**GEOG 281: THE WORLD & BIG DATA**

**Instructor:** Hedda R. Schmidtke  
**Lecture:** 2 x 1.5-hour lecture/week  
**Labs:** 1 x 1-hour lecture/week

**TEXTBOOKS**  


**COURSE DESCRIPTION**  
Today’s world is driven by data. Political, economical, and environmental decision making is increasingly based on larger data sets, and advertisers and services like Google *personalize* what you see and are offered based on data about you and “customers like you.” Moreover, social media allow the public to directly participate in decision making, disaster reporting, and humanitarian interventions. In order to do this, massive amounts of data need to be collected, separated into relevant and irrelevant forms, and processed into classification mechanisms able to categorize incoming new data in real-time. Location, time, and geographic parameters are among the most important indicators within these data spaces and among the most ancient. Maps, in fact, can be seen as powerful visualizations storing and conveying large amounts of geographic data, and geographic information systems were among the first problem domains for research in “Very Large Databases.” Moreover, geography research continues to create requirements and solutions for spatial and temporal representations in the next generation of Big Data systems.

This course is an introduction to the field of Big Data research from a socio-technical systems perspective. It explains historical roots, social and economic application areas, as well as technical fundamentals. The area is explored along the dimensions of the “five Vs”: volume, variety, velocity, veracity, and visualization.

**EXPECTED LEARNING OUTCOMES**  
After completing the course students

- Understand the historical background and conceptual framework behind the big data movement and the role geography played in its emergence
- Are able to visualize spatial and temporal dimensions of a big data set and to critically evaluate what insights can be gained from a big data set or service
- Are able to apply fundamental computational methods for developing a geospatial big data project
- Understand the specific technical difficulties big data systems have to tackle
- Are able to conceptualize and explain spatial, temporal, and data volume scales outside of human everyday experience
ESTIMATED STUDENT WORKLOAD

The course contains lectures, reading assignments, and in-class activities/quizzes, as well as lab assignments, including a final project. Students spend three hours in lectures and one hour in labs per week. Each lecture consist of 50 mins presentation by the instructor and 30 mins in-class activities. Presentations are interleaved with in-class activities, in order to allow students to actively engage with concepts and to make theoretical material tangible with hands-on experience. In-class activities include short quizzes, discussion group exercises, computational exercises in the conceptualization of scale, practical explorations of data sets and big data services on the web, and individual exercises in the development and evaluation of data sets, data services, and visual presentations. Assignments deepen the practical part of the learning experience enabling students to apply the presented concepts so as to reach learning objectives. Assignments practice the main steps for developing a geospatial big data service and give hands-on experience with components of geospatial big data systems and their web interfaces. Students are expected to spend about six hours per week on assignments: one hour in labs and five hours outside of classroom. Another three hours outside of classroom are required for course readings. There are ten assignments, which constitute together 50% of the grade. Assignments are one week in length, i.e., require on average six hours of work to complete.

GRADING

Grades will be determined according to the following schema:

- Examinations (30%): mid term (10%), final (20%)
- Lab assignments (50%): ten lab assignments
- In-class activities (20%)

GRADING RUBRIC

Grading criteria follow http://gradeculture.uoregon.edu:

A+ Only used when a student’s performance significantly exceeds all requirements and expectations for the class. Typically very few to no students receive this grade.

A Excellent grasp of material and strong performance across the board, or exceptional performance in one aspect of the course offsetting somewhat less strong performance in another. Typically no more than a quarter of the students in a class receive this grade, fewer in lower-division classes.

B Good grasp of material and good performance on most components of the course. Typically this is the most common grade.

C Satisfactory grasp of material and/or performance on significant aspects of the class.

D Subpar grasp of material and/or performance on significant aspects of the class.

F Unacceptable grasp of material and/or performance on significant aspects of the class.

Grading in basic activities in examinations and lab assignments, evaluates in how far an answer reflects that the question with its background was understood and solved following the methods to be applied in the specific answer. Grading of advanced activities with higher degrees of freedom additionally evaluates the suitability of the choices made, e.g. the method chosen for analysis. Students should make sure that they seek guidance early for these tasks so as to actively discuss alternatives and should justify their choices in write-ups.

Reading assignments should be followed accompanying the lectures and are due Friday evening of the listed week the latest.
COURSE SCHEDULE AND ASSIGNMENTS

WEEK 1

Lecture 1
INTRODUCTION,
Reading: Meier (2015), Chapter 1, pp. 1-24

Lecture 2
THE WORLD AND THE WEB,
Reading: Peterson (2014), Chapter 1, pp. 1-17

Lab 1
A Map Gallery, Peterson (2014), Chapter 4, pp. 61-75

WEEK 2

Lecture 3
THE FIVE VS OF BIG DATA,
Reading: Meier (2015), Chapter 2, pp. 25-47

Lecture 4
PROCESSING MAP DATA,
Reading: Peterson (2014), Chapter 2, pp. 18-45

Lab 2
What is Big Data?

WEEK 3

Lecture 5
HOW THE MIND STORES BIG DATA,
Reading: Peterson (2014), Chapter 3, pp. 46-60

Lecture 6
A BRIEF HISTORY OF MAPS,
Reading: Peterson (2014), Chapter 5, pp. 76-95

Lab 3

WEEK 4

Lecture 7
CROWD COMPUTING SOCIAL MEDIA,
Reading: Meier (2015), Chapter 3, pp. 49-72

Lecture 8
CROWD COMPUTING SATELLITE AND AERIAL IMAGERY,
Reading: Meier (2015), Chapter 4, pp. 73-94

Lab 4
Programming the Web, Peterson (2014), Chapter 8.4-8.8, pp. 148-157

WEEK 5

Lecture 9
MAPS ON THE WEB,
Reading: Peterson (2014), Chapter 6, pp. 96-114

Lecture 10
Midterm,
Reading: from weeks 1-4

Lab 5
Map Mashups, Peterson (2014), Chapter 10.1-10.3.4, pp. 176-191
WEEK 6

Lecture 11
MAP SCALE AND ABSTRACTION,
Reading: Peterson (2014), Chapter 7, pp. 115-135

Lecture 12
MAP DIGITIZING AND GPS,
Reading: Peterson (2014), Chapter 9, pp. 158-175

Lab 6
Map Mashups, Peterson (2014), Chapter 10.3.4-10.9, pp. 191-204

WEEK 7

Lecture 13
POINTS AND POINT DATA,
Reading: Peterson (2014), Chapter 11, pp. 205-219

Lecture 14
THE ONLINE MAP,
Reading: Peterson (2014), Chapter 13, pp. 250-266

Lab 7
Point Maps and Graduated Symbols, Peterson (2014), Chapter 12, pp. 220-249

WEEK 8

Lecture 15
MAP LAYERS AND GIS,
Reading: Peterson (2014), Chapter 15, pp. 296-310

Lecture 16
DATABASES, MYSQL, AND PHP,
Reading: Peterson (2014), Chapter 17, pp. 325-334

Lab 8
Line and Area Mashups, Peterson (2014), Chapter 14, pp. 267-295

WEEK 9

Lecture 17
ARTIFICIAL INTELLIGENCE FOR HANDLING BIG DATA,
Reading: Meier (2015), Chapter 5, pp. 95-110

Lecture 18
ARTIFICIAL INTELLIGENCE IN REMOTE SENSING,
Reading: Meier (2015), Chapter 6, pp. 111-124 , Reading:

Lab 9
Map Layer Mashups, Peterson (2014), Chapter 16, pp. 311-324

WEEK 10

Lecture 19
VERIFYING BIG DATA,
Reading: Meier (2015), Chapter 7 & 8, pp. 125-154

Lecture 20
Final Examination Q & A,
Reading: from weeks 1-10

Lab 10
Mapping from a Database, Peterson (2014), Chapter 18, pp. 335-360