Geospatial Multistate Archive and Preservation Partnership

GIS Definitions

SAA 2009-- 1:30-3pm Saturday 15 August Session

" Mash-up: Archivists and GIS Practitioners: Capturing Earth's Footprints in a Changing World"

For an extended list of GIS terms and definitions, please visit: http://www.opengeospatial.org/ogc/glossary

Attribute Data – generally defined as additional information about each spatial feature housed in tabular format.

ESRI (Environmental Systems Research Institute) - providers of GIS Software (ArcGIS).

Feature - natural and man-made geographic features represented by points/symbols, lines, and areas on a map. Object in a geographic or spatial database with a distinct set of characteristics. A defined aspect of the earth's surface that is not further subdivided. For example, a road segment, manhole, building, or area designated having the same soil type. *(Kansas Association of Mappers: http://www.kansasmappers.org/kam/services/gisdictionary.cfm#F)*

FGDb (File Geodatabase) – a native file structure for ArcGIS. It shows as a folder in the Windows file system. Each dataset is held as a file that can scale up to 1 TB in size. It is a folder of binary formats, and is supported on any operating system platform. Preferred to a personal geodatabase. Supports individual datasets containing over 300 million features and that can scale beyond 500 GB. *(ESRI definition).*

FGDC (Federal Geographic Data Committee) – promotes sharing of the nation's geospatial resources. The FGDC metadata standard was adopted in 1994, and is a standard developed to determine the robustness, the method of accessing, and the successful transfer of geospatial data *(http://www.fgdc.gov/metadata).*

Geodatabase – collection of geographic datasets (features class [a collection of features or a table of rows, where each row has a geographic column], raster data, attribute tables). This is a native structure for ESRI ArcGIS. There are three types of geodatabases:

- Personal Geodatabase

- File Geodatabase (see definition above)
- ArcSDE – stored in enterprise DBMS (database management system), and supports multi-user editing via versioning. It requires ArcEditor or ArcInfo to edit.

Geospatial Metadata – provides a way to describe geospatial data and other electronic records. It contains such information as the coordinate system, when the data was created, when it was last updated, who created it and how to contact them and definitions for any of the code attribute data.

GIS (Geographical Information Systems) – incorporates graphical features with tabular data in order to assess real-world problems (e.g., prioritizing sensitive species habitat to determining optimal real estate locations for new businesses).

At the simplest level, GIS can be thought of as a high-tech equivalent of a map. The key word to this technology is Geography – this usually means that the data (or at least some proportion of the data) is spatial, in other words, data that is in some way referenced to locations on the earth. Coupled with this data is usually tabular data known as attribute data. Attribute data is generally defined as additional information about each of the features, which then can be tied to spatial data *(GIS Lounge: http://gislounge.com/what-is-gis/)*.

For example:

Location: John Smith Elementary School

Spatial data – actual location of the school (x, y coordinates)

Attribute data – school name, level of education taught, school capacity, etc.

(*United States Geological Survey definition*) 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 location. Practitioners also regard the total GIS as including operating personnel and the data that go into the system.

(*ESRI definition*) A geographic information system (GIS) is a computer-based tool for mapping and analyzing things that exist and events that happen on earth. GIS technology integrates common database operations such as query and statistical analysis with the unique visualization and geographic analysis benefits offered by maps.

(*NASA GIS definition*) GIS is an integrated system of computer hardware, software, and trained personnel linking topographic, demographic, utility, facility, image and other resource data that is geographically referenced. If you've ever used an Internet mapping program to find directions, congratulations, you've personally used GIS. The new supermarket chain on the corner was probably located using GIS to determine the most effective place to meet customer demand.

ISO (International Standards Organization) Geographic Information Standards – ISO/TC 211. The core concept of ISO Geographic Information Standards is ISO 19101 Reference Model, which has four

conceptual components: Conceptual Modeling, the Domain Reference Model, the Architectural Reference Model, and Profiles. ISO/TC 211 is for standardization in the field of digital geographical information.

Latitude & Longitude (Geographic Coordinate System) - the most common coordinate system in use to describe a location on the earth's surface. Lines (lines running horizontally are called parallels) of *latitude* run parallel to the equator and divide the earth into 180 equal portions from north to south (or south to north). Measures how far north or south an object is from the equator.

The reference *latitude* is the equator and each hemisphere is divided into ninety equal portions (lines running vertically are called meridians), each representing one degree of latitude.

The lines of *longitude* run perpendicular to the equator and converge at the poles. Lines of longitude run perpendicular to the equator and converge at the poles. The reference line of longitude (the *prime meridian*) runs from the North Pole to the South Pole through Greenwich, England. Subsequent lines of longitude are measured from zero to 180 degrees east or west (values west of the prime meridian are assigned negative values for use in digital mapping applications) of the prime meridian *(http://www.socialstudiesforkids.com/articles/geography/latitudelongitude.htm)*.

The degrees are further divided into minutes (') and seconds ("). There are sixty minutes in a degree, and sixty seconds in a minute.

For example:

Austin, TX

Latitude: 30° 16' 1"

Longitude: -97° 44' 35"

Oblique Imagery – *(from http://aims.jocogov.org/AIMSData/Oblique.aspx)* is aerial photography that is captured at approximately a 45 degree angle with the ground. The angle which is inherent to oblique imagery allows viewers to see and measure not only the top of objects but the sides as well. Oblique Imagery more closely resembles how people normally view their landscape compared to traditional orthogonal (straight down) imagery.

Orthophotography – digital imagery in which distortion from the camera angle and topography have been removed, thus equalizing the distances represented on the image. A rectified copy of a photograph (typically an aerial photograph), showing image features corrected for variations in scale and height displacements. (*From http://www.websters-online-dictionary.org/or/orthophotography.html)* Aerial photographs that more precisely show the features of the landscape, including those that might be important for agriculture such as slope or size of gullies, because they are corrected for distortion caused by tilt, curvature, and ground relief.

Raster Data – cell-based data such as aerial imagery and digital elevation models. Raster data is characterized by pixel values. Basically, a raster file is a giant table, where each pixel is assigned a specific value from 0 to 255. The meaning behind these values is specified by the user- they can represent elevations, temperatures, hydrography, etc. Satellite imagery uses raster data to record different wavelengths of light. Raster data is advantageous to vector data in constructing 3D images, as the values for every pixel are calculated through a process called *interpolation* (http://www.umich.edu/~ipcaa/GIS/General%20GIS%20Concepts.htm).

Remote Sensing System - Remote sensing can be defined as the study of something without making actual contact with the object of study.

Four basic components of a remote sensing system include a target, an energy source (illuminates or provides electromagnetic energy to the target), a transmission path, and a sensor (remote device that collects and records the electromagnetic radiation). The resulting data set must be transmitted to a receiving station where the data are processed into a usable format, which is most often as an image *(http://chesapeake.towson.edu/data/all_intro.asp*).

Example: weather satellite imaging of the Earth

SDE – see description above under Geodatabase.

.shp file ("shapefile") – (*from ESRI*) stores non-topological geometry and attribute information for the spatial features in a data set (digital vector storage format created in ArcView). The geometry for a feature is stored as a shape comprising a set of vector coordinates. Because shapefiles do not have the processing overhead of a topological data structure, they have advantages over other data sources such as faster drawing speed and edit ability. Shapefiles handle single features that overlap or that are noncontiguous. They also typically require less disk space and are easier to read and write.

Consists of:

- Main file – direct access, variable-record-length file in which each record describes a shape with a list of its vertices. Basically the file that stores the feature geometry (.shp).
- Index file -each record contains the offset of the corresponding main file record from the beginning of the main file (.shx). Basically the compiled file that stores the index of the feature geometry.
- dBASE table – contains feature attributes with one record per feature (.dbf)

Spatial Data - also known as *geospatial data* or *geographic information* it is the data or information that identifies the geographic location of features and boundaries on Earth, such as natural or constructed features, oceans, and more. Spatial data is usually stored as coordinates and topology, and is data that can be mapped. Spatial data is often accessed, manipulated or analyzed through Geographic Information Systems (GIS). *http://www.webopedia.com/TERM/S/spatial_data.html*.

Spatial data = Spatial (Where) + Data (What)

Tabular Data - consists of *attribute tables* that define the parameters of the map features. There is really no limit to what the tables can contain, whether Boolean strings (True/False), Text, or Numeric data. For example, a Boolean entry in a cities table may define whether or not each city is a national capital. A text entry may have the city's name, or the archaeological period in which it flourished. A numeric entry could have population figures or lat/long coordinates. The advantage of the relational database system is that the different columns can be sorted and selected according to the user's need. These selections then appear highlighted on the map (*http://www.umich.edu/~ipcaa/GIS/General%20GIS%20Concepts.htm*).

Temporal – existing for a time only. As an example, in the Dublin Core elements, temporal can be defined by a date, date range or a named period.

[Getty] Thesaurus of Geographic Names (TGN) – structured (controlled) vocabulary that improves access to information about places. Places in TGN include administrative political entities (e.g., cities, nations) and physical features (e.g., mountains, rivers). Current and historical places are included. (http://www.getty.edu/research/conducting_research/vocabularies/tgn/)

Vector Data –spatial data represented as points, lines and polygons. This system of recording features is based on the interaction between arcs and nodes, represented by *points, lines, and polygons.* A point is a single node, a line is two nodes with an arc between them, and a polygon is a closed group of three or more arcs. With these three elements, it is possible to record most all necessary information (*http://www.umich.edu/~ipcaa/GIS/General%20GIS%20Concepts.htm*).