GIST 7128
ArcGIS 1: Introduction

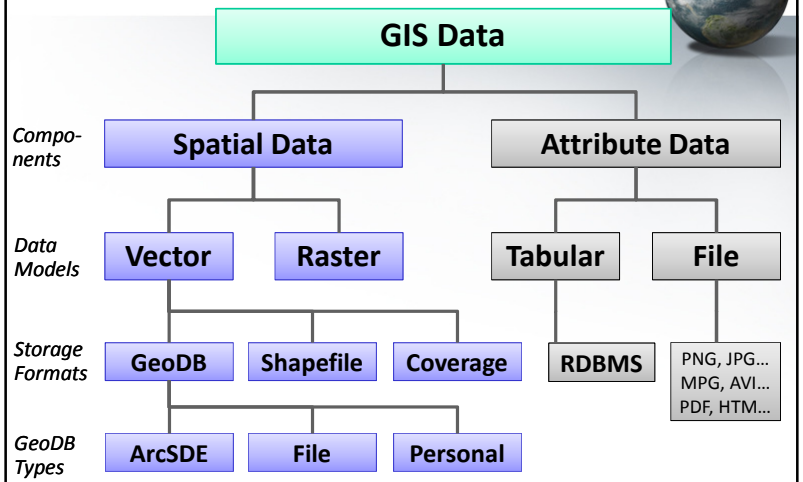
Lecture 5
Geodatabases and Features

Module 5 Topics

- **Lecture**
  1. Geographic Data Models
  2. ArcGIS Data Storage
  3. Topology
  4. Raster Data
  5. Data Input Techniques
- **Lab**
  - Chapter **11**. Building Geodatabases
  - Chapter **12**. Creating Features
- **Project**
  - EastCity – Part 3: East City Maps (*continued*)
- **Next Module**
  - MidTerm Exam: 30 questions, 15 marks (to chap 12 & lec 5)

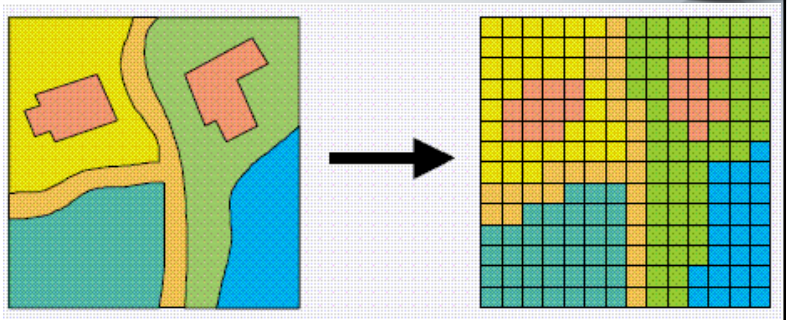
2

1. Geographic Data: Overview



3

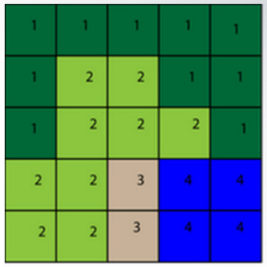
Data Models: Raster & Vector



Spatial data example for Vector & Raster models

4

Data Models: Raster & Vector



Values	Name	Count
1	Forest	10
2	Grass	9
3	Beach	2
4	Water	4

FID	Name	Area	Public	Owner
1	Water	8 ha	Yes	BC
2	Beach	4 ha	Yes	BC
3	Grass	18 ha	Yes	BC
4	Forest	20 ha	No	Jones

Spatial and attribute data for Raster & Vector models

5

2. ArcGIS Data Storage

- **Coverages** (being phased out)
  - INFO database for spatial data
  - Oracle, Informix, Access, etc. for other database tables
- **Shapefile**
  - Single Feature Class
  - Point, line, or polygon features (no annotation)
  - dBase IV (.dbf) file format (1980's)
  - Common interchange format with other GIS packages
- **Geodatabase**
  - Multiple Feature Classes
  - Can store annotation
  - Includes attribute domains (i.e. set or range of valid values)
  - Multiuser, File, and Personal geodatabase types
  - Use Oracle, DB2, SQL Server, Folder, or Access to store *all* data
  - *More details on next 3 slides* →

6

Shapefile Data Storage

- Very popular GIS data exchange file format
- contains one feature class, single geometry type
- does not store topology – calculated 'on the fly'

Shapefile consists of a minimum of three files which must be in the same directory: .shp, .shx, .dbf

- **.shp** — actual feature geometry
- **.dbf** — attribute data (in dBase IV format, 1980s)
- **.shx** — geometry index

7

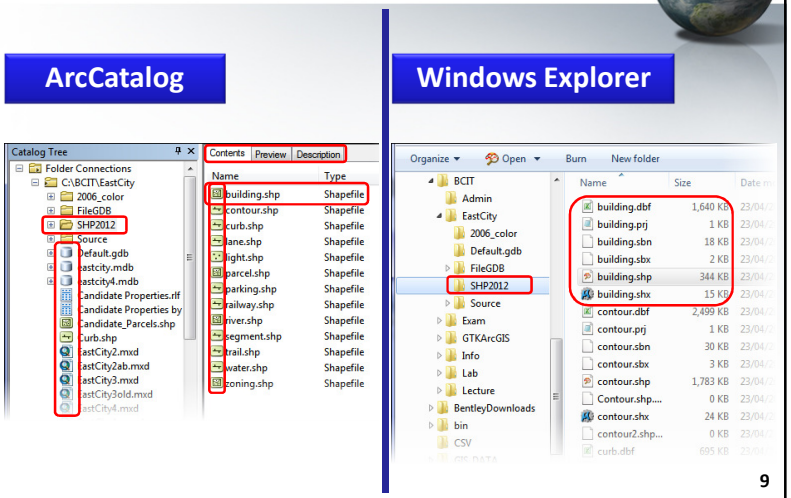
Shapefile Data Storage

Optional related files:

- **.prj** — coordinate system and projection information, text file
- **.shp.xml** — geospatial metadata in XML format
- **.sbn** and **.sbx** — spatial index of the features
- **.fbn** and **.fbx** — spatial index for read-only shapefiles
- **.ain** and **.aih** — attribute index of active fields in a table
- **.ixs** — geocoding index for read-write shapefiles
- **.mxs** — geocoding index for read-write shapefiles (ODB format)
- **.atx** — attribute index for .dbf fields: *shp_name.col_name*.atx
- **.cpg** — specify the character encoding (for .dbf)

8

Shapefile Views (Review - week 1)



9

ArcGIS Geodatabase Types

	ArcSDE Geodatabase	File Geodatabase	Personal Geodatabase
A.K.A.	Multiuser geodatabase Enterprise geodatabase	n/a	Access geodatabase
Storage	Tables in a relational database: Oracle, Microsoft SQL Server, IBM DB2, IBM Informix, PostgreSQL	Folder containing a binary file for each dataset; folder name ends in ".gdb"	All data in a single Microsoft Access database (.mdb) file
Recommended for	Large organizations with resources to support an enterprise relational database management system	Medium and small organizations that do not have resources to support an enterprise relational database management system	Small projects

10

ArcGIS Geodatabase Types

	ArcSDE Geodatabase	File Geodatabase	Personal Geodatabase
Number of users	100s Many readers Many writers	10s Single user and small workgroups: several readers or one writer *per feature class*	1s Single user and small workgroups with small datasets: few readers and one writer *per GeoDB*
Size limits	Several TBs Up to DBMS limits	1 TB for each dataset in geodatabase; can be raised to 256 TB for extremely large image datasets	2 GB but effective limit before performance degrades is approx. 500 MB
Platforms	Windows, UNIX, Linux and direct connections to DBMSs on other platforms	Windows, UNIX, Linux	Windows only

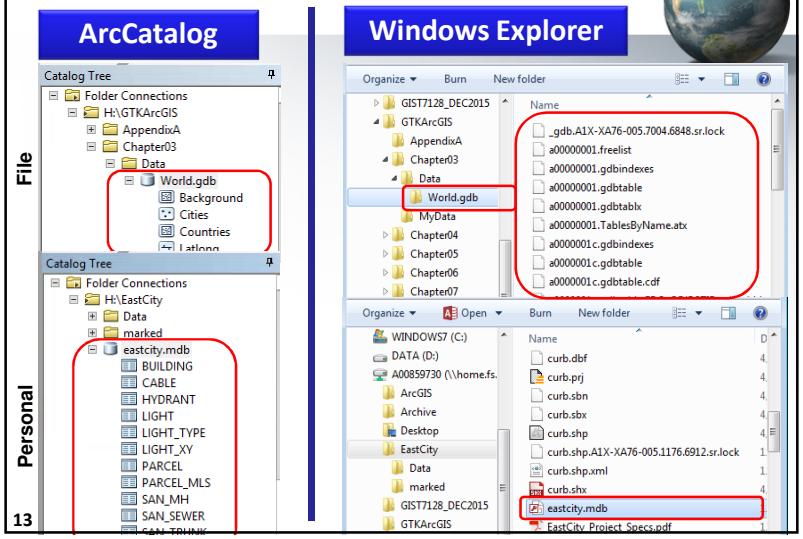
11

ArcGIS Geodatabase Types

	ArcSDE Geodatabase	File Geodatabase	Personal Geodatabase
Security	Provided by DBMS	Operating system file security	Windows file security
Administration tools	Full DBMS functions for backup, recovery, replication, SQL support, security, and so on	File system management	Windows file system management
Notes	Requires ArcSDE technology; included for SQL Server Express; Includes versioning support	Optionally store data in read-only compressed format to reduce size	Often used for attribute tables

12

Geodatabase Views



13

Geographic Data – Components (week 1)

Component	Description	Example
Spatial Data a.k.a. geography, map data, graphics	Position/location Shape Size	Fire Hydrant is point feature located at: • X = 654,321mE • Y = 5,234,567mN
Attributes a.k.a. table, properties, fields	Characteristics Textual information Facts describing the feature	• Capacity = 25 gal/min • Manufact = ACME • ValveID = V02468 • MainID = W13579 • Installed = 4 Dec 1954 • LastTest = 9 Jul 2009
Topology explicit (stored) or implicit (calculated dynamically)	Relationships with other spatial data (adjacent, contained, overlapping, touching, coincident, near, intersecting, etc.)	• At corner of Ontario St and East 4th Ave • Closest building is 24 East 4th Ave • In Mount Pleasant ward

14

Spatial Component

- Geographic data can be represented as vector points, lines, and polygons, and as rasters (later)
- **Point**
  - Discrete location showing object too small to be shown as line or polygon, or an imaginary point (e.g. a mountain peak), or label
  - One coordinate (x,y)
- **Line**
  - Series of connected coordinates that represents the shape of a linear object too small to be shown as a polygon, or an imaginary line (contour)
  - Start and end points (nodes) are significant in topology; others not
  - Mínimum of 2 coordinates (x1,y1, x2,y2)
- **Polygon**
  - Closed homogenous area
  - Typically containing a point feature (centroid/label) to identify
  - Mínimum of 3 coordinates (x1,y1, x2,y2, x3,y3)

15

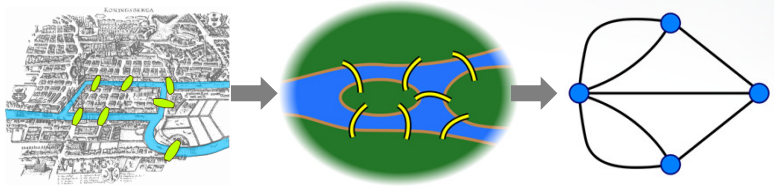
Attributes Component

- Descriptive data about a feature
- Two types of attributes:
  1. Spatial-related values
     - Line: Length, Endpoints coordinates, Direction
     - Polygon: Area, Perimeter, Centroid coordinates
  2. Non-spatial values
     - *For example, attributes for a road could be:*
     - Jurisdiction, Name
     - Surface Type, Number of Lanes
     - Construction, Maintenance Date
- Linking Features to Attributes
  - Every feature has unique internal ID used to link spatial data with attribute data.

16

3. Topology Component

- Topology defines the spatial relationships between (adjacent) features
- Seven Bridges of Königsberg
  - Mathematical problem in 1735 laid the foundations of graph theory and prefigured the idea of topology
  - Cross all bridges exactly once on a single trip

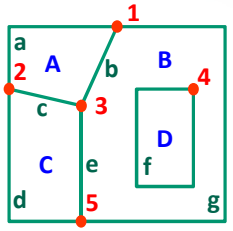


- Abstraction

17

What is Topology?

- Spatial relationship between connecting or adjacent geographic features
- *Examples:*
  - street centerlines and census block polygons share common geometry
  - adjacent parcels share common boundaries



1 – Node (or Vertex)

a – Edge (or Arc)

A – Face (or Polygon)

OBSERVATIONS

- polygons A & B share edge b and nodes 1 and 3
- edges c & e share node 3
- polygons B & D share edge f and node 4
- edges d, e, g share node 5

18

Topology

- *More Examples:*
  - adjacent parcels share common boundaries
  - street centerlines & census blocks have common geometry
- Enables user to perform analysis, such as:
  - find adjacent features
  - remove coincident boundaries
  - dissolve polygons with equal attribute values
  - traverse linear networks
- Answer questions like:
  - Is the park next to the lake?
  - Is Broadway connected to Granville?
  - How many hospitals are located in Vancouver?

19

ArcGIS Topology

- Two types of topology
  - Map Topology (temporary)
  - Geodatabase Topology (persistent)
- ArcGIS includes topology layers in ArcMap to display topological relationships, errors, and exceptions
- ArcMap also includes a rich set of tools to query, edit, validate, and repair/correct topologies
- ArcToolbox includes a comprehensive set of geoprocessing tools to build, analyze, manage, and validate topologies.

20

Map Topology

- Simple topology during an edit session
- Created for the current display extent only
- Edit features that overlap or touch each other
  - Edit coincident nodes and edges
  - Use tools from Edit and Topology toolbar
- Features can be in one or more feature classes
  - e.g. road and bridge
- Have different geometry
  - e.g. road (line) and bridge (points)
- ArcView (Basic) license required

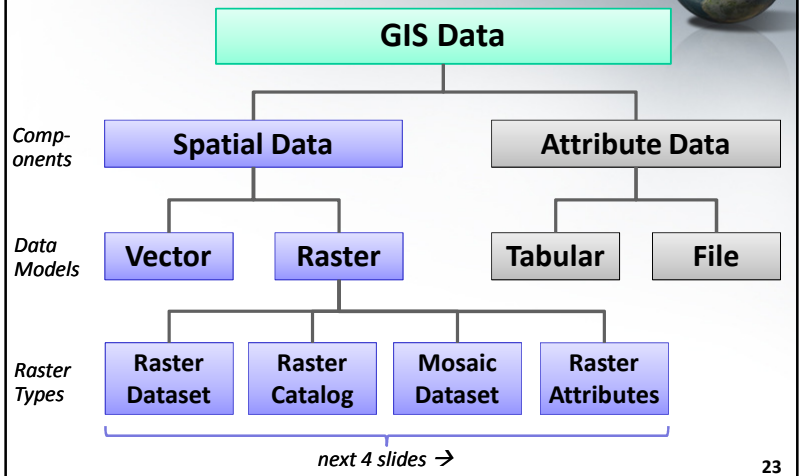
21

Geodatabase Topology

- Models the behavior of feature types
  - Collection of rules and tools
  - User defines relationships between feature classes
  - More accurately model geometric relationships in reality
- Topology defines and enforces data integrity rules
  - *Examples:*
  - Polygons must not overlap or have gaps
  - Lines must not intersect
  - Points must be covered by line
- Created and managed in ArcCatalog
- ArcEditor (Standard) or ArcInfo (Advanced) license

22

4. Raster Data: Overview



23

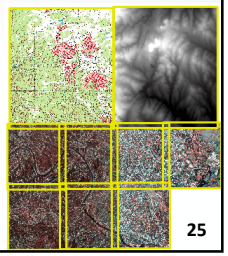
Raster Dataset

- Covers a spatially continuous area
  - single original image
  - many images joined (mosaicked) into a single image
  - if overlapping, *only one set of data is retained*
- Homogeneous data: single format, data type, & file
- Any valid raster format (50+ supported formats)
- Organized into one or more bands
  - Each band consists of a "layer" of pixels (cells)
  - Each pixel has a value.
- Many geoprocessing tools available
- Saved as files on disk, or within a geodatabase

24

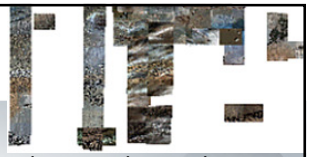
Raster Catalog

- Collection of raster datasets
  - defined by records in a database table, each of which represents an individual raster dataset, including extents
  - displayed as a single layer
- Heterogeneous data:
  - multiple formats, data types, file sizes, & coordinate systems
  - dynamically reprojected when displayed if different projections
- Typically used to display multiple raster datasets
  - no need to mosaic multiple images into one large raster dataset
  - *overlapping areas are preserved*
- Useful for:
  - massive image repositories
  - tiled raster layers
  - series of rasters over time
  - any related assortment of rasters



25

Mosaic Dataset



- Collection of raster datasets
- Stored as a raster catalog, *but newer and more advanced:*
  - faster to display at any scale
  - use in Image Analysis window
  - create mosaic image without data loss
  - input to geoprocessing and analysis tools
  - publish as an Image Service via ArcGIS Server
  - generate an overview for entire mosaic dataset
- Useful for:
  - managing, visualizing, querying raster data
  - can be unconnected datasets (not adjoining or overlapping)
  - can be captured over different dates
    - query on date/time and display results

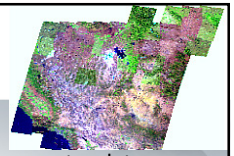
26

Raster Attributes

- Rasters can also be an attribute of a feature in a feature class
- A field of type raster can be included within a geodatabase feature classes (only)
- Similar to having a hyperlink of a file-based image in a field, except the raster image is stored and managed within the geodatabase
- Examples:
  - photograph of a property and building as an attribute of a parcel feature
  - scanned construction drawing as part of an asset record (bridge, building, manhole, well, etc)

27

ArcGIS Raster Tools



Built-in Tools

- **Image Analysis Window** can apply functions to your raster data or mosaic dataset to perform this processing on-the-fly
- Or you can use the **geoprocessing tools** and models to generate secondary products from your data

Extensions

- **ArcGIS Spatial Analyst**—powerful geoprocessing tools to perform raster-based modeling and analysis
- **ArcGIS 3D Analyst**—work with rasters as surfaces and other 3D data representations
- **ArcGIS Geostatistical Analyst**—advanced statistical tools for surface generation and for analyzing and mapping continuous datasets
- **ArcScan for ArcGIS**—raster-to-vector conversion of scanned documents, including raster editing, raster snapping, manual raster tracing, and batch vectorization

28

Raster Formats

- Over 50 raster formats supported by ArcGIS
- Most formats are read-only, about 20% are read/write*
- Some popular examples (especially **bold** ones):
  - .bil—ESRI Band Interleaved by Line*
  - .bip—ESRI Band Interleaved by Pixel*
  - .bsq—ESRI Band Sequential*
  - .bmp/.bmw—Windows Bitmap*
  - .dem—Digital Elevation Model
  - **.ecw—Enhanced Compression Wavelet**
  - .gif/.gfw—Graphics Interchange Format*
  - .img/.igw—ERDAS Imagine*
  - .jpg/.jgw—Joint Photographic Experts Group*
  - **.jp2—JPEG 2000***
  - **.sid/.sdw—Multi-resolution Seamless Image Database (MrSID)**
  - .png—Portable Network Graphics*
  - **.tif/.tfw—Tagged Image File Format***

29

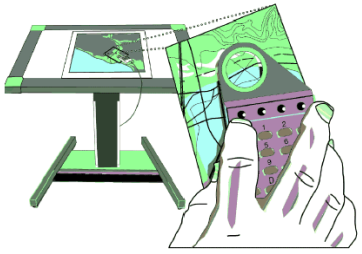
5. Data Input Techniques

- Collect digital data from existing sources
  - GIS/CAD/survey data in a format ArcGIS can read
  - Or use a data conversion tool (e.g. FME)
- Scan map sheets to create raster data
  - then convert to vector - automated or manual:
    - Trace digital raster air photos or satellite images
    - Digitize from a paper map or printed photo/image
- Enter coordinate values (e.g. COGO)
- Purchase existing data in correct format

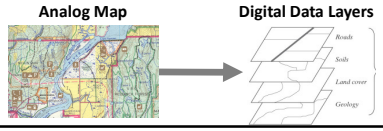
30

Digitizing

- Manual method of analog → digital conversion
- Spatial and attribute data from a paper/mylar map
- Map features are captured by tracing them with either:
  - Digitizing table(t) & cursor
    - Table digitizing: almost obsolete
  - Scanned image of map
    - Screen or heads-up digitizing
- Output:
  - digital vector data
  - some attribute data



1. Digitizer cursor transmits a pulse from an electromagnetic coil under the view lens.
2. Pulse is picked up by nearest grid wires under tablet surface.
3. Result is sent to computer after conversion to x and y units.



31

Thank You

End of Lecture 5