /, *) allows for addition, subtraction, division, multiplication of two raster maps or numbers or combination of the two.



b. Boolean operator: The operator (AND, NOT, OR, XOR) uses Boolean logic (TRUE or FALSE) on the input values. Output values of true are written as 1 and false as 0.

~~AND~~

AND

A	B	output
0	0	0
0	1	0
1	0	0
1	1	1

OR

A	B	output
0	0	0
0	1	1
1	0	1
1	1	1

XOR

A	B	output
0	0	0
0	1	1
1	0	1
1	1	0

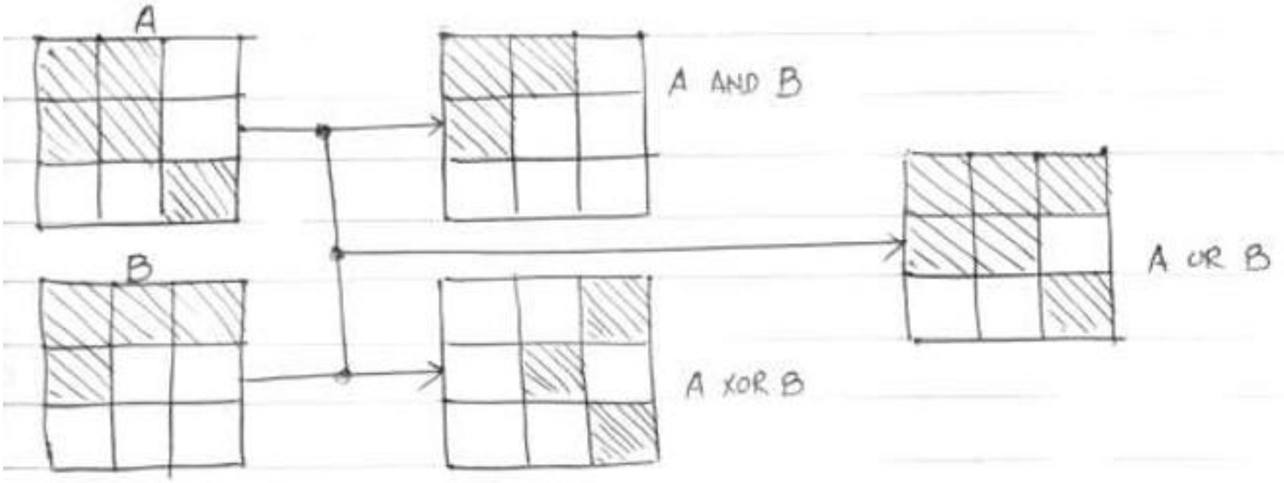
NOR

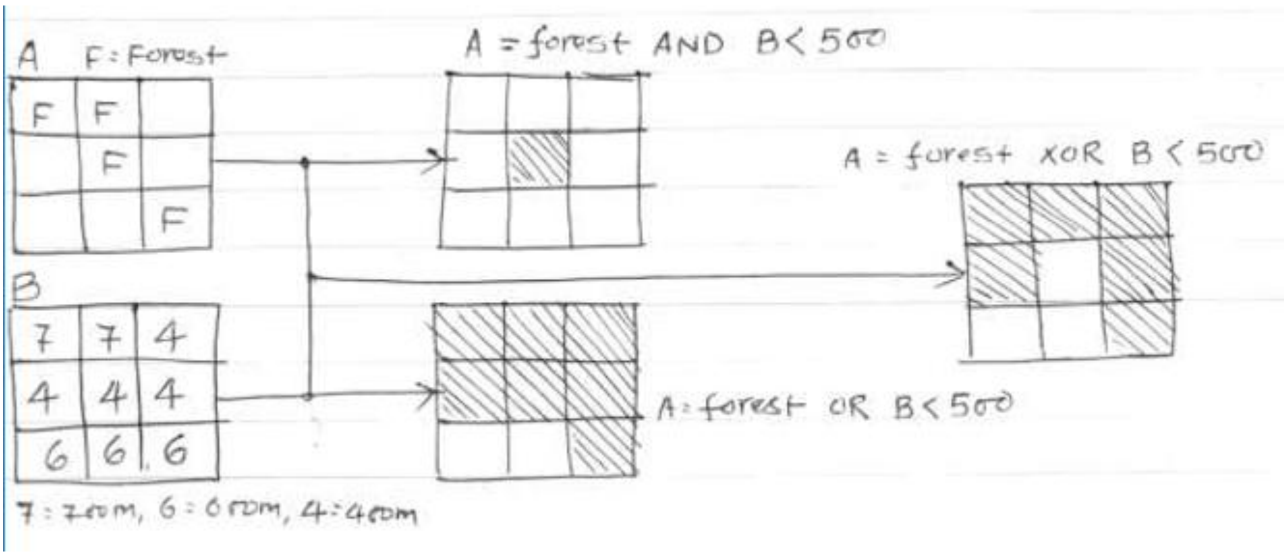
A	B	output
0	0	1
0	1	0
1	0	0
1	1	0

NAND

A	B	output
0	0	1
0	1	1
1	0	1
1	1	0

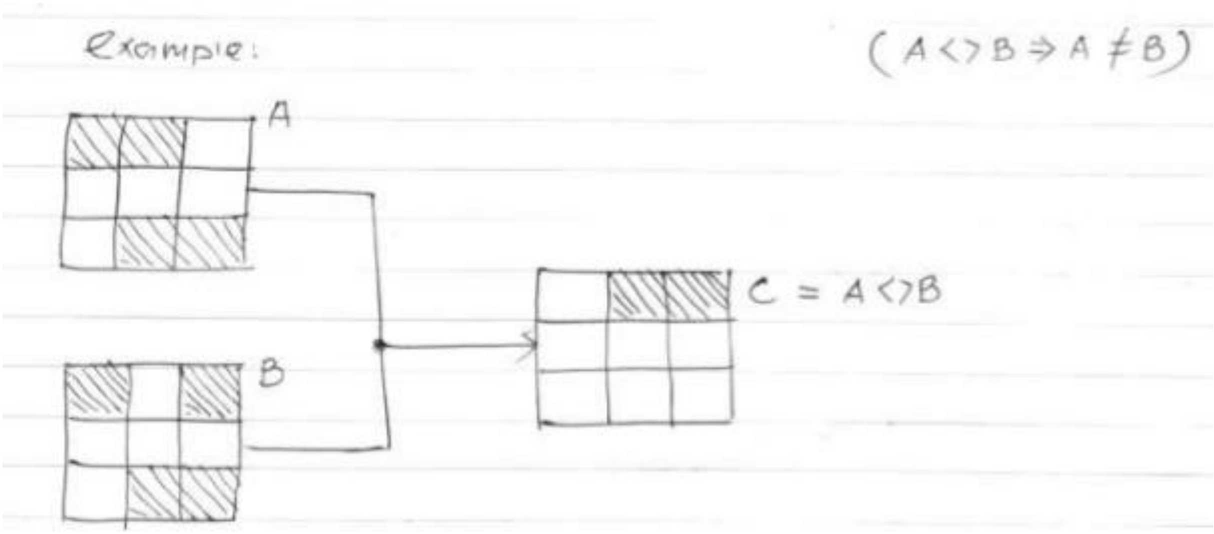
Examples of Boolean operators:





c. **Relational/Comparison operator:** This operator (<, < =, <>, =>, =) evaluate specific relational conditions. If the condition is TRUE the output is assigned 1, if the condition is FALSE output is assigned 0.

For example:



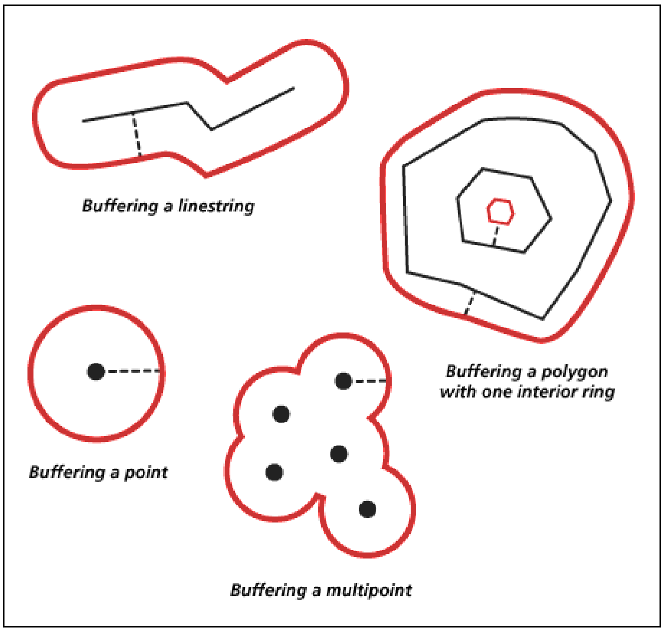
Buffering:

It is defined as GIS operation that creates zones consisting of area within a specified distance of selected feature.

Based on the concepts of proximity, buffering creates two areas:
i. One area that is within a specified distance of selected features.
ii. Other area that is beyond.

Features of buffering:

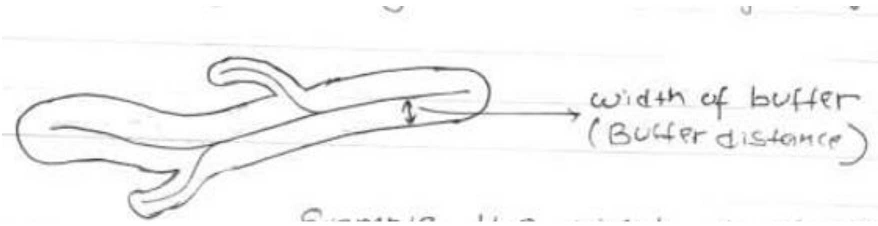
i) **Points:** Buffering around points creates a series of elongated buffer zones.
ii) **Polygon:** Buffering around polygon creates buffer zones that extend outward from the polygon boundaries.
iii) **Lines:** Buffering around lines creates circular buffer zone.



New polygons have the attributes of original objects, can be given new attributes for each elements coordinate are calculated for buffer zone limit.

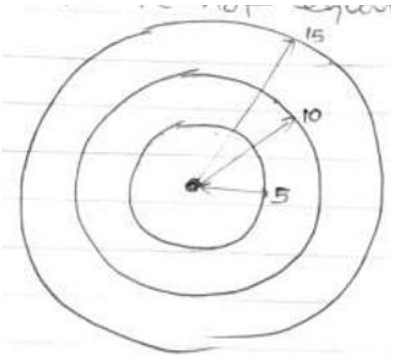
Variation in buffering:

The buffer distance of buffer size does not have to be constant. It can vary according to the values of given field.



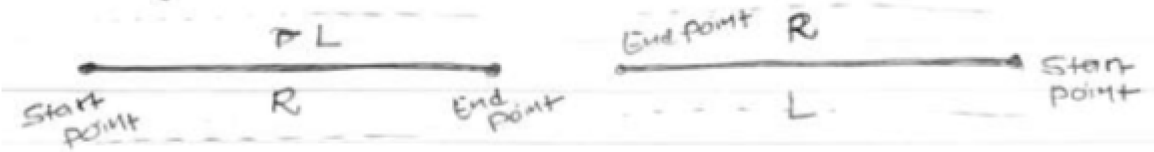
In the example, the width of riparian buffer can vary depending on its expected function and the intensity of adjacent land uses as shown in figure above.

A feature may have more than one buffer zone. Example, A nuclear power plant may be buffered with distance of 5, 10, 15 miles, thus forming multiple rings around the plant are not equal in area as shown below:

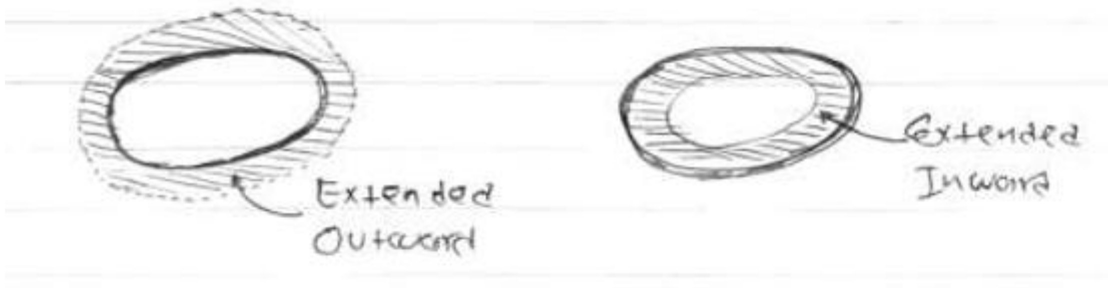


Buffering around line features does not have to be both sides of the lines. It can be on either left or right side of line feature.

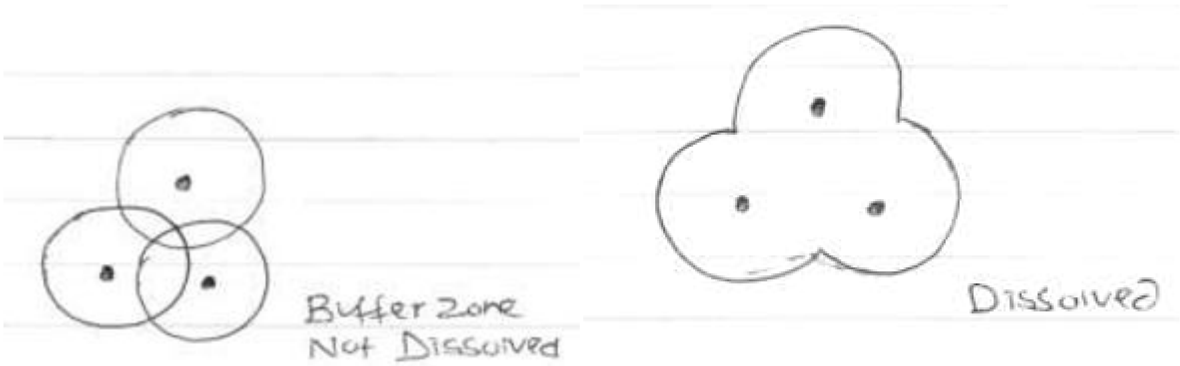
The left or right side is determined by the direction from the starting to end point of a line.



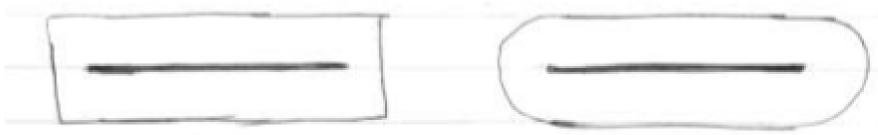
Buffer zone around polygon can be extended either outward or inward from the polygon boundaries.



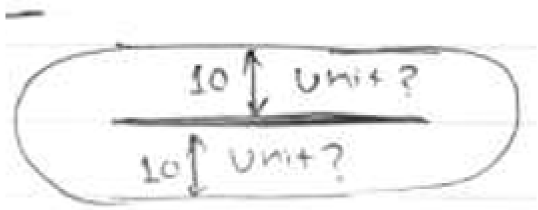
Boundaries of buffer zones may remain intact so that each buffer zone is a separate polygon or these boundaries may be dissolved so that there are no overlapped areas between buffer zones.



The end of buffer zone can be either round or flat but round is best.



We must know measurement unit (Eg. Meter, feet) as buffering uses distance measurement.



Accuracy of buffer zones is determined by positional accuracy of spatial feature, if buffering uses distance measurements from spatial feature.

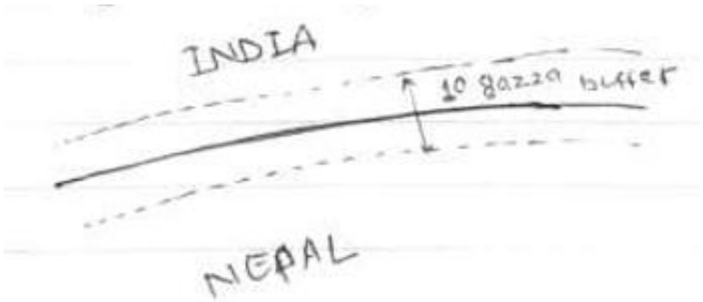
Most GIS package offers buffering as:

- Analysis tool
- Buffer tool
- Editing tool
- Multiple ring buffer tool

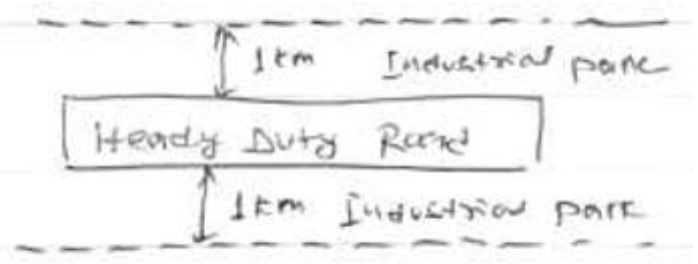
Application of Buffering:

i) Used in planning or regulatory purposes, as buffer zone is often treated as protected zone.

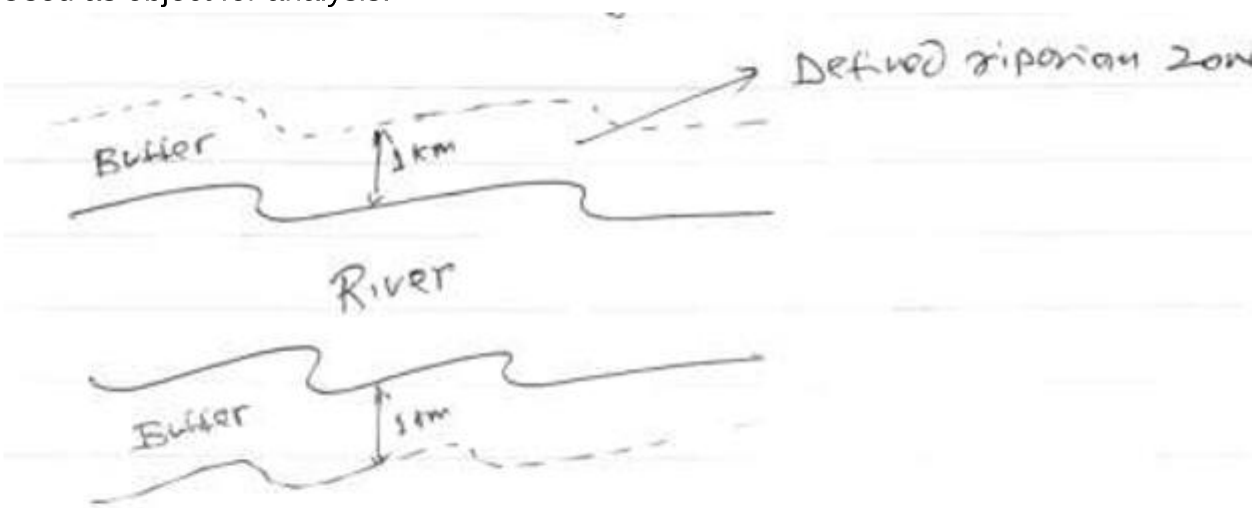
ii) Used in conflict resolution by treating buffer zone as neutral zone.



iii) Used to represent inclusion zone, for example, industrial zone around heavy duty road.



iv) Used as object for analysis.



We can evaluate riparian habitat of wild life by creating buffer zone of defined riparian zone.

v) Used as sampling method with multiple buffering ring.

5.2 Map outputs and its basic elements

Visualization:

Visualization is considered as the translation or conversion of spatial data from a database into graphics. These are in the form of maps, enabling the user to perceive the structure of the phenomenon or the area represented. The visualization process is guided by the saying "How do I say what to whom, and is it effective?" "How" refers to the cartographic methods that are used for making the graphics or the map. "I" refers to the cartographer, or the GIS user who is preparing the map for exploring the data or for presentation. "Say" refers to the semantics that represent the spatial data. "What" refers to the spatial data and its characteristics. "Whom" refers to the map audience and the purpose of the map. The usefulness of a map depends upon the following factors.

Who is going to use map?
The map audience or the users will influences how a map should look like. A map made for school children will be very different from one made for scientists. Similarly, tourist maps and topographic maps of the same area are very much different in their contents and look as they are made for different users.

What is their content?
The usefulness also depends upon the contents of a map. The contents can be seen as primary content (main theme), secondary content (base map information) and supportive content (legends, scale, etc).

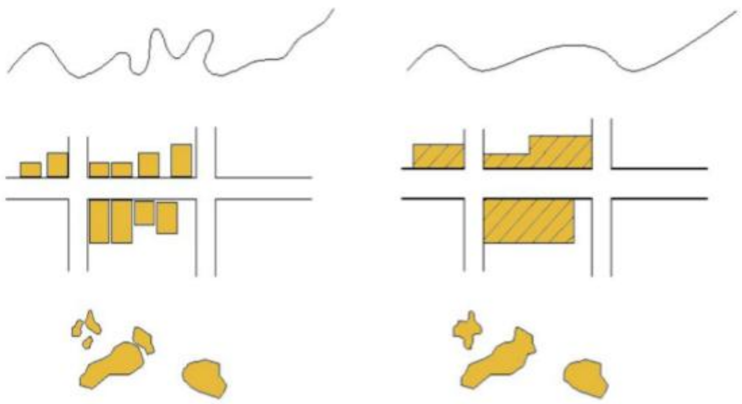
What is the scale of the map?
The map scale is the ratio between a distance on a map and the corresponding distance in the terrain. Scale controls the amount of detail and extent of area that can be shown. Scale of the output map is based upon considerations such as - the purpose of the map, needs of the map user, map content, size of the area mapped, accuracy required etc.

What is the projection of the map?
Every flat map of a curved surface is distorted. The choice of map projection determines how, where and how much the map is distorted. Normally, the selected map projection is that which is also used for topographic maps in a certain country.

Accuracy:
GIS has simplified the process of information extraction and communication. Combining or integrating various data sets has become possible. However, this has created the possibilities of integrating irrelevant or inconsistent data. The user should be aware with the aspects of data quality or accuracy, such as, "What is the source of data? Are the places at correct locations? Are the attribute values correct? Are the themes correctly labelled? Is the data complete?"

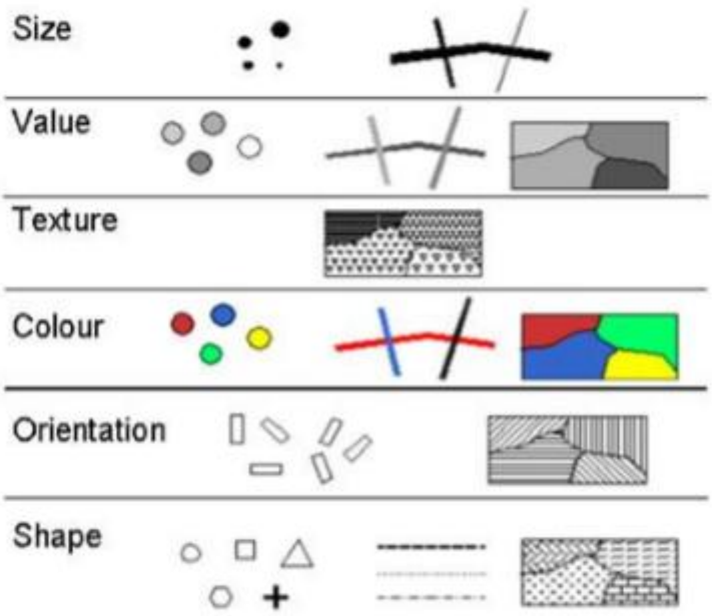
Map design:
Map making is both science and art. A beautiful map may become more popular even if it is less accurate than a plain map. Maps influence people’s perception of space. This influence is partly because of convention and partly because of the graphics used. People understand the world differently, express this understanding differently in maps, and gain different understanding from the maps.

Generalization:



Maps contain a certain level of detail depending upon its scale and purpose. Large scale maps usually contain more detail than small scale maps. Cartographers often generalize the data simplifying the information so that the map is easier to read. The process of reducing the amount of detail in a map in a meaningful way is called generalization. Generalization is done normally when the map scale has to be reduced. However, the essence of the contents of original map should be maintained. This implies maintaining geometric and attribute accuracy as well as the aesthetic quality of the map. There are two type of generalization –graphical and conceptual generalization. Graphic generalization involves simplification, enlargement, displacement or merging of the geometric symbols. Conceptual generalization mainly deals with the attributes and requires knowledge of the map contents and the principles of the themes mapped.

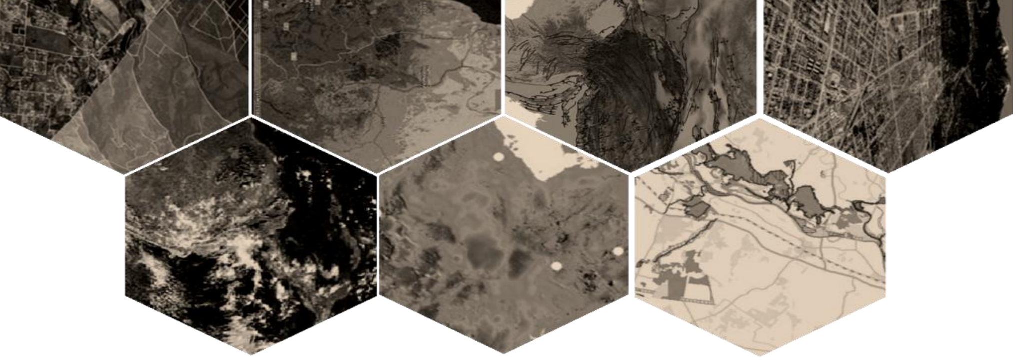
Graphic variables:



The differences in the graphic character of symbols give different perceptions to the map reader. These graphic characteristics are termed as graphic variables which can be summarised as size, Lightness or grey value, Grain or texture, Colour, Orientation, and Shape or form. Knowledge of these basic graphical variables and their perceptual characteristics help the map designers in selecting those variables that provide a sensation which matches the data or the objectives of the map.

Use of colour:

Colour perception has psychological, physiological and conventional aspects. It has been noted that it is difficult to perceive colour in small areas, and more contrast is perceived between some colours than between others. In addition to distinguishing nominal categories, colour differences are also used to show deviations or gradation.



CHAPTER 6

Introduction to Spatial Data Infrastructure

CHAPTER OUTLINE

6.1 SDI concepts and its current trend

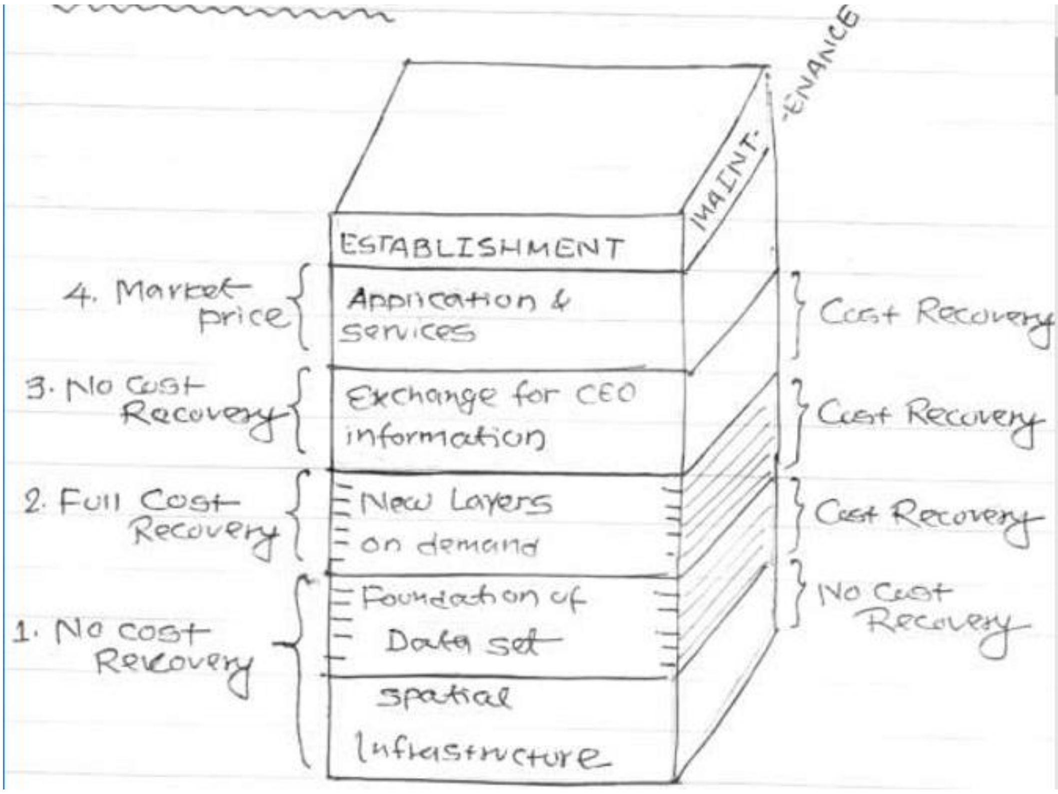
Spatial data infrastructure (SDI) can be defined as data series and electronically based services that satisfy common needs of different user group for accurate positioning and georeferrenced data. Technical development such as internet and GPS have made this type of work very relevant.

National Spatial Data Infrastructure (NSDI)

The definition of NSDI varies from country to country, but it normally includes elements such as:

- **i. Institutional framework:** Defines policy, financial arrangement, human resources needed and legislative and administrative arrangements and cooperation.
- **ii. Geodetic Network and Positioning Services:** Develops and maintains geodetic network and accurate position the services.
- **iii. Standardization:** Develops national standards for geographical data.
- **iv. Fundamental data:** Collects and manages fundamental data sets.
- **v. Technical Frameworks:** Enables users to identify, access and use fundamental data sets.
- **vi. Customer services:** Distribute data and application services.

Model of NSDI:



Above figure shows possible model for establishment of NSDI, which will encourage the production and use of fundamental data and which will also be user driven. The model is based on "No cost recovery" for basic spatial infrastructure and establishment of the foundation data sets, where as maintenance of foundation data sets should be based on Cost Recovery.

6.2 The concept of metadata and clearing house

Data that provides information about geospatial data is metadata. It is also called data of data. Metadata are integrate part of GIS data prepared and entered during data production process.

Metadata are used for GIS project:

i. It let us know if the data meet out specific needs for area coverage, data quality, data concurrency.
ii. It shows how to transfer, process and interpret geospatial data.
iii. It includes the contact for additional information.

Federal Geographic Data Committee (FGDC):

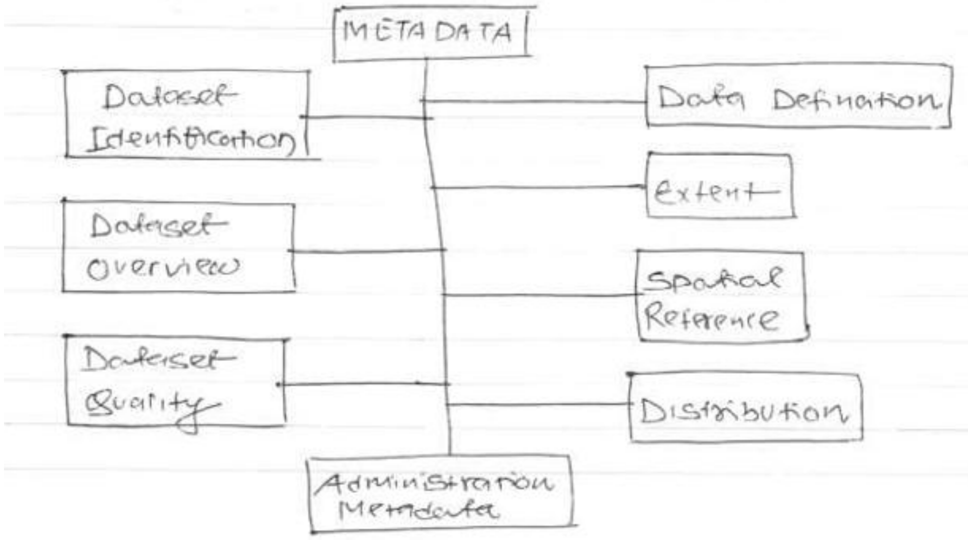
DGDC has developed content standards for metadata and provides detailed information about the standard in following categories:

i) **Identification information:** Data set, title, geographic data covered, currency.
ii) **Data quality information:** Quality of data set, positional and attribute accuracy.
iii) **Spatial data organization information:** Data representation in data set (raster, vector) number of spatial objects.
iv) **Spatial reference Information:** Reference frame, encoding coordinate, coordinate system, datum, map projection.
v) **Entity attribute Information:** Content of data set, entity type.

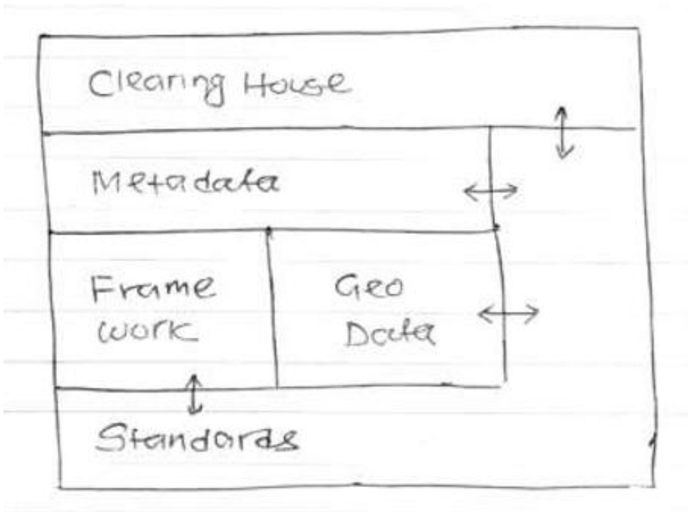
vi) **Distribution information:** Information about obtaining data set.

vii) **Metadata reference information:** Information on the currency of metadata information and responsible party.

Data model for metadata base:



Clearing House:



Goal: To provide access to digital spatial data and related online services for data access, visualization or order.

Objective: To minimize unnecessary duplication of effort for data capture and maximize the benefit of geographic information sharing.

Function: To detailed catalogue service with support for links to spatial data and browse graphics.

Includes: metadata include hyperlinks to online resources (Map services, data download location, data access services, applications) within their metadata entries.

Advertising: It provides low-cost advertising for spatial data providers via internet.

Band data together: It allows geographically defined individuals or communities to band digital data together through federated metadata service.

Processing: User can query on internet with minimum transactional processing with 'peer' clearing house server within clearing house activity.

Institutional view of clearing house: People and infrastructure to facilitate discovery of who has what geographic information.

Technical view of clearing house: A set of information services that uses hardware, software, telecommunication network provide searchable access to information.

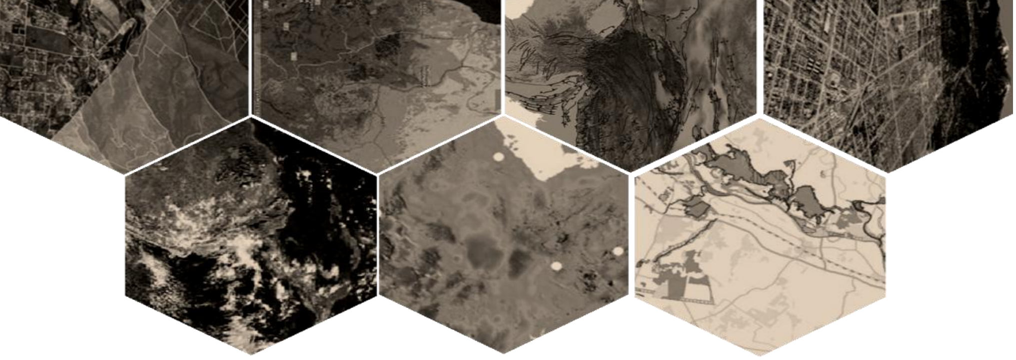
Working of Clearing House:

i. A clearinghouse allows data providers to register their geographic data sets, the quality of these data and also instructions for accessing them.
ii. Each data provider provides on electronic description of each spatial data set. In addition access to spatial data set itself.
iii. Thus, function as detailed catalogue services with support for links to spatial data browsing capabilities.
iv. The data described in the clearinghouse may be located at site of data producers or at sites of designated data disseminators located elsewhere in the country.
v. Computer network facilitates for successful data transmission.

6.3 Critical factors around SDIs

Geographic data that are collected by different data provider can be exchanged and shared by advance technology of Data handling & Data communication. This data sharing give rise to problem which is critical factor in SDI.

i. **Data standard:** Exchange of data between database is difficult if they support different data standard or different query language.
ii. **Heterogeneity:** The inconsistency among data set that exists in variety of location, managed by variety of database, collected by different methods for different purposed, stored in different structure creates many problem when data is shared.
iii. **Communication problems:** Best suit relevant communication technology for transfer of huge amount of spatial data in secure and reliable way is challenging. Efficient tools and communication protocols are important to provide search, browse, delivery mechanism.
iv. **Institutional and economic problems:** The absence of policy for pricing, copyright, privacy, liability, data quality, data standard is essential to create for data sharing in right environment.



CHAPTER 7

Open GIS

CHAPTER OUTLINE

7.1 Introduction of Open concept in GIS

Open source technology is a growing trend in GIS. Open source software is software in which the source code used to create the program is freely available for the public to view, edit, and redistribute. Any type of software program can be open source, including operating systems (e.g., Linux), databases (e.g., PostgreSQL), applications (e.g., OpenOffice.org), games, and even programming languages (e.g., Python).

License

Open source software is identified by the type of license it is released under. These licenses include the Apache 2.0 license, the Microsoft Public License, and the GNU General Public License. While there are some variations, most open source licenses require that the source code be freely available and users are free to modify the source code and redistribute the software and derived works.

Shared community approach

Non-open source software is called closed source or proprietary software. The open source license encourages a shared community approach to the development, extension, and patching of open source software. Most open source projects have a dedicated group that moderates and directs the core software development and ensures that needed new features are being developed, bugs are being fixed, and the supporting documentation remains current.

Why open source software?

- **User demand**: The last 20 years have seen dramatic developments in GIS technology and geographical information science. Fierce competition and growing user demand has resulted in a number of high-quality solutions, which are largely responsible for the vast increase in the GIS marketplace. However, the vast majority of the industry solutions are aimed at supporting basic needs of capture, archival and visualization of spatial data.

- **User-friendly interfaces**: Recent technological advances have concentrated in issues such as user-friendly interfaces, interoperability across data repositories and spatial extensions of database technology. These developments have largely ignored recent advances in GI Science, which include research areas such as spatio-temporal data models, geographical ontologies, spatial statistics and spatial econometrics, dynamic modeling and cellular automata, environmental modeling, and neural networks for spatial data.
- **Integration of spatial data**: Moreover, GIS software development is bound to witness substantial change in the upcoming years, induced by technological advances in spatial databases. Current and expected advances in database technology will enable, in the next few years, the complete integration of spatial data types in data base management systems. This integration is bound to change completely the development of GIS technology, enabling a transition from the monolithic systems of today (that contain hundreds of functions) to a generation of spatial information appliances, small systems tailored to specific user needs. Coupled with the data handling capabilities of new generation of database management systems, rapid application development environments will enable the construction of "vertically-integrated" solutions, directly tailored to the users' needs. Therefore, an important challenge for the GIS community is finding ways of taking advantage of the new generation of spatially-enabled database systems to build "faster, cheaper, smaller" GIS technology.
- **Co-operative development network**: One of the possible responses to this challenge would be to establish a co-operative development network, based on open source technology. In a similar approach as the Linux-based solutions, the availability of GIS open source software would allow researchers and solution developers access to a wider range of tools than what is currently offered by the commercial companies.
- **Resolve the "knowledge gap"**: A second important reason for developing open-source spatial analysis tools is the need to resolve the "knowledge gap" in the process of deriving information from images and digital maps. This "knowledge gap" has arisen because our capacity to build sophisticated data collecting instruments (such as remote sensing satellites, digital cameras, and GPS) is not matched by our means of producing information from these data sources. To a significant extent, we are failing to exploit the potential of the spatial data we collect. For example, there are currently very few techniques for image data mining in remote sensing archives, and thus we are failing to use the information available in our large earth observation data archives.
- **General open source GIS library**: Therefore, the geographical information community would have much to benefit from the availability of a general open source GIS library. This resource would make a positive impact by allowing researchers and solution developers access to a wider range of tools than what is currently offered by the commercial companies. In a similar approach to the Linux and subsequent open source software efforts, we recognize that such development

does not happen by spontaneous growth, but needs a core set of technologies from which further developments may happen.

- **Research**: This co-operative GIS and image processing software environment would allow researchers to share their results with the EO community, thus reducing the "time to market" from academia to society. As an example of such products, a group of R&D institutions in Brazil is currently developing TerraLib, an open-source GIS library that enables quick development of custom-built applications for spatial data analysis.

Open GIS:

Open GIS is the full integration of geospatial data into mainstream information technology. What this means is that GIS users would be able to freely exchange data over a range of GIS software systems and networks without having to worry about format conversion or proprietary data types.

An open source application by definition is software that you can freely access and modify the source code for. Open source projects typically are worked on by a community of volunteer programmers. Open source GIS programs are based on different base programming languages. Three main groups of open source GIS (outside of web GIS) in terms of programming languages are: "C" languages, Java, and .NET.

Fundamental requirements:

- **Interoperable application environment** – a user environment that is configurable to utilize the specific tools and data necessary to solve a problem irrespective of the data structure origin or software.
- **Shared data space** – a generic data model supporting a variety of analytical and cartographic applications.
- **Heterogeneous resource browser** – a method for exploring and accessing the information and analytical resources available on a network – this is becoming an especially important goal with the rise of geo-based Internet sites.

Open GIS Consortium (OGC):

The OGC is a consensus-based association of public and private sector organizations to meet the following three objectives. Its purpose is to create and manage an industry-wide architecture for interoperable geo-processing.

1. The first objective is the creation of an interest group to consolidate the Geographic Information System's industry activities and establish a channel to communicate interoperability issues within the Open GIS Consortium.
2. The second focus is the identification and resolution of interoperability issues and their introduction into the global Open GIS specification process.
3. The third focus is the informative role an Open GIS Consortium would perform to inform the GIS industry about the Open GIS process.

Benefits of Open GIS:

The establishment of standards for an Open GIS system will benefit all users of georeferenced data. Standardization of data structures and processes will help to optimize data exchange between agencies as well as increase the accessibility of GIS to mainstream users.

7.2 Open source GIS and freeware GIS applications

List of open source GIS software:

1. **FlowMap**:
   FlowMap is a freeware application designed to analyze and display flow data. This application was developed at the Faculty of Geographical Sciences of the Utrecht University in the Netherlands.
   Platforms: Windows OS
2. **GMT Mapping Tools**
   GMT is a free, public-domain collection of ~60 UNIX tools that allow users to manipulate (x,y) and (x,y,z) data sets (including filtering, trend fitting, gridding, projecting, etc.) and produce Encapsulated PostScript File (EPS) illustrations ranging from simple x-y plots through contour maps to artificially illuminated surfaces and 3-D perspective views in black and white, gray tone, hachure patterns, and 24-bit color.
   Platforms: UNIX, Macintosh
3. **GRASS**
   Geographic Resources Analysis Support System (GRASS) is the public domain GIS software application originally developed by the US Government. GRASS is probably the most well known open source and original GIS software applications. GRASS is a raster-based GIS, vector GIS, image processing system, graphics production system, data management system, and spatial modeling system. GRASS can be downloaded for free.
   Platforms: Linux, Macintosh, Sun Solaris, Silicon Graphics Irix, HP-UX, DEC-Alpha, and Windows OS
4. **gvSIG**
   gvSIG is an open source GIS application written in Java.
   Platforms: Windows, Macintosh, Linux, UNIX
5. **MapWindow GIS**
   MapWindow GIS is open source GIS application that can be extended through plugins. The application is built using Microsoft's .NET
   Platforms: Windows
6. **OpenJUMP GIS**
   OpenJUMP GIS is an open source GIS written in Java through a collaborative effort by volunteers. Formerly known as JUMP GIS, the application can read shapefiles and GML format files.
   Platforms: Windows, Macintosh, Linux, UNIX
7. **Quantum GIS**
   Also referred to as QGIS, Quantum GIS is an Open Source Geographic Information System (GIS). More: Getting Started With QGIS: Open Source GIS

Platforms: Linux, Unix, Mac OSX, and Windows.

8. **SPRING**

   SPRING is a GIS and Remote Sensing Image Processing system with an object-oriented data model which provides for the integration of raster and vector data representations in a single environment.
   Platform: Windows, Linux, UNIX, Macintosh

9. **TNTLite**

   TNTLite MicroImages, Inc. provides TNTlite as a free version of TNTmips , the professional software for geospatial data analysis. The free TNTlite product has all the features of the professional version, except TNTlite limits the size of Project File objects, and TNTlite enables data sharing only with other copies of TNTlite (export processes are disabled). Can either be downloaded or ordered on CD. Platforms: Windows

10. **uDig GIS**

    uDig GIS is a free, open source GIS desktop application that runs on Windows, Linux and MacOS. uDig was designed to use OGC's OpenGIS standards such as WMS, WFS and more. One-click install allows you to view local shapefiles, remote WMS services and even directly edit your own spatial database geometries. Platforms: Windows, Linux, Macintosh

11. **GeoMajas**

    Written in java, GeoMajas is an open source GIS framework for the web.

12. **GeoServer**

    Java based open source server software that allows users to edit and share geospatial data and uses open standards to spublish GIS data.

13. **MapGuide Open Source**

    First introduced as open source by Autodesk in 2005, MapGuide Open Source allows for the development of web based mapping.

14. **MapFish**

    An open source mapping development framework for web mapping applications based on the Pylon Pythons web framework.

15. **MapServer**

    MapServer is an Open Source development environment for building spatially enabled Internet applications. The software builds upon other popular Open Source or freeware systems like Shapelib, FreeType, Proj.4, libTIFF, Perl and others.

16. **OpenLayers**

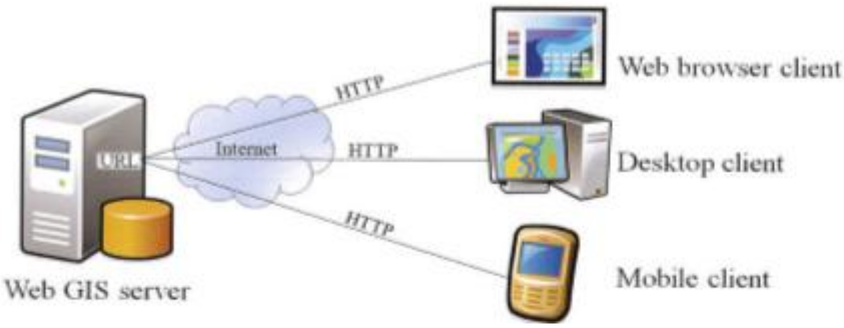
    Javascript library that is open source for displaying GIS data within a browser environment. OpenStreetMap uses OpenLayers for its main map display (aka the "Slippy Map").

17. **TileMill**

    Built on open source libraries (Mapnik, node.js, backbone.js,express and CodeMirror). The Chicago Tribune included TileMill in a series entitled Making Maps using PostGIS, Mapnik, TileMill, and Google Maps.

7.3 Web based GIS system

- Web GIS started off as GIS running in Web browsers and has evolved into Web GIS serving desktop and mobile clients.
- Web GIS is any GIS that uses Web technology to communicate between components: server(identified by URL) and client (a browser, a desktop application, or a mobile application). The communication is via HTTP. The format of the response can be an HTML, binary image, XML(Extensible Markup Language), or JSON(JavaScript Object Notation).
- The simplest architecture: 2-tier client-server (C-S) system. Server program runs on the Web or in the cloud. C-S can be on one computer.
- Multi-tier system: 3 tier (including a data tier), >3 tiers in mashup Web GIS.
- GeoWeb is not identical to WebGIS. GeoWeb is the merging of geospatial information with non-geospatial information (e.g., Webpages, photos, videos, and news). GeoWeb is related to the geotagging and geoparsing research area of WebGIS.



Characteristics of web GIS:

- Global reach by HTTP.
- Support a large number of users simultanuously: requires high performance and scalability.
- Better cross-platform capability:
    - Different Web browsers: IE, Firefox, G. Chrome for diverse OSs (Win, Linux, Mac OS, iOS).

- Web GIS relying on HTML clients supports different operating systems(OSs).
- Web GIS relying in Java, .Net, and Flex can run on multiple platforms.
- However, Web GIS for mobile clients is far from being cross-platform b/c of the diversity in mobile Oss and the incompatibility of mobile Web browsers.

- Easy to use for end users. “If I do not know how to use your site, it is your fault”.
- Unified system update.
- Diverse applications. Neogeopgarphy,("new geography“), is commonly applied to the usage of geographical techniques and tools used for personal and community activities or for utilization by a non-expert group of users

Functions of web GIS:

- Mapping/visualization and query (attribute or spatial).
- Collaborative collection of geospatial information. E.g., wikimapia, OpenStreetMap, VGI (volunteered geographic information).
- Geospatial analysis: measurement, optimal driving path, routing, pollution dispersion modeling, retail site selection etc.

Uses of web GIS:

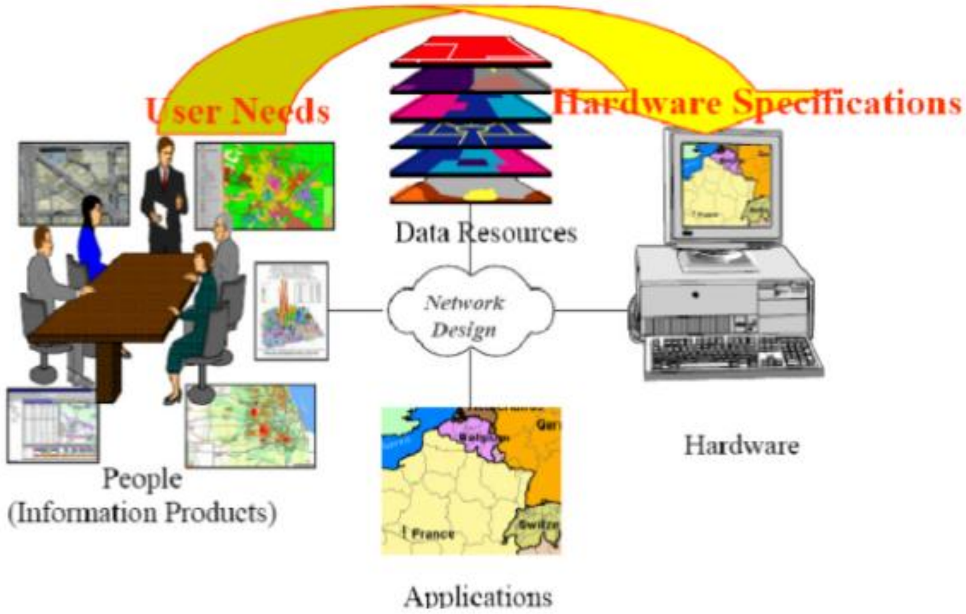
Web GIS as a new business model and a new type of commodity: Location-specific advertising based on map mapping, e.g., Google Map. SaaS business model: Web GIS can also be provided as a commodity. E.g., ESRI Business Analyst Online (BAO). • Web GIS as an engaging and powerful tool for e-government.

A new infrastructure for e-science: E-Science (or eScience) is computationally intensive science that is carried out in highly distributed network environments, or science that uses immense data sets that require grid computing (the federation of computer resources from multiple administrative domains to reach a common goal); Web GIS provides an infrastructure for geo-science research collaboration.

Web GIS in daily life: location-based service (LBS) supported by mobile Web, smart phones and tablets. LBS include services to identify a location of a person or object, such as discovering the nearest banking cash machine or the whereabouts of a friend or employee. LBS include parcel tracking and vehicle tracking services. LBS can include mobile commerce when taking the form of coupons or advertising directed at customers based on their current location. They include personalized weather services and even location-based games.

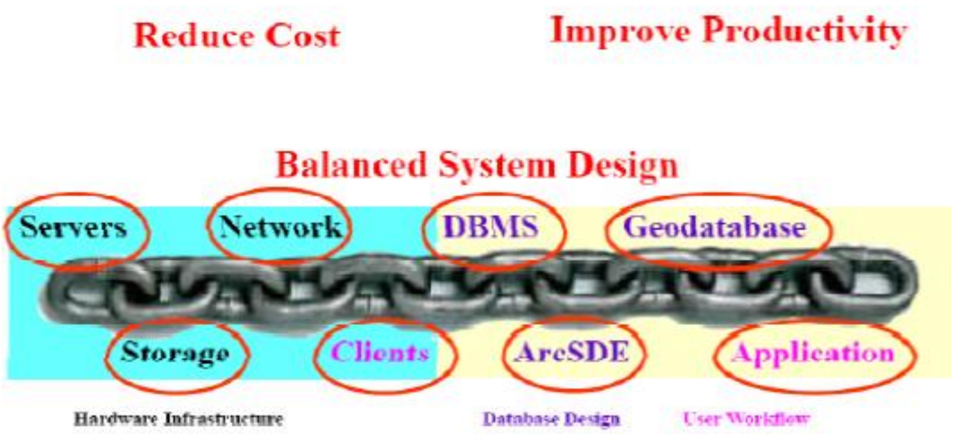
7.4 System analysis and design in GIS

System architecture design is a process developed by ESRI to promote successful GIS implementations. This process supports existing infrastructure requirements and provides specific recommendations for hardware and Application requirements, data resources, and determining the optimum hardware solution as shown in figure:



Importance of system design:

A distributed computer environment must be designed properly to support user performance requirements. The weakest "link" in the system will limit performance. The system architecture design process develops specifications for a balanced hardware solution. Investment in hardware and network components based on a balanced system load model provides the highest possible system performance at the lowest overall cost as shown in figure below:



System architecture design provides a framework for supporting the implementation of a successful enterprise GIS. User workflows must be designed to optimize interactive client productivity and efficiently manage heavier geoprocessing loads. The geodatabase design and database selection should be optimized to support performance requirements. The selected system platform components (servers, client workstations, storage systems) must perform adequately and have the capacity to support peak user workflow requirements. The system architecture design strategy must address performance needs and bandwidth constraints over distributed communication networks— technology and configuration must be selected to conserve shared infrastructure resources. System architecture design provides a solid foundation for building a productive operational environment.

☺☺☺