Geomatics Canada Géomatique Canada

Canada

Introduction to Geographic Information Systems (GIS)

GeoAccess Division
Canada Centre for Remote Sensing
Natural Resources Canada
and
Centre for Topographic Information
Natural Resources Canada

September/October 1997

Natural Resources Canada / Ressources naturelles Canada

Louiselle St-Laurent

Spatial Systems Section, GeoAccess Division
Canada Centre for Remote Sensing
615 Booth St., Ottawa. K1A 0E9

Tel: (613) 943-1312
Fax: (613) 952-0916
E-mail: louiselle.st-laurent@geocan.nrcan.gc.ca

Natural Resources Canada / Ressources naturelles Canada

Maria Glieca

Client Services Section, GeoAccess Division
Canada Centre for Remote Sensing
615 Booth St., Ottawa. K1A 0E9

Tel: (613) 992-1949
Fax: (613) 952-0916
E-mail: maria.glieca@geocan.nrcan.gc.ca

Canada Centre for Remote Sensing / Centre canadien de télédétection

Schedule

8:30 - 10:00 a.m.	What is GIS?
10:00 - 10:15 a.m.	Break
10:15 - 12:00 a.m.	The GIS Database
12:00 - 1:00 p.m.	Lunch
1:00 - 2:15 p.m.	GIS Analysis
2:15 - 2:30 p.m.	Break
2:30 - 3:15 p.m.	GIS Applications
3:15 - 3:30 p.m.	Summary and Conclusion

Workshop Objectives and Requirements

Objectives
- To understand the fundamental principles of GIS
- To acquire tools to further the learning and use of GIS

Requirements
- Some basic knowledge of Geography and RDBMS
- Some basic understanding of current trends in information technology

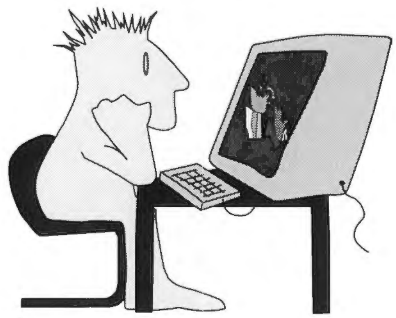
Module 1

What is GIS?

Module 1

Objective:

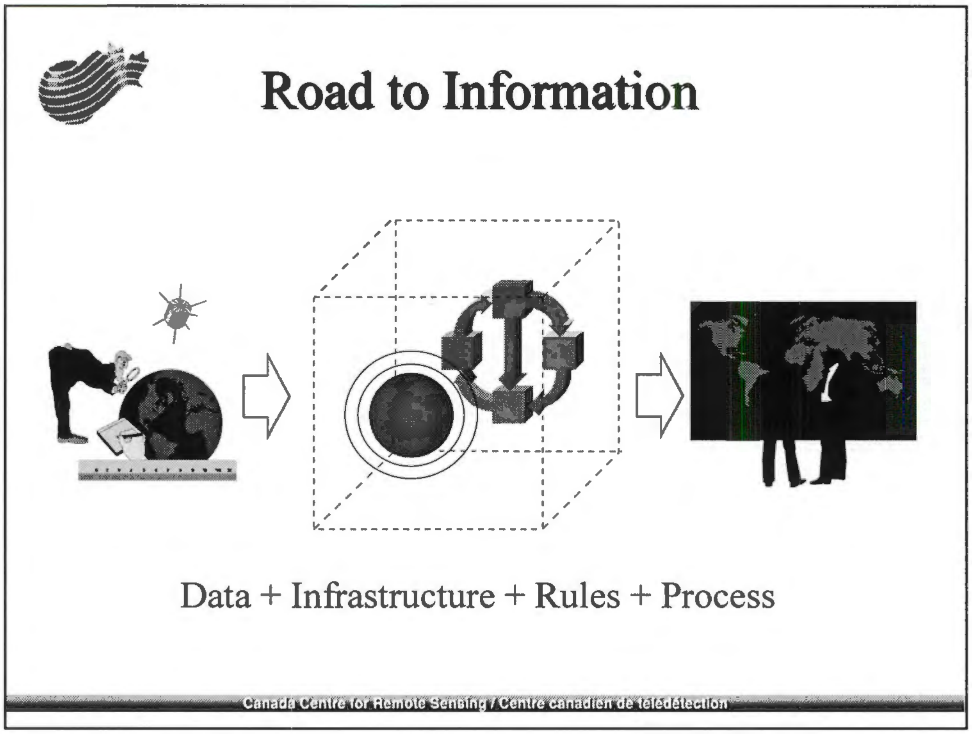
- To become familiar with the components of GIS and the GIS database.



Contents:

- Information framework and GIS
- Components of GIS
- The nature and organization of geographic data

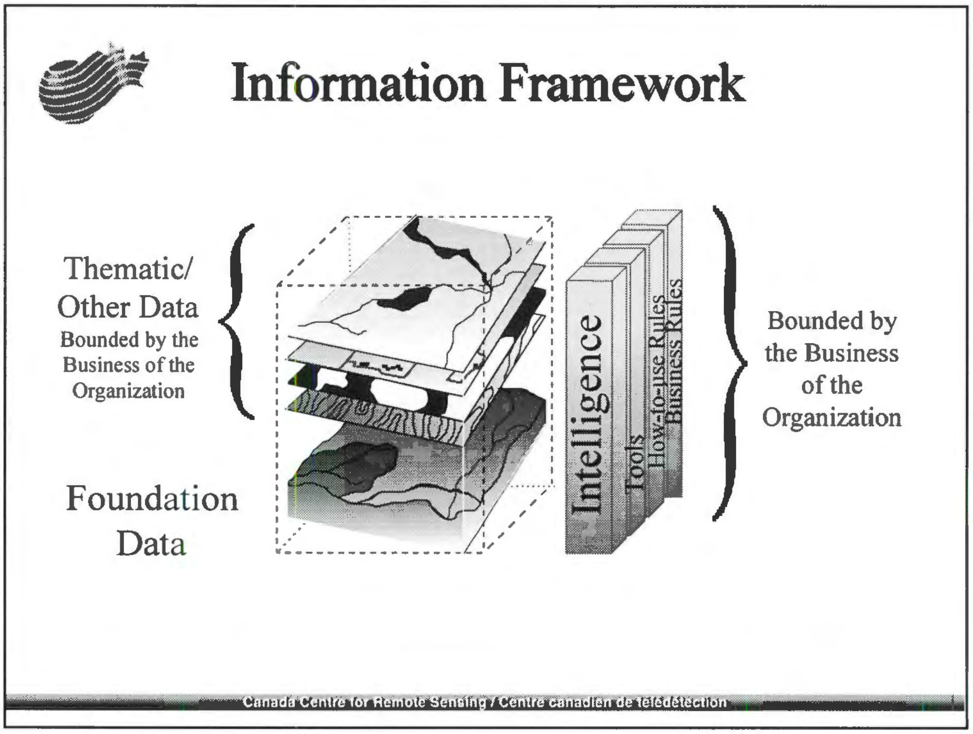
Canada Centre for Remote Sensing / Centre canadien de télédétection



In a decision-making process, information is derived from the manipulation and analysis of specific data elements extracted from real world observations.

The manipulation stage involves structuring and organization of the collected data. The process is supported by a “geospatial” infrastructure.

The analysis stage requires that sets of rules, dependent on the geospatial nature of the data and the context of the issues involved, be applied to the organized data.



In the geospatial domain, the decision-making process relies on a framework, the geographic reference frame plus the components specific to the application domain.

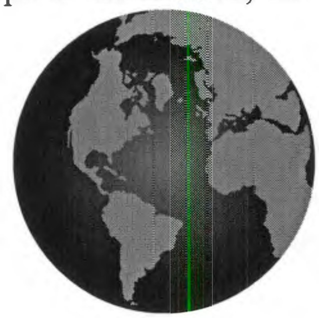
Geographic information systems are used within this framework.

Geomatics and GIS

Geomatics is the Science and Technolgy dealing with the character and structure of Spatial Information, its methods of

- Capture
- Organization
- Classification
- Qualification
- Analysis
- Management
- Display
- Dissemination

} GIS



as well as the infrastructure neccessary for the optimal use of the information.

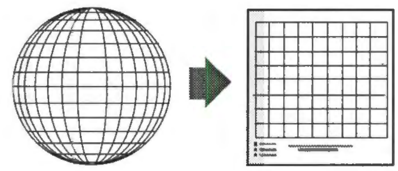
Canada Centre for Remote Sensing / Centre canadien de télédétection

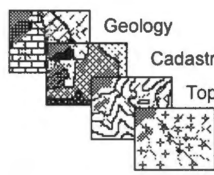
GIS provides the tools to deal with geospatial data.

Why GIS?

Restrictions with Paper Maps

- Representation is limited to specific map projections and scales
- Complex analysis is difficult
- Lengthy manual process is involved for extraction or merging of themes

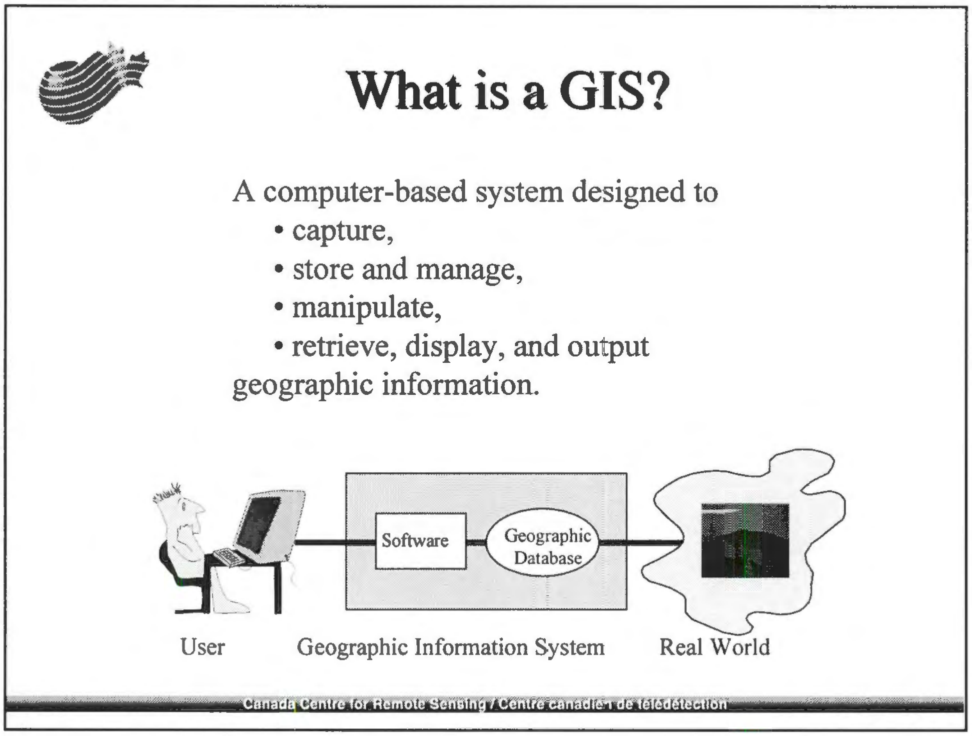




Canada Centre for Remote Sensing / Centre canadien de télédétection

GIS supports automated processes to answer questions on a specific feature or group of features of the earth, such as -

- Where is it?
- What are its characteristics?
- What's around it?
- How does it relate to other features?



In the strictest sense, a GIS is a computer system capable of assembling, storing, manipulating, and displaying geographically referenced information, i.e. data identified according to their locations. Practitioners also regard the total GIS as including operating personnel and the data that go into the system.

A GIS shares elements from other technologies: computer cartography, computer graphics, computer aided design (CAD), database management systems, image analysis systems and desktop mapping. It is often confused with these technologies; however, it differs from any one in particular in that it has the ability to synthesize varied sources and forms of spatial information.

Additionally, a GIS can link spatial information with attributes, as well as, model numerous scenarios. These qualities make GISs a valuable component in problem-solving / decision-making processes for a wide range of applications.

GISs are sometimes 'relabeled' to reflect specific applications. Some of these terms include:

- Land Information System (LIS)
- Environmental Information System (EIS)
- Natural Resource Management Information System

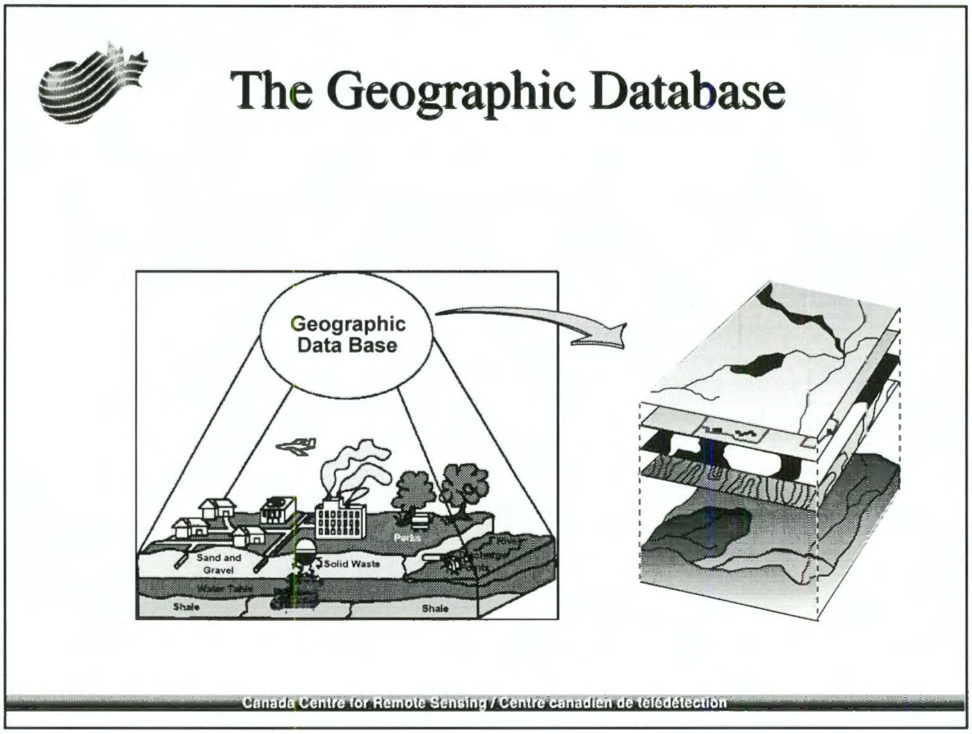
Components of a GIS

- **Geographic Database**
  - integrated digital representation of the real world, manipulated to produce information to address a wide range of needs.
- **GIS Software**
  - used to control and perform operations on the database.
- **Hardware**
  - used to store, process and display.
- **Expertise**
  - the human element required to operate the system (ask questions, interpret results).

Canada Centre for Remote Sensing / Centre canadien de télédétection

This workshop addresses only the first three GIS components.

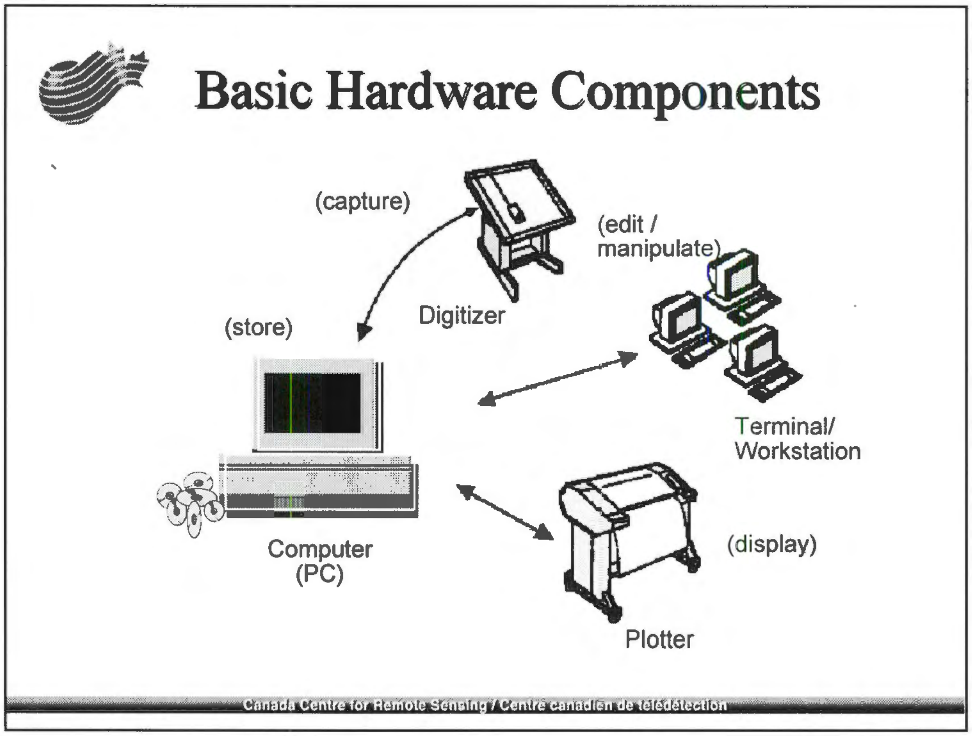
The human element encompasses the people responsible for overseeing GIS implementation and operation - managers, specialists, digitizers/key operations, programmers, users, etc., and the experts who contribute to the knowledge base in the fields of GIS application.



The contents of the geographic database depends on the application. For example: to locate potential sites for a new water well, a municipal planner might need information on the water service areas, land use and land cover (to select undeveloped areas), water quality of the streams (to eliminate non-drinkable water sources), point sources of pollution and description of pollutants (to avoid toxic areas), surficial geology (to show earth materials above bedrock in order to select the most probable water storage - sand, gravel), and, other components such as roads and streams. A geographic database includes features and attributes of these features that are considered important to the decision making process.

A GIS, which can use information from many different sources, in many different forms can help with such analysis. The primary requirement for the source data is that the variables be geographically referenced.

A simple representation of the geographic database is illustrated by the thematic layer model. Each feature of the geographic area is modelled into a "map-like" layer called a theme. Traditional maps are abstractions of the real world, a sampling of important elements portrayed on a sheet of paper with symbols to represent physical objects. GISs also require an abstraction of the real world. However, GISs emphasize the spatial relationships among the objects being mapped. While a computer-aided mapping system may represent a road simply as a line, a GIS may also recognize that road as the border between wetland and urban development, or as the link between Main Street and Blueberry Lane. One of the main component of the model is the reference base - that layer to which all the coordinates and shape of all other layers must match - usually topographic features.



Hardware refers to the equipment required to operate a functional system. Standard hardware to meet minimal requirements includes:

- a computer terminal with extended memory, disk storage (in the gigabytes), graphics capabilities, and network connections.

- a colour graphics monitor

- extra storage drives (tape, CD,other..)

- a digitizing tablet

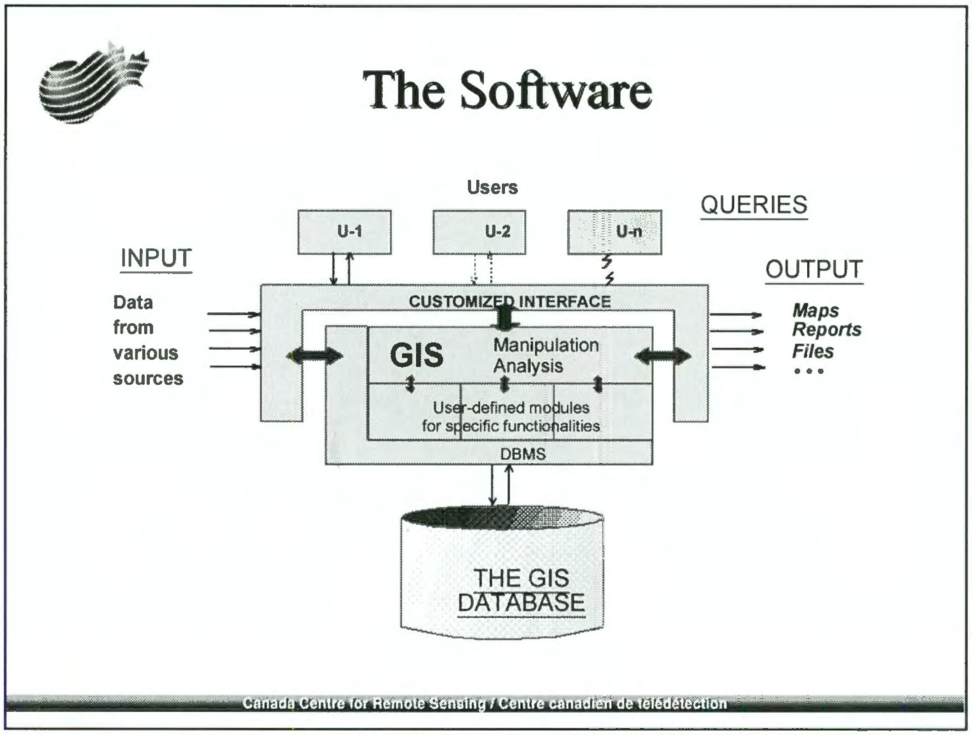
- a plotter

- a printer

The selection of hardware will vary according to the application and data requirements.

Data entry might require only a digitizing tablet. Often, a scanner is used to transform analog data rapidly into digital format. Archived digital data can be transferred from tape drives, optical disks or even over a network.

Data output, such as hardcopy (maps, lists, tables) is achieved using printers (laser, ink-jet, wax-transfer) and plotters (pen, ink-jet, electro-static). Softcopy output is done through a colour graphics monitor. Electronic output requires peripherals such as tape drives, floppy disk drives or optical disk drives and/or telecommunications media.

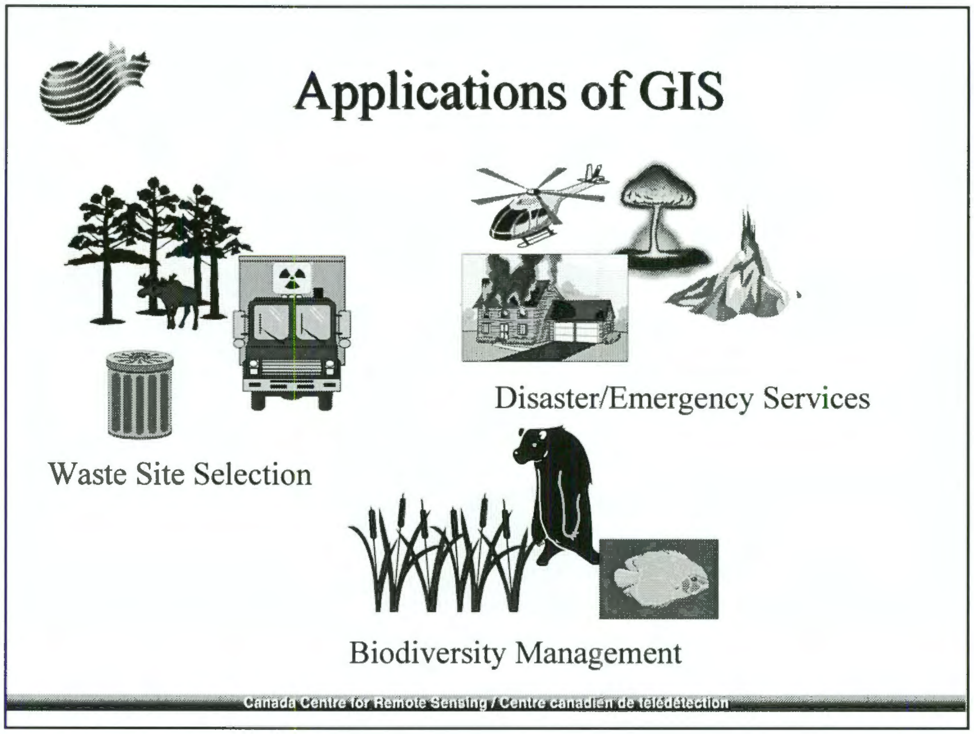


The GIS software is a collection of computer programs (procedures and functions) which are used to create and manage the geographic database, manipulate the data and produce output.

For the user, the main view of the software is through the interface. The majority of modern GIS software uses a graphic interface with windowing capabilities, widgets and point-and-click interactions. With a GIS, the decision maker can "point" to a location, object, or area on a screen and retrieve recorded information about it.

The main functions of the software are:

- **Data Entry and Editing**
- **Data Management**
- **Data Manipulation**
  - -Mathematical Analysis
  - -Geometric Analysis
  - -Network Analysis
  - -Statistical Calculations
  - -Query
- **Data Display**
- **Data Output**



GISs can be used for scientific investigations, resource management and development planning. For example, a GIS might allow emergency planners to easily calculate emergency response times in the event of a natural disaster, or a GIS might be used to find wetlands that need protection from pollution. GIS is particularly useful when several physical factors must be considered and integrated over a large area.

Typical uses are:

- **Site selection.** Locating a potential site for a new water well within one kilometer of a water company service area.
- **Transportation planning and waste disposal site location**.
- **Emergency response planning**. Road network information can be combined with earth science information to analyze the effect of an earthquake on the response time of fire and rescue squads.
- **Environmental impact analysis**. An oil-spill simulation model can be applied to a database of ecological information to determine the probable effects, and, to support decision makeing.
- **Biodiversity Management**. Measurements and observations on species, vegetation communities, specific wildlife combined with geology, protected areas, soil types, slope classes, roads, etc. can be used to produce maps to assist decision making for habitat protection.
- **Other uses**. Municipal applications - infrastructure management, bus route planning, forest resource management, etc...

Nature of Geographic Data

Geographic Data is characterized by:

- Position
- Attributes
- Spatial relationships
- Time

Canada Centre for Remote Sensing / Centre canadien de télédétection

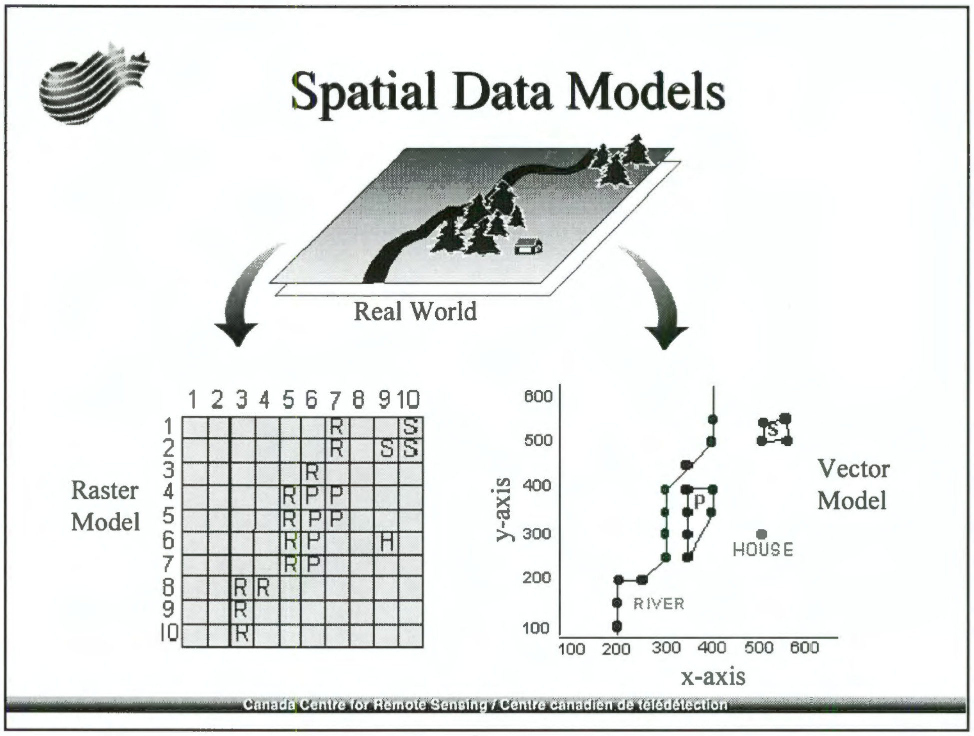
More simply, the four components provide answers to the following questions: 1) where is it?, 2) what is it?, 3) what is its relationship to other spatial features?, 4) and when did the condition or feature exist?

Locations are recorded in terms of a coordinate system like the Latitude/Longitude or a cartesian system (in UTM coordinates). A GIS requires that a common coordinate sysem be used for all the datasets that will be used together. To some extent, the locational data are always imprecise at some level of detail.

The attributes are the non-spatial components of geographic features. For example, the attributes of a feature representing a forest stand might include the species composition, average tree height, the crown closure, and the date it was last logged.

Spatial relationships are generally very numerous, may be complex, and are important. It is not possible to store information about all the possible spatial relationships. Instead, only some are explicitly defined in the GIS, and the remainder is either calculated as needed or is not available.

The representation of time is an added level of complexity that is difficult to handle. As a result, the time factor has generally not been addressed in a sophisticated manner.



There are two fundamental approaches to the repesentation of the spatial component of geographic information: the vector model and the raster model. In the above representation, pine forest stand (P) and spruce forest stand (S) are area features. The river (R) and house (H) are line and point features, respectively.

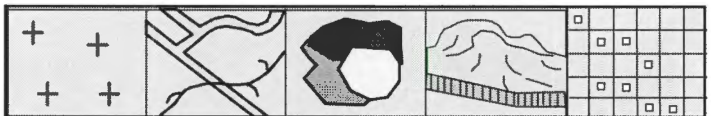
In the vector model, objects in the real world are represented by the points and lines that define their boundaries. The position of each object is defined by its placement in a map space which is defined by a coordinate reference system. A point may represent a meteorological station, a line may represent a road and an area (a polygon) may represent a forest stand.

In the raster model, the space is regularly subdivided into cells, usually square in shape. The location of geographic objects is defined by the row and column position of the cells they occupy. The area of a cell defines the spatial resolution. The position is only recorded to the nearest cell. The value stored for each cell indicates the type of object that is found at that location.

The quadtree data model is a variation on the raster representation by using a variable-sized grid cell. Finer subdivisions are used in those areas with finer detail.

Raster data files can be manipulated quickly by the computer, however, they are often less detailed and may be less visually appealing than vector data files, which can approximate the appearance of more traditional hand-drafted maps.

Types of Spatial Data



POINTS	LINES	AREAS (POLYGONS)	SURFACE	RASTER/MATRIX
FORMAT • single coordinate pair • no length or area	FORMAT • string of coordinates with beginning and endpoints • has length but no area	FORMAT • string of coordinates with same beginning and endpoints • has length and area	FORMAT • area with vertical coordinates • has area, length and elevation	FORMAT • matrix of digital values • has area
EXAMPLES • rainfall • elevation • labels for areas • well sites	EXAMPLES • roads • streams • utility lines • fault lines • boundary of area	EXAMPLES • socio - economic zones • roads • land use • soils boundaries • eco districts	EXAMPLES • slope map • terrain surface	EXAMPLES • DEM • remote sensing data

Canada Centre for Remote Sensing / Centre canadien de télédétection

Some examples of existing digital databases which supply spatial data:

National Topographic Data Base (NTDB)
Canadian Road Network (CRN)
Digital Terrain Elevation Data (DTED)
Canadian Geographical Names Data Base (CGNDB)

[http://www.NRCan.gc.ca/ess/]

Current developments in spatial data access aim at implementing geospatial data infrastructures. These infrastructures provide access to the digital databases through concepts such as "clearinghouses" and exploit the information highway.

The Canadian initiative is called the Canadian Geospatial Data Infrastructure. [http://cgdi.gc.ca/]

Attribute Data

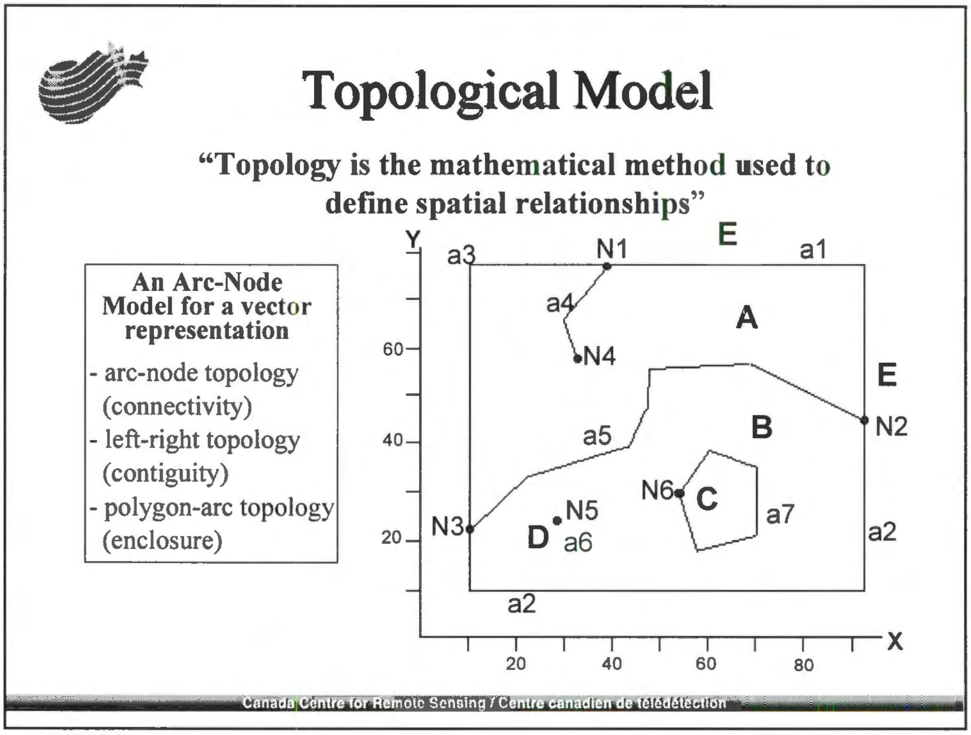
Real World Object	Spatial Entity	Example of Attributes
Tree	Point	Species = PINE Diameter = 18 cm Status = Disease free
Road	Line	Type = two lanes Max_speed = 60 km/hr Surface = paved
Lake	Area	Name = Jasper Lake Boat_type = 25 hp Max_depth = 20 m

Canada Centre for Remote Sensing / Centre canadien de télédétection

Attribute data is the non-spatial component of geographic data. It represents the information which describes the geographic object. It can take different formats: text, reports, numbers, bearings, shades, symbols.

For example, the attributes for a bird colony, aside from its locational attribute (longitude and latitude) could be numbers representing the population, the percentage of adults, the year it was surveyed, or, text identifying the name of the colony, the nearest water body, the name of the person who did the survey, the generic name of the species occupying the colony, etc.

In a GIS, attribute data can be stored with the spatial data to which it is related or it can be stored in an external database system to which the spatial data can be linked through a common identifier.



In the Arc-Node model, the ***arc*** represents a series of points that start and end at a node. A ***node*** refers to an inersection point where two or more arcs meet. A ***polygon*** is comprised of a closed chain of arcs that form the boundary of the area. In the diagram: N1 to N6 are nodes; a1 to a7 are arcs; A to E are polygons.

A Node Topology Table stores each node and the arcs to which it belongs.

E.g.: Node N3 is related to arcs a2, a3,and a5.

An Arc Topology Table defines the relationship of the nodes and polygons to the arcs (the start and end nodes, the left and right polygons).

E.g.: Arc a5 starts with node N3 and ends with node N2, has polygon A on its left and polygon B on its right.

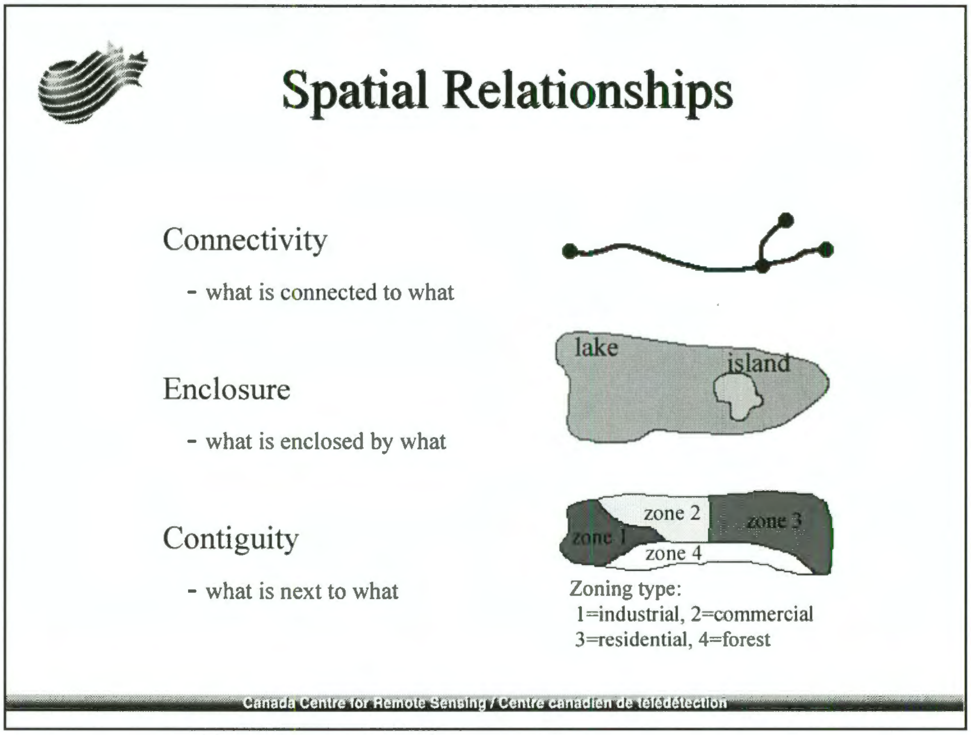
A Polygon Topology table defines the arcs that make up the boundaries of each polygon.

E.g.: Polygon B is defined by arcs a2, a5, 0, a6, 0, a7. In this example polygon B has 2 islands (polygon C and D).

In this model, a ***point*** is a polygon with no area. The boundary of polygon D is defined by the arc a6; arc a6 starts and ends with the same node N5.

An arc coordinate table is maintained where each arc is represented by one or more straight-line segments defined by a series of coordinates.

Arc	Start X,Y	Intermediate X,Y	End X,Y
a2	94, 4	94, 44; 10, 10	10, 23
a6	28, 24		28, 24



The relationship among linear features such as pathways and networks is maintained in a GIS database. This relationship is referred to as connectivity and is derived from the Node Topology table in the Arc-Node model. Connectivity relationships enables a user to perform numerous functions commonly referred to as network analysis. Examples of such analyses include finding the optimum routings through a network, such as the most efficient delivery route or the fastest travel route. Network analysis can also be used to optimize transportation scheduling or to predict loading at critical points in a network (e.g.: water flow at a bridge crossing from heavy storm runoff).

The contiguity or adjacency of polygon features is also explicitly recorded in a GIS database. The relationship among area features provide answers to what polygons are adjacent to each other. This information provides left-right topology and is derived from the Arc Topology table in the Arc-Node model. Contiguity analysis is applied to a wide range of applications. For example, biologists can determine which habitats occur next to each other; city planners can resolve resolve zoning conflicts such as industrial zones bordering recreational areas.

Enclosure or containment refers to area definition. This is a more complex relationship which requires set operations to determine union/intersection of two or more topologically structured layers. This type of relationship gives rise to the overlay analysis functions.

Module 2

The GIS Database

Module 2

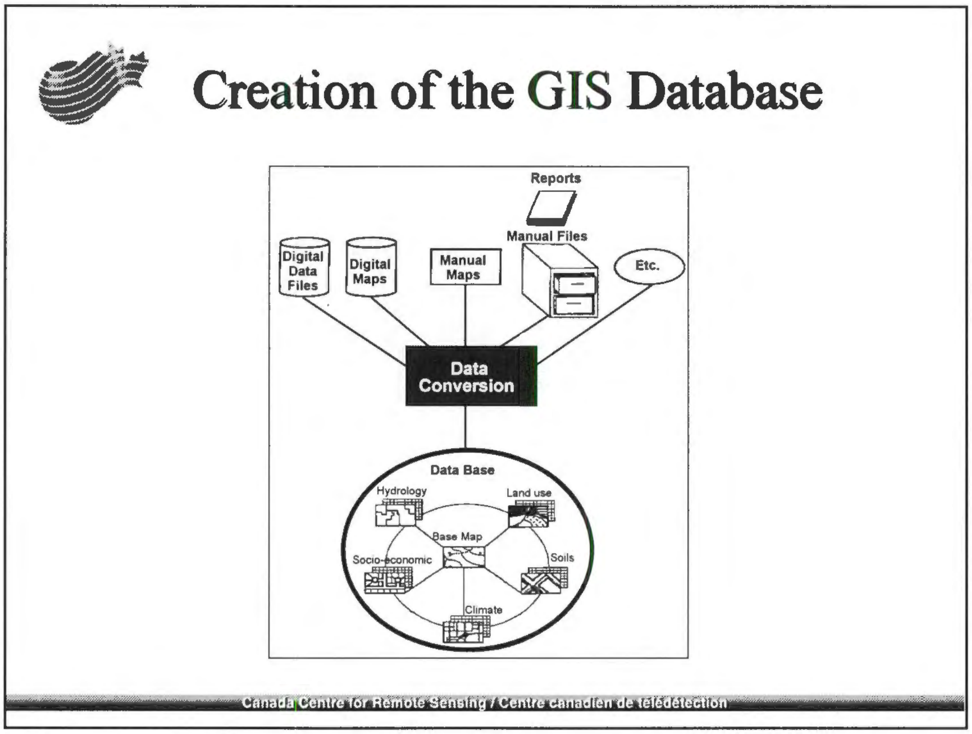
Objective:

• To become familiar with the steps required to build a GIS database

Contents:

• Data sources
• Data integration
• Data structuring

Canada Centre for Remote Sensing / Centre canadien de télédétection



Data sets are not usually available in a format suitable for immediate or direct input into a GIS database. This factor makes constructing the database the most time consuming and expensive aspect of GIS operation. Each data type presents unique problems of data capture.

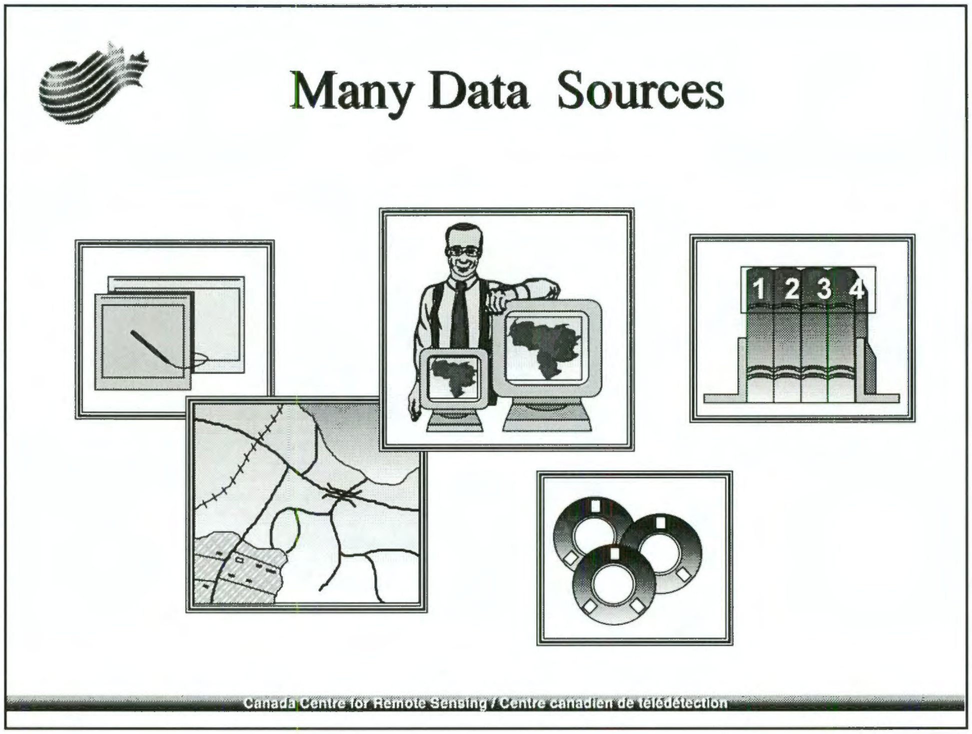
Image Data: Often image data must be digitized (air photos) or pre-proceesed through image analysis systems (satellite imagery) prior to GIS input.

Paper Maps: They must be digitized and analyzed to ensure accuracy or to extract the desired parameters (eg. DTM generation).

Digital Data Sets: They are rarely available in a format compatible with GIS input formats. Conversion software is frequently required (ensure availability when planning).

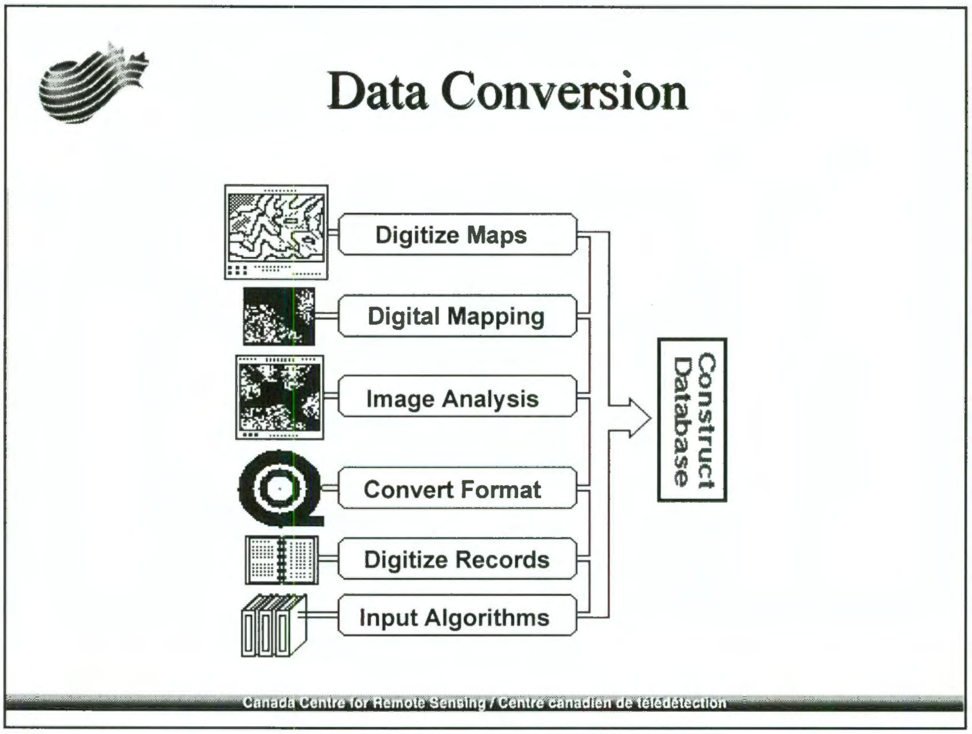
Analog Records: Field observations or analog data recordings must be structured, digitized and formatted for compatibility with GIS databases.

Algorithms: Algorithms derived from current research and required for special processing, often require user installation within a generic GIS software package.



When building a GIS database, it is essential that the problem being addressed be thoroughly understood. Data needs must be analyzed and assessed prior to any input. Failure to do so can cause numerous problems and drive the user further away from addressing the issue at hand. Estimating resources and time limits are **key elements** to the successful use of a GIS for problem-solving issues. Data capture is a time-consuming component. Identities of the objects on the map must be specified, as well as their spatial relationships. Editing of information can also be difficult. Electronic scanners record blemishes on a map just as faithfully as they record the map features. For example, a fleck of dirt might connect two lines that should not be connected.

Geographic data are organized within a GIS so as to optimize the convenience and efficiency with which they can be used. The form of organization chosen will be influenced by the types of data to be used, the types of analyses to be performed, and the methods used to encode the data. On a paper map, geographic information is usually organized as a set of themes. In a GIS, themes are treated as GIS data layers. A data layer consists of a set of logically related geographic features and their attributes. The features to be grouped in a single data layer are chosen for the convenience of the users. The principle may be to group similar feature types. For example, roads and railways might be combined as a single transportation data layer and streams and lakes as a hydrology data layer. The organization of the data layers will also depend on the restrictions imposed by the GIS software used. It may be necessary or more convenient to store point, line, and area features in separate data layers.



The main objective of data conversion is to be able to build a database where each spatial component can be processed digitally and be related to each other to permit analysis.

Practically, whether the data is converted into a raster or vector structure depends mainly on the software that is being used. There are advantages and disadvantages to each of these structures:

Advantages

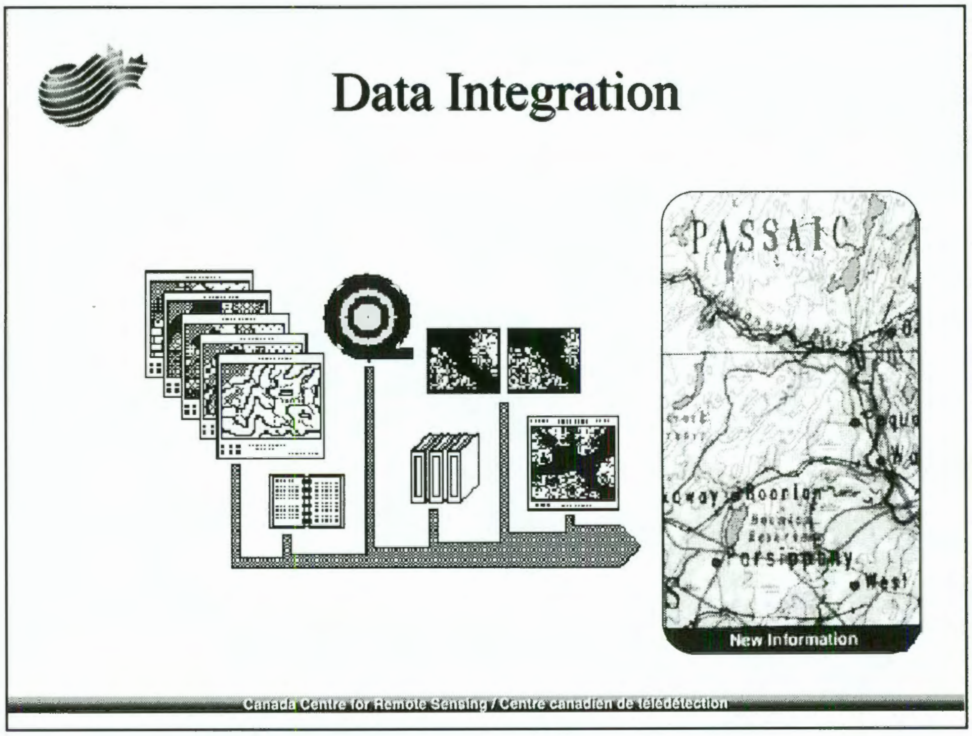
Vector: is a more a compact data structure; has the capability for high resolution and precision; allows easy map manipulation and updating. The size of a vector database is directly related to data density.

Raster: is a simple data structure; allows easy map combination and spatial analysis; allows efficient data capture in large volumes. Spatial relationships are inherent in the data structure.

Disadvantages

Vector: has complex data structures; spatial relationships are not inherent in the structure; requires complex map-overlay functions; has low tolerance for topological errors.

Raster: requires display in a non-traditional way; map precision is limited by cell size of input data; results in large volumes of data and inefficiencies in data storage since variations in data density and homogeneity are not recognized.



Data Integration is the process of making different data sets compatible with each other, so that they can be displayed reasonably on the same map and that their relationships can be analysed. The more ambitious the application and the more data sets that are needed, the more likely it is that data integration will be a problem.

The GIS database should have the following characteristics:

- All themes (layers) shoud be converted to the same reference system.
- Features of the same theme originating from different sources should match at edges.
- There should be consistency between each component; attribute values should be valid across themes. For example, a database may contain a layer on land use, showing in a given location, a corn field. In the same database, for the same location, a layer on production should show crop yield and not fishery yield.

Data Manipulation

Reprojection
adjusts data to a common base

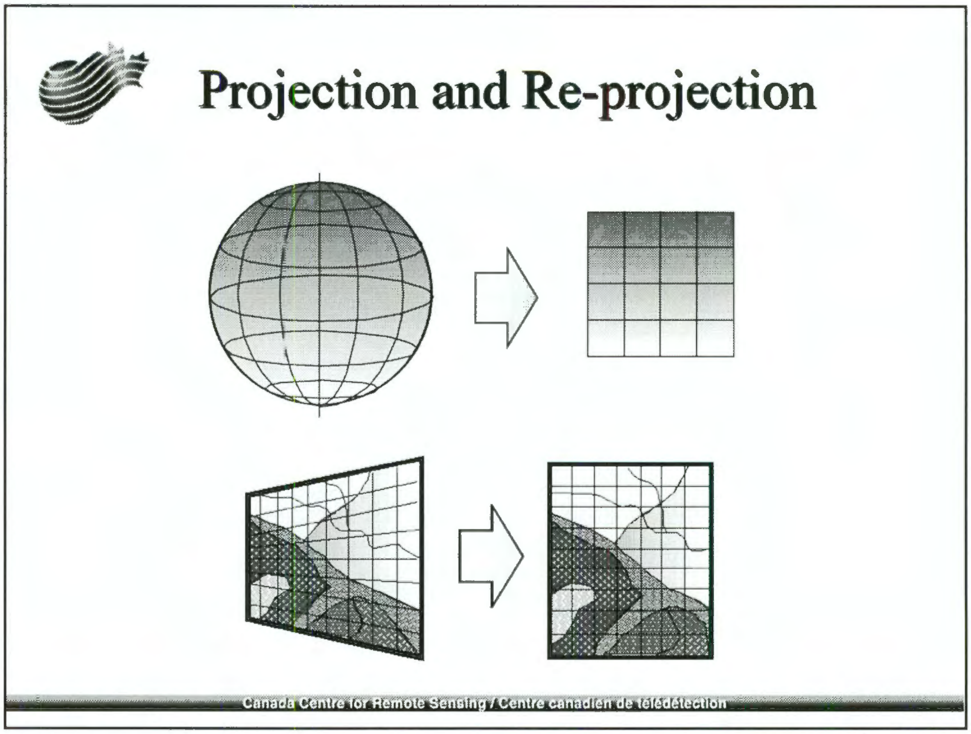
Rubber sheeting
removes distortions

Edge matching
joins features to form a seamless view

Restructuring
derives spatial relationships mathematically

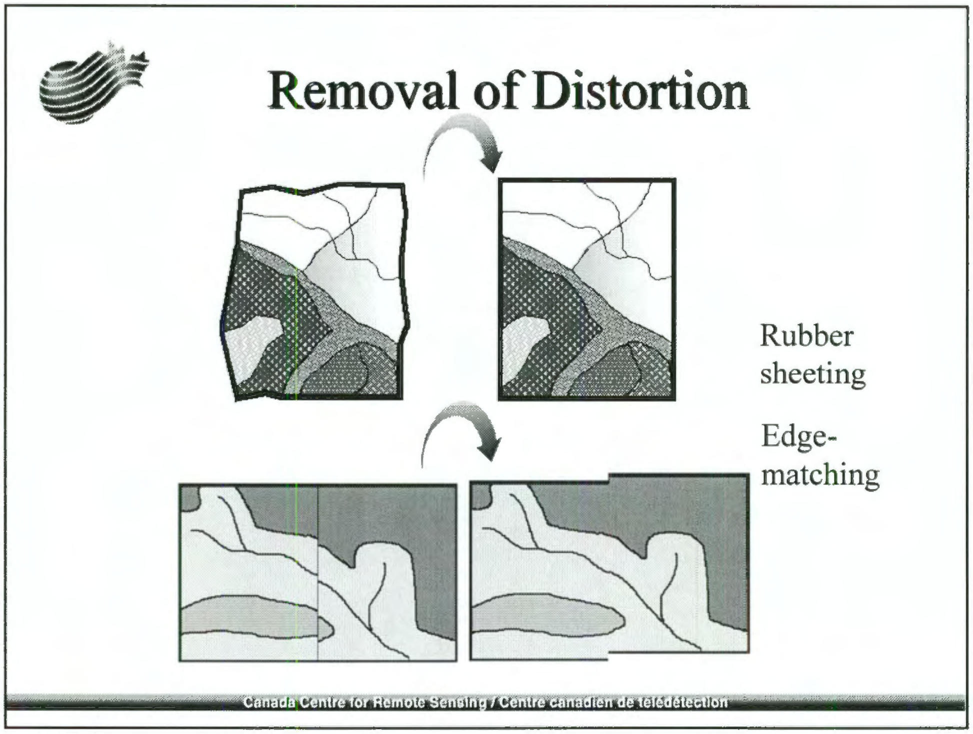
Canada Centre for Remote Sensing / Centre canadien de télédétection

It is often necessary to transform the original spatial data set into a geometry which is more manageable, accurate and consistent with the base data. The base data refer to the features of a map or image upon which a study area is based and to which the coordinates and shape of all other map features must match. Available data sets are rarely consistent in scale, projection, spatial accuracy, orientation and coverage, and thus the manipulation of data is required.



Projection is a fundamental component of mapmaking. A projection is a mathematical representation of the Earth's three-dimensional, curved surface onto a two-dimensional medium - paper or a computer screen. Different projections are used for different types of maps because each projection is specifically suited for certain uses. For example, a projection that accurately represents the shapes of the continents will distort their relative sizes. In GIS, the type of analysis which will be performed, guides the selection of the projection system. A projection which preserves distances will be prefered for applications where measuring of distances is important (e.g.: bus routing - shortest path), while projection which preserves areas will be suited for those where areas are measured (e.g.: determining crop yield per hectare).

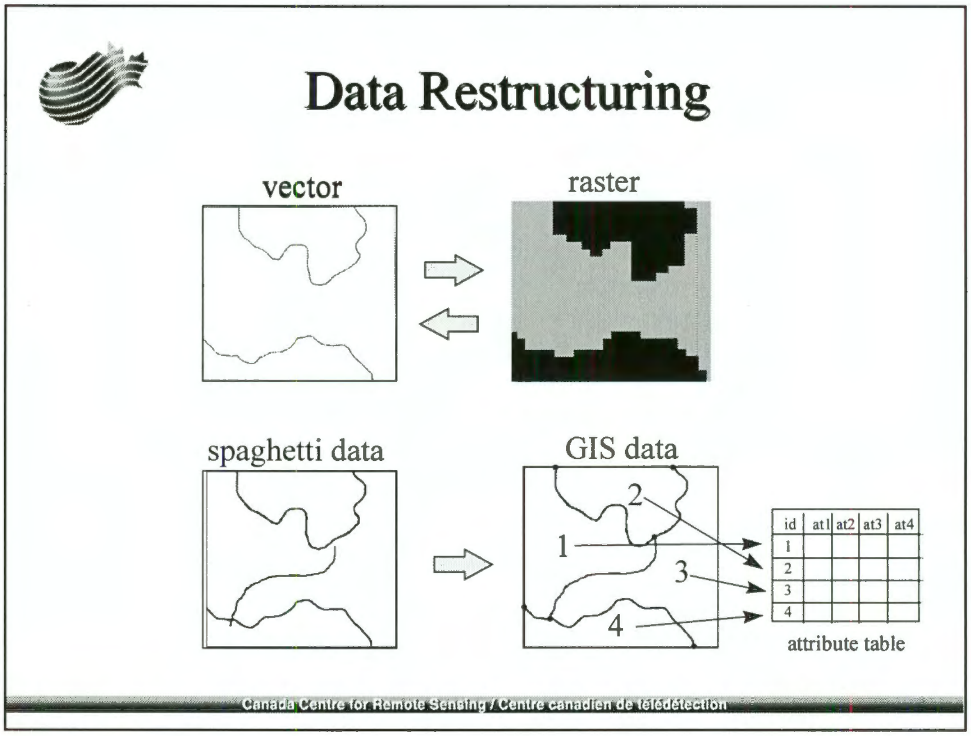
Since much of the information in a GIS comes from existing maps, a GIS uses the processing power of the computer to transform digital information gathered from sources with different projections to a common projection. A GIS commonly supports several projections and has software to transform data from one projection to another.



Geometric transformations are used to assign ground coordinates to a data layer within a GIS or to adjust one data layer so it can be correctly overlayed with another layer covering the same area. This procedure is termed **registration** where different data layers are registered to a common coordinate system or to one data layer that is used as a standard.

Rubber sheeting is a procedure identifying points (master) on a layer (slave) for which the positions are known. Then the GIS calculates a mathematical function to transform the coordinates of the layer to be corrected to fit the coordinate system of the master points.

Edge-matching is the process used to reconcile the position of features that extend onto an adjacent map but are not correctly aligned at the map boundary. The discrepancies may be difficult to correct because the features along the map edge may not be shifted to the same degree or in the same direction. Some GISs provide software to reconcile these differences by making adjustments to the position of features on one or both maps. Edge-matching must be done for the geographic information from several adjacent maps to be represented as a single continuous data layer.

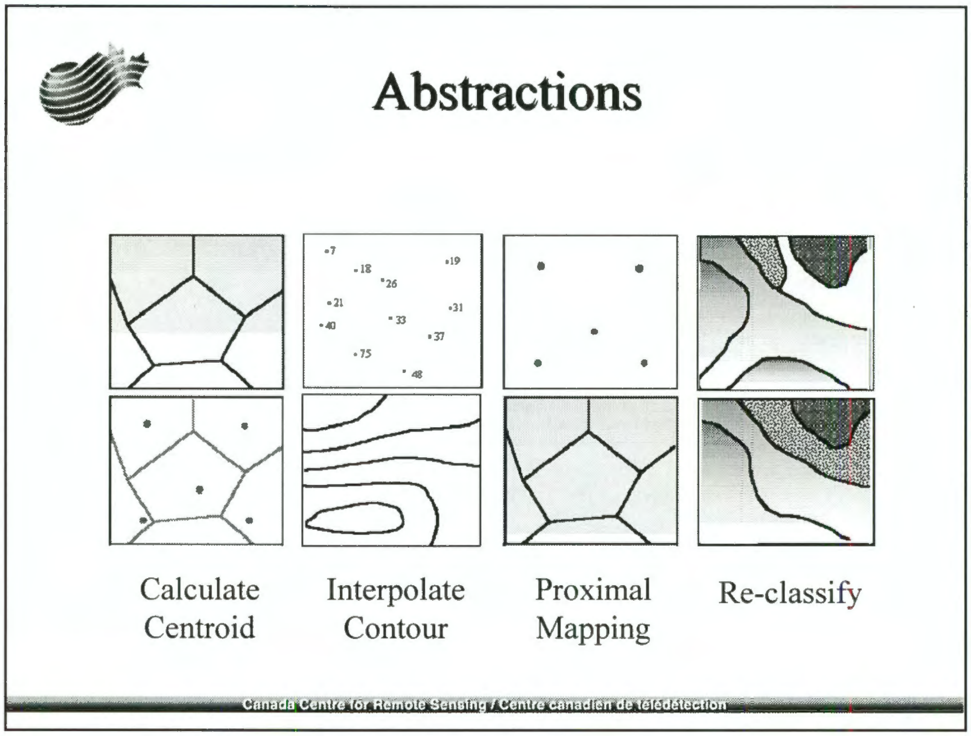


An essential step in data integration is the restructuring of spatial data into a form that is compatible with the software.

Data restructuring can be performed by a GIS to convert data into different formats. For example, a GIS may be used to convert a satellite image map to a vector structure by generating lines around all cells with the same classification, and, determining the spatial relationships, such as adjacency or inclusion.

Data restructuring also involves the transformation of digital data into a topological format. For example, vector data require cleaning to remove dangling arcs and to close polygons; when this process is complete the intersections and topological tables can be created.

Data restructuring also creates the links between the spatial component and the attribute data through database management functions.



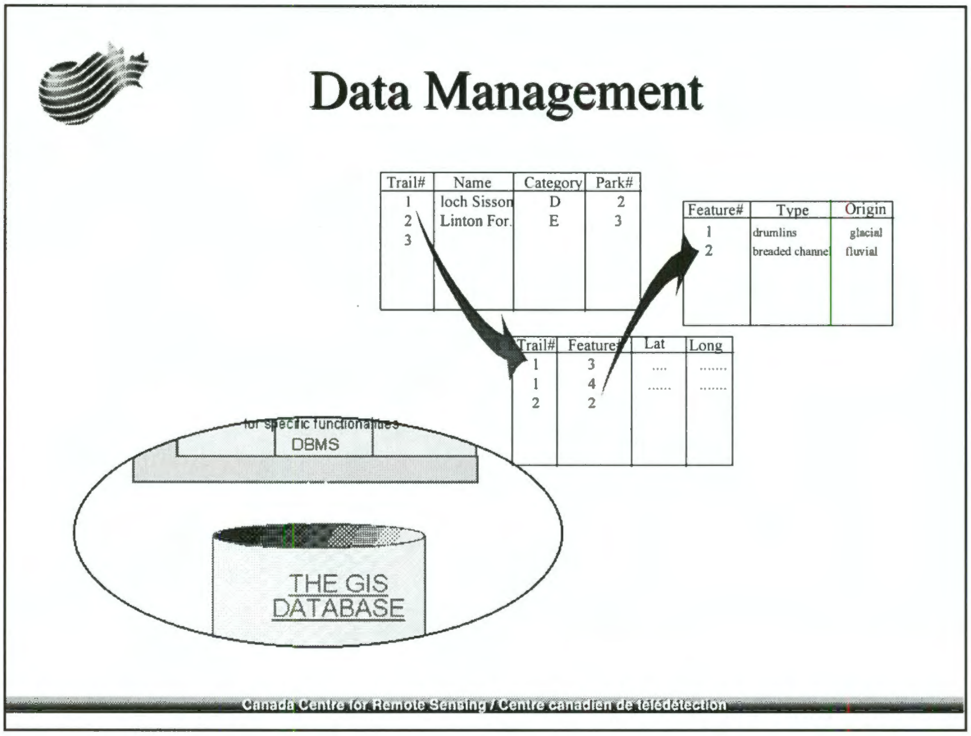
The interpretation and integration of data sets can often be simplified and made more efficient through database abstractions. Common abstractions include: contouring, proximal mapping, centroid calculation, and reclassification or generalization.

Contouring refers to the generation of isolines from point data, such as rainfall, elevation, and barometric pressure. Contouring is advantageous for visual interpretation of patterns and trends. Contouring requires interpolation. Interpolation methods are usually part of GIS functions. Generation of a digital elevation model from point elevation values is a common operation.

Proximal mapping is used to generate polygons which form proximal neighbourhoods around centroids.

Centroid calculation is used to describe polygons as point locations as in the creation of distance weighted models.

Reclassification or Generalization is used to regroup existing attribute data to produce a more generalized map or one which displays the combined influence of two or more interdependent spatial distributions.



Many GISs are implemented over a traditional DBMS (Data Base Management System). Almost all GISs use a relational database management system (RDBMS) such as ORACLE, Dbase IV, etc. In a RDBMS, tables are used as the basic data model.

Some GIS use a one-database approach where a single database is used to store both the spatial and non-spatial data. In other GIS, a dual approach is used where one database is used for spatial data and the other for non-spatial data. Most GIS software can also create links with external databases which carry attribute information. This is very useful when these data are maintained by other parties.

The use of a DBMS has its advantages: reduced data redundancy; enhanced data sharing, data security and data recovery; data integrity and consistency.

Also the use of DBMSs enforces standards: it is easier to enforce standards on data representation and description, thus helping data interchange among users.

A DBMS can also provide a uniform interface and efficient manipulation procedures as a package. Data query, retrieval and update can be executed in a standard way. Most GISs offer database access using SQL (Standard Query Language). The GIS software usually hides the complexity of DBMS through packaged procedures (data management module) which facilitates data management to the data manager and data access for the decision-maker.

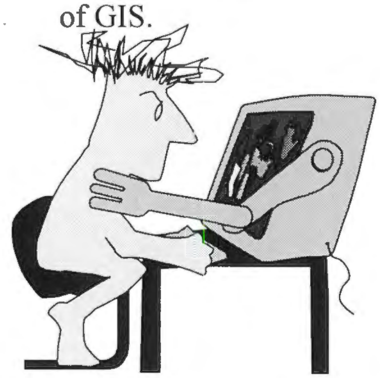
Module 3

GIS Analysis

Module 3

Objective:

- To become familiar with the spatial analysis functions of GIS.



Contents:

- GIS functions
- Query
- Classification
- Overlay
- Proximity/Connectivity
- Advanced analysis
- Output

GIS Functions

- Data Entry and Editing
- Data Management
- Data Manipulation and Analysis
  - Mathematical Analysis
  - Geometric Analysis
  - Network Analysis
  - Statistical Calculations
  - Query
- Data Visualization and Output

Canada Centre for Remote Sensing / Centre canadien de télédétection

The data entry and editing component converts data from their existing form into one that can be used by the GIS.

The data management component includes those functions needed to store and retrieve data from the data base.

The data manipulation and analysis functions determine the information that can be generated by the GIS.

The data visualization functions include screen display as well as hardcopy output like maps.

Query

Data retrieval functions extracts existing data that satisfy certain conditions.

- Queries can be simple or complex.
- The conditions can be spatial, non-spatial, or mixed.

Canada Centre for Remote Sensing / Centre canadien de télédétection

A spatial condition:

Find all houses within a user-defined region.

A non-spatial condition:

Find all countries with a population larger than 20 million.

A mixed condition:

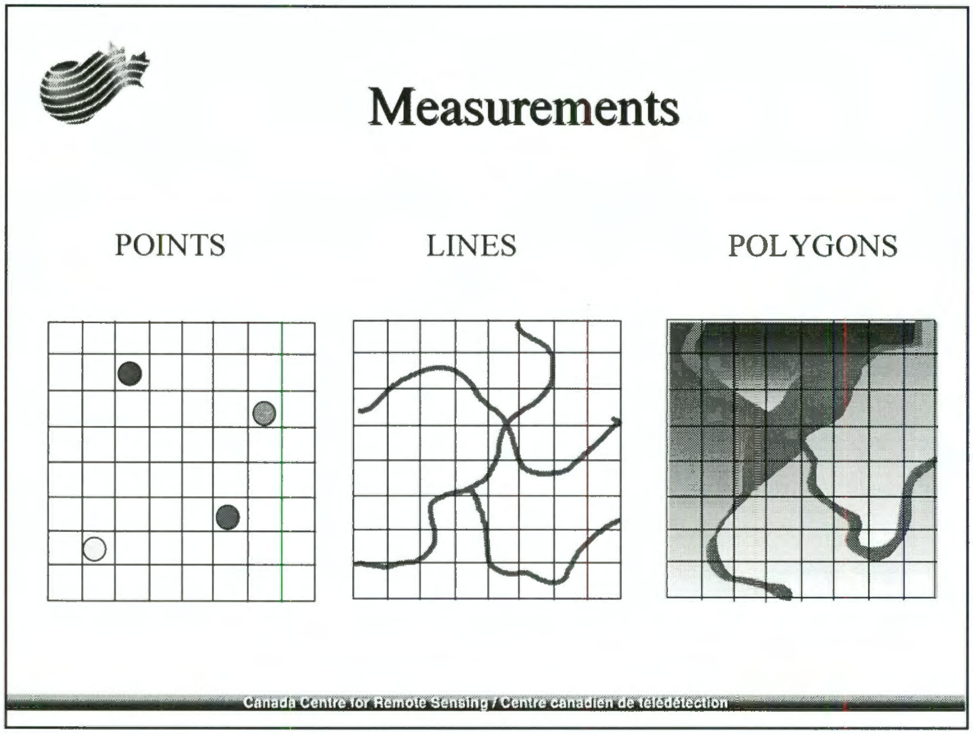
Find all properties on sale and facing Blue Lake.

Retrieval is conceptually a simple operation. It involves scanning all data record one by one and determining if any of them match the given condition. The difficulty occurs when the data volume is large. It is a challenge to develop a retrieval mechanism to find the required record without going through all the data.

When the query becomes complex, spatial operations need to be performed.

An example of a complex query:

Identify the shortest route from the school to the hospital.



The most basic function of GIS is measurement. Measurement allows you to calculate the number, length, area and volume of spatial elements. This function usually includes the capability to list and summarize results of calculations in some meaningful manner.

Number: How many waste sites are in a given watershed?

Length: What is the distance between Montréal and Quebec following highway 132?

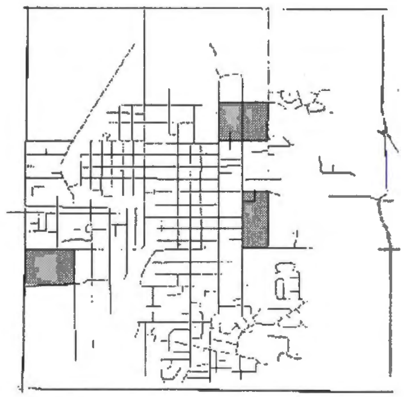
Area: What percentage of the LaVerendry Park is covered by water?

Volume: What is the fluctuation volume of water in a lake knowing the variability in the lake water level and contour?

Simple Query - Example

"Locate all forest areas larger than 100 hectares that are suitable for a conservation area."

Result of query ⇨ ■



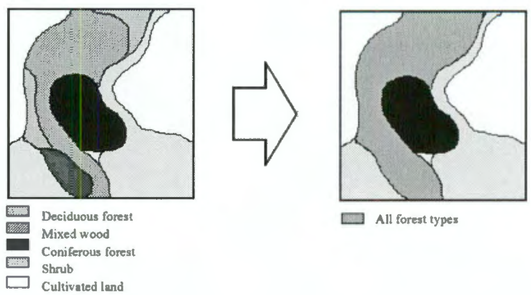
Canada Centre for Remote Sensing / Centre canadien de télédétection

This simple query utilises the organization of the database: features represented as polygons linked to attribute information describing the area - utilised or vacant, forest cover, grassland or rock, presence of wilderness, etc...

Measurement function is applied to calculate the area. The result is displayed using the graphic component of the database and the display characteristics of the system.

Classification

Classification is the process of assigning an entity to one of several pre-defined classes.



Canada Centre for Remote Sensing / Centre canadien de télédétection

Reclassification changes the classification scheme to a different one, normally to one which is more general. Both classification and reclassification require functions. The function could be spatial or non-spatial.

A spatial function could be based on area or length.

(e.g.: classify all islands smaller than one square kilometre as invisible on a map of Canada)

A non-spatial function could be based on population density.

(e.g.: classify all countries with a population density lower than 100 persons/sq.km. as sparsely populated)

Reclassification often requires merging of entities. In a vector system, the merging of polygons involves the removal of arcs. This is not necessary in a raster system. Polygons will be merged automatically after reclassification.

Complex Query - Example

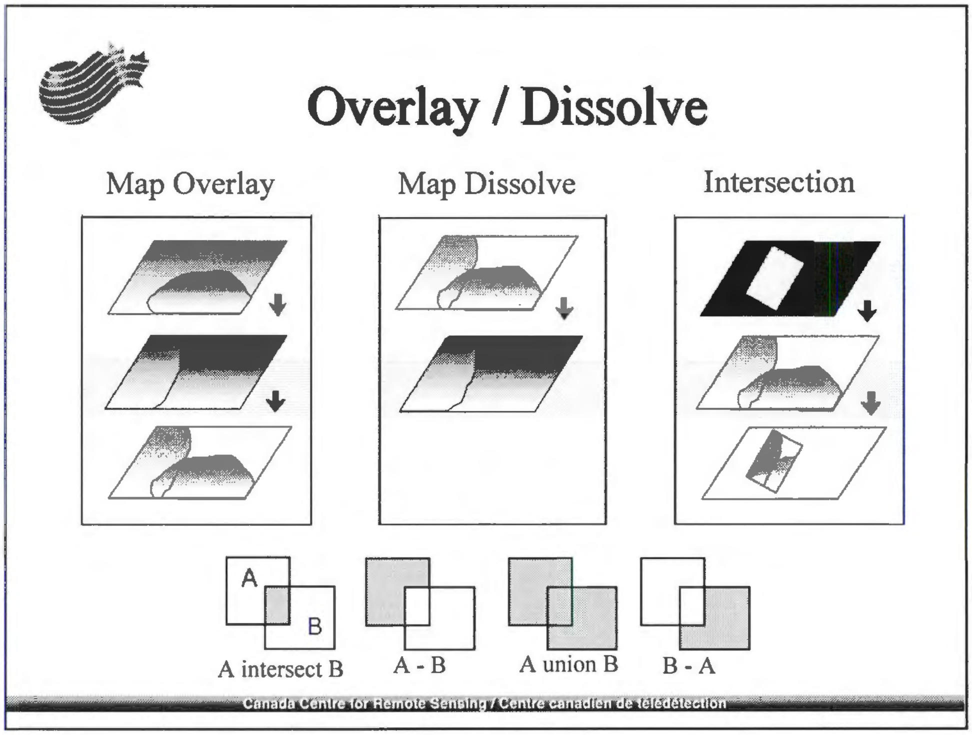
"Locate areas of the city with highest crime rate."

Frequency of break-ins

- high ⇨
- medium ⇨
- low ⇨

Canada Centre for Remote Sensing / Centre canadien de télédétection

This query uses a spatial function - classification of areas on a range of values for the number of break-ins.

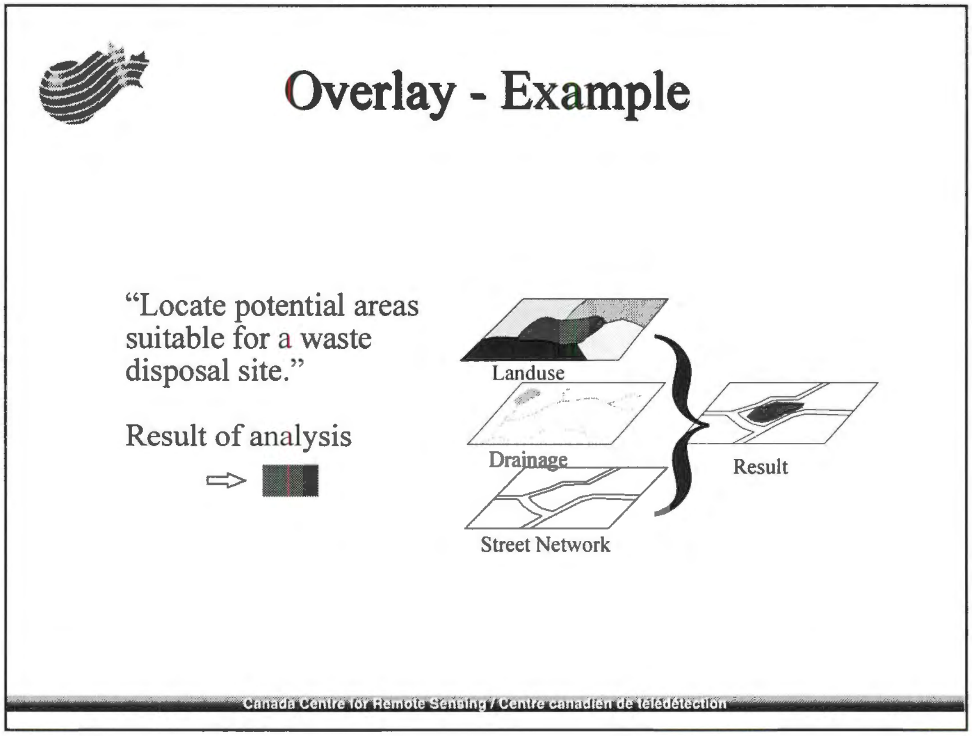


Arithmetic and logical overlay operations are part of all GIS software packages. These operations allow the combination, the intersection and the containment of data layers. Overlay operations create new spatial data layers. Features on each layer, when overlayed, create new output features with attributes combined from the original features. The GIS overlay functions are used to answer questions such as:

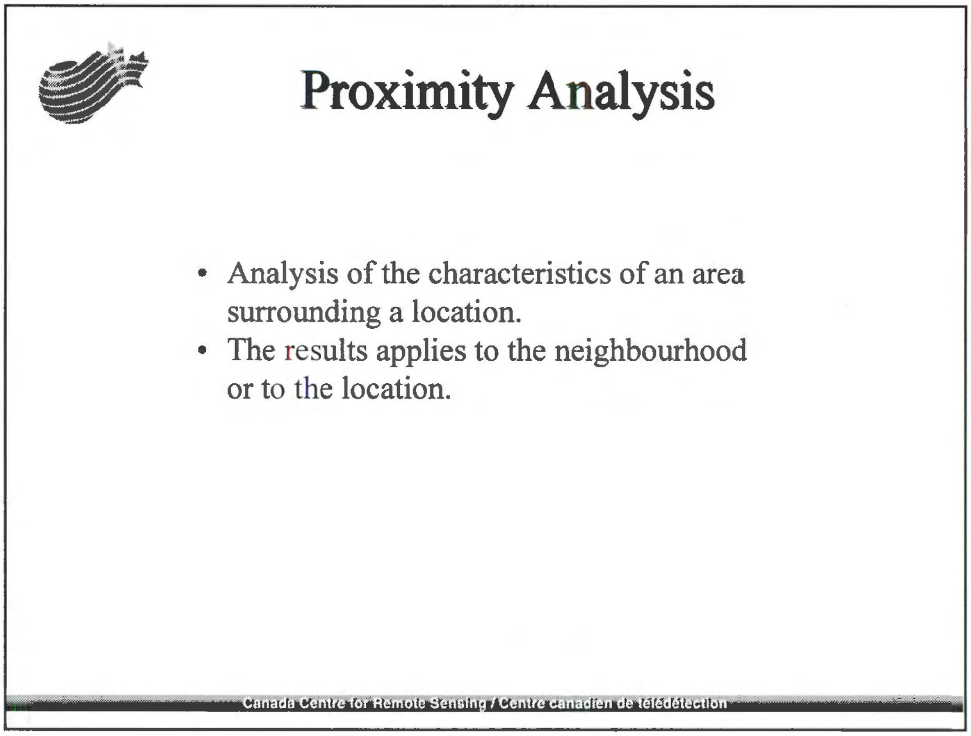
- Given coverages on land-use and soils, identify the regions where agriculture is carried on clay soils.
- Given coverages on land-use and population statistics, identify the urban areas where French is the native language in greater than 85% of the population.

Dissolve operations and classification are often used in combination with overlay operations. The dissolve function permits the deletion of boundaries between adjacent polygons having the same attribute values which meet given criteria (equal values, or within a certain tolerance). This creates a generalized map layer. Usually only the attributes for the items used to perform the dissolve operation are retained in the output map layer.

Note: coverage = map layer



Themes from the database that describe the use of the land, the drainage characteritics and the accessiblity by road are overlaid. Criteria for the selection of the area are applied (classification) on the result of the overlay operation.



Proximity analysis requires

- a location
- characteristics of the neighbourhood, and
- functions to be performed on entities within the neighbourhood.

Eg.: Find the average size of land parcels within 10 km of City Hall.

Examples of neighbourhood characteristics:

- Shape (rectangular, circular, arbitrary polygon)
- Position (centered on location, lower left corner touching location, etc.).

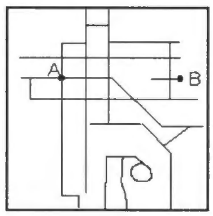
Examples of neighbourhood functions:

- Highest value
- Lowest value
- Averge value
- Sum of values.

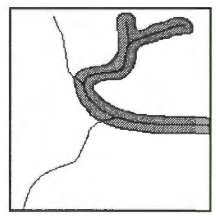
Connectivity Analysis

Analysis of the characteristics of the spatial linkage among locations.
Distance measurement is the simplest form of connectivity analysis

Network analysis



Buffer zone generation



Canada Centre for Remote Sensing / Centre canadien de télédétection

Connectivity analysis requires:

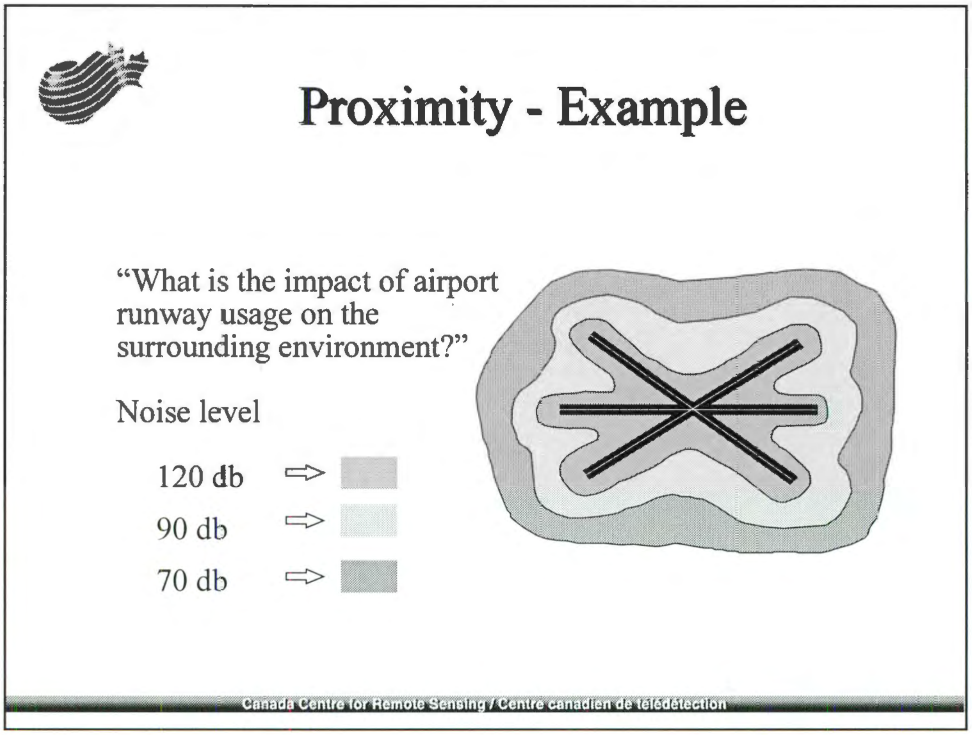
- a specification of linkage
- functions or rules for movement along the linkage, and
- a unit of measurement such as length or travel time.

Topological structures of networks are specifications of a linkage. Examples of movement along the specified linkage can be unrestricted, or, have speed decreasing with distance. In distance measurements, the linkage between the two points is a straight line and the movement along the straight line is unrestricted. The unit of measurement is length.

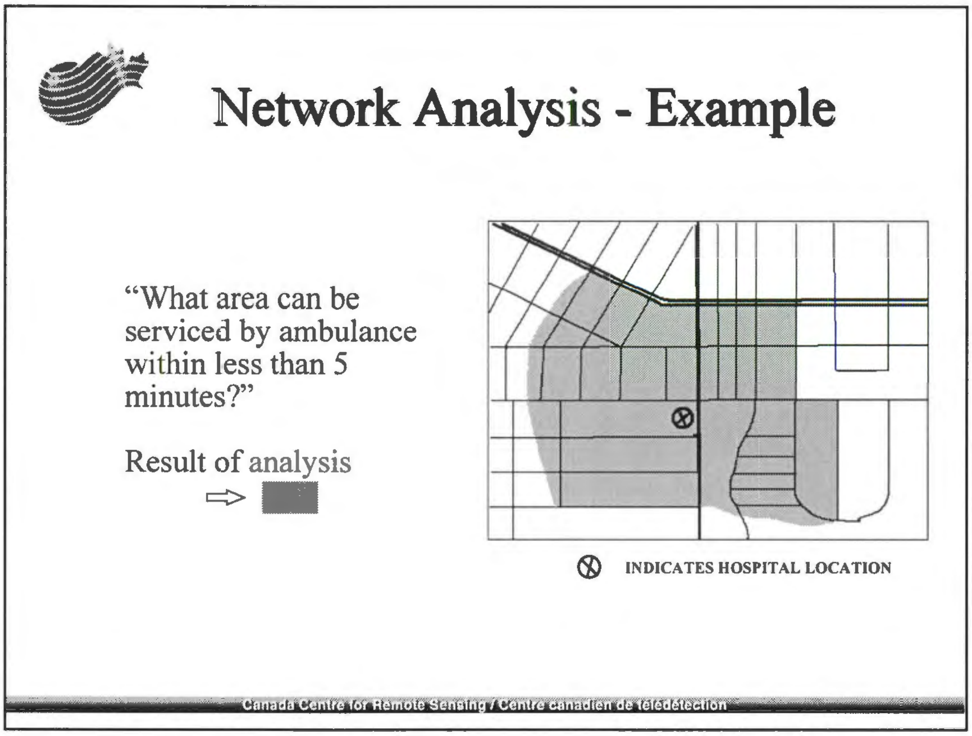
Buffer zone generation is a generalization of distance measurement. The distance measurement is from a line instead of from a point. The linkage could be orthogonal to the line. The movement could be unrestricted. The unit of measurement is length.

Example of buffer zone applications:

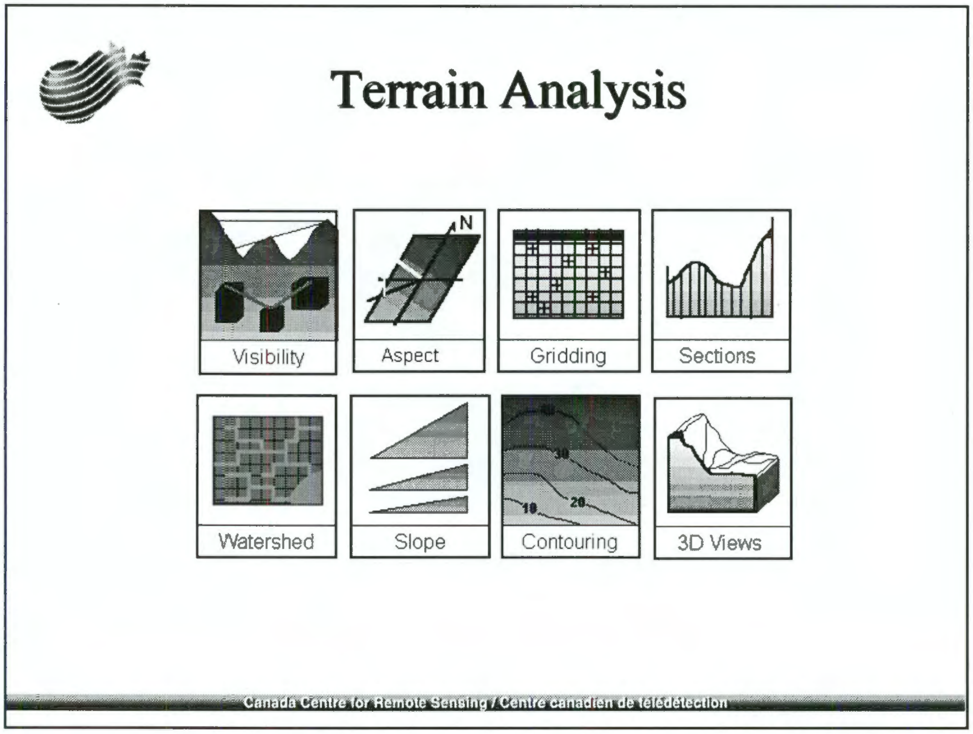
- Restrict forest cuts to areas at least 200 m from streams where the slope is more than 3%.



Using the spreading characteristics of sound waves and the geo-referenced information of airport features, a buffering function is applied. The result of this type of analysis can be used to determine the impact of new development on the environment.



Taking into consideration - driving speed, street directions, and other parameters such as traffic lights, density of traffic, climate conditions, and others, proximity and connectivity functions are applied to calculate shortest route and time of travel to define such things as area service by ambulance etc...



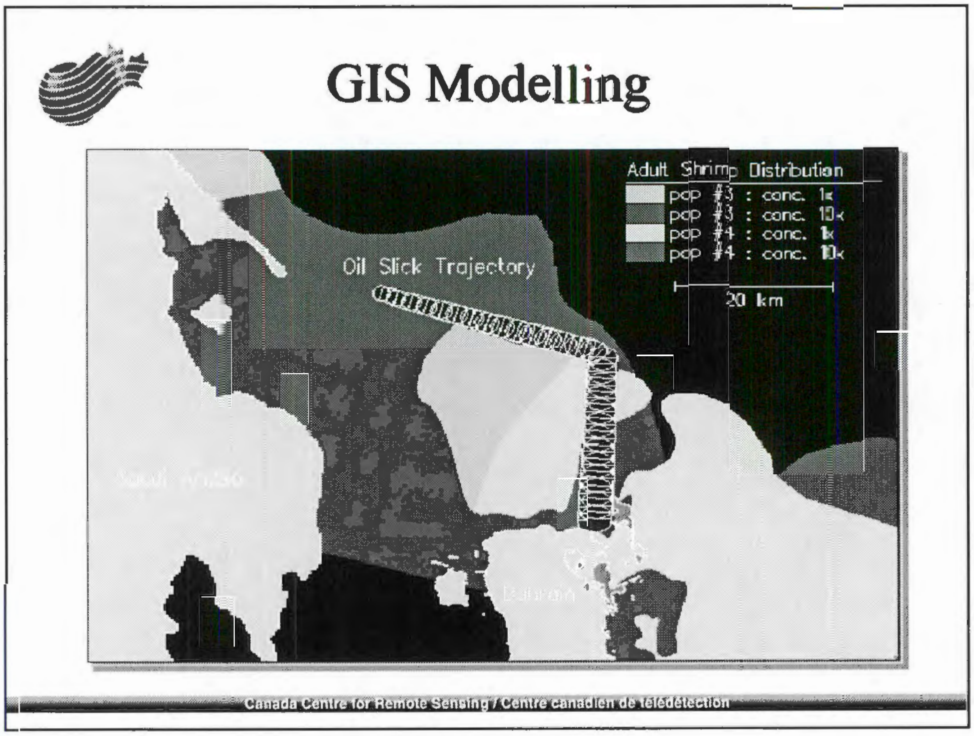
By permitting the display and analysis of complex surface relations, a GIS can facilitate many common terrain analysis requirements.

Visibility - Perspective point visibility assessment is important in applications where "line of sight" is essential (eg.: landscape planning, communications engineering).

Slope/Elevation/Aspect - Digital Elevation Models produced from contours or other source data can provide vital information in a wide range of applications, from geology and engineering to microclimate assessment.

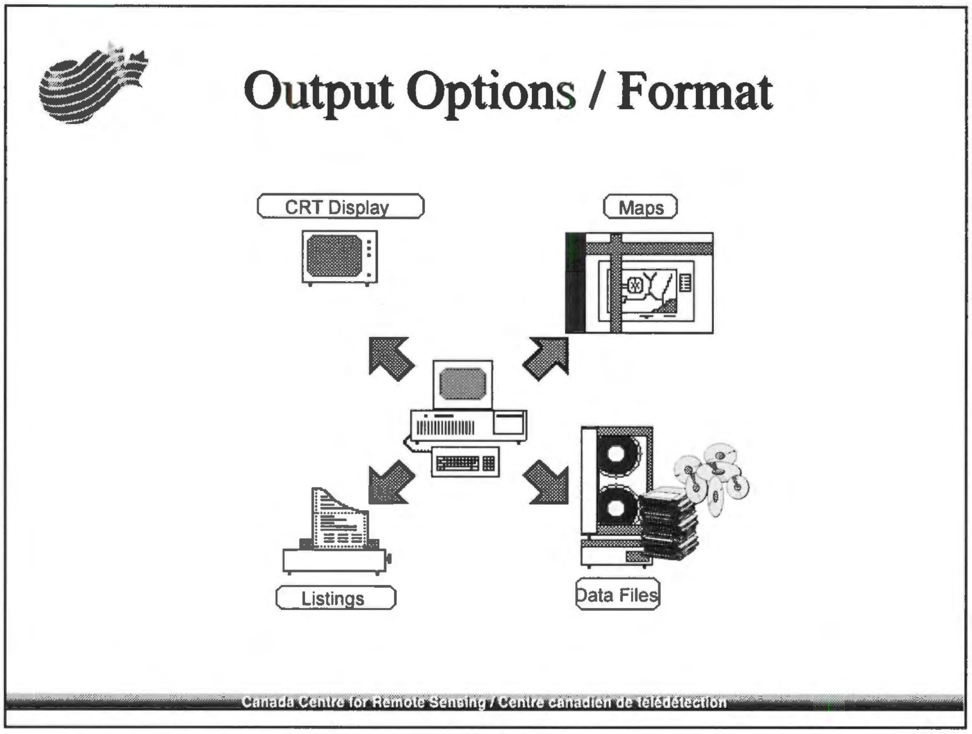
Network Analysis - The ability to analyze network characteristics together with other environmental variables represents a powerful tool in many applications such as watershed management and transportation planning.

Perspective Viewing - The possibility to view the results of GIS operations in 3D enables these relationships to be presented in a perspective which is both striking and easily understood.



Using a combination of various GIS functions, the simulation of real world phenomena can be performed. The models are coded and run using the GIS software. The results of the computation are visualized - either as a static map or as a dynamic presentation of the results as the simulation progresses.

In this example, the simulation of the trajectory of an oil slick is visualized as a function of time. The model integrates factors such as wind, water currents, salinity, temperature and type/characteristics of the oil spill. The effect on the environment can be computed by linking the model to a natural resources database.



Data output

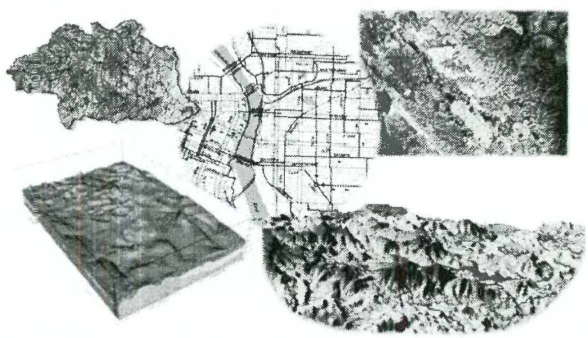
A critical component of a GIS is its ability to produce graphics on the screen or on paper that convey the results of analyses. Wall maps and other graphics can be generated, allowing the viewer to visualize and thereby understand the results of analyses or simulations of potential events.

Data Display - Interface to Information Retrieval

What do you know about the swampy area at the end of your street? With a GIS you can "point" to a location, object, or area on the screen and retrieve recorded information about it from database files.

Using scanned aerial photographs as a visual guide, you can ask a GIS about the geology or hydrology of the area or even about how close a swamp is to the end of a street. This kind of analytic function allows you to draw conclusions about the swamp's environmental sensitivity.

GIS Output - Examples



Canada Centre for Remote Sensing / Centre canadien de télédétection

The result of GIS manipulation is best visualized a in a conventional way - a map. GIS provides new and exciting tools to extend the art and science of cartography. Conventional maps have been used for millennia. GIS adds new capabilities; with the use of terrain models, perspective views can be created and combined with images. Map displays can be integrated with reports, three-dimensional views, photographic images, and other output, such as multimedia.

Data output

A critical component of a GIS is its ability to produce graphics on the screen or on paper that convey the results of analyses. Wall maps and other graphics can be generated, allowing the viewer to visualize and thereby understand the results of analyses or simulations of potential events.

Data Display - Interface to Information Retrieval

What do you know about the swampy area at the end of your street? With a GIS you can "point" to a location, object, or area on the screen and retrieve recorded information about it from database files.

Using scanned aerial photographs as a visual guide, you can ask a GIS about the geology or hydrology of the area or even about how close a swamp is to the end of a street. This kind of analytic function allows you to draw conclusions about the swamp's environmental sensitivity.

Module 4

GIS Applications

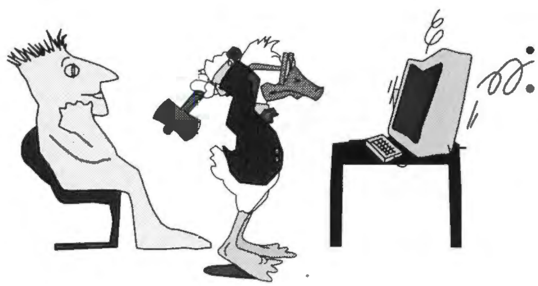
Module 4

Objective:

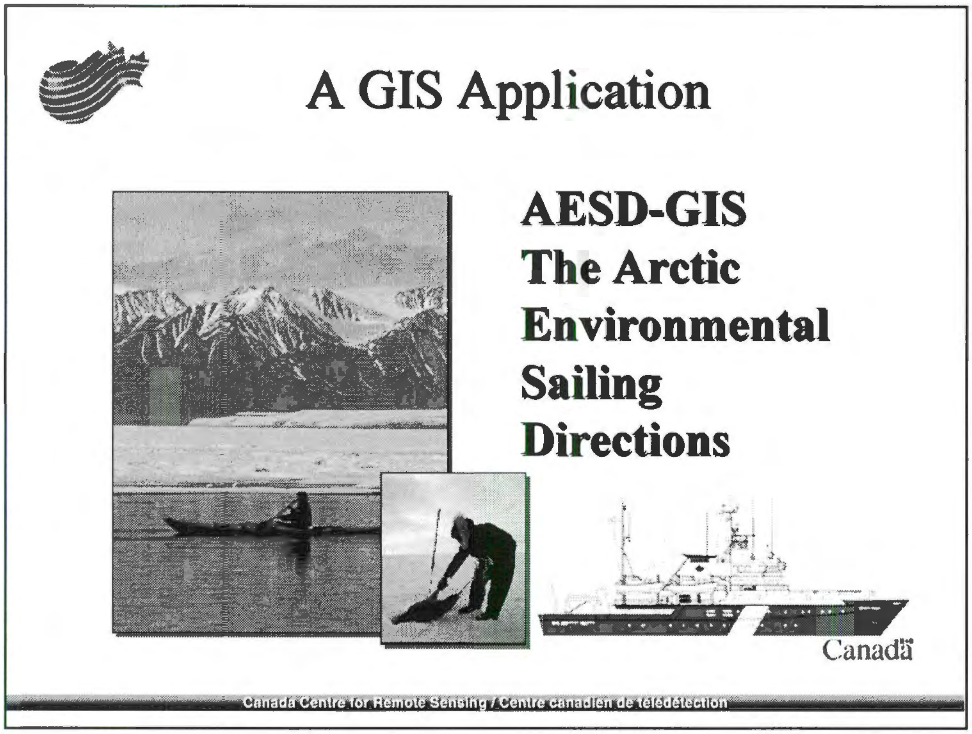
- To become familiar with the development of a GIS application.

Contents:

- A GIS Application
- GIS Implementation - Costs and benefits



Canada Centre for Remote Sensing / Centre canadien de télédétection



What is the AESD?

The environmental component of the Sailing Directions - Arctic Canada is a set of directives to assist the mariners in the selection of their route through the waters of Northern Canada. The ultimate objective is to minimise the impacts of shipping operations on the biophysical and human environment.

Objectives of the project:

To produce sensitivity-to-navigation maps of the arctic environment for the Canadian Coast Guard and the Canadian Hydrographic Services and to build the GIS Data Base required for future updates of the Sailing Directions.

The area:

Marine areas which include the navigable routes of the Canadian Arctic - the North-West Passage.

The database:

Data of special interest and legal restrictions.

Valued Ecosystem Components (VECs): biological and socio-economic

Base: reference data

The analysis:

Application of a model of the sensitivity of the environment based on seasonal variabilityof the ice regimes, sensitivities and vulnerabilities of the biological species on ship and aircraft activities, local population activities and legal restrictions.

Output:

Sensitivity maps, VECs maps and descriptive textual/image material.

GIS Implementation

GIS Implementation viewed in 6 phases:

- Awareness.
- Development of system requirements.
- System evaluation.
- Development of implementation plan.
- System acquisition and start-up.
- Operational phase.

Aronoff, S. (1989). *Geographic Informaltion Systems: A Management Perspective* WDL Publications, Ottawa.

Canada Centre for Remote Sensing / Centre canadien de télédétection

Awareness - is developed by acquiring basic knowledge about GISs, doing pilot projects, or feasibility studies.

System Requirements - User requirements should be analysed and detailed specifications of the system should be prepared as a Request For Proposal (RP) document sent to selected vendors.

System Evaluation - Replies from vendors are reviewed and ranked. Benchmarking tests are run by the short list of vendors and the results are evaluated for system selection.

Implementation Plan - This plan describes how technology, information, and people will be organized into an operating information system. GIS capability can be acquired in several ways: contracting for all services to purchasing a complete system, to developing the entire set of H/W and S/W components in-house.

System Acquisition and Start-Up - Database development is the first stage of implementation but requires the largest expenditure. It is advantageous that it be started as early as possible. The development of in-house expertise early in the project is also an asset.

Operational Phase - This stage is reached when end-users are making effective use of the system. Responsibility should be clearly assigned for maintenance, up-grading systems and databases, keeping aware of new developments, and setting and maintaining data quality standards.

Costs of Using GISs

- Hardware and software purchasing
- Personnel
- Site preparation
- Software customization
- Data conversion
- Maintenance and upgrading

Canada Centre for Remote Sensing / Centre canadien de télédétection

H/W - Platforms vary from high end (workstations) to low end (personal computers - PCs) installations. GIS data require large disk storage volume. Usually, one should allow at least 4x size of the database.

S/W - Over 200 GIS and related software are listed in the International GIS Sourcebook from GIS World, Inc. The cost ranges from ~500US$ to over 50K US$. The low end packages offer less functions, some being limited to functions such as measurement and query. These are used mainly in management applications. High end packages are broken down by modules, thus cutting the cost of acquisition.

Personnel - Training is required before system installation; the extent can vary from computer literacy, database management, to strictly GIS operations.

Site Preparation - Preparation involves creating an air conditioned, dust controlled environment, with computers protected from power failures and surges, as well as creating an operating environment for users, with proper lighting, desks and chairs; cabling needs and access to site have to be considered.

Software Customization - Customization is often required and takes the form of a modified GIS or an application written over a GIS.

Data Conversion - Conversion can cost 70 to 80% of system cost. However, more data is becoming available and accessible in digital format (e.g.: DCW).

Maintenance and Upgrading - Maintenance is often forgotten. For H/W and S/W, the cost is approximately 12% of the purchasing cost per year.

Benefits

- New and better analytical capabilities
- More accurate
- Less data redundancy
- Better revenue performance
- Higher staff productivity
- Reduced costs
- Better decision support
- Higher quality information

Canada Centre for Remote Sensing / Centre canadien de télédétection

Other benefits of GIS:

- Better sharing of data resources
- Less redundancy of data
- Consistency of data content and format (standardization)
- Spatial and descriptive data are explicitly linked
- Lower cost for data capture, storage and use
- Different "user views" of the same data are efficiently supported
- Flexibility of data retrieval, analysis and reporting
- Faster development of user applications
- Micro-level economical analysis
- Temporal Changes

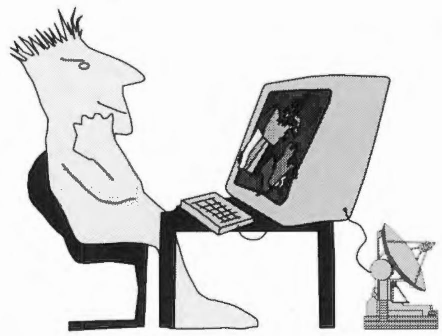
Conclusion

Trends
References
Other GIS Sources

Module 5

Objective:

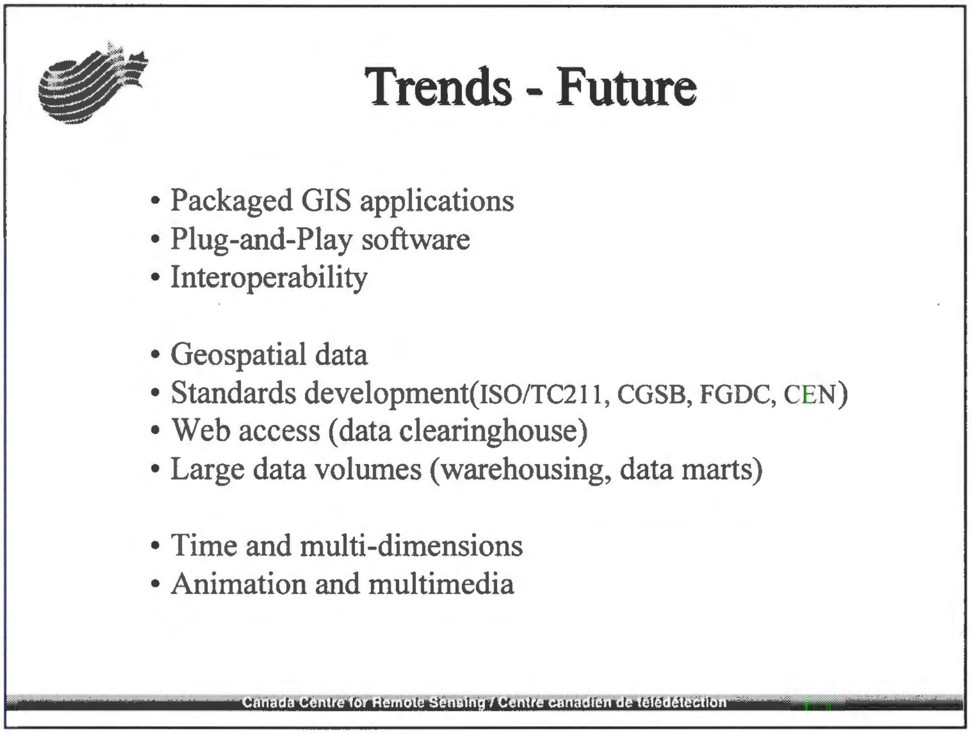
- To gain an understanding of current and future trends in GIS.



Contents:

- Trends
- Sources of GIS information
- GIS on the Web

Canada Centre for Remote Sensing / Centre canadien de télédétection

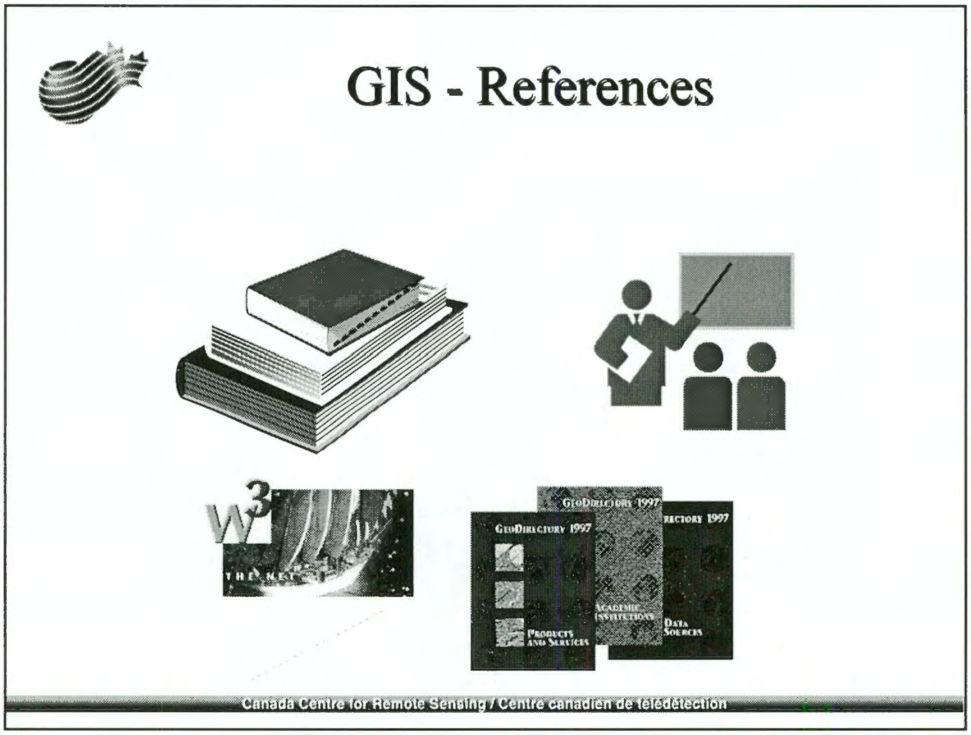


An active GIS market has resulted in lower costs and improvements in the H/W and S/W components of GIS. These, in turn, have led to a much wider application of the technology throughout government, business, and industry.

As the scientific community recognizes the environmental consequences of human activity, GIS technology is becoming an essential tool in the effort to understand the process of global change. The interactions of complex natural systems can be simulated and a variety of images and cartographic products, some animated, allow researchers to view their subjects in new ways.

By feeding satellite data into a GIS and adding the element of time, animations can simulate Earth processes over days, months, and years and give researchers the ability to examine the variations in these processes.

GIS and related technology will allow for a better understanding of terrestrial processes and better management of human activities.



Books

- Aronoff, Stan. 1989. *Geographic Information Systems: A Management Perspective*, WDL Publications, Ottawa. 294 pages. (A classic)
- Davis, Bruce 1996. *GIS: A Visual Approach,* OnWord Press, 386 pages. (For beginner, non-technical person).
- Maguire, David J., Michael F. Goodchild and David W. Rhind, editors, 1991. *Geographical Information Systems, Vol. 1: Principles, Vol. 2: Applications*, John Wiley & Sons, Inc., New York, 447 pages.
- *GeoDirectory 1997: Vol. 1,2 & 3. Gayle Rodcay, Ed., GIS World Inc., 1997. A comprehensive list of international GIS resources. The GeoDirectory replaces the GIS World Sourcebook.*

On the Internet: "Internet shopping" - [http://warehouse.geoplace.com/]

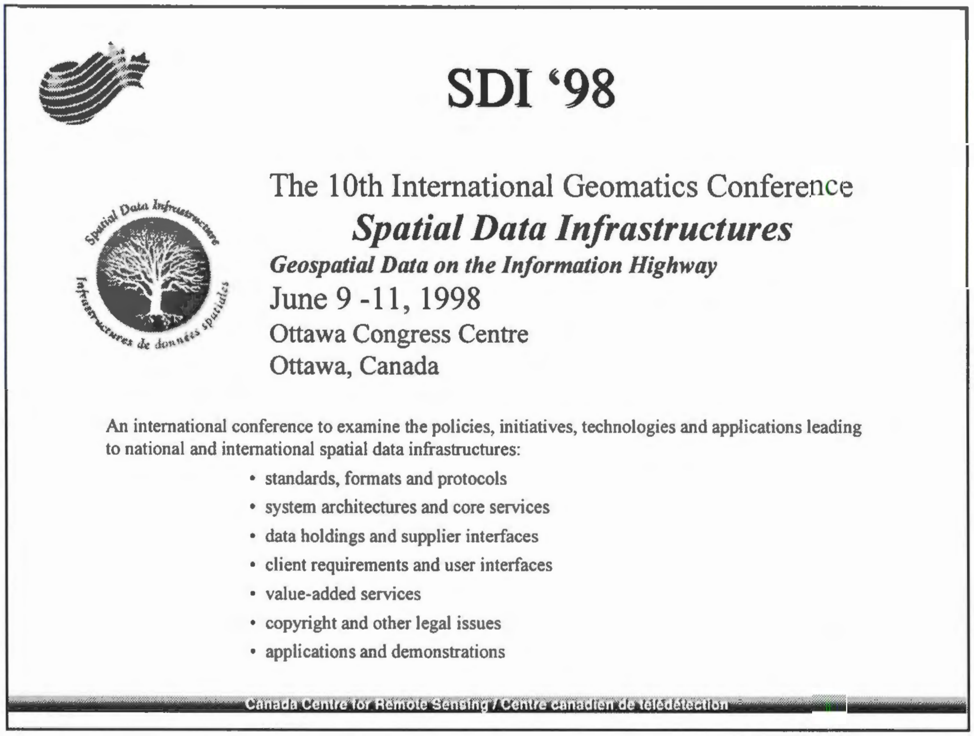
- *Internet Book Shop* [http://www.bookshop.co.uk/]
- *GIS BOOKS from Around the World,* GIS World, Inc. [http://www.geoplace.com/]
- *Internet GIS and RS Information Sites* by Michael McDermott, Queen's University [ftp://gis.queensu.ca/pub/gis/docs/gissites.html]
- *Nice Geography and GIS Servers* by the Faculty of Geographical Sciences, Utrecht University, The Netherlands [http://www.frw.ruu.nl/nicegeo.html]
- *Finding information Places which index lots of useful GIS-related pages:* [http://enva2.env.uea.ac.uk/gwww.html]
- *Great GIS Net Sites!* by Jim Aylward of HDM [http://www.hdm.com/gis3.htm]

GIS Journals & Magazines

Canada Centre for Remote Sensing / Centre canadien de télédétection

Scientific Journals and GIS Magazines

- International Journal of Geographical Information Systems (G.B.)
- Geographical Systems, The International Journal of Geographical Information, Analysis, Theory and Decision (GESYEO) (The Netherlands)
- Geo Carto International. A Multi-disciplinary Journal of Remote Sensing & GIS (Europe)
- Cartography and Geographic Info Systems. Journal of American Congress on Surveying & Mapping
- GEO-INFORMATIONS-SYSTEME Journal for Cross-disciplinary Exchange of Knowledge in the Geo-Sciences
- GEOMATICA, Canadian Institute of Geomatics/Association canadienne des sciences géomatiques
- GIS World (The World's leading Geographic Information Systems Publication)
- Gis Europe (Europe's Geographic Information System Magazine)
- Geo Info Systems (Applications of GIS and related spatial information technologies) [http://www.geoinfosystems.com/gis.htm]
- EOM, The Magazine for Geographic, Mapping, Earth Information



Other Conferences:

Geomatics Atlantic, 27-29 October 1997, Halifax, Nova Scotia

GIS/LIS '97, 28-30 October, 1997, Cincinnati, Ohio

GEOMATIQUE VI / GEOMATICS VI, 13-14 November 1997, Montreal, Que

GIS '98/RT '98, 06-09 April, 1998, Toronto, Ontario

URISA '98, 19 - 23 July, 1998, Charlotte, North Carolina

(Urban Regional Information Systems Association)

BGC '98 , 11-14/05/98, Dallas, Texas, Business Geographics

19th International Cartographic Association (ICA) General Assembly and Conference, 14-21 August 1999, Ottawa, Ontario

XIX. ISPRS Congress , 16 - 22 July 2000, Amsterdam, The Netherlands

(19th International Congress of Photogrammetry and Remote Sensing)

etc...

See [http://www.ccrs.nrcan.gc.ca/ccrs/comvnts/events/confere.html] for a current list of up-coming conferences.

Hands-On GIS

GIS for the CURIOUS

Multimedia guide on GIS

- vector and raster representations of geographic features geographic data retrieval and database operations
- topology
- spatial analysis
- cartographic editing
- and many others

Canada Centre for Remote Sensing / Centre canadien de télédétection

A hands-on guide to Geographic Information Systems

by Y.C. Lee, University of New Brunswick

for Geomatics Canada in collaboration with Universal Systems Limited and the University of New Brunswick

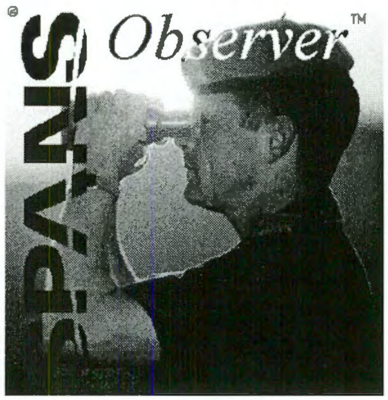
- Interactive GIS software and tutorial dataset (on CD)
- GIS handbooks and animation (additional)

This educational software program includes a text with hands-on exercises that a user can follow using the accompanying GIS software. It provides a good, hands-on introduction to the fundamentals of GIS .

There are animations to help you understand GIS concepts and operations of the software.

[http://universal.ca/documents/products/curious_main.html]

GIS Data Browser



SPANS Observer

- royalty-free licensed software
- to display and query existing SPANS data

Canada Centre for Remote Sensing / Centre canadien de télédétection

SPANS Observer

This product is royalty-free licensed software which allows you to display and query geographic data. The funding of this development project was provided in part by the Canadian Federal government through programs such as the Geographic Information Systems Development Program (GISDP) of Geomatics Canada.

The suite of functions it contains are outlined in its comprehensive on-line help which meets the same high level standards as TYDAC Research Inc.'s commercial software.

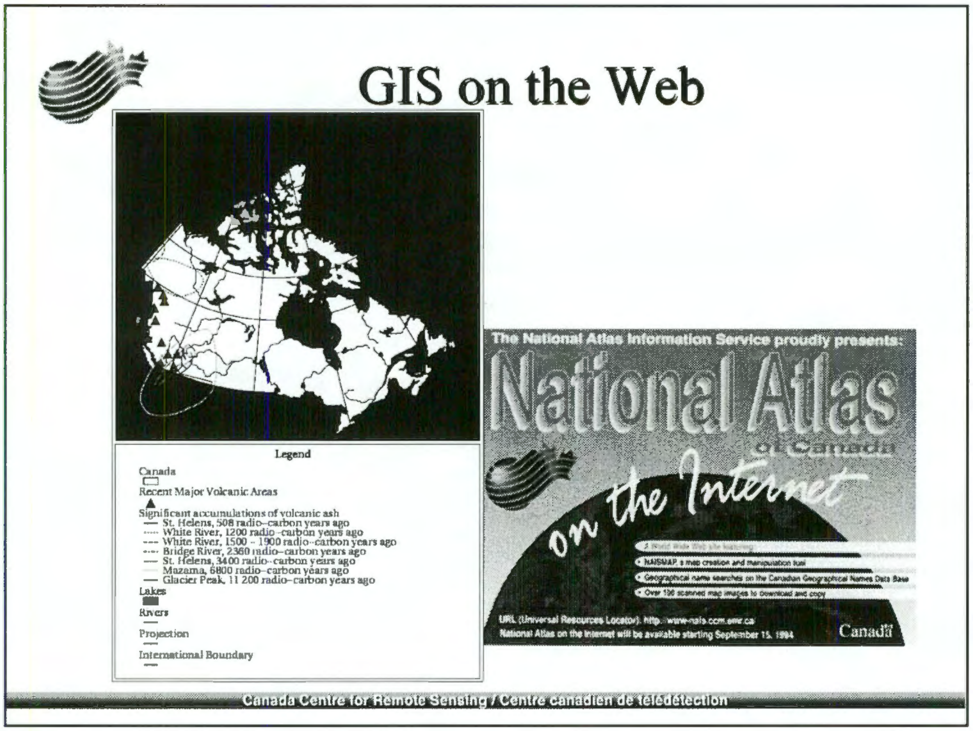
You can display existing Map Compositions or create new ones containing any SPANS-format entities; such as, points, lines, areas, quadtrees, and/or rasters. Existing SPANS data and Arc Info export files can both be imported.

You can annotate your Maps with legends, titles, scale bars and north arrows.

You can perform spatial, non-spatial and "nested" queries; that is, subsequent queries which evaluate only the selected records of previous queries.

Finally, you can save slides in PCX format for insertion into documents or presentations.

[http://www.tydac.com/develop/products/modules/observ.html]



A user need only have a Web browser and access to the internet to use the large spatial database and powerful GIS. It provides display and analysis tools to facilitate data sharing and cooperation between environmental planning agencies, public action groups, citizens, and private entities.

NAISMap - [http://ellesmere.ccm.emr.ca/naismap/naismap.html]

Interactive mapping tool for geographic study of Canadian issues.

Created by geographers and cartographers of the GeoAcess Division.
Data layers: small scale maps of Canada; single theme or topic; text describes significance of layer.

Issues: groupings of layers that may be used to illustrate/explore an issue (E.g.: Aging population; natural hazards; endangered species: birds, fish, marine mammals, reptiles, plants; languages; wetlands; boreal forest; agriculture; base map components)

GRASSLinks 3.0 - [http://www.regis.berkeley.edu/grasslinks/]

Public Access GIS: offers public access to environmental data.

Other GIS on the Web: through [http://www.hdm.com/gis3.htm] and other URLs given previously.

A recent review: [http://www.geoplace.com/] Also appears in GIS World September 1997. Four client applications to access data over the Web are reviewed.

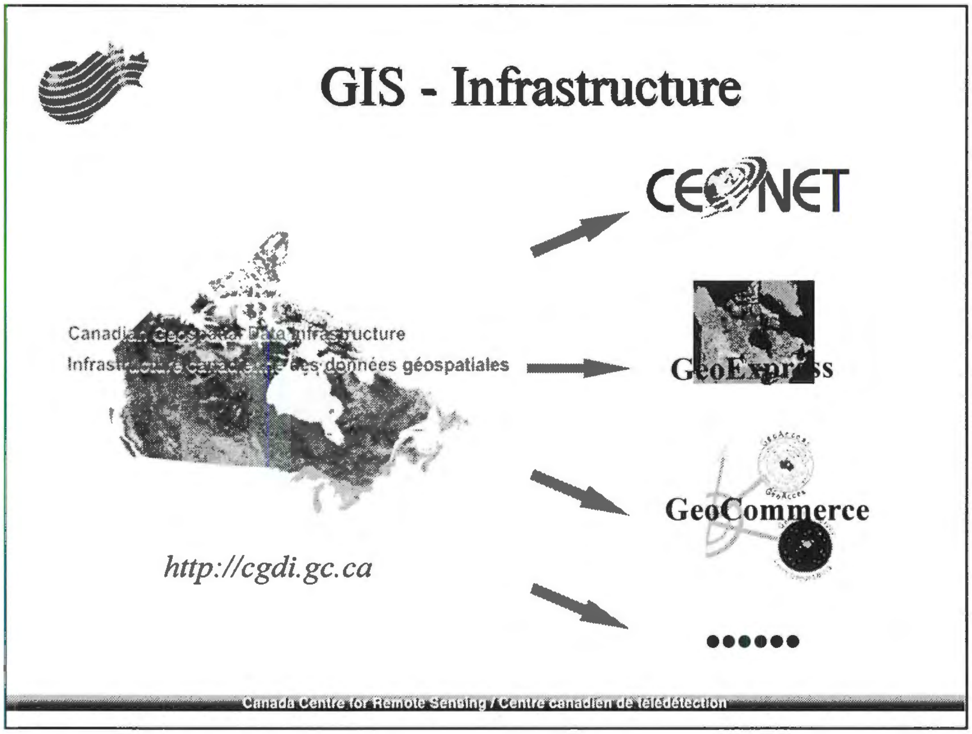
1997 Federal Election Map: [http://ccrs_gad1.ccrs.nrcan.gc.ca/]

This Web mapping tool allows a user to plot demographic variables with the 1997 election results to determine the influences of voting patterns.

CARIS Internet Server: [http://caris0.universal.ca/demo/login/]

A web service designed on top of the CARIS GISallow users to search, query, manipulate and display geographic data from a GIS database.

Bennett, D and K. Davie, **"Accessing Spatial Data Over the Internet"**
[http://www.uwin.siu.edu/ucowr/updates/99/bennett.html] ...and others....



The CGDI is a web site designed, operated and maintained for the geomatics community by the GeoAccess Division Earth Sciences Sector, Natural Resources Canada on behalf of the Interagency Committee on Geomatics and the Canadian Council on Geomatics.

Similar initiatives are being undertaken around the world - the National Spatial Data Infrastructure from the US is one example. The NSDI is accessible through the Federal Geospatial Data Committee home page [http://fgdc.er.usgs.gov/].

These infrastructures will become part of a global initiative - the Global Spatial Data Infrastructure.

The objective of Geomatics Canada is to provide up-to-date geographical information on Canada's landmass.

Specifically:

- to provide a reliable system of surveys, maps, remotely sensed data and geographically referenced information covering the Canadian territory, in support of national sovereignty, defence, the environment, socio-economic development and the governing of Canada;

- to promote the development of technologies for surveying, mapping, remote sensing and geographic information systems, and to foster the growth of related expertise in both the public and private sectors; and

- to contribute to Canada's international competitiveness and to support trade in geomatics through international agreements and export market development, in cooperation with the Canadian geomatics industry.

L'objectif de Géomatique Canada est de fournir de l'information géographique à jour sur la masse territoriale du Canada.

Plus précisément :

- fournir une série d'informations fiables - levés, cartes, données de télédétection et information à référence géographique - sur l'ensemble du territoire canadien, en vue de contribuer à l'exercice de la souveraineté nationale, à la défense, à la protection de l'environnement, à la croissance socio-économique et à l'administration du pays;

- favoriser la mise au point et le perfectionnement de techniques relatives aux levés, à la cartographie, à la télédétection et aux systèmes d'information géographique et encourager l'enrichissement des connaissances en ces domaines dans les secteurs public et privé; et

- contribuer à accroître la compétitivité du Canada à l'étranger et appuyer les activités commerciales en géomatique par l'entremise d'ententes internationales et du développement des marchés d'exportation, en collaboration avec l'industrie canadienne de la géomatique.