GIS for Research II:

GIS Research, Data Management, and Visualization

James Whitacre
GIS Specialist
Scholarly Commons
Instructor: James Whitacre

• GIS Specialist
  Scholarly Commons, Main Library at the University of Illinois at Urbana-Champaign

Previously:
• GIS Manager
  Carnegie Museum of Natural History

• Master of Science in Geography, concentration in GIS and Cartography
  Indiana University of Pennsylvania

• Bachelor of Arts in Zoology
  Ohio Wesleyan University
GIS at Scholarly Commons

• Provide GIS research consultations and services

• Provide GIS workshops and training

• Manage the University Library’s geospatial data and GIS software resources

• Act as a central resource for the University’s GIS community and promote the use of GIS in research and teaching
GIS for Research Core Workshops

• GIS for Research I: Introduction to GIS Concepts, Software, and Data

• GIS for Research II: GIS Research, Data Management, and Visualization

• GIS for Research III: Geoprocessing, Analysis, and Web GIS

Slides available at http://guides.library.illinois.edu/gis
Special Topics GIS Workshops

• ArcGIS Online and Story Maps
• ModelBuilder and Python in ArcGIS
• Field data collection techniques and Collector for ArcGIS
• Geocoding and georeferencing
• Spatial Analysis Techniques: Interpolation, Clusters, Statistics, and more
• Lidar and 3D mapping techniques
Learning Objectives

• Understand how to frame research questions and methods for GIS

• Overview of geospatial data management best practices

• Use ArcGIS Desktop to manage and visualize geospatial data
Framing Research for GIS

1. Acquire Geospatial Data
2. Analyze Geospatial Data
3. Present Geospatial Data
Defining Research Problems

• Key is to clearly frame research to be adapted to GIS

• What is your specific research question?
  – What quantitative and qualitative parameters are needed?
  – What is the final product needed to answer your research question?
Determining Data Requirements

• **Knowing your data situation**
  – What data is desired/required vs. available to answer this question?
  – How does this change your research question?

• **Data Considerations**
  – What formats and data model will you use?
  – What scale of data is available?
  – What coordinate system will you use?

• **How will you organize your data?**
  – Determine how you will name and organize files
  – Documentation and metadata
  – Come up with a system
Analysis and Visualization

• **Simple vs. Complex uses of GIS**
  – Simple: Making a map to visualize data
  – Complex: Using geoprocessing tools to analyze multiple dataset to produce new data

• **What analysis is needed to answer your research question?**

• **How will you visualize your final product(s)?**

• **Framing research for GIS is a process**
  – Exploring data, techniques, and the output will impact the process
  – Does not need to all be in place
  – May be helpful to work backwards
  – How can you best represent the real world with your data
Exercise 2

Framing Research for GIS

Acquire Geospatial Data

Analyze Geospatial Data

Present Geospatial Data
**Scenario: Simple Research with GIS**

- You are doing research on international students at the University of Illinois with your professor.

- You have a presentation coming up and you want to include a simple map in your PowerPoint presentation comparing the number of international students from each country.

- Your professor also wants to know the gender breakdown from each country.
Scenario: Simple Research with GIS

• What data will you need to make your map?
  – World country boundaries
  – Number of students from each country

• What analysis will you need to perform?
  – Table Join of the students from countries to the world countries boundary

• How will you visualize and present your data? What is the final product?
  – Map showing the variation of the number of students from each country
Geospatial Data Management
Importance of Data Management

- Meet funding agency requirements
- Protect federal investment in research and development
- Expedite the scientific process; saving time and resources in the long run
- Use or re-use the value, the uniqueness, and the importance of data
- Ensure that research data and records are accurate, complete, authentic and reliable
- Ensure research integrity and replication
- Increase research efficiency
- Enhance data security and minimize the risk of data loss
- Prevent duplication of effort by enabling others to use your data
- Comply with practices conducted in industry and commerce

What is Data Management in GIS

• Planning for acquisition, storage, and analysis of geospatial data and associated metadata

• Implementing the plan using GIS and data management best practices

• Regarded as ‘one of the biggest issues confronting GIS professionals’
What are Best Practices?

• Describes methods, procedures, techniques, and workflows

• Driven by results, quality, and standards

• Designed to be cost and time effective

• Formal or informal
Practical GIS Data Management Tips

• Track and document data sources and metadata

• Name files descriptively, unambiguously, and systematically (e.g. NOT Test1, Test2, etc.)

• Minimize the number of folder levels in your project working folder

• Keep an original copy of the downloaded or acquired geospatial data

• Export or copy data to a working folder or geodatabase for geoprocessing and analysis

• 3-2-1 Rule: 3 copies, 2 different media, 1 copy offsite
What is Metadata?

• Data about data

• Describes quality, condition, origin, and many other characteristics

• For GIS data, projection, scale, resolution, and accuracy are important

• Should include a data dictionary or code book – an explanation of attributes and codes
Geospatial Metadata

• Includes information such as geographic location or extent, coordinate system, and scale or resolution

• Two primary standards in the US
  – FGDC Content Standard for Digital Geospatial Metadata (CSDGM)
  – International Organization for Standardization (ISO) 19115 (and other related in the 19100 group)
Assessing Geospatial Data Quality

• Is there clear metadata or other documentation?

• Is the source or originator clearly identified?

• Are the attribute fields defined or clearly named? Are field values appropriate and consistent?

• Is the coordinate system clearly defined? Is it an appropriate coordinate system?

• Is the scale or resolution clearly indicated? Is accuracy and precision of the data addressed?
Can’t Find the Right Data?

• Do you expect the data to be available in the form you imagine?

• Are you thinking about your research problem correctly?

• Can you find others working on similar issues?

• Can the data be collected or created?
Exercise 2

Geospatial Data Management

Acquire Geospatial Data

Analyze Geospatial Data

Present Geospatial Data
Joining Tables in GIS
A (Brief) Note on Database Design...

• It is not good practice to duplicate data

• Data tables should be focused in scope (i.e. topical or thematic)

• Tables should be designed to facilitate relationships between each other

• Linking, or joining, tables is one way tables can relate to each other
What is a Table Join?

• Appending the attributes from two tables (or from a feature class to a table)

• Both tables must have a common, or key, field with exact matching values

• The primary table or feature class is the target, or origin, table

• The secondary table is the join, or foreign, table
Types of Table Joins

• One-to-one: the origin table will have exactly one match to the foreign table

• Many-to-one: the origin table will have one or more records that can match to one record of the foreign table
Exercise 3

Table Joins
Creating Maps in ArcGIS
Elements of a Map

- Title
- Map Symbols
- Legend
- Map Scale
- North Arrow
- Map Area
- Labels
- Borders and Neatlines
- Graticule or Coordinate Grid
Map Design

• Colors and Shading

• Typography/Fonts

• Focus of attention

• Map element hierarchy/Figure-ground considerations

Exercise 4

Design a Map Layout for PowerPoint
GIS for Research Workshops

• GIS for Research I: Introduction to GIS Concepts, Software, and Data
  – Thurs. Sept. 21 1-3pm & Fri. Nov. 3, 1-3pm, rm. 314

• GIS for Research II: GIS Research, Data Management, and Visualization
  – Thurs. Sept. 28 1-3pm & Weds. Nov. 8, 1-3pm, rm. 314

• GIS for Research III: Geoprocessing, Analysis, and Web GIS
  – Thurs. Oct. 5 1-3pm & Fri. Nov. 11, 1-3pm, rm. 314
Special Topics GIS Workshops

• Sharing Research with Story Maps
  – Tues., October 17, 3-4pm, room 314
  – Thurs., November 2, 2-3pm, room 314

• Geocoding
  – Thursday, October 19, 2 – 3 pm, room 314
Special Topics GIS Workshops

• Introduction to Python for ArcGIS
  – Friday, December 1, 1 – 4pm, room 314

• Advanced Python for ArcGIS
  – Friday, December 8, 1 – 4pm, room 314
Library Resources

• Scholarly Commons GIS Services:
  – GIS data discovery and research services
  – GIS consultations by appointment
  – http://www.library.illinois.edu/sc/datagis
  – http://guides.library.illinois.edu/gis

• Map Library:
  – Geospatial datasets, GIS reference books and journals, aerial photos, paper maps, etc.
  – www.library.illinois.edu/max
Questions?
Thank You!