SIG
Geographical information systems

Pau Fonseca i Casas

pau@fib.upc.edu

Contents

- Definition
- SIG structure
  - Organization
  - Human resource
  - Tecnical resource
  - Data
- Interoperability
- SIG clients

SIG introduction

DEFINITIONS

What is a SIG?

- ARONOFF (1989): “A set of manual or computer procedures used to store and process data geographically referenced.”.
- NCGIA (1990): “System hardware, software and procedures designed for the capture, storage, manipulation, analysis, modeling and presentation of spatially referenced data for solving complex problems of planning and management”

What is a GIS?

- “An information system that is designed to work with data referenced by spatial or geographic coordinates. In other words, a GIS is both a database system with specific capabilities for spatially-referenced data, as well as a set of operations for working with the data.” (Star and Estes, 1990, p. 2).

What is a GIS?

- “Automated systems for the capture, storage, retrieval, analyze, and display of spatial data.” (Clarke 1995, p13).

What is a SIG?

- “A powerful set of tools for storing and retrieving at will, transforming and displaying spatial data from teh real world for a particular set of purposes”. (Peter Burrought 1986 p. 6)

What is a GIS?

- http://www.geog.ucsb.edu/~good/

- GOODCHILD (1985): “A GIS is best defined as a system witch uses a spatial data base to provide answers to queries of geographical nature”.

What is a GIS?

- This definitions can lead people to the misconception that a SIG is only a database, for that Cebrian(1994) explain the concepts SIG and SGBD to establish a clear distinction:
- “In a GIS, the information contained in the data base can be dissect first by spatial localization or by the context.”.
- “In a SGBD, the items can be spatially addressed if, and only if, a correspondence is defined between the geographical localizations and the information registers (memory positions)”.

What is a GIS?

- A geographical information management method or technique, allowing the effective combination of basic information to obtain derived information. For that the system rely on the information sources and a set of informatics tools (hardware and software) to simplify this task; all of this is framed in a project defined by a team and controlled (managed) by technicians in charge of its development and implantation.

What is a GIS?

- A GIS is a tool capable of combine graphical information (maps) and alphanumeric information (statistics…) to obtain derivative information over the space.

Question for a GIS

- Where is the object identified by “A”?
- Where is the object “A” in relation to object “B”?
- How many occurrences of “A” exists in a distance “d” of “B”?
- What is the value of the Z function in the position X?
- What is the dimension of “B” (area volume, perimeter, frequency)?
- What is the result from the intersection between different kinds of information?
- What is the short path (less resistance and les cost) over the terrain from a point (X1, Y1) following a set of paths P until the point (X2, Y2)?
- What are in the point (X, Y)?
- What objects are near to some objects wit a specific characteristics combination?
- What is the result from a classification some sets of spatial information?
- Using a model, defined in the real world, simulate the effect of the process “P” in time “t” on a scenery “S”.

Application areas

- GIS is an applied technology is applied. Its possibilities are related to any activity related with spatial information processes.
- Traditional areas of application are:
  - Land registry management and urban property.
  - Environment.
  - Urban planning.
  - Big networks management (telecommunications, gas, water, electricity, etc.).
- New application areas:
  - Business management, archaeology, historical analysis, epidemiology, criminology, among others. That means, any scientific, technical or business area, working with spatial data.

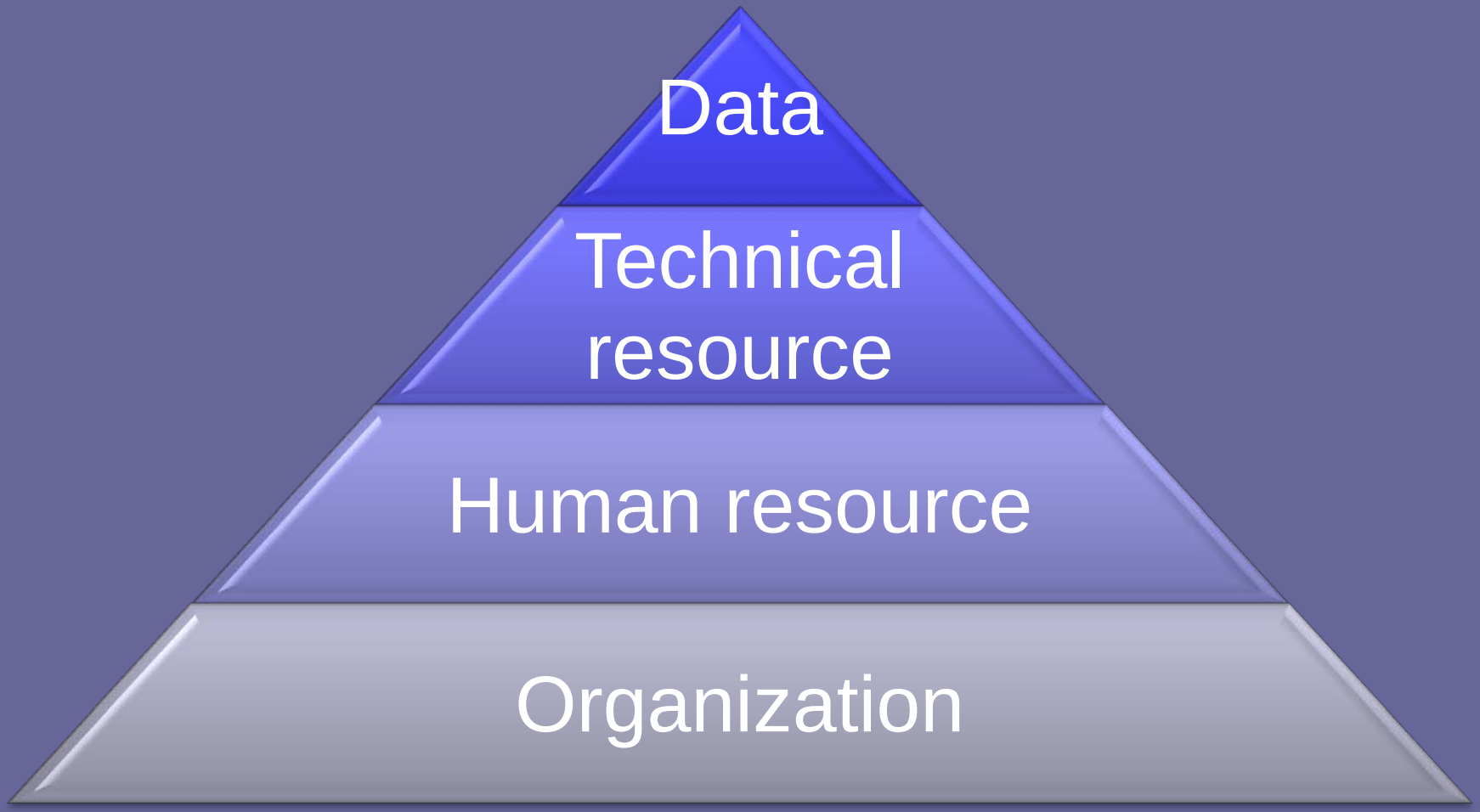
Data viability

- More and more responsible organisms of the geographical infrastructure of the more industrialized countries, have digital information and databases accessible for the public, that are, at least the “food” for the GIS.
- This avoids one of the bottle necks that suppose the data loading process in a GIS, that some times makes unviable its use in small organizations or in projects with low resources.

GIS introduction

GIS STRUCTURE

GIS structure



GIS structure

ORGANIZATION

Organization

- GIS implantation can be developed successfully under the protection of a strong institutional will, deeply decided and convinced of the implications derived of adopt this technology

Organization

- http://www.teleatlas.com/index.htm
- http://www.digitalglobe.com/
- http://www.navteq.com
- http://www.nasa.gov/
- http://www.icc.cat

SIG structure

HUMAN RESOURCE

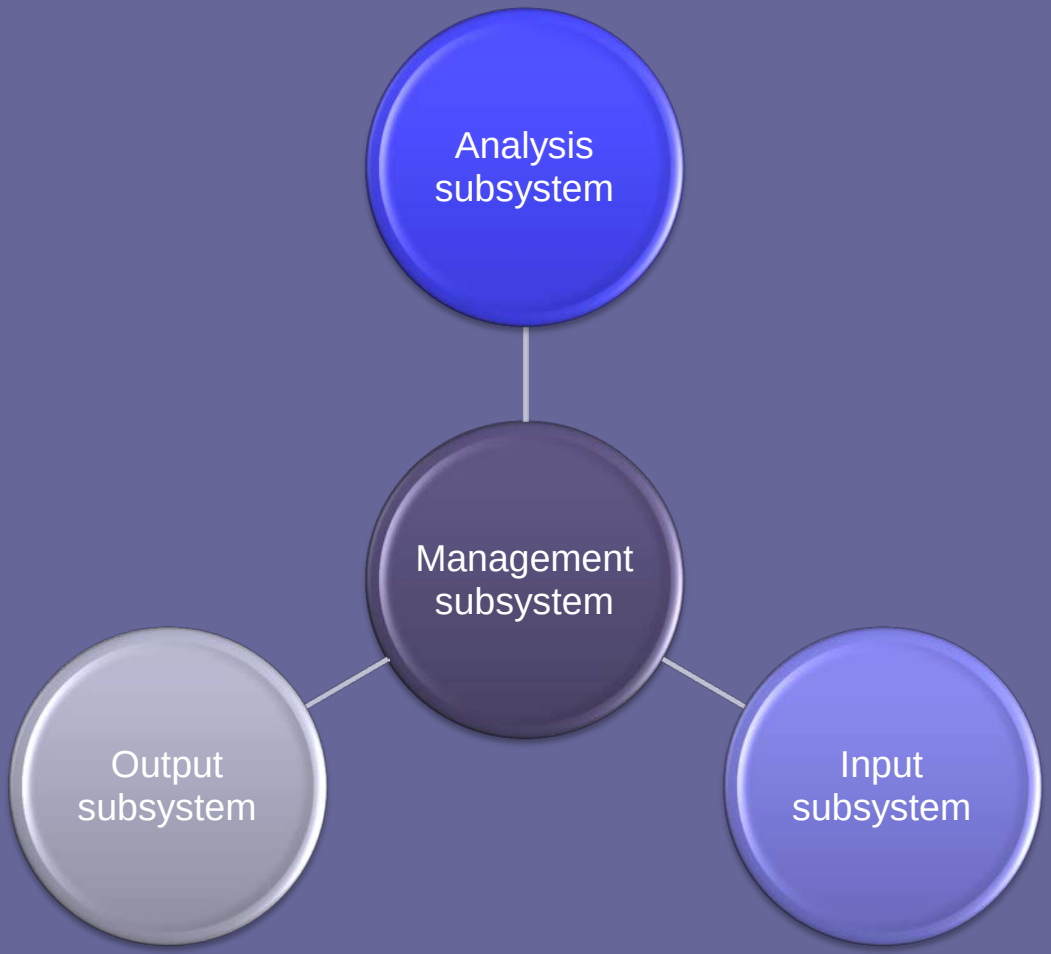
Human resource

- A GIS needs a human team, with knowledge not only of GIS but also a reasonable knowledge related to the analysis and application areas.
  - People skilled to conceptualize and manage the utilities of GIS technology
  - People that acts as a client.

GIS structure

TECHNICAL RESOURCE

Technical resource



Maps digitalization systems

- Cartographic databases of GIS are build from maps, photography or images that represents the starting data.
- To load this information in the GIS database is needed to transform all the information to the system digital format. To do that usually digitalization programs and format conversion programs exists.
  - Digitalization of analogical maps can be done with a scanner or with a tablet PC among other methods.

Maps digitalization system

- Scanners can be used to obtain digital information in raster format from maps or photos. This information can be used directly by the family of GIS that uses this kind of data (raster GIS).
- Sometimes the digitalization tools provided by the GIS are not good enough, for that is needed to use other programs to transform the information from the analogical maps to the digital format. Then, we can use the transformation or importation tools of the GIS to use this new digital maps. Usually the GIS imports a previous digitalization of a map, for example digitalized using AutoCad® (obtaining a DXF, Digital Exchange File).

Input subsystem

- This systems performs the capture and the transformation of the analogical data, like printed maps, alphanumerical registers in paper or champ observations.
- Transforms the digital information that comes form different remote sensors or other systems, to a a compatible platform with the GIS computational language

Input subsystem

- Among the different input devices are:
  - Graphic tablets devices (Wacom).
  - Scanners.
  - Laser, magnetic readers.
  - Keyboards.
  - Terminals and ports.
  - Internet.
  - …

Management subsystem

- Is the subsystem that allows the storage, sort and recovery of data.

- This organization is possible thanks to programs known as database management systems (DBMS) that allows manage digital spatial data.

Management subsystem

- Using databases and DBMS to the management of the data permits the query, management of derived data and retro alimentation.

- Among the devices used to storage the data are:
  - Hard discs, magnetic tapes and compression units and zip-back up.

Analysis subsystem

- This subsystem allows to relate spatial data obtaining new spatial data using the a defined relation.
- A simple example consists to obtain a map with the areas with a set of spatial features, as an example we can obtain a map containing the residential areas over radioactive geological area.

Analysis subsystem

- In that case first is needed to establish a classification of the different geological material depending on the radioactive emissions, obtaining a map that can be superposed with a map of the residential areas. This leads to obtain a map of land uses. The spatial relation that we establish between these two maps are based in the coordinates coincidence, representing only the areas that accomplish this condition. This is a typical problem that cannot be solved using a typical database system, because cannot perform spatial superposition.

Analysis subsystem

- Different GIS analysis can be performed, starting from a easy comparison of objects depending on the attributes, until complex analysis of optimal routes on space and time.
- In a GIS package are usual the spatial analysis, proximity analysis, analysis of networks and the analysis in the third dimension, among others.

Analysis subsystem

- The exit of this operations lays on the quality and the preparation of the information that must be analyzed. Is needed a correct conceptualization of the different tasks and an analysis previous its running.

- As an example, in Google Earth®, queries to determine the optimal route between two cities can be performed.

Cartographic representation subsystem

- Allows to draw maps from selected elements of the databases, perform different cartographical compositions and send these maps to the output devices, like printers and plotters.
- Maps are the most usual way to represent the results of an analysis performed by a GIS. Often tables and diagrams are used to complement the information. This subsystems implements tools to create this kind of documents.

Output subsystem

- Is the subsystem that understand how to represent the data and the results derived from the analysis subsystem.
- The output data can have a graphical structure (maps and graphics) or an alphanumeric structure (tables, reports).

Output subsystem

- Output can be generated in analogical or digital formats. These formats usually can be exported to other GIS directly or using a specific software.
- Among the output devices we can find:
  - Terminal and output ports, printers, plotters, magnetic tapes, hard discs, optical devices, …

GIS structure

DATA

Data

- Maybe the more important part of a GIS is its data.
- A GIS works with geospatial data.

Data

- Related to a geographical space with a known position (under a coordinate system).
- GIS stores the localization of the data, its spatial relation with other data (topology) and a description through its attributes.

Spatial database

- An spatial database is a set of data structured in a way that allows a set of GIS applications working with it. This data are spatial and are geo referenced related acting as a model of the reality, classified and grouped depending on its features. The data is under a redundancy control and integrated to allows the development of applications and analysis over the information.

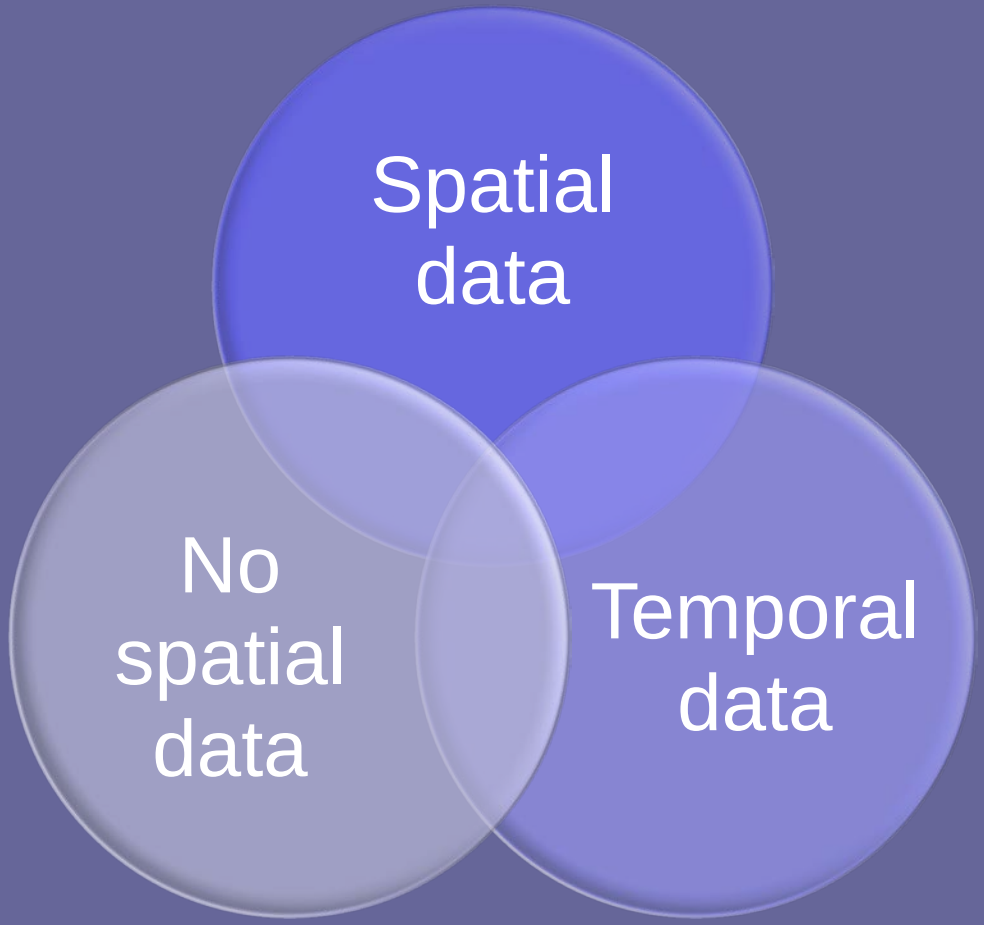
Database management systems (DBMS)

- A database management systems (DBMS) is a kind of software used to manage and analyze the data stored in a computer.
- From these systems the data can be stored in tables, the relations between the data can be represented and new tables can be created with the results. This tables can be related with the spatial database representing the results in a thematic map.

DBMS structure

- The central nucleus of the system is based on the spatial databases and thematic databases. These databases stores, structured, the cartographical objects (its position, size and shape) and its no geometrical features (attributes) respectively.
- In a plot map, the shape and the situation of the plots are represented in the geographical database, while the information related to the owners, land use, etc, will be in the thematic database.
- In some systems both databases are clearly separated and have different structure. In others are represented by only one entity and can use other external databases like Dbase or Access to store thematic attributes.

Geographical data



Spatial data

- The spatial data or attributes, are the geographical features of the described objects (placement, dimension, shape). That means, as an example, that the points that constitutes the perimeter of a population are stored in some kind of files that are understood by the geographical applications that are in the market.

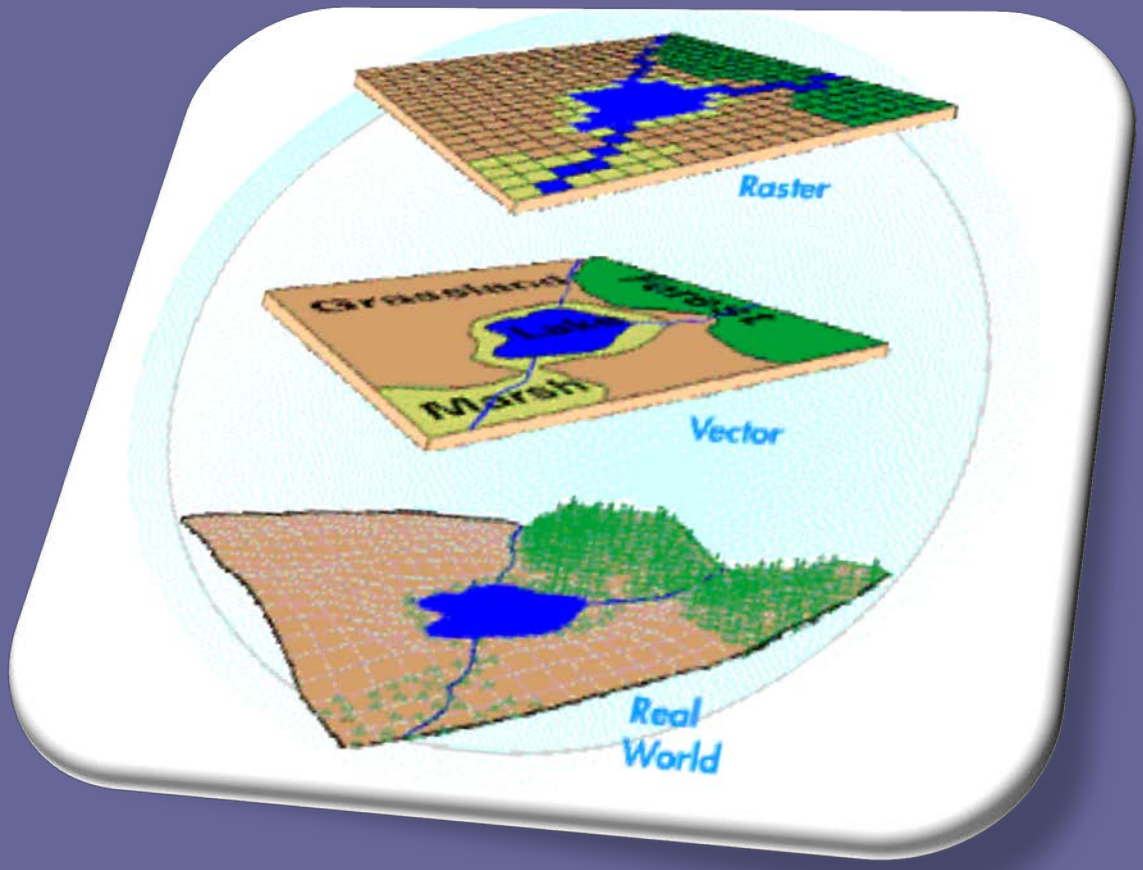
No spatial data

- No spatial data or attributes are the quantitative features related to the object that must be described, usually are stored in tables and are managed by a DBMS. These data are also named descriptive data.

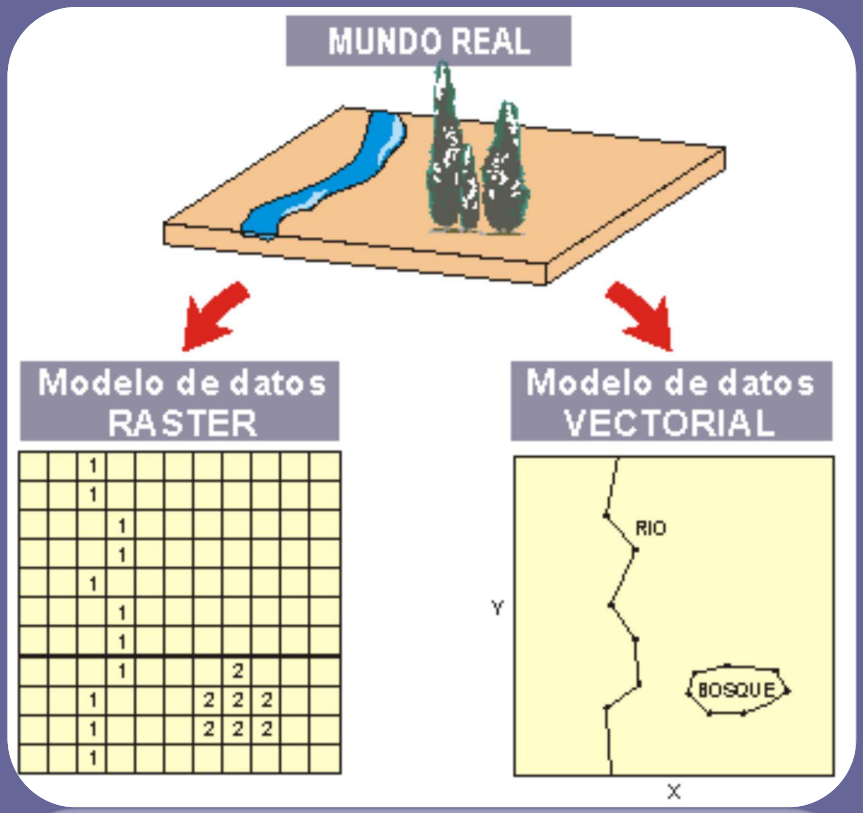
Temporal data

- Temporal data are used to inform about the time related to each data stored in the GIS. Although this data is not common as the other data, more and more are increasing its relevance.

Maps representation files



Maps representation files



Raster data example

- Example of “rasterization” of a cartographic element.

20	30	10	90	178	12	89	12
34	23	56	12	342	56	34	38
12	12	34	15	76	87	12	32
...							

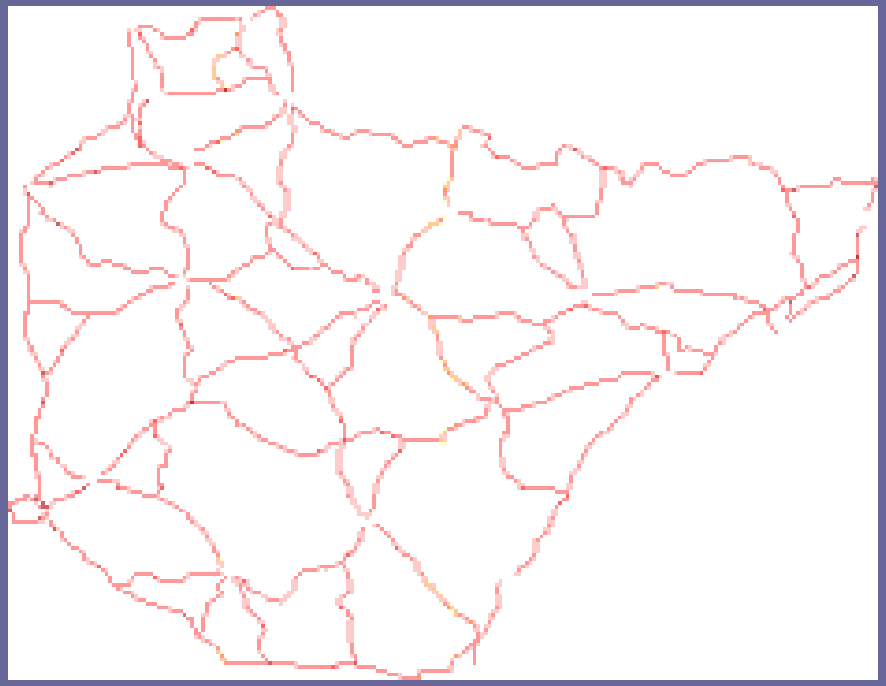
Rasterization

- **Rasterization** or Rasterisation is the task of taking an image described in a vector graphics format (shapes) and converting it into a raster image (pixels or dots) for output on a video display or printer, or for storage in a bitmap file format.

Vector codification

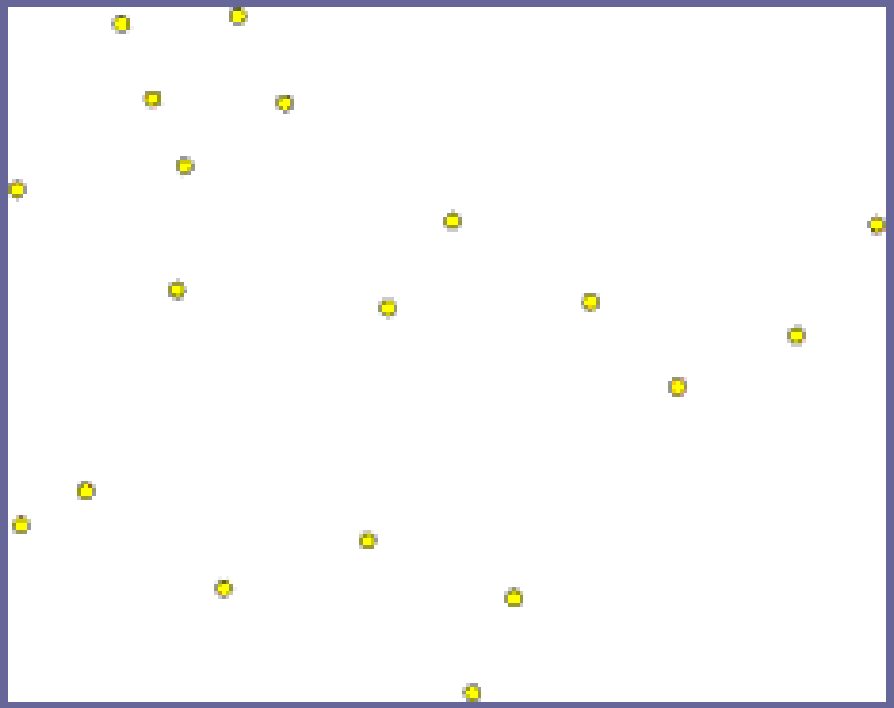
- Final desired represetation.

Vector codification



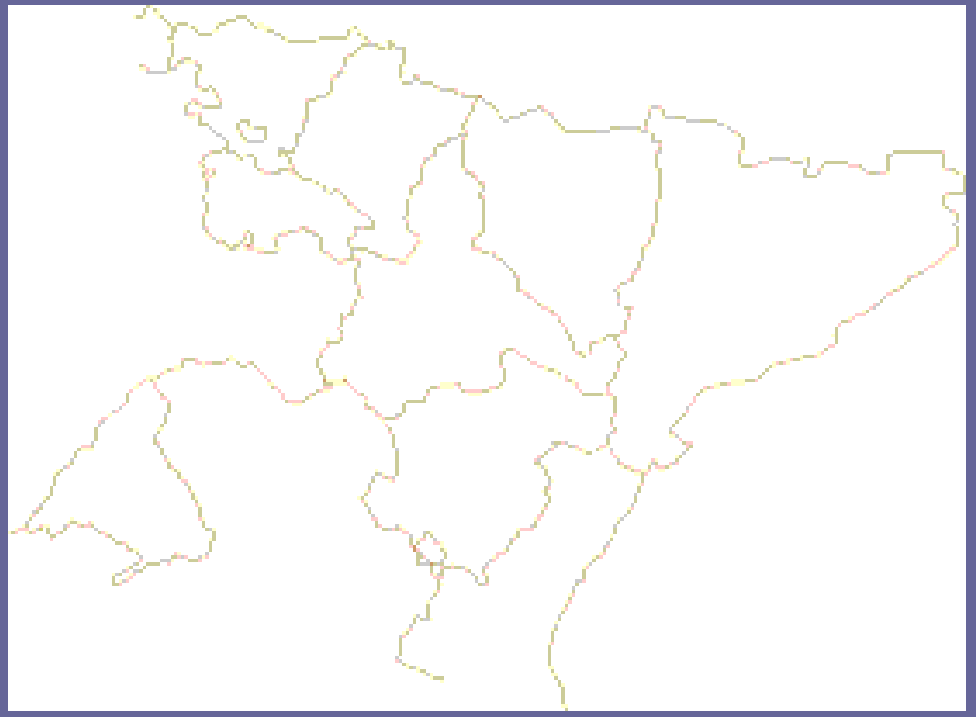
- Road codification.
- Polilines.

Vector codification



- City codification.
- Points.

Vector codification



- Province codification.
- Poligons.

Raster model (advantages)

- Simplest data structures.
- Direct representation of the remote sensing data.
- Superposition operations simplest.
- Spatial operations like distribution, density and surface, more efficient.
- The spatial unity keeps the same shape and size, simplifying the simulations.
- Useful for the analysis of big extensions with low detail of spatial features.

Raster model (disadvantages)

- Big volume of graphical data:
  - Using big cells to reduce the volume of data implies that the representation of the phenomenon can be lost causing a lost of information.
- High level of error in estimations related to area, perimeter and longitude calculus.
- Storing space wasted for the spare spatial data.

Vector model (advantages)

- Good representation of the data structure, compact data structure.
- Good similarity with the real shapes.
- Calculus more precise of areas, perimeters and longitudes.
- Network analysis more consistent.
- Can be adapted under object oriented databases.

Vector model (disadvantages)

- Complex data structure.
- The combination of different maps with vector polygons and raster maps is difficult.
- The superposition needs more errors verifications and can be time demanding.
- The different topology of the spatial units difficult the simulation exercises.

Raster ambit

- Raster files GIS, are more suitable to work with data that presents a continuous variation in the space, such as the topographic surfaces, the temperature maps, the maps of substance concentration, etc.
- Raster are also used when satellite images are available as a source of data, more and more usual in environmental problems. The application scope is broad and raster is used in disciplines like biology, geology, medicine or environmental sciences, among others.

GIS structure

DATA INTEROPERABILITY

Interoperability

- To simplify and promote the use of geo referenced information of multiple sources in needed interoperability.
- Interoperability lies on agreements and standards establishing the geospatial concepts, the communications protocols and the format of the data.

Organisms

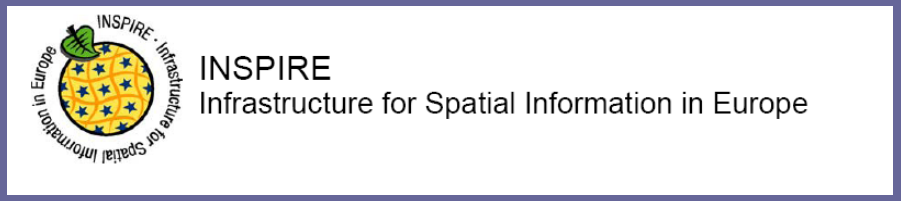
- The more important organizations that establishes and maintains the standards and specifications for the geographical information are
  - *Open Geospatial Consortium Inc.* (OGC)
  - *International Organization for Standardization* (ISO) .

Organisms

- Spatial data infrastructures (SDI) are a set of technologies, politics and institutional agreements, with the objective of simplify the production, the share and the use of the spatial information. Also have the objective of promote the social economic and environmental development of the territory.

Organisms

- The European project INSPIRE (*Infraestructure for Spatial Information in Europe*)



The initiative intends to trigger the creation of a European spatial information infrastructure that delivers to the users integrated spatial information services.

The target users of INSPIRE include policy-makers, planners and managers at European, national and local level and the citizens and their organisations.

Metadata

Is advisable to no invent a proprietary standard.

Always possible, its urged to chose a international standard, trying to keep inside the structures defined in it. Subtle changes in a international standard, like a collapse of its composite elements, can, in long term, be expensive:

- In that case you cannot use standard tools and the metadata cannot be directly interchanged or analyzed by other “software”.

XML

- The "Extensible Markup Language" (XML) gives two solutions to this metadata problem:
  - First, adds a language with solid structural rules through a file to control and validate the structure of the document.
  - Second, through other standard ("XML Style Language", or XSL), an XML document can be used with a style sheet to produce standard presentations of the content, allowing to the user change the order of the elements, changing labels or shown nothing more that selected information fields.

XML

- Using XML with style sheets allows a structured exchange format and for flexible presentation.
- Thus, a metadata entry may occur in many ways from one and the same structured coding.

Tools for metadata

- There are currently some tools to edit, store and search metadata ... most correspond to the ISO 19115 standards FDGD and there are also others that follows the Standard NEM.

ArcCatalog (ESRI)

- Create and edit metadata.
- Facilitates the synchronization of data (non-proprietary format) and metadata.
- Metadata is data move and there is continuous updating thereof.
- Calculation of coordinates (if you know the projection system)
- Field control
- Are filled with a series of forms
- Collect all the tags defined for the standard for iso11995

ArcCatalog (ESRI)

- Edition of 16 December 2005 extended the NEM Metadata Editor
  - Integrated into a GIS tool such as ArcCatalog that allows the automatic loading of a number of fields (size, name, date, etc.) and automatic synchronization of data editing.
  - Configurable based on the XML configuration file to allow editing by users.
  - You can define mandatory fields and sections, default values, domains based on the configuration file.
  - Metadata validation that meets the specification required fields NEM.
  - The friendly interface for editing in ArcCatalog.
  - Consultation and leveraging metadata search functionality ArcCatalog
  - Management and storage of metadata with the data file both as Relational Database

CatMdEdit (IDEE)

- Application deployed under the request of IDEE support the creation of metadata. Create a digital file containing the metadata properly organized
- Check the syntactic structure of the file
- Better error handling (eg date field in the order ISO)
- Create multilingual version: Replicates the metadata record by changing the language. The new record remains associated with the original registration because they are different versions of the same record.
- Edit ISO 19115 / edit NEM
- The tool provides CatMDEdit thesauri, with a friendly interface for use. This lack of direct relationship between data and metadata makes the simple version of the tool CatMDEdit (which is distributed from IGN), can not automatically extract certain information from the data as boundary coordinates, the system reference coordinate system, the character sets of data and metadata, etc..
- It is a free resource download website. (www.idee.es).

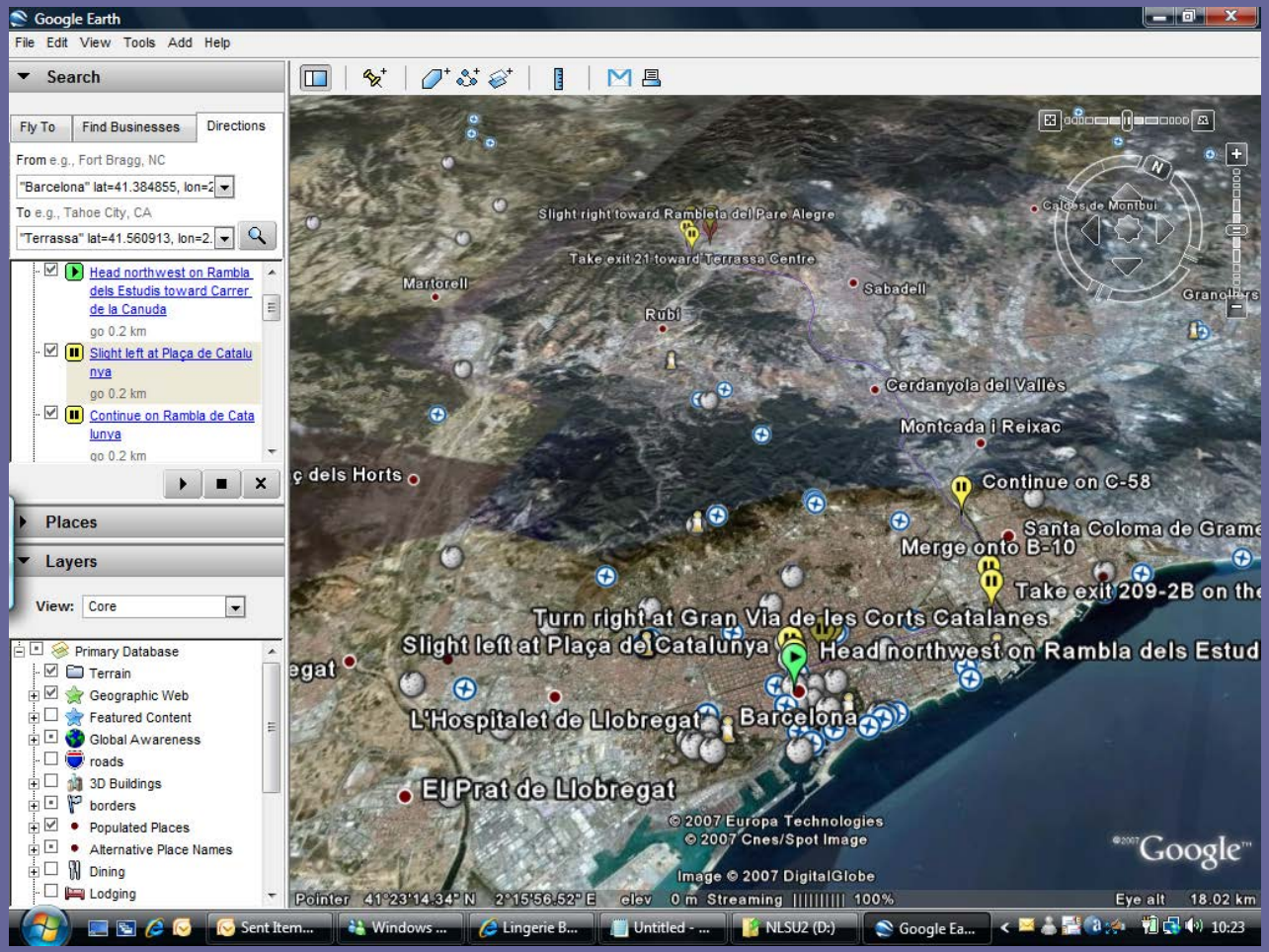
MetaD (IDEC)

- IDEC capture utility and maintenance of geographic metadata standard as ISO-19115
- Properly organized.
- Check the syntactic structure of the file.
- VALIDATE option.
- Better error handling that CatMDEdit.
- Language version in English, Spanish and Catalan.
- Download free resource under registration. (www.geoportal-idec.net)

Remotely access to a GIS

GIS CLIENTS

Google Earth



Google Earth

- Input subsystem
- Output subsystem
- Management subsystem
- Analysis subsystem

Google Earth

- Input subsystem
  - KML files.
  - WMS.
  - Keyboard.

Google Earth KML

- **“KML** is a file format used to display geographic data in an Earth browser, such as Google Earth, Google Maps, and Google Maps for mobile. KML uses a tag-based structure with nested elements and attributes and is based on the XML standard.”

Google Earth KML

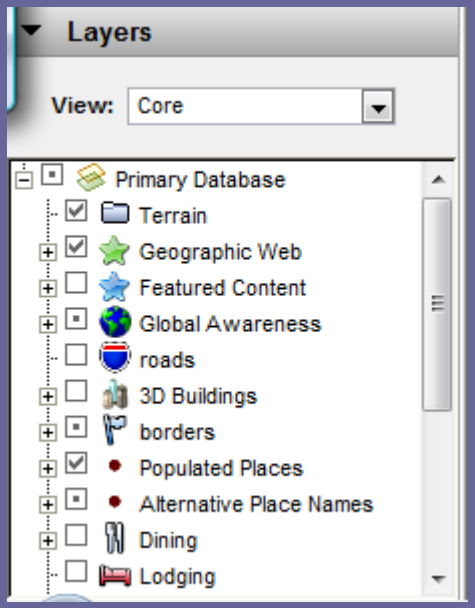
- <?xml version="1.0" encoding="UTF-8"?>
  <kml xmlns="http://earth.google.com/kml/2.1">
  <Placemark>
  <name>Simple placemark</name>
  <description>Attached to the ground. Intelligently places itself at the height of the underlying terrain.</description>
  <Point>
  <coordinates>-122.0822035425683,37.42228990140251,0</coordinates>
  </Point>
  </Placemark> </kml>

Google Earth

- Output subsystem
  - Mainly graphic.
    - DirectX, OpenGL
  - Disc.
    - KML files.
  - Printer.

Google Earth

- Management subsystem.
  - Local information storage.



Google Earth

- Analysis subsystem.

Other clients

- A lot of different clients allows to access GIS data:
  - Via Michelin
  - Microsoft Virtual Earth
  - ESRI Explorer
  - World Wind
  - Yahoo Maps
  - ...

Thanks!

- Pau Fonseca i Casas
- pau@fib.upc.edu

- Universitat Politècnica de Catalunya
- Laboratori de càlcul de la Facultat d'informàtica de Barcelona
- Jordi Girona 1-3
- (+34)93401773