INFO 520 - Introduction to Spatial Data Management - Syllabus

Karl Benedict

Revision History

Revision 2014.001	2013-07-28	KB

Table of Contents

Many data products are inherently spatial. Obviously spatial data include data collection locations, but many other data may also be considered spatial: locations in space that documents pertain to and locations of historic or literary events illustrating just a few. While maps are a familiar product derived from spatial data, there is significant understanding of the underlying data — the processes to which it has been subjected, the actual values within the data, the originator of the data, any limitations in the appropriate use of the data, and the nature of the dataset itself (format, scale, coordinate system, units) — that is required before it can be productively used for research or applications. This course is designed to provide graduate-level students with the necessary skills and knowledge to meet this challenge through hands-on work in *discovering*, *creating*, *managing*, *using*, *documenting* and *sharing* spatial data. After completing this course students will be better prepared to develop a plan for the management of their spatial data, locate and evaluate data sources that they need for their research, create and structure data that they collect for maximum value both during and after their research project, and document their data throughout their research projects, maximizing the impact of their research and the value of the data they generate and share with other researchers.

1. Course Instructor

Karl Benedict

University of New Mexico

- Director, Earth Data Analysis Center
- Assistant Research Professor, College of University Libraries and Learning Sciences
- Assistant Research Professor, Department of Geography
- Adjunct Professor, Department of Anthropology

email: kbene@unm.edu [mailto:kbene@unm.edu]

Phone: (505) 277-3622 x234

Office Hours: **TBD**

2. Course Description and Objectives

An understanding of core spatial data concepts and principles is increasingly important in the current world of collaborative, spatially enabled research and applications. We are no longer working in a vacuum as individual researchers that only need to understand and use the data that we create and use in our separate research projects. Successful research depends upon being able to integrate data generated by others with our own and by extension being able to share our data with others, both during our research projects and also for posterity (and to meet the requirements of funding agencies). This class will focus on the following aspects of spatial data management that relate to this need for effective integration, use, collaboration and sharing:

- The *Research* and *Data Lifecycles*
- Types of spatial data
- Spatial database design and management
- Working with and managing gridded data
- Spatial data documentation standards and practices
- Ethical, legal and privacy issues as they relate to spatial data
- Data management planning
- Emerging topics

Upon completion of the course students will have improved their knowledge and skills in the following areas:

- Locating and evaluating spatial data based upon knowledge of formats, content models and documentation standards
- Structuring data (both in terms of format selection and content) from a variety of sources to enable integrated research
- Evaluate data products to determine which elements of a dataset might raise ethical, legal or privacy issues if released or shared with others
- Documenting data as an ongoing process throughout the research cycle
- Producing machine- and human-readable documentation for data to support discovery, understanding, and use of data that they produce

3. Course Format

TBD

4. Readings

Nikos Mamoulis (2012), Spatial Data Management. Synthesis Lectures on Data Management #21. Morgan & Claypool Publishers. DOI10.2200/S00394ED1V01Y201111DTM021. SDM [http://libproxy.unm.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00503a&AN=unm.b7199537&site=eds-live&scope=site]

Michael J. Hernandez (2003). Database Design for Mere Mortals: a Hands-on Guide to Relational Database Design. 2nd ed. Addison-Wesley. DBD [http://libproxy.unm.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00503a&AN=unm.b3649820&site=eds-live&scope=site]

5. Evaluation and Grading

TBD

While students are encouraged to collaborate in their work on the project and homework assignments, submitted work must be original and written and submitted by each individual student.

Please refer to the Pathfinder [http://pathfinder.unm.edu/] for detailed student conduct policies, and in particular the following Policy on Academic Dishonesty [http://pathfinder.unm.edu/policies.htm#academicdishonesty].

> Each student is expected to maintain the highest standards of honesty and integrity in academic and professional matters. The University reserves the right to take disciplinary action, up to and including dismissal, against any student who is found guilty of academic dishonesty or otherwise fails to meet the standards. Any student judged to have engaged in academic dishonesty in course work may receive a reduced or failing grade for the work in question and/or for the course.
>
> Academic dishonesty includes, but is not limited to, dishonesty in quizzes, tests, or assignments; claiming credit for work not done or done by others; hindering the academic work of other students; misrepresenting academic or professional qualifications within or without the University; and nondisclosure or misrepresentation in filling out applications or other University records.
>
> — Adopted by the UNM President on June 15 1992 *UNM Policy on Academic Dishonesty*

6. Technical Requirements

6.1. Software

- GIS - Quantum GIS [http://www.qgis.org/]
- Spatial Database - SpatiaLite
  - Mac OS X [http://www.kyngchaos.com/software/frameworks]
  - Windows & Source (Linux) [http://www.gaia-gis.it/gaia-sins/]

6.2. Hardware

TBD

6.3. Network Connectivity

TBD

7. Weekly Schedule

Week	Date	Topic	Lab	Reading	Assignment	Project
Week 1	1/21/2014	Course Overview - Introduction to the Data Lifecycle	Class Introduction	-	-	-
Week 2	1/27/2014	Types of Spatial Data - Vector	-	-	Domain specific literature review	Define data management focus for term

Week	Date	Topic	Lab	Reading	Assignment	Project
Week 3	2/3/2014	Types of Spatial Data - Raster	-	-	-	-
Week 4	2/10/2014	Database design I	-	-	Post literature review to discussion	-
Week 5	2/17/2014	Database design II	-	Locate and describe data and review for documentation, usability and understanding	-	-
Week 6	2/24/2014	Geodatabase design	-	-	-	-
Week 7	3/3/2014	Managing raster data	-	-	-	-
Week 8	3/10/2014	Data formats for Analysis and Archiving	**Presentations** of data review	-	-	Enumerate specific data (>= three datasets) to be used in the project
Week 9	3/17/2014	Spring Break	-	-	-	-
Week 10	3/24/2014	Documenting data - the interview	-	-	-	Create initial data
Week 11	3/31/2014	XML Document creation, editing and validation	-	-	-	-
Week 12	4/7/2014	Metadata Standards - FGDC	-	-	-	Document Data
Week 13	4/14/2014	Metadata Standards - ISO and Dublin Core	Data management planning process Q&A	-	Create a data management plan	-
Week 14	4/21/2014	Data management planning	-	-	-	-
Week 15	4/28/2014	Ethical, legal and privacy issues	-	-	Data management plan peer review	Project data and documentation peer review
Week 16	5/5/2014	Emerging concepts	**Project Presentations**	-	-	Present project results and

Week	Date	Topic	Lab	Reading	Assignment	Project
						peer review outcome
Week 17	5/12/2014	Finals Week	-	-	-	-