OILS 515 - Introduction to Spatial Data Management - Syllabus

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.

Course Instructor

Karl Benedict

- Associate Professor, College of University Libraries and Learning Sciences
- Affiliated Faculty, Department of Geography
- Adjunct Professor, Department of Anthropology

Email: kbene@unm.edu

Phone: (505) 277-5256

Office Hours: By appointment.

Office Location:
Centennial Science and Engineering Library - CSEL L259

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
- Data management planning
- Ethical, legal and privacy issues as they relate to spatial data
- 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

Course Format

The course is structured as a combination of short lectures/demonstrations (presented as part of our concurrent collaboratory sessions) that set the stage for the technical topics covered in the readings, hands-on work with data and data documentation, and data management planning exercises. While offered as an online course, several online web conferences (collaboratory sessions) are required as part of the class participation. These sessions are plananed for Wednesday evenings from 5:00-6:30 pm and listed under the "Collaboratory" column in the class calendar.

Readings

Nikos Mamoulis (2012), Spatial Data Management. Synthesis Lectures on Data Management #21. Morgan & Claypool Publishers. DOI10.2200/S00394ED1V01Y201111DTM021. http://unm.worldcat.org/title/spatial-data-management/oclc/767844616&referer=brief_results[SDM]

Michael J. Hernandez (2003). Database Design for Mere Mortals: a Hands-on Guide to Relational Database Design. 3rd ed. Addison-Wesley. http://unm.worldcat.org/title/database-design-for-mere-mortals-a-hands-on-guide-to-relational oclc/872560697?ht=edition&referer=di[DBD]

Additional online readings will also be assigned over the course of the semester.

Evaluation and Grading

Course grades will be based on a combination of participation in live and online discussions and peer-review, the smaller assignments (listed under the "Assignment" column in the class calendar), and the semester-long class project. The grade for the class will be weighted according to the following breakdown:

- Class Participation: 20%
- Small Assignments: 40%
- Class Project: 40%

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 for detailed student conduct policies, and in particular the following Policy on Academic Dishonesty.

> 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.

Technical Requirements

Software

- Recent Windows, Mac or Linux Operating System
- GIS - Quantum GIS http://www.qgis.org/
- Spatial Database - SpatiaLite
  - http://www.kyngchaos.com/software/frameworks[Mac OS X]
  - http://www.gaia-gis.it/gaia-sins/[Windows & Source (Linux)]
- Python (possible, based upon interest)

Hardware

Relatively recent laptop or desktop computer with at least 4GB ram and 50 GB free hard disk space

Headset with integrated microphone and headphones (strongly recommended) or microphone and speakers (not recommended but can work in some circumstances) for participation in remote collaboration sessions

Network Connectivity

Broadband internet connectivity. Some wireless networks may not be sufficient and should be tested prior to participation in web conferences.

Weekly Schedule

Week	Date	Topic	Collaboratory	Assignment	Project
1	1/12	Course Overview - Introduction to the Data Lifecycle	1/14-Class Introduction	-	-
2	1/19	Types of Spatial Data - Vector	-	Domain specific literature review	Define data management focus for term
3	1/26	Types of Spatial Data - Raster	-	-	-
4	2/02	Database design I	2/04-Discuss literature review results	Post literature review to discussion	-

5	2/09	Database design II	-	Locate and describe data and review for documentation, usability and understanding	-
6	2/16	Geodatabase design	-	-	-
7	2/23	Managing raster data	-	-	-
8	3/02	Data formats for Analysis and Archiving	3/04-*Presentations* of data review	-	Enumerate specific data (>= three datasets) to be used in the project
9	3/09	Spring Break	-	-	-
10	3/16	Documenting data - the interview	-	-	Create initial data
11	3/23	XML Document creation, editing and validation	-	-	-
12	3/30	Metadata Standards - FGDC	-	-	Document Data
13	4/06	Metadata Standards - ISO and Dublin Core	4/08-Data management planning process Q&A	Create a data management plan	-
14	4/13	Data management planning	-	-	-
15	4/20	Data management planning (continued)	-	Data management plan peer review	Project data and documentation peer review
16	4/27	Emerging concepts/ Ethical, legal and privacy issues	4/29-*Project Presentations*	-	Present project results and peer review outcome
17	5/04	Finals Week	-	-	-