# Spatial Databases

## Course Outline

### Contact Information

**Instructor:**      Peter Glenday,
                     PhD Candidate, Department of Geography and Environmental Management
**Office:**          EV1-219
**Office Hours:**    TBD
                     If you need to schedule an appointment outside of these drop-in hours, please contact me.
**Phone:**           (519) 888-4567 ext. 30381
**E-mail:**          pjglenda @ uwaterloo.ca
                     From Monday to Friday, I make every effort to answer emails within 24hrs. Email sent on the
                     weekend will normally be answered on the following Monday.

### Teaching Assistants

| Assistant | Email | Lab/Office hrs |
|-----------|-------|----------------|
| TBD | TBD | TBD |
| TBD | TBD | TBD |
| TBD | TBD | TBD |

### Course Presentation

**Lecture:**  PAS 2083
              Monday, 2:30-5:20 pm.

**Reserved lab time:** EV2-1002A Geddes
              Lab 101  Thurs. 9:30-11:20 am
              Lab 102  Thurs. 11:30-1:20 pm
              Lab 103  Thurs. 6:30-8:20 pm
              Lab 104  Thurs. 4:30-6:20 pm

### Overview

Geospatial data and information are used in a wide variety of planning, facilities management, business, resource management, location based services, and academic research. A key component of any spatially enabled solution is the underlying geospatial database that must be designed to support efficient data maintenance and analysis in a heterogeneous operating environment i.e., different users sharing common data whilst using a variety of client technologies to create, access, extract, transform, load and manage the data.

We are in an era of Big Data with exabytes of structured and unstructured georeferenced and geospatial data collected each day by citizens, public agencies and private enterprise. Consequently, managing, processing and using the data to derive useful, actionable knowledge/information presents significant challenges, usually referred to as the four V's of big data: Volume; Velocity; Variety, and; Veracity. Traditional GIS and CAD methods of storing and processing data involved Shape, DXF, or LAS files or a proprietary geodatabase. Files and tiles (each feature

having its own file replicated for multiple map sheets or tiles) are ill suited to storing, managing and processing large volumes data in multiuser, technologically heterogeneous environments.

With regards to conducting geospatial analysis, big data is big and it's more efficient to move small algorithms to the big data than moving large data volumes to the algorithms. In response to the desire to better manage geospatial data and move the algorithms to the big data, commercial Database Management Systems (DBMS) software vendors (Oracle, Microsoft, IBM, etc.) and open source e.g., PostGreSQL PostGIS geospatially enabled the RDBMS with Open Geospatial Consortium (OGC) native spatial datatypes and SQL to query the data. Moreover, the data model evolved from a file (.shp) containing geometry linked to attributes in a data file (.dbf) to an object model wherein the geometry is an attribute of the object (or record/row in a database) with multiple data management, processing, and end-user benefits.

This course is an introductory-level database design, implementation, management, and access course for students interested in organizing spatial and non-spatial data in geospatially enabled enterprise Relational Database Management Systems (RDBMS). Students will gain a basic understanding of database systems terminology, design principles, and issues facing database managers and analysts. We will also present a high-level discussion of Not Only SQL (NoSQL) databases and Hadoop Distributed File System (HDFS) concepts as they relate to geospatial data. They will gain practical database experience utilizing commercial database management system software and geographic information systems software to design and use spatial and non-spatial data.

This course is a prerequisite to the Geography / Planning 487 courses and part of the requirements for the *Diploma of Excellence in GIS*. <u>Prerequisite</u>*:* Successful completion of Geography or Planning 281.

## Objectives

The overarching goals of this course are: to develop a sound understanding of the principles and techniques of database design and development as they apply to geospatial data; apply these principles and techniques in designing and building geospatial databases; use geospatial databases to perform common types of queries and geospatial analyses; and publish spatial data to the internet.

Specifically, the objectives of this course are as follows:

1. To provide students with an introduction to the fundamental social and managerial dimensions of geospatial data creation and use.
2. To introduce students to enterprise geospatial database management systems, principles, terminology, architecture and technology,
3. To provide students with experience creating ,documenting, editing, managing a database using a variety of software systems and tools, including:
    - ArcGIS, ArcCatalog, QGIS
    - Safe Software's Feature Manipulation Engine (FME)
    - Oracle Spatial & Graph , Oracle SQL Developer, Oracle SQL Modeler
    - Oracle Mapviewer API, MapBuilder, Oracle JDeveloper
4. To introduce students to ArcGIS' and Oracle's commonly-used spatial analysis, query and information production procedures when connected to an RDBMS.

Upon completing this course, students will be able to:

- Develop and document a conceptual database design for a specific applied geospatial problem.
- Populate a database through a variety of means (e.g., data migration, conflation,  and creation via analysis)

- Build and use key database components (e.g. attribute domains, topology rules, relationship classes) to maintain data integrity and facilitate database queries
- Possess knowledge regarding the development, operation and application of Enterprise Geospatial Database Management Systems for managing geospatial data originating from a variety of sources (CAD, GIS, GPS, Imagery & Open Source).

## Text and Readings

There is no text for this course. Course notes, lecture presentations, links to readings and tutorials will be provided in this syllabus and on the **Desire2Learn** course web site (http://learn.uwaterloo.ca).  Students are expected to make extensive use of the ArcGIS, Oracle Spatial and FME on-line help and web resources.  Selected links will be provided on the course website.

## Schedule

The class meets from 2:30-5:20pm on Fridays in PAS 2083.  Class time will be used to introduce concepts and techniques related to spatial database design and development, demonstrate methods needed for the lab assignments, and deal with questions related to the lab exercises. Table 1 details the course schedule, lecture topics, required readings, resources, tutorials, assignment hand out and submission dates, and a final test with both a take-home and in-class portion.

**N.b.** The instructor reserves the right to make changes to the schedule and content as necessary.

## Learning modes and course format

This course builds upon the understanding of GIS concepts you gained in Geography / Planning 281 through a series of compulsory lectures, tutorials and lab sessions. Lectures will be used to discuss concepts, principles and techniques of geospatial database analysis, design and implementation coupled with access, query, and reporting processes using data held in a database.  The third hour of the class will be used for workshops, lab demonstrations and/or for you to work on your assignments.

Hands-on work with ArcGIS, FME and Oracle Spatial, SQL Developer/Modeler will take place in the Galileo lab (EV1-240). The lab assignments are designed to build your skills in using GIS and database software systems and tools to strengthen your understanding of how geospatial analytics can be properly applied to real world problems. The Galileo lab (EV1-240) has been reserved GEOG/PLAN 387. See above for dates/times.

Note that developing a strong understanding of geospatial database concepts and the corresponding practical skills necessary to complete the assignments requires a significant investment of time.  **In addition to the scheduled class and lab time, students should expect to spend at least 5 hours per week working on course assignments. *These time requirements will vary from student-to-student***. Students who are unwilling or unable to make this time commitment should consider other courses.

## Resources

### Computer Labs

Make use of your scheduled lab time in the Geddes (EV2-1002A).  You can also use the Galileo (EV1-240) lab for practical work when they are not booked for other courses.

**Note:** No food or drink is allowed in the labs. Failure to abide by this rule may result in your computer accounts being suspended.

## Course Website

A website for this course has been created as part of the **Desire2Learn** (D2L) system. Students in the course can access the course website by going to the D2L homepage ([http://learn.uwaterloo.ca](http://learn.uwaterloo.ca)) and entering their UW userid and password in the logon form displayed on this page.  Once you are logged on to D2L, you will see a list of courses that you are registered in.  Click on GEOG/PLAN 387 to select this course.

Table 1. Course Schedule

| Date | Lecture Topics | | Workshops/Demos | Readings & Resources | Tutorials | Assignments |
|---|---|---|---|---|---|---|
| Sept. 09 | Course Overview<br><br>Evolution of GIS & Future Trends<br><br>3 tier architecture GIS | Structured & unstructured data<br><br>Big data<br><br>Assignment 1 briefing | ArcGIS 10.4 Creating a Geodatabase in ArcCatalog Loading External Data FME Workbench | Newell, R. & Theriault, DG (C.1990). Ten difficult problems building a GIS. Avail on course website<br><br>Hart G. and Dolbear C. (2007). **What's So Special about Spatial?** In: The Geospatial Web. Springer, , pp.39-44. | ESRI Building a GeoDB 1-4 Introduction to Editing 1-2<br><br>ESRI Introduction to Editing 4-5 Building a GeoDB 6-8 | Hand out. Assignment 1. (10%) ESRI Geodatabase Basics |
| Sept. 16 | Geospatial Data Models<br><br>Geospatial Data Modeling Projections & Coordinate Reference Systems | Geospatial Data Standards<br><br>Geospatial Metadata | | Goodchild, M. (2001). **A Geographer Looks at Spatial Information Theory.**<br><br>Güting, R (1994). **An introduction to spatial database systems** The VLDB Journal,3(4),pg.357-399 | | |
| Sept. 22 | | | | | | Assignment 1 Due 11:55pm |
| Sept. 23 | Spatial vs Non spatial databases<br><br>Database Types<br><br>SQL Introduction | Database management system architecture: conceptual, logical, and physical | Assignment 2 briefing<br><br>FME Workbench Introduction to Oracle Spatial SQL Developer SQL Developer DM | Book: Chapter 1 of **Pro Oracle Spatial 11g** Avail. online via uwaterloo lib.<br><br>Book: Preface & Chapter 1 of **Applying and Extending Oracle Spatial** Avail. online via uwaterloo lib. | FME Workbench Chapter 2 Format Trans. Chapter 3 Data Restructure | Hand out. Assignment 2. (20%) Building a database |
| Sept. 30 | Database Design Strategies System Dev Life Cycle (SDLC) CRUD | Data Models Entity-Relationship (E-R) Models Normalization Data Dictionary | | | FME Workbench Chapter 4: Content Trans. Chapter 5: Data Reproject | |
| Oct. 06 | | | | | | Assignment 2 Due 11:55pm |

Table 1.Continued... Course Schedule

| Date | Lectures | | Workshops/Demos | Readings & Resources | Tutorials | Assignments |
|---|---|---|---|---|---|---|
| Oct. 07 | Advanced Spatial SQL - Spatial Operators Spatial Indexes Spatial Filters | Short & Long Transactions Version Management | Oracle SQL Developer SQL Modeler FME .shp <-> Oracle | Smallworld Technical Paper No. 4 - Version Management in GIS - Applications and Techniques Oracle Technology Network http://www.oracle.com/technetwork/index.html | Tutorial: Getting Started with Oracle Spatial SQL Reference tool W3 Schools – SQL ref. | In class test (10%) T/F & Written answer 50 minutes<br><br>Hand out. Assignment 3. (20%) |
| Oct.14 | No Lecture | | | | | |
| Oct. 21 | Rasters in RDBMS | | Oracle Georaster | | | |
| Oct. 27 | | | | | | Assignment 3 Due 11:55pm |
| Oct. 28 | 3D Terrain and Surface Modeling | TINs, Point Clouds | Assignment 4 briefing Searching and retrieving Open Source Geospatial Data | OSGeo   http://www.osgeo.org/ OpenStreetMap http://www.openstreetmap.org/ | | Hand out. Assignment 4. Part A (10%) Enterprise Geospatial Database & Dashboards |
| | | | | | | |
| Nov.04 | Network Utility Modeling Schematics | Guest Lecture: TBD | Enterprise Geo DBMS for Network Utilities | | | |
| Nov.10 | | | | | | Assignment 4 Part A due 11:55pm |
| Nov.11 | Modeling the Built Environment | Building Info Models (BIMs) 3D City GML | | CityGML   http://www.citygml.org/ 3D City GML Database   http://www.3dcitydb.net/ | | Hand out. Assignment 4. Part B (20%) Enterprise Geospatial Database & Dashboards |
| Nov.18 | Spatial Data Infrastructures (SDI) | Guest Lecture: TBD | TBD | | | |
| Nov.25 | Geospatial Web HTML5 CSS Javascript | Citizen Sensor Big data | | Text: The Geospatial Web Avail. online via uwaterloo lib. http://www.springerlink.com/content/uu3847/#section=26957&page=1 | Javascript HTML5 CSS GML | |
| Dec.01 | | | | | | Assignment 4 Part B due 11:55pm |
| Dec.02 | Brief Course Summary | Course Evaluation Forms | | | | In class test, 10% of overall grade |

### Getting Help

Students are expected to get into the habit of using the on-line resources as the first source of problem solving and help.  I will be available in my office for consultation during regular office hours or by appointment.  Additional help is available during the scheduled lab sessions and from the MAD help desk.

## Evaluation

Lab assignments will account for 80% of the final mark and will be weighted as indicated in the course schedule table.  There will be a test at the end of the term which will count for 20% of the final mark in the course.

Assignments must be submitted to the appropriate digital drop box on the course website by 11:55 p.m. on the specified due date.  <u>Late assignments will not be accepted and will receive a mark of 0</u>. The assignments have been designed to be completed in the allotted timeframe.  Exceptions may be made for documented medical reasons following the UW Academic policies listed below.

The end of term test will consist of 2 parts and count for 20% of your final mark. The first part will consist of a take-home test (10%) that will be assigned on Monday March 21 2015 and must be submitted by 5:00pm Friday, Mar.25, 2015. The second part will be an on-line quiz (10%) completed during the last hour of regular class time on Monday, Mar.28, 2015 with the class split evenly between the Galileo and Geddes labs.

## UW Academic Policies

<u>Academic Integrity:</u> To create and promote a culture of academic integrity, the behaviour of all members of the University of Waterloo is based on honesty, trust, fairness, respect and responsibility.

<u>Grievance:</u>  A student who believes that a decision affecting some aspect of his/her university life has been unfair or unreasonable may have grounds for initiating a grievance. Read Policy 70 - Student Petitions and Grievances, Section 4, http://www.adm.uwaterloo.ca/infosec/Policies/**policy70**.htm

<u>Discipline:</u> A student is expected to know what constitutes academic integrity, to avoid committing academic offenses, and to take responsibility for his/her actions. A student who is unsure whether an action constitutes an offense, or who needs help in learning how to avoid offenses (e.g., plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course professor, academic advisor, or the Undergraduate Associate Dean. When misconduct has been found to have occurred, disciplinary penalties will be imposed under Policy 71 – Student Discipline. For information on categories of offenses and types of penalties, students should refer to Policy 71 - Student Discipline, http://www.adm.uwaterloo.ca/infosec/Policies/policy71.htm

<u>Appeals:</u> A student may appeal the finding and/or penalty in a decision made under Policy 70 - Student Petitions and Grievances (other than regarding a petition) or Policy 71 - Student Discipline if a ground for an appeal can be established. Read Policy 72 - Student Appeals, http://secretariat.uwaterloo.ca/Policies/policy72.htm

<u>Ethics:</u> The University of Waterloo requires all research conducted by its students, staff, and faculty which involves humans as participants to undergo prior ethics review and clearance through the Office of Research Ethics (http://iris.uwaterloo.ca/ethics/). The ethics review and clearance processes are intended to ensure that projects comply with the guidelines established through Tri-Council Policy Statement (TCPS).

**Note for students with accessibility needs:**   AccessAbility Services, located in Needles Hall, Room 1132, collaborates with all academic departments/schools to arrange appropriate accommodations for students with accessibility-related needs without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with AccessAbility Services at the beginning of each academic term.

**Illness or other emergency:**   Students with an illness, or other emergency, that prevents them from handing in an assignment on time or attending a test, must contact you to notify you of this problem prior to the deadline or test date (or as soon afterwards as is reasonably possible). Students with an illness must document evidence of that illness with a note from a doctor. For other emergencies, such as the death of a family member, students should be asked to meet with you as soon as possible in order to make arrangements to make up any missed assignment or test. Any questions or concerns that you have regarding these matters should be directed to the Department Chair or Associate Chair.

**Religious Observances:**   Please inform the instructor at the beginning of term if special accommodation needs to be made for religious observances that are not otherwise accounted for in the scheduling of classes and assignments.

**Turnitin Software:**   Plagiarism detection software (Turnitin) may be used to screen assignments in this course. This is may be done to verify that use of all materials and sources in assignments is documented. Students will be given an option if they do not want to have their assignment screened by Turnitin. In the first week of the term, details will be provided about arrangements and alternatives for the use of Turnitin in this course.

**Annotated bibliography:**

For advice on how to prepare an annotated bibliography, see:

**http://www.lib.sfu.ca/help/writing/annotated-bibliography**