Week 2 Cookbook: What is GIS?

Week 2 Overview

1. What is GIS?
2. What is Geospatial Data?
3. Vector vs. Raster Data
4. Spatial vs. Attribute Data
5. Datalayer and Database Concept
6. The Concept of Topology
7. Your Turn: Online Research. See separate instructions in the Week 2 Assignment Sheet.

Simple Question: What is GIS? Not a simple answer!

We will explore and discuss this question this week. As always, the Week 2 Cookbook contains all the information, resources you need for this week.

The Week 2 Assignment Sheet contains all the information you need for this week’s assignment.

Please note our What is GIS? discussion forum
Here we will share and discuss our GIS definitions. But, please do not just ‘post-and-run’ – that’s boring. Instead, read what other people wrote, make comments, suggestions, think about useful analogies or metaphors, and find examples. That’s how we engage with each other to form an online learning community. Once we combine everyone's definitions and ideas we'll have a much better idea what this GIS thingy is all about.

The same applies to our ongoing Map Lounge discussion forum. Don’t just post a cool map and disappear. Check back often, see what other folks have posted – you are sure to find something interesting! You can also post more than just one map!

If you are unsure about discussion forums: re-visit the Getting Started folder and re-read the Do's and Don'ts of Online Communication as a start, then check The Kitchen Sink and the Discussion Forum Grading Rubric.

Consider printing this cookbook for your convenience
1) What is GIS?

This is what you have officially signed-up for…does that help?

Does that explain what GIS is?

No!

GIS = Geographic Information System(s)

Let’s look at this acronym word-by-word:

Geography = ‘Writing about the Earth’

Geography is the scientific study of the spatial variations in physical and human phenomena on Earth.

- Basically we try to figure out where, why, and how things happen on the Earth.
- I prefer this explanation: Geography is to Space what History is to Time

Geographic Information

= Data that include the location of the data in some kind of systematic mathematical format.

This mathematical format can be simple street addresses, latitude/longitude, or any other kind of (X,Y) coordinates. We use geographical information all the time!

- For example: Telling someone that it was 90° yesterday makes no sense at all, unless you also tell the person where on Earth it was 90°F yesterday.
- For example: Inviting someone to your 4th of July BBQ makes no sense at all, unless you tell the person where you are having your BBQ.

In a GIS we can specify the location of a feature using street addresses or latitude/longitude – whatever works better for a given application (see Section 2)

Geographic Information System

= Geographic information that is systematically organized (in a database).

➤ If we have geographical information, organized in a systematic manner, then we can use a computer for quantitative geographic analysis. Remember: a computer is dumb, but a computer can perform repetitive tasks really fast. Thus, if you have geographic information in a systematic format, a computer can help you display it as a map and perform mathematical calculations for you.

FYI: database is really only a fancy world for a digital table, for example a MS Excel spreadsheet.
The figure above illustrates what a GIS is and what the components of a GIS are.

A GIS allows us to visualize and analyze data in ways that reveal spatial relationships, patterns, and trends in the form of maps, reports, and charts.

One of the big concepts in GIS is the so-called layer or sandwich concept: When we overlay different types of information on top of one another we can see patterns and relationships that we could not see by looking at each type of information alone.
Consider the BLT Sandwich Analogy
Bacon, lettuce, and tomatoes are delicious on their own. But, when you layer them onto top of one another, between two slices of toast, you get a whole new dimension of flavor!

That’s the same with geographic data. A wetlands map is interesting, a map of gas stations is interesting, and a map of drinking water wells is interesting. But when you overlay them you begin to see patterns and relationships, for example where are the gas stations that are within 100 feet of a wetland and a drinking water well?

Therefore, we need to distinguish between the Big Picture and the Small Picture.

<table>
<thead>
<tr>
<th>Big Picture: GIS is a Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>The location of objects and their interactions are (often) not random or independent. Instead, the world is characterized by spatial patterns and relationships. A GIS allows us to visualize (map) and analyze these spatial patterns and relationships.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Small Picture: GIS is hardware, software, data, people, and methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware: PCs, servers, GPS, printers, scanners, Smart-phones, PDAs, etc.</td>
</tr>
<tr>
<td>Software: ArcGIS, Google Earth, Google Maps, Mapquest, IDRISI, etc.</td>
</tr>
<tr>
<td>Data: Shapefiles, geodatabases, air photos, DEMs, etc. (create yourself or download)</td>
</tr>
<tr>
<td>People: Scientists, planners, environmental consultants, police, military, business, etc.</td>
</tr>
<tr>
<td>Methods: Spatial statistics, mapping, surveying, remote sensing, etc.</td>
</tr>
</tbody>
</table>

2) What is Geospatial Data?
That’s a big word – geospatial data, but what does it mean? We already covered it in Section 1, but it is worth repeating and expanding upon.

Geospatial Data or Information
= Data that include the location of the data in some kind of systematic mathematical format.

We commonly use two different mathematical formats for locations:

1. Street, City, State, Zip Code.
   That’s the system that we are all used to.
   One problem: the address system is not uniform between different countries of the world, which makes a little bit tricky when you deal with different countries.

2. Latitude and Longitude = the Geographic Grid.
   The geographic grid is great because it works the same way no matter where you are on the planet!

But, ArcGIS (our GIS software) allows us to use both systems and even converts easily between them!
Example: Carsten’s House

<table>
<thead>
<tr>
<th>Street Address</th>
<th>Geographic Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>110 Haywood Street</td>
<td>Latitude: 42.598631°N</td>
</tr>
<tr>
<td>Greenfield, MA 01301</td>
<td>Longitude: -72.586614°W</td>
</tr>
<tr>
<td>USA</td>
<td></td>
</tr>
</tbody>
</table>

Both formats are useful. If you want to send me a present with the U.S. Postal Service you would use my street address. If you want to find my house in a GIS you can use both format!

Your Turn: Find the latitude and longitude for your house or dorm (use Google Earth)

In GIS, we distinguish between three types of objects – these are called spatial features.

1. Point Features
   A point feature is an object that is a single point or that we can represent as a single point. Examples: a tree, a well, the location where we took a soil sample, utility pole, etc. It gets a bit trickier in some cases. For example, the location of a city may be shown on a map as a point, even though the city limits are really a polygon. That depends of the scale and detail of your map.

2. Line or Polyline Features
   Line = straight (start point, end point)
   Polyline = has wiggles (start point, end point, and many points or nodes in-between)
   In reality, every line is actually a polyline when you look close enough. These are features that are long and skinny, for example rivers, roads, and railroad tracks. Again, it depends a bit on the scale and detail of your map. On a map of the entire USA we would show the Connecticut River as a line. On a map of downtown Springfield, on the other hand, the width of the Connecticut River is important and we would use a polygon.

3. Polygon Feature
   In reality, everything is a polygon when you show it in enough detail. Even a tree or utility pole has an area. A polygon can be as simple as a triangle (3 sides, 3 nodes) or extremely complex (many side, many nodes). Examples: Lakes, state outlines, county lines, oceans, building foot prints, etc.

3) Vector vs. Raster Data

In a GIS we distinguish between two different data types: Vector Data and Raster Data

**Vector Data**
Vector data uses points and their (x, y) coordinates to construct spatial features (such as points, polylines, polygons).

The information about the spatial feature stored as attributes in a linked attribute table.

Vector data are best for showing discrete features (e.g. roads, rivers, utility poles, wetlands, states, etc.)
**Raster Data**

Raster Data uses a grid and grid cells to represent the spatial variations of a parameter.

The Information about the parameter is the grid cell value.

Raster data are suitable for representing continuous parameters (e.g. elevation, temperature, wind speed, etc.)

The most commonly used raster are air photos, satellite images, and digital elevation models.

A digital elevation model is a uniform grid of elevation data.

Each grid cell has an x-coordinate (or column number), a y-coordinate (or row number), and a z-coordinate (= grid cell content = elevation).

The GIS now uses the grid cell content to assign different colors for display.

Of course, the real-world consists of both…vector and raster data. In other words, we can show the real world in vector or raster format.
**For example: a lake.**

You can show a lake as vector data, namely as a polygon. You can get the lake shore very detailed if you add many nodes and sides to your polygon. Or, you can simplify things because you are not interested in the details of the lake shore. You can even show the lake as a point if the actual shape of the lake is not relevant for your map or analysis.

But, you can also show a lake as raster data. Here you break the lake apart into squares. If you use really small squares (e.g. 1 foot) you can show the lake in great detail. If you use larger squares (e.g. 100 yards) your lake shore will be much simplified.

Real map making, for example the topographic maps produced by the United States Geological Survey, always uses both vector and raster data. Why? Makes sense! You use whatever data is best suited for your specific map!

**ArcGIS can handle and analyze both vector and raster data. It can even convert one into the other!**

### About Making Maps

In map making you always simplify and generalize spatial features, because you cannot show every little detail. Plus, you may not want to show every little detail – it may distract and clutter your map. Maps are always smaller than the real world, so we have to shrink reality to make the map…and therefore we have to simplify and generalize the shapes and geometries of spatial features.

In theory we could make a 100 percent accurate map that shows every detail of the real world, but that would be pointless, since the map would be the same size as the real world!

### 4) Spatial vs. Attribute Data

**Spatial Data**

Spatial data refers to the geographic location and geometry of a feature (for example a point, polyline, or polygon). Therefore, the spatial data provides the location of a feature in some kind of systematic mathematical format

- Examples of spatial data: the location and geometry of a wetland (polygon), a river (polyline), or utility pole (point).
- Spatial data tells us where something is located, but not what it is.

**Attribute Data**

Attribute data tells us what a feature is (i.e. its characteristics), but not where it is located. Attribute data is stored in a table, hence the term attribute table is used in GIS.
• Examples of attribute data: the height of a mountain and its rock type, the bio-diversity of a wetland, and whether a road is paved, dirt, private, etc.

The power of a GIS is that it can handle, display, and analyze both types of data together. You can ask “Where is something?” and “What does it do?”

The GIS stores and link the spatial and the attribute data. All of that happens automatically and you do not really need to think about it.

5) Datalayer and Database Concept

Data Layer Concept
In a GIS or on a paper map, the geographic information is represented by a series of layers. For example, one layer might be topography, another layer might be roads, and a third might be wetlands. These different layers are displayed laid on top of each other.

• In a paper map, the map maker decided what layers to overlay and how.
• In a GIS, you as the user have the choice and can optimize the overlay for display and analysis for a specific purpose.

GIS = dynamic or intelligent map
Paper Map = static map
In a GIS, you make the choice what layers that you want to show for a given task. With a paper map, the map maker and map publisher decided for you as in the famous quote from Henry Ford: *A customer can have a car painted any color that he wants so long as it is black.*

**Data Base Concept**

In a GIS, the attribute data are stored in a relational database, which is series of tables interconnected by a common field. This means that you can change the attribute of a feature (for example the value of a property parcel) in one table, and the change “cascades” through the entire linked data base and the associated map.

In the example on the right, four tables are interconnected using the common field LOC_ID, whereas two tables are interconnected using the common field SAMP_ID.

Even better: Your database tables are dynamically-linked to your maps! So, once you change the values of a property parcel in the database tables your map will automatically change accordingly.

**MS Excel Analogy**

Let’s say you have a simple spreadsheet with the cities in Hampden County as rows and you have, for each city, the total population in second column. In MS Excel you can now easily make a bar graph that shows the height of each bar as a function of the population of the respective city. If you change the value for population in the spreadsheet (e.g. enter 500,000 for Westfield) the bar representing Westfield in your bar graph will get much higher because the spreadsheet and the graph are dynamically-linked.

**Have a look at the Jings that compare MS Excel to a GIS!**
6) The Concept of Topology

To recap: A Geographic Information System (GIS) is a computer system for collecting, storing, querying, analyzing, and displaying geospatial data. Geospatial data describes the location and characteristics of objects (e.g. roads, timber stands, lakes, utility poles, schools, etc.)

In short, a GIS knows:
- Where an object is located (= its location)...spatial data
- What an object is (= its attributes)...attribute data
- What is located around the object (= topology)...adjacency/incidence

Topology
= The relationship between objects that remains constant under certain transformations, such as bending and stretching.

Image you have a map printed onto a rubber sheet. You can stretch and bend the rubber sheet, thereby distorting the map. But the basic spatial relationships between objects on the map will always stay the same:
- Objects next to each other will stay next to each other.
- Objects that connect/intersect will continue to do so.
- Objects that are contained within each other will remain so.

For example, with topology the GIS not only knows the location, geometry, and characteristics of the Connecticut River.

But also knows what is located to the right and left of the river (for example Vermont and New Hampshire) and what it is connected to (for example the Westfield River or the Atlantic Ocean).

That’s topology – a bit of a weird concept at first!

Here’s another example: Imagine you have a map of properties in your town. Each property line represents a border to another property, as well as the property line of that adjacent property.

For example: My property is a simple rectangle. 3 of my property lines are the borders with my neighbors, thus are also their property lines. 1 of my property lines is the road I live on, owned by the town. That’s topology!
7) Your Turn: Online Research

Okay, now that we have covered a few of the basics: it’s your turn. Have a look at the Week 2 Assignment sheet for detailed instructions.

This activity/assignment consists of three inter-connected parts:

- Part 1 is the actual online research where you will further investigate topics related to GIS and summarize your research as a written report.
- Part 2 is a discussion forum where we will share and discuss our GIS definitions.
- Part 3 is our weekly Map Lounge – this week for sharing cool maps we found on the web.