Characteristics of Geospatial Data

Workshop 2: Technical Issues Towards Effective Applications of Geospatial Technologies and Data in DRM

Date : Monday 6-8 September 2021

Presented by Dr. Jacob Opadeyi,
Disaster Risk Management Consultant

Objectives

- Introduce participants to geospatial data storage models

Outline

- What are GIS data models?
- Implementation of GIS Data Models
- Types of Geospatial Data
- Vector Model
- Raster Model
- Data Modelling Process

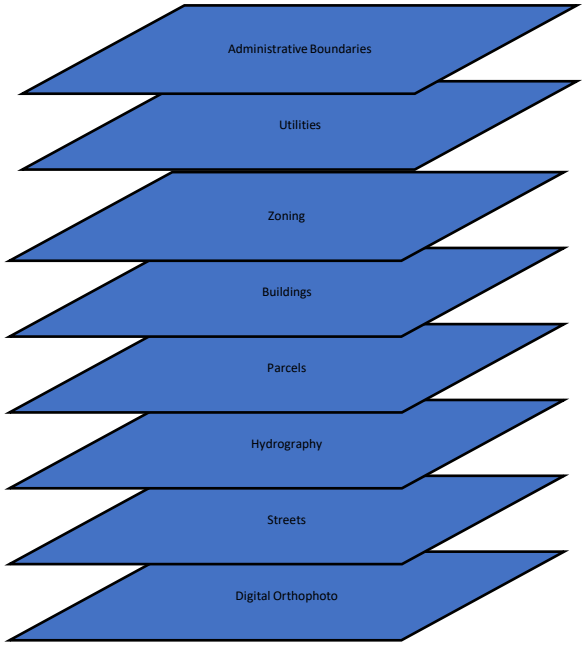
GIS Data Models

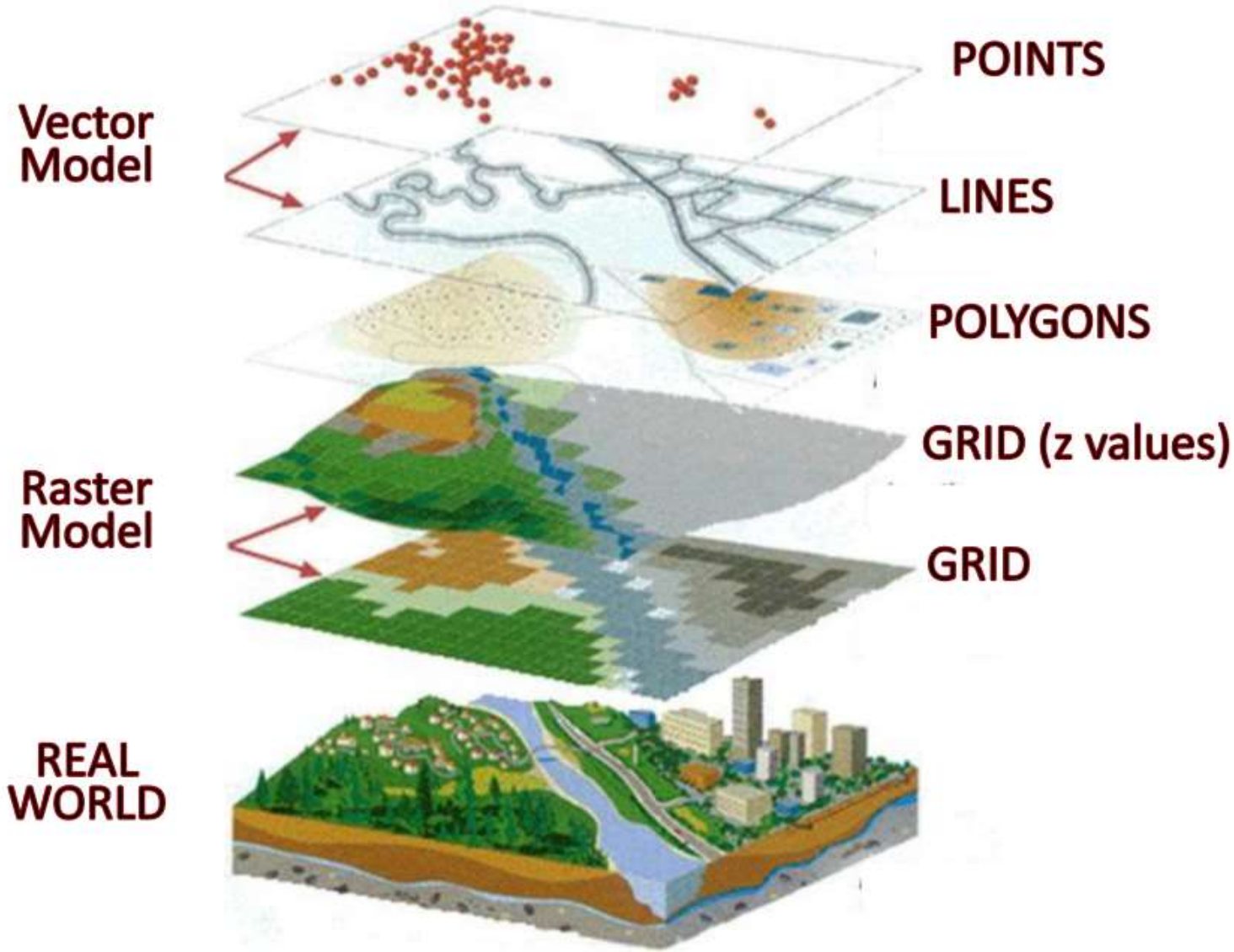
Facilitate the representation of **geographic features** in **real world locations** digitally so that they can be stored in a database.

Allows abstract representation in **map** form and can also be worked with and **manipulated** to address a **problem.**

Implementation

- Data is organized by layers or themes *with each layer representing a common feature.*
- Layers are integrated using explicit location on the earth's surface
- A spatial reference system becomes critical for ensuring accurate overlaying of layers




Vector Model
POINTS
LINES
POLYGONS
Raster Model
GRID (z values)
GRID
REAL WORLD

Types of geospatial data

- Cartographic/Spatial - observations on spatially distributed features:
  - Points
  - Lines
  - Areas (Polygons)
- Non-Cartographic/Attribute - descriptive information about the cartographic features:
  - Attributes
  - Attribute Values

Raster Models

Area is covered by grid with (usually) equal-sized, square cells

Attributes are recorded by assigning each cell a single value based on the majority feature (attribute) in the cell, such as land use type

Image data is a special case of raster data in which the "attribute" is a reflectance value from the geomagnetic spectrum. cells in image data often called *pixels* (picture elements)

Vector Models

All geographic features in the real world are represented either as:

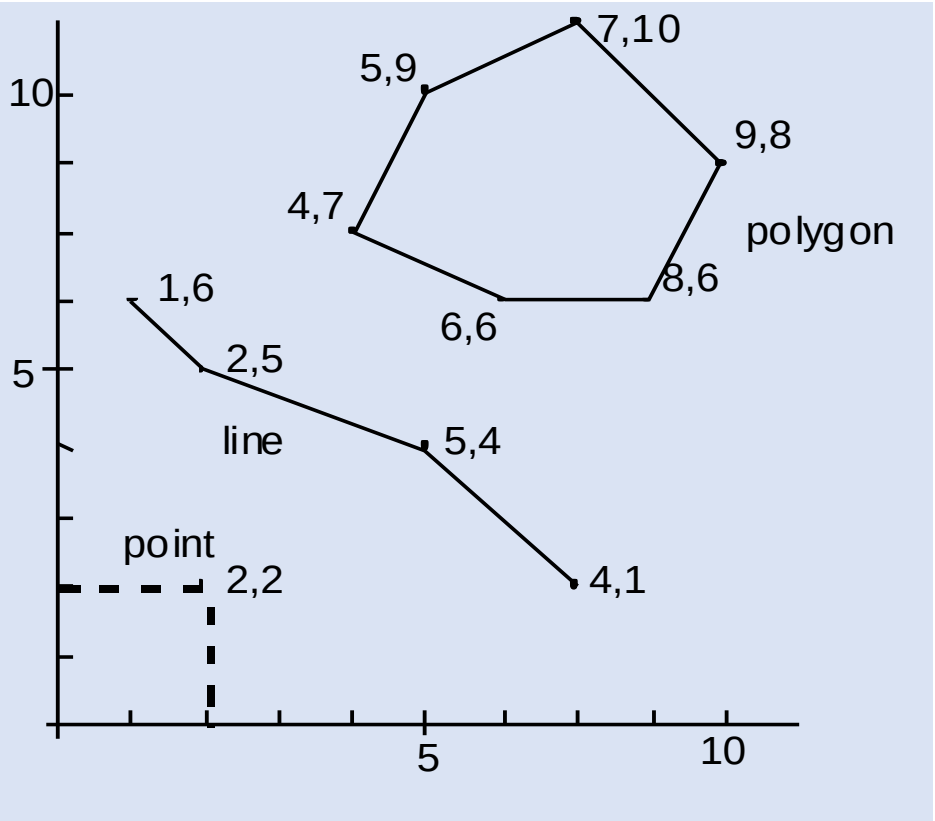
points or dots (*nodes*): trees, poles, fire plugs, airports, cities

lines (*arcs*): streams, streets, sewers,

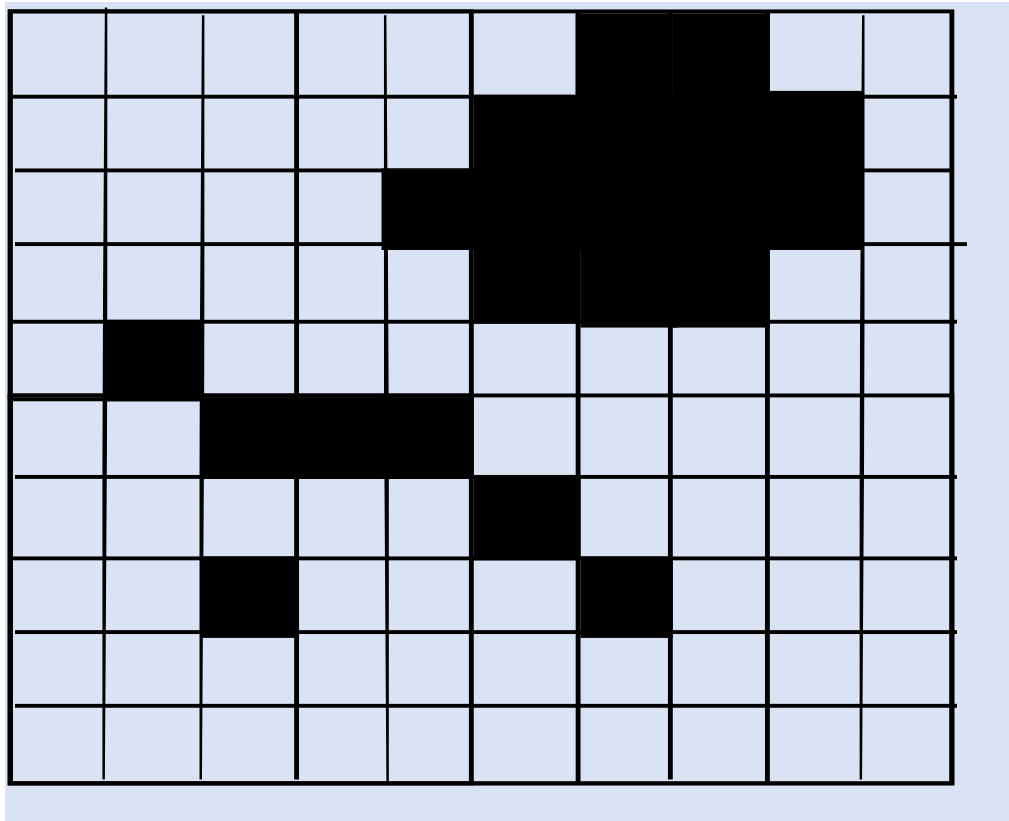
Areas (*polygons*): land parcels, cities, counties, forest, rock type

Because representation depends on shape, these files are commonly referred to as *shapefiles*

Spatial Data Structures

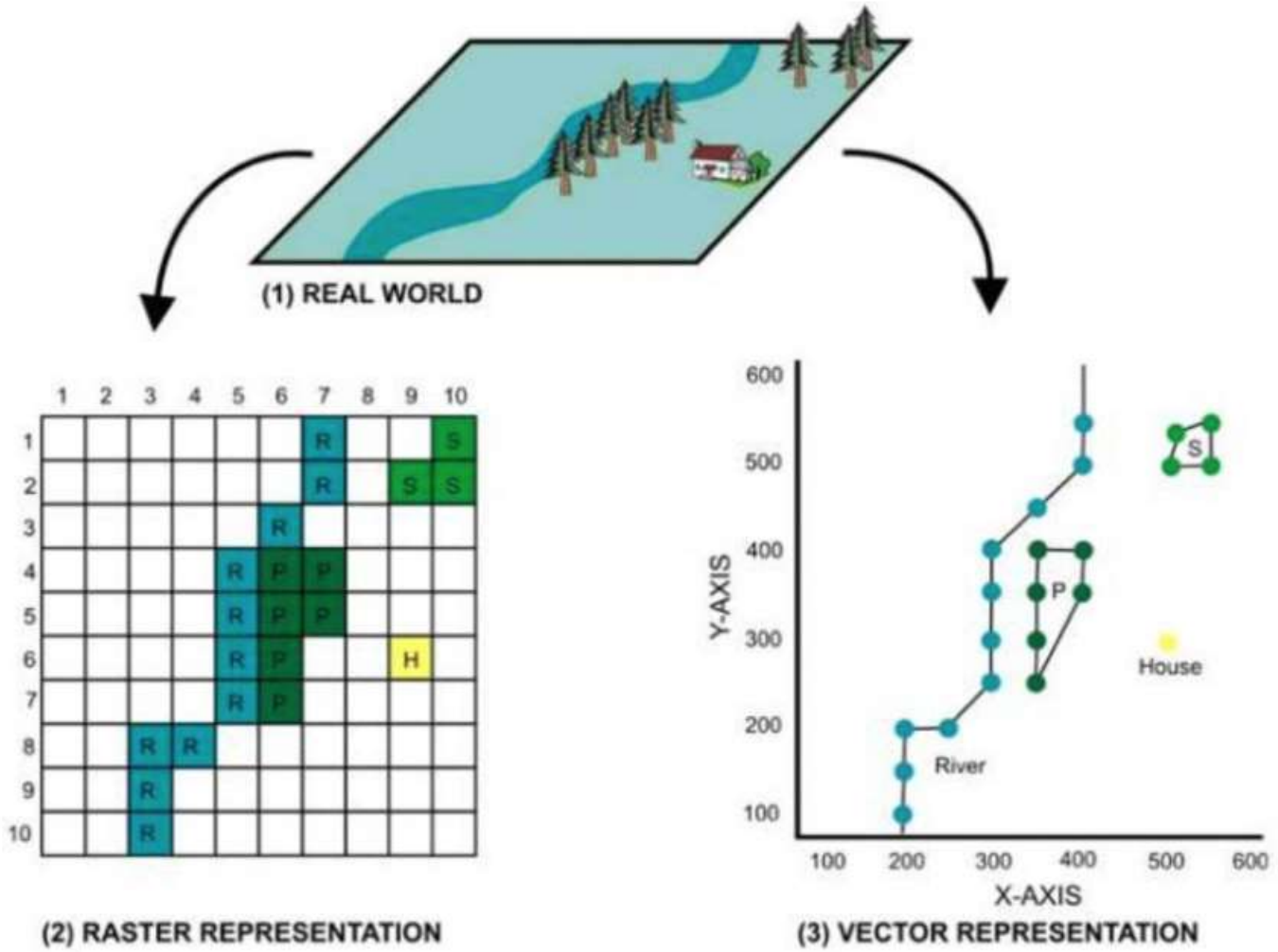


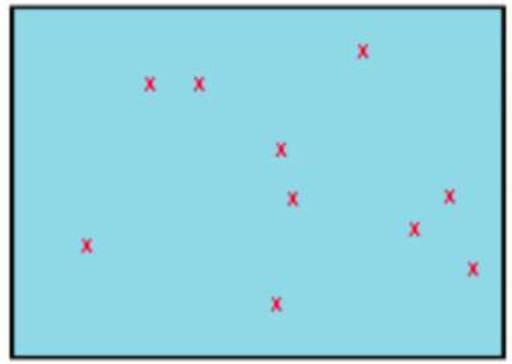
Vector Data: stored as geometric objects – points, lines, polygons



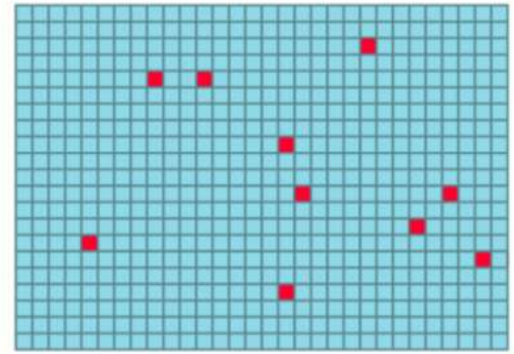
Raster Data: stored as image files composed of grid-cells (pixels)

Spatial Data Structures (cont)



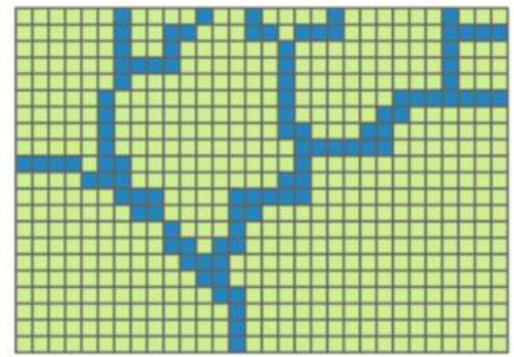


Vector Point Features

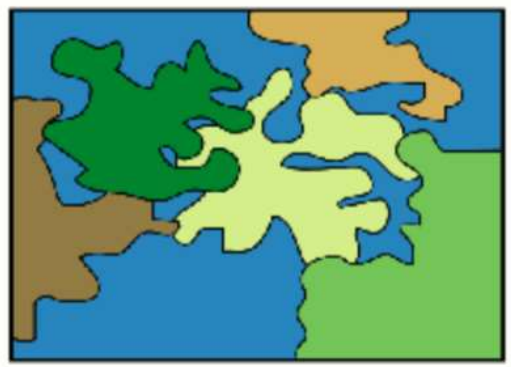


Raster Point Features

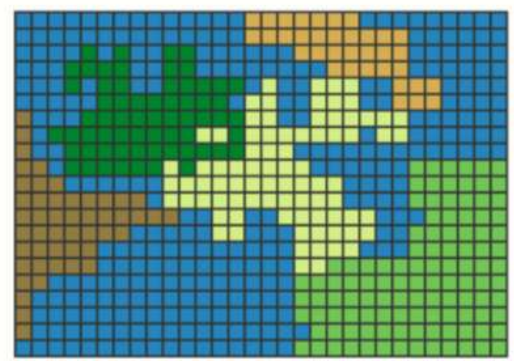
Vector Line Features



Raster Line Features



Vector Polygon Features



Raster Polygon Features

Vector Model

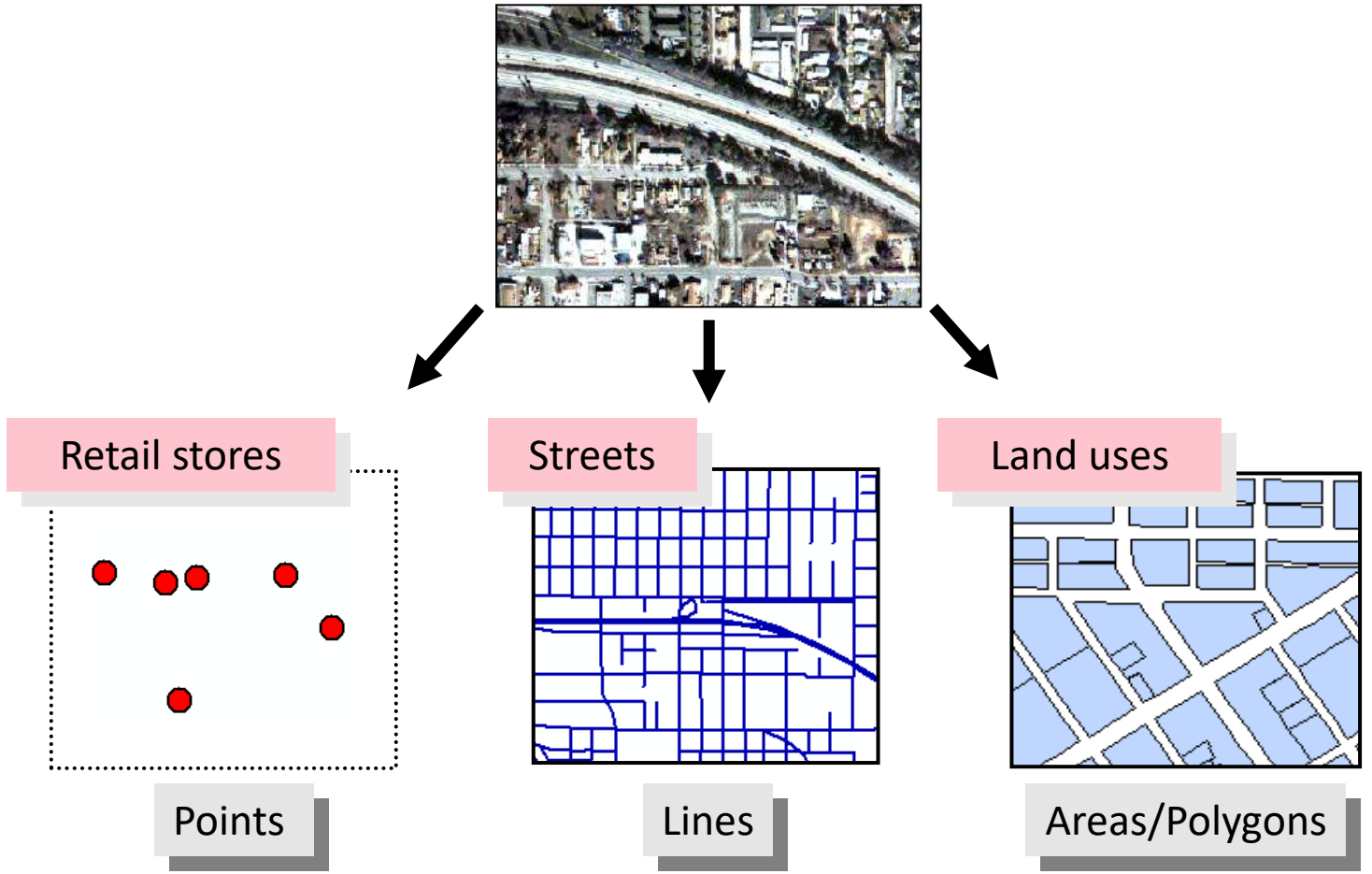
- The fundamental concept of vector GIS is that all geospatial features in the real world can be represented either as:
  - points (nodes): trees, poles, fire plugs, airports, cities
  - lines (arcs): streams, streets, sewers,
  - polygons (areas): land parcels, cities, counties, forest, rock type

- Vector model is best used to represent features with discrete boundaries (roads, buildings, lakes, rivers, administrative boundaries)

- This model tells "**where everything occurs**", i.e., it gives location to every object.

Representing features in vector data

- Real-world entities are abstracted into three basic shapes:
  - Points
  - Lines
  - Polygons



Impact of map scale

- Map scale determines the size and shape of features

1:500

Large scale

Smaller area

More detail

1:25000

1:25000

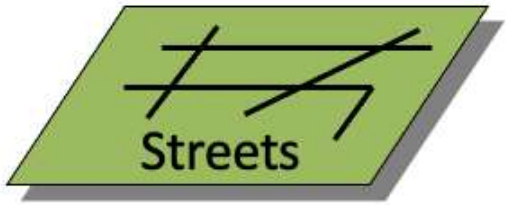
Small scale

Larger area

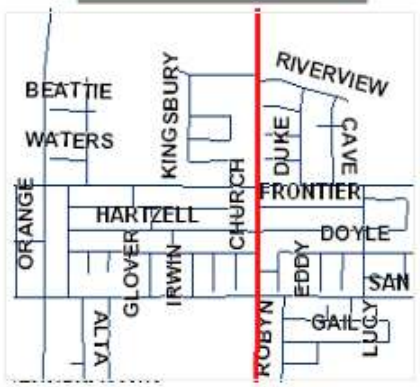
Less detail

1:250000

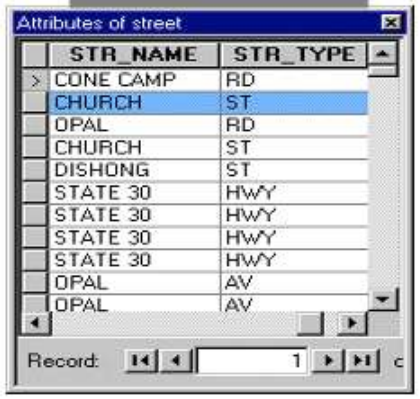
Components of vector data



Geometry



Attributes



Behavior

Rules:
Streets and highways may not intersect

- Each feature corresponds to a record in the attribute table

Vector Models

Advantage:

- analysis tasks that require accurate positioning,
- for defining spatial relationship (i.e. the connectivity and adjacency) between coverage features (topology),
- important for network analysis (for example to find an optimal path between two nodes in a complex transport network)

Disadvantage:

- boundaries of the resulting map polygons are discrete (enclosed by well-defined boundary lines), whereas in reality the map polygons may represent continuous gradation or gradual change, as in soil maps

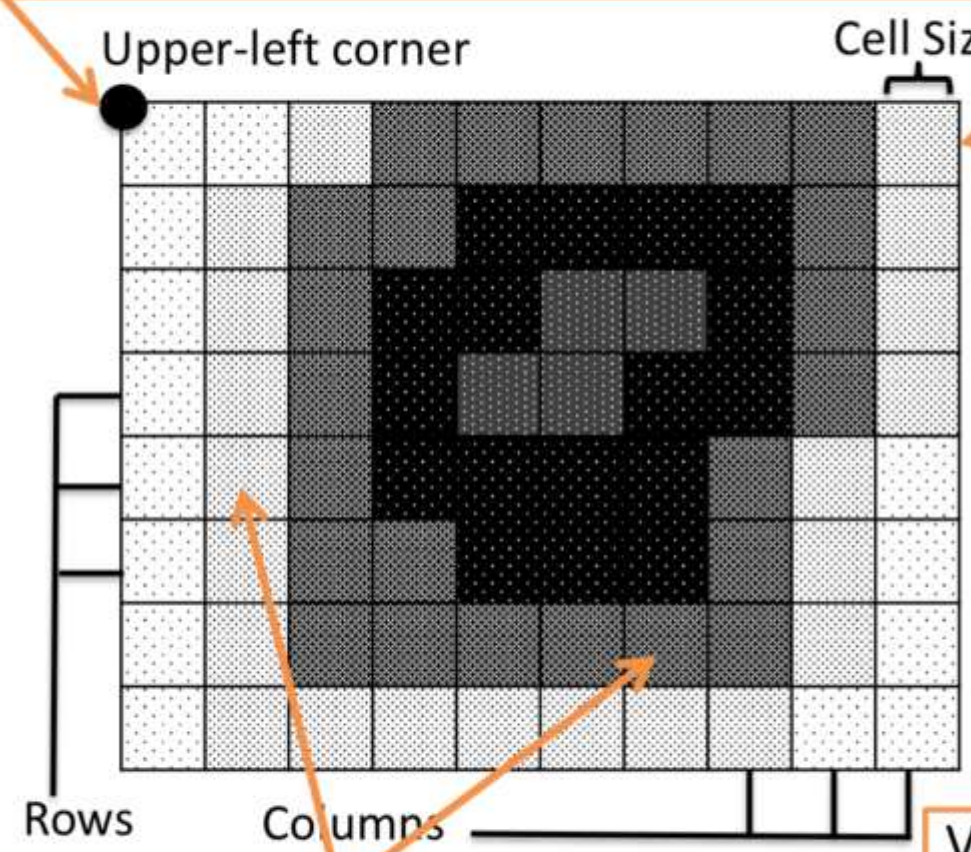
Raster Model

- Area is covered by grid with (usually) equal-sized, square cells organized in rows (horizontal) and columns (vertical)
- The model tells "**what occurs everywhere**" (space filling), i.e. everywhere has a value (even if that value is NoData)
- good for representing indistinct boundaries
  – thematic information on soil types, soil moisture, vegetation, ground temperatures
- *Image data* is a special case of raster data in which the "attribute" is a reflectance value from the electromagnetic spectrum

  - cells in image data often called pixels (picture elements)

A raster's position (spatial location) is defined by assigning x,y coordinates to one or two corners. The positions of the cells themselves flow from this reference point.

Spatial resolution of a raster is defined by its cell size (1m, … 20m,…, 50m, …, 100m, …)



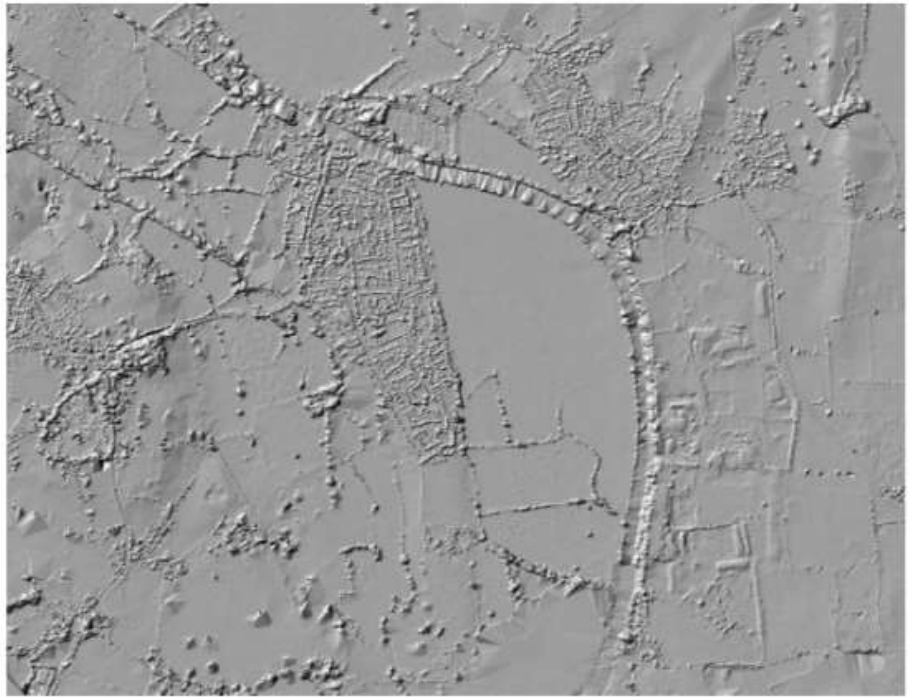
Value	Fire Spread
1	Day 1
2	Day 2
3	Day 3
4	Day 4
5	Day 5

Each cell in the raster contains a value. Values can be numeric: integer or decimal numbers.

Values can also be nominal or ordinal when something is not measured but classified into categories or themes.

Raster Spatial Resolution

- Size of grid cell used to represent the surface being modelled
- It relates to the size of the smallest feature that can be detected

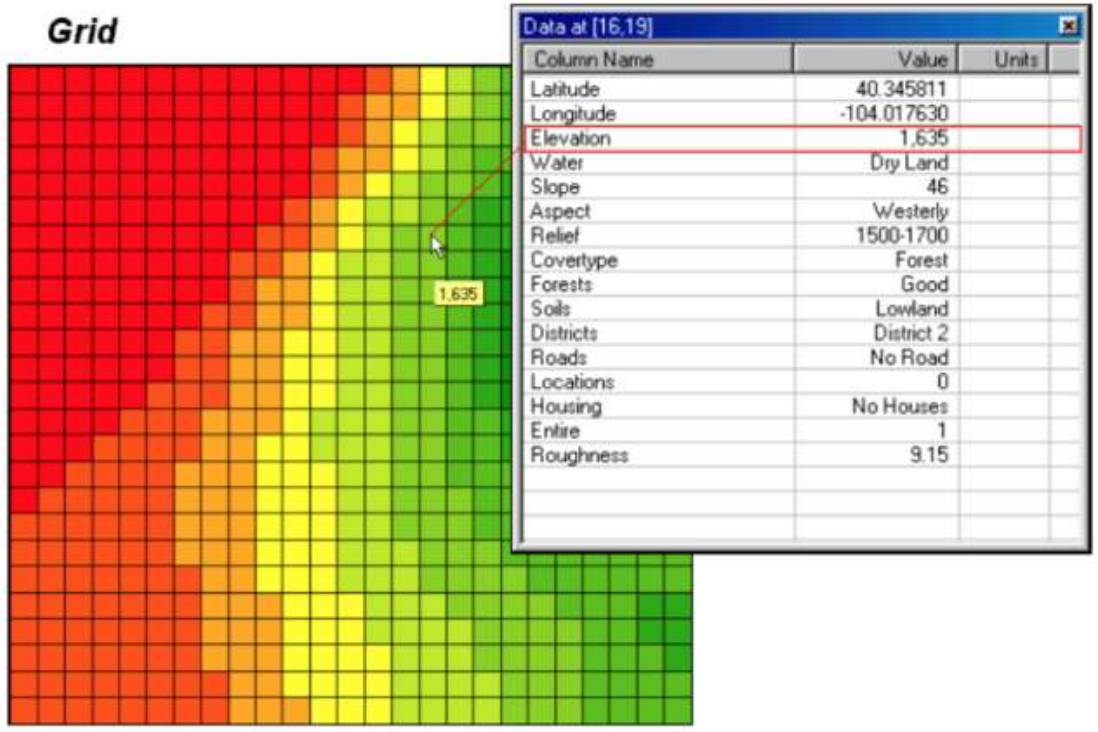


1 m Grid

10 m Grid

Components of raster data

- Attributes are recoded by assigning each cell a single value.



Raster models

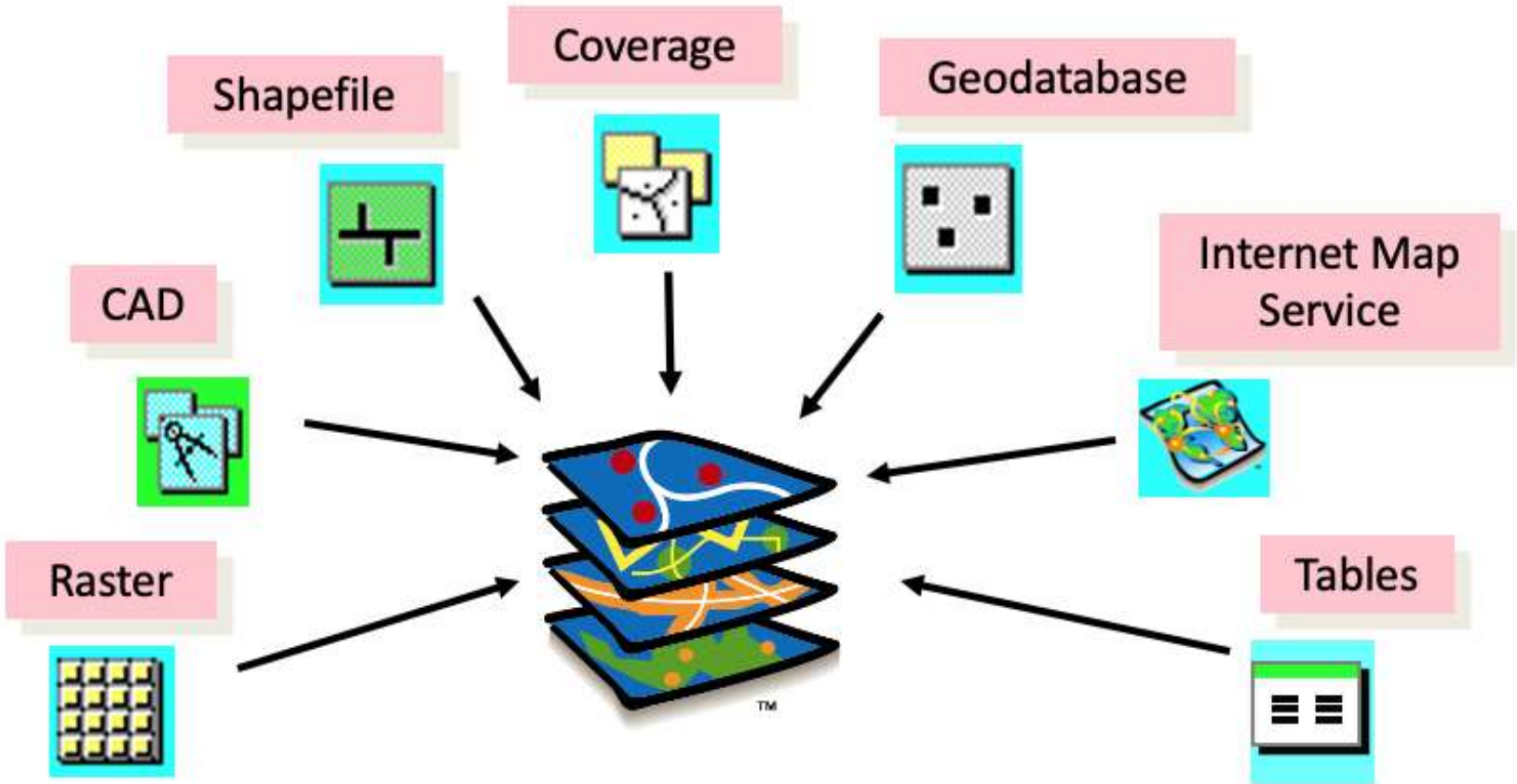
Advantages:

- good for representing indistinct boundaries such as thematic information on soil types, soil moisture, vegetation, ground temperatures
- as reconnaissance satellites and aerial surveys use raster-based scanners, the information (i.e. scanned images) can be directly incorporated into GIS

Disadvantage:

- the higher the grid resolution, the larger the data file is going to be
- generalisation
- loss of unique features

Spatial data formats



Data Modelling Process – Step 1

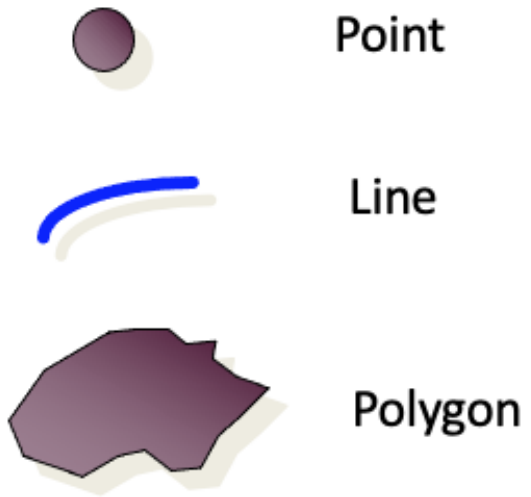
- Identify possible features from the real world (vector data for the example below)

- Buildings
- Road centrelines
- Lamp columns
- Gas pipes
- CTV Access covers
- Road surfaces

Data Modelling Process – Step 2

- Select a method for representing the feature in the GIS

  - Buildings
  - Road centrelines
  - Lamp columns
  - Gas pipes
  - CTV Access covers
  - Road surfaces



Data Modelling Process – Step 3

- Use the selected method to represent the feature in the GIS

Feature: Building
Shape: Polygon

Data Modelling Process – Step 4

- Identify appropriate attributes for the feature

Name :	Next Store
Address:	5 Market Place
Town:	Kingston
Owner:	Ms J Shore
Tel. No:	0181 547 1245
Floor space	1300 sq m

Projection, Scale, Accuracy and Resolution

the key properties of spatial data

- **Projection:** the method by which the curved 3-D surface of the earth is represented by X,Y coordinates on a 2-D flat map/screen
  - distortion is inevitable
- **Scale:** the ratio of distance on a map to the equivalent distance on the ground
  - in theory GIS is scale independent but in practice there is an implicit range of scales for data output in any project
- **Accuracy:** how well does the database info match the real world
  - *Positional*: how close are features to their real world location?
  - *Consistency*: do feature characteristics in database match those in real world
    - is a road in the database a road in the real world?
  - *Completeness*: are all real world instances of features present in the database?
    - Are all roads included.
- **Resolution:** the size of the smallest feature able to be recognized
  - for raster data, it is the *pixel* size

The tighter the specification, the higher the cost.

Metadata Standards

Objectives

- Introduce participants to the principles of metadata usage in GIS
- Familiarize participants with common metadata standards in GIS

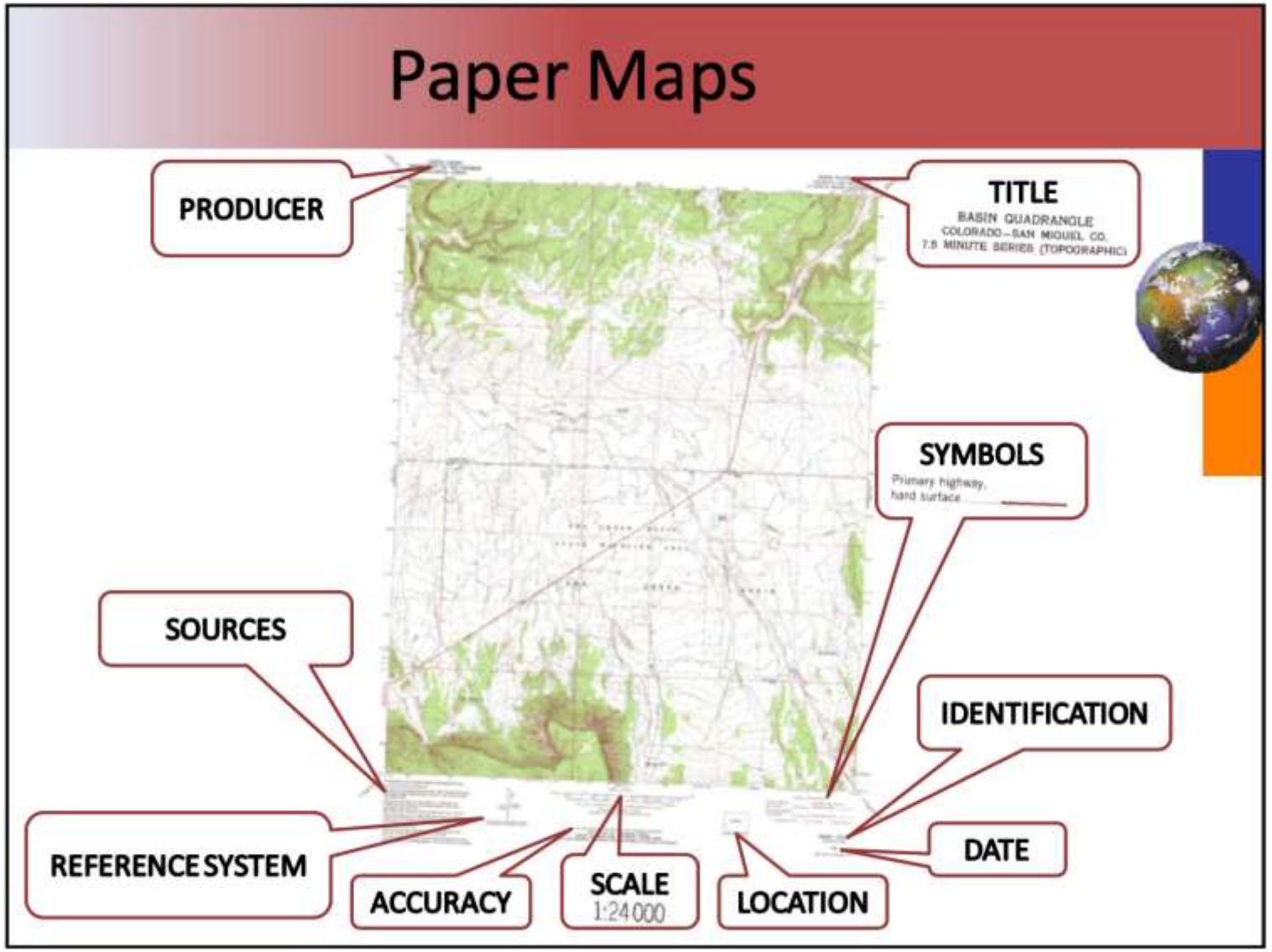
Outline

- What is Metadata
- Uses of Metadata
- Importance of Metadata
- Metadata Standards
- Metadata Elements

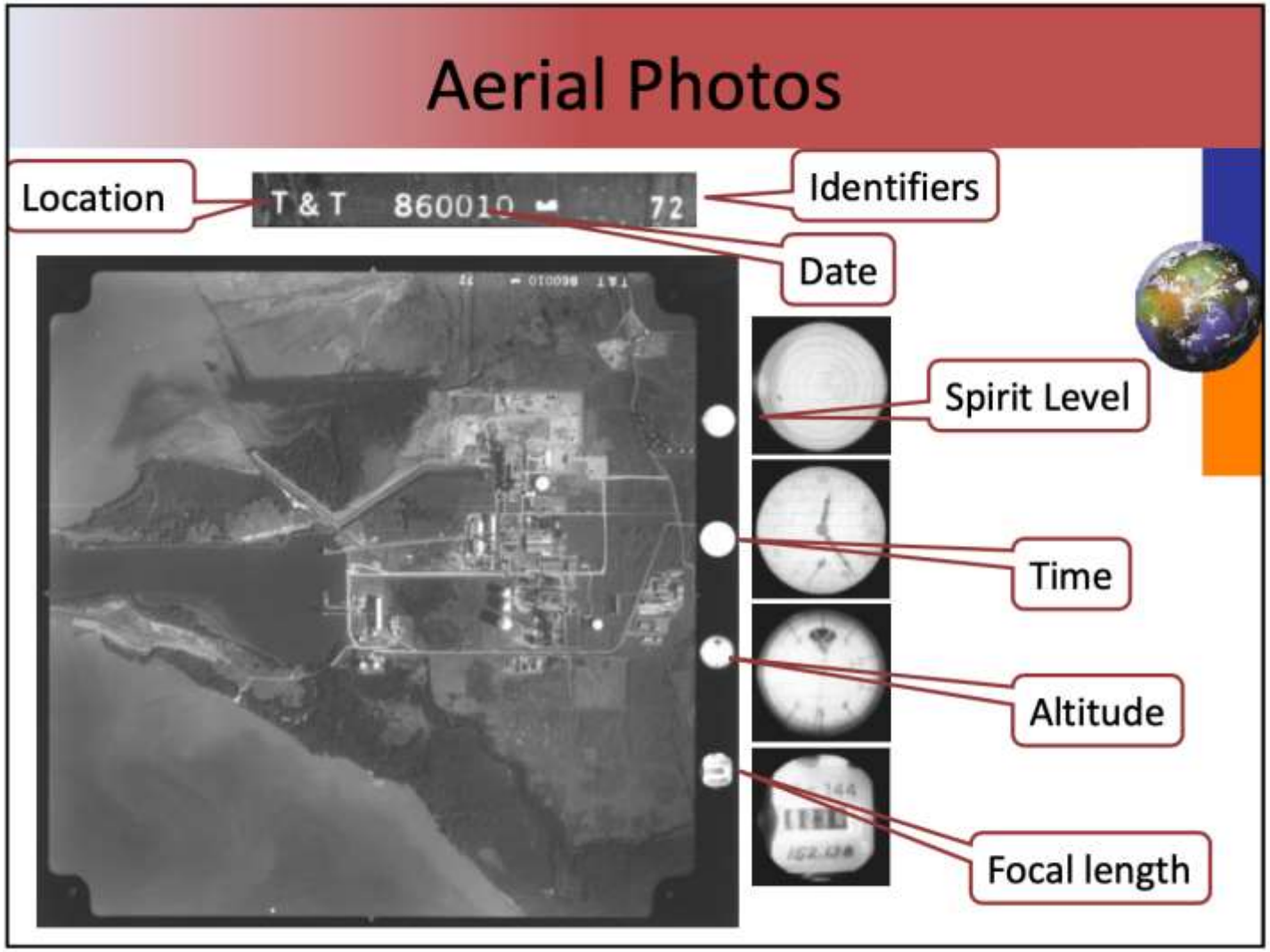
What is metadata?

- “Metadata provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference and the distribution of digital geographic data” – ISO 19115:2003
- Summary document that describes the content, quality, format, creation, source and other characteristics of data.
- Metadata answers who, what, where, when and how about every facet of the datasets that are being documented.
- **Metadata exists to provide potential users with the information which they require to make an informed decision about the adequacy of data for a particular purpose.**

Metadata on Paper Maps



Metadata on aerial images



Uses of metadata

Organise and maintain an organization’s data

Provide information to data catalogs and portals

Provide information to aid data transfer

Importance of metadata

1. Organize and maintain and organization’s investment in data.

   - As personnel change or time passes, information about an organization’s data will be lost and the data may lose their value. Later workers may have little understanding of the content and uses for a digital data base and may find that they can’t trust results generated from these data.

   - Complete metadata descriptions of the content and accuracy of a geospatial data set will encourage appropriate use of the data. Such descriptions also may provide some protection for the producing organization if conflicts arise over the misuse of data.

Importance of metadata

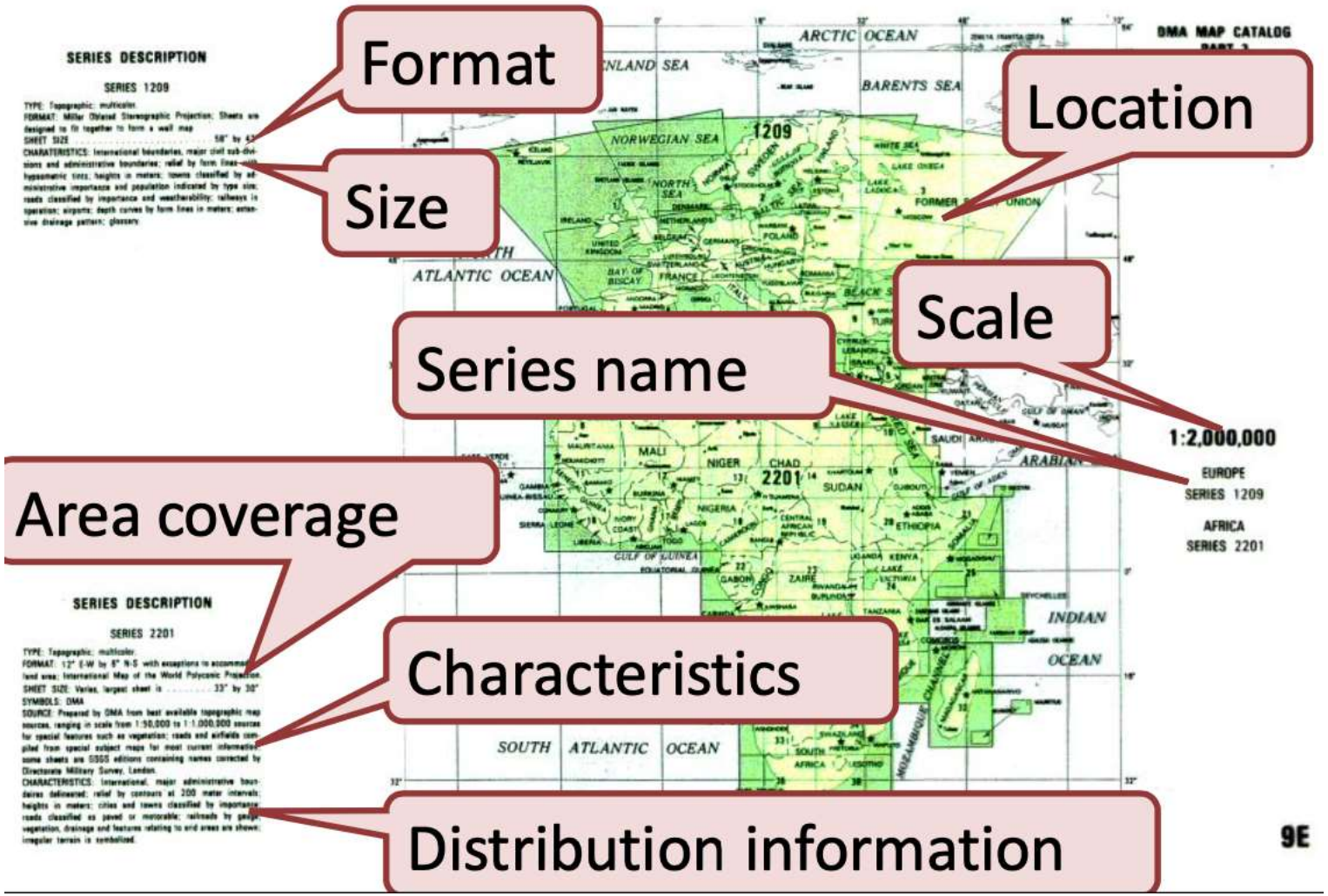
2. Provide information to data catalogs and portals.

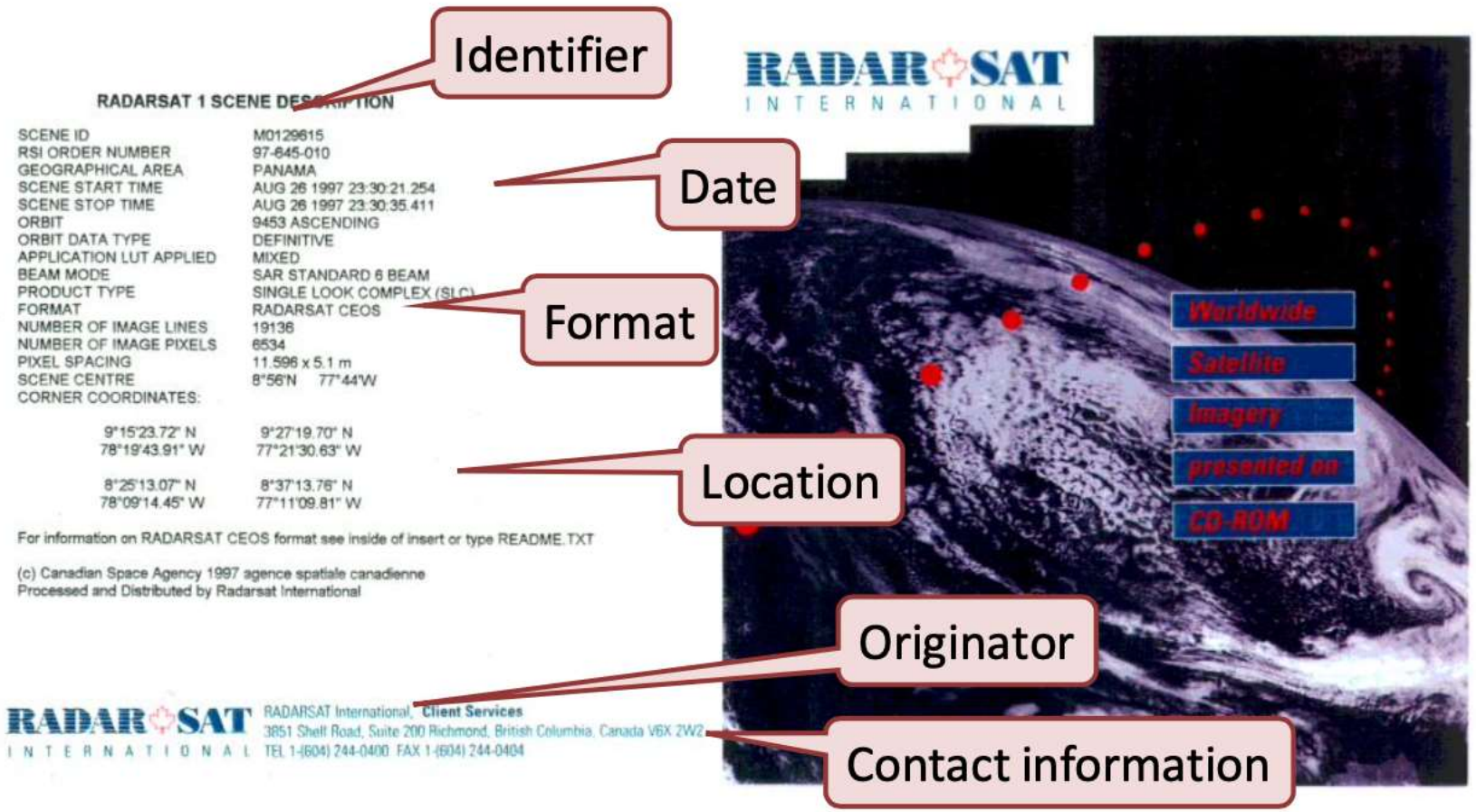
- Applications of geographic information systems often require many themes of data. Few organizations can afford to create all data they need. Often data created by an organization also may be useful to others.
- By making metadata available through data catalogs and clearinghouses, organizations can find data to use, partners to share data collection and maintenance efforts, and customers for their data

Importance of metadata

3. Provide information to aid data transfer

- Metadata should accompany the transfer of a data set. The metadata will aid the organization receiving the data to process and interpret data, incorporate data into its holdings, and update internal catalogs describing its data holdings.


Format
Size
Location
Scale
Series name
Area coverage
Characteristics
Distribution information
SERIES DESCRIPTION
SERIES 1209
SERIES DESCRIPTION
SERIES 2201
DMA MAP CATALOG
1:2,000,000
EUROPE
SERIES 1209
AFRICA
SERIES 2201
9E


RADARSAT 1 SCENE DESCRIPTION
SCENE ID M0129615
RSI ORDER NUMBER 97-645-010
GEOGRAPHICAL AREA PANAMA
SCENE START TIME AUG 26 1997 23:30:21.254
SCENE STOP TIME AUG 26 1997 23:30:35.411
ORBIT 9453 ASCENDING
ORBIT DATA TYPE DEFINITIVE
APPLICATION LUT APPLIED MIXED
BEAM MODE SAR STANDARD 6 BEAM
PRODUCT TYPE SINGLE LOOK COMPLEX (SLC)
FORMAT RADARSAT CEOS
NUMBER OF IMAGE LINES 19136
NUMBER OF IMAGE PIXELS 6534
PIXEL SPACING 11.596 x 5.1 m
SCENE CENTRE 8°56'N 77°44'W
CORNER COORDINATES:
9°15'23.72" N 9°27'19.70" N
78°19'43.91" W 77°21'30.63" W
8°25'13.07" N 8°37'13.76" N
78°09'14.45" W 77°11'09.81" W
For information on RADARSAT CEOS format see inside of insert or type README.TXT
(c) Canadian Space Agency 1997 agence spatiale canadienne
Processed and Distributed by Radarsat International
RADARSAT INTERNATIONAL
RADARSAT International, Client Services
3851 Shell Road, Suite 200 Richmond, British Columbia, Canada V6X 2W2
TEL 1-(604) 244-0400 FAX 1-(604) 244-0404
RADARSAT INTERNATIONAL
Worldwide
Satellite
Imagery
presented on
CD-ROM
Identifier
Date
Format
Location
Originator
Contact information

Metadata benefits

- Provides an inventory of data assets
- Helps determine and maintain the value of data
- Helps you determine the reliability and currency of data
- Supports decision making
- Documents legal issues
- Helps keep data accurate and helps verify accuracy to support good decision making and cost savings
- Helps determine budgets because it provides a clearer understanding of when or if data needs to be updated or repurchased

Metadata standards

- Official standards organizations define metadata standards.
- By adhering to common metadata standards, organizations are more able to share data.
- An important standard in the United States is the FGDC Content Standard for Digital Geospatial Metadata.
- The International Organization for Standardization (ISO) has also created a spatial metadata standard

Metadata standards

Organisation	Standard	Clearing house	About the Standard
Federal Geographic Data Committee (FGDC), *USA*	Content Standard for Digital Geospatial Metadata (CSDGM)		Provides a common set of terminology and definitions for the documentation of digital geo-spatial data. All US federal agencies that produce geospatial data are required to use the CSDGM standard.
European Committee for Standardization (CEN) and Information Society Standardization System (ISSS), *International group*	Dublin Core		A metadata element set that is designed to describe digital resources. It supplements existing methods for searching and indexing electronic resources on the internet. It is much simpler than other standards. This standard is to be mapped to the ISO 19115 and the FGDC standard.
International Organisation for Standardization (ISO)	ISO 19115 International Metadata Standard		This standard was adopted in March 2003. It defines the schema required for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference and the distribution

Metadata standards

Organisation	Standard	Clearing house	About the Standard
			of digital geographic data.
Association of Geographic Information (AGI), *UK*	Discovery Metadata Specifications (previously known as the National Geospatial Data Framework	GIgateway	
Australia New Zealand Land Information Council (ANZLIC)	ANZLIC Metadata guidelines		The standard consists of 41 core elements grouped into ten categories. This standard is consistent with the CSDGM and ISO 19115.

Metadata profile

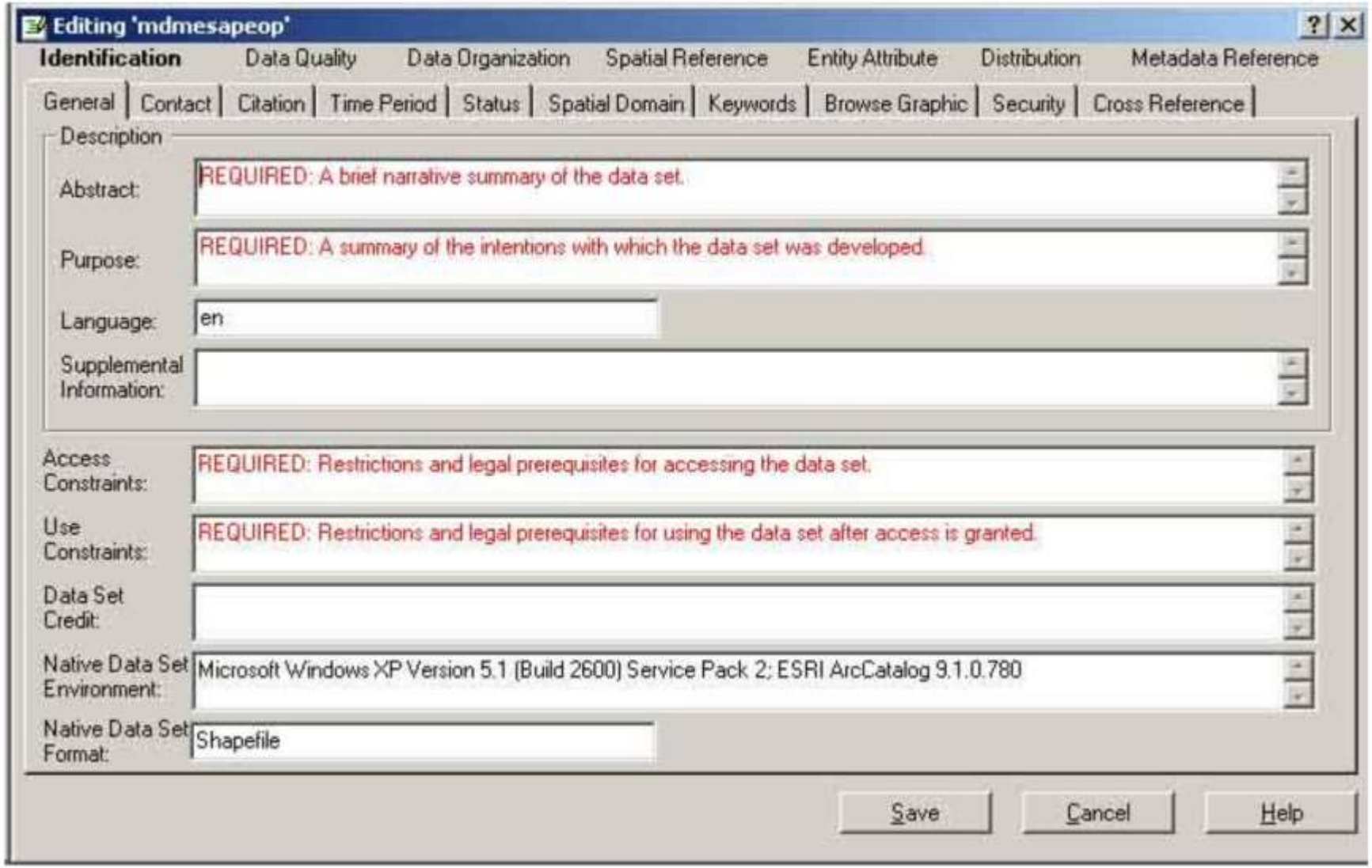
- A metadata profile is a document that modifies a metadata standard.
- A profile may reduce the overall number of metadata elements defined by a standard.
- A profile may further restrict the optionality of a metadata element, making it mandatory where before it was optional; however, a profile cannot make mandatory elements optional.
- Metadata profiles can be adopted by a standards body, agency, or organization in place of a metadata standard.
- One example of a metadata profile is the North American Profile of ISO 19115:2003 that has been jointly created and adopted by the United States and Canada.

CORE META-DATA ELEMENTS CATEGORIES	DEFINITIONS OF DATA ELEMENTS CATEGORIES
1. IDENTIFICATION	basic information about the data set. Examples include the title, the geographic area covered, currency, and rules for acquiring or using the data.
2. DATA QUALITY	- an assessment of the quality of the data set. Examples include the positional and attribute accuracy, completeness, consistency, the sources of information, and methods used to produce the data.
3. SPATIAL DATA ORGANISATION	the mechanism used to represent spatial information in the data set. Examples include the method used to represent spatial positions directly (such as raster or vector) and indirectly (such as street addresses or county codes) and the number of spatial objects in the data set.
4. SPATIAL REFERENCE	description of the reference frame for, and means of encoding, coordinates in the data set. Examples include the name of and parameters for map projections or grid coordinate systems, horizontal and vertical datums, and the coordinate system resolution.
5. ENTITY AND ATTRIBUTE	information about the content of the data set, including the entities types and their attributes and the domains from which attribute values may be assigned. Examples include the names and definitions of features, attributes, and attribute values.
6. DISTRIBUTION	- information about obtaining the data set. Examples include a contact for the distributor, available formats, information about how to obtain data sets online or on physical media (such as cartridge tape or CD-ROM), and fees for the data
7. METADATA REFERENCE	information on the currency of the metadata information and the responsible party. Examples include currency and information about the organization that provided the metadata.

Metadata elements

CATEGORY	ELEMENT		DESCRIPTION OF ELEMENT	OBLIGATION
Data Quality			a general assessment of the quality of the data set	
	Attribute Accuracy		an assessment of the accuracy of the identification of entities and assignment of attribute values in the data set	**M**
	Logical Consistency Report		an explanation of the fidelity of relationships in the data set and tests used.	**O**
	Completeness Report		information about omissions, selection criteria, generalization, definitions used, and other rules used to derive the data set	**M**
	Positional Accuracy		an assessment of the accuracy of the positions of spatial objects. horizontal and vertical coordinate measurements and a description of the tests used	**M**
	Lineage		Information about the events, parameters and source data which constructed the data set and information about the responsible parties.	**M**
		Source	Information about the source data used in creating the data specified	**M**
		Process Steps	Information about tan event or transformation in the life of a dataset including the process used to maintain the dataset.	**O**
	Cloud Cover		area of a data set obstructed by clouds, expressed as a percentage of the spatial extent.	**O**

Metadata editor



- In United States, FGDC (**Federal Geographic Data Committee**) Content Standard for Digital Geospatial Metadata
  - Published in 1998.
- The International Organization for Standardization has also created a spatial metadata standard-- ISO

What made all of these possible?

Access to

- **Technology (currency)**
- **Data (currency, resolution, accuracy)**
- **Policy (relevance, monitoring)**
- **Human resources (skills and benefits)**
- **Funding (driven by applications and results)**
- **Political support (motivation for success)**

Thank you!

- ----------END----------