GEOG 384: Principles of the Geospatial Web

Renee Sieber, renee.sieber@mcgill.ca

Geospatial Data Administrator, Ian Tattersfield, ian.tattersfield@mcgill.ca

Sieber Office Hours Tuesday, Thursday, 2:30-3:30pm, 411 Burnside Hall
TA: Sam Lumley, sam.lumley@mail.mcgill.ca

Introduction
Most often when students think of geographic information systems (GIS), it is a map making tool, a complete yet expensive, desktop-bound software package requiring considerable computational power. We now work in a software "stack" of interoperable and "mashed up" web 2.0 applications, many of which are free and require no more processing power than being able to run a basic web browser. The learning curve has changed as well. Where once a GIS technician entered, managed and visualized data as part of a specific learning curve, these functions are being replaced by smarter software and the skillset is shifting to coding by neogeographers.

The nature of geographic data has changed as well. Large and streaming datasets, called big data and are derived from sources such as social media and mobile sensors, are commonplace. Entire industries have sprung up to harness the power of interpreting patterns revealed by the data. These industries also have the capability to influence policy and society as they monetize what has been termed the "ambient awareness" of pervasive space-time data collection. The sheer amount of data with spatial attributes, which is produced in a day, could not have been created a mere ten years ago. Whereas the underlying technology has rapidly evolved, our understanding of the information we generate has not kept pace. This calls to question issues of accuracy, rights, and privacy. Should we be able to scrape for user generated content and trust in its validity? Or should we use uncritically utilize crowdsourced asserted data (i.e., location-based sentiments and observations from non-experts called volunteered geographic information {VGI})? The data is far more unstructured and is sourced to a myriad of multimedia (e.g., videos, graphs, and photos). This geospatial data constitutes a departure from traditional authoritative datasets of latitude, longitude, projection and attributes, and introduces varied data structures, potentially disputed sources and one record of information at a time. The user is simultaneously the producer of new forms of spatial data, as well as user and commodity.

In this course we will cover the principles of this paradigm, called the Geospatial Web or Geoweb. Like GIS, the Geoweb has broad applicability that extends well beyond the traditional discipline of geography. Numerous examples can be seen in health provision (e.g., fitness tracking), wildlife and natural resources (e.g., with sensors), criminology (e.g., predictive mapping of likely offenders), and transportation (e.g., Uber, Lyft, as well

as docked/dockless vehicles). The Geoweb can be used in retail marketing, political mobilization, tourism development, journalism and the humanities. The Geoweb also can be fully linked to social networking platforms. Perhaps the most lucrative use of the Geoweb, however, has been in location-enabled mobile apps that we use every day. Have you ever used a physical map to navigate space or do you seek directions to a new location via Google or Siri/Alexa?

The course will offer a combination of theory and practice of the Geoweb. In terms of theory, we will cover subjects like VGI and big data and how it changes our ideas about spatial data accuracy. In practical lab sessions students will design and develop Geoweb "apps". An additional goal is not to learn specific skills, but to 'learn how to learn'. Geoweb software changes monthly so students learn of resources and strategies to effectively use emerging technologies and anticipate innovations in geospatial technologies. The course will cover the following topics:

- Exploring the underlying methods of digital earth architectures, including georeferent systems. These architectures underlie most platforms (e.g., Google Maps/ Earth, OpenStreetMap and OpenLayers)
- Exploring political, sociological, economic and legal issues in using VGI and platforms
- Critically analyzing the concept of VGI, for example, spatial data accuracy and uncertainty of heterogeneous data sources (spatial data quality)
- Understanding the infrastructure of the Geoweb, including the geospatial software stack, Application Program Interfaces (APIs), and the Document Object Model (DOM)
- Comparing and contrasting GIS and the Geoweb
- Repurposing geographic digital content (secondary data), for example via web scraping
- Learning underlying concepts of server/cloud geospatial applications
- Understanding issues related to real time streaming data (e.g., changes in sampling and geostatistics)
- Working with location-based services and mobile platforms
- Identifying basic problem-solving requirements for geospatial apps. For proposed solutions, critically evaluating and justifying various existing and emergent geospatial technologies and enabling software stacks.

We also have the opportunity to hear from experts in the field. We anticipate having guest lectures from representatives of Geoweb firms like Mapbox and Stamen Design representatives of traditional GIS firms like ESRI, coordinators of VGI sites like OpenStreetMap, and faculty members working in the Geospatial Web (e.g., researching copyright laws for geospatial data).

Finally and time willing, we are planning some workshops outside class time to assist in learning. Among possible topics are HTML and CSS, the Document Object Model and Scalable Vector Graphics. Stay tuned for times and locations.

Course Prerequisites
GEOG 201 and COMP 202 are required (COMP 202 can taken in parallel). OR permission of the instructor.

Books and Other Reading Material
During the course, students will be required to read a variety of articles, white papers and other material. These will be posted on the course website (the course has its own website). The student is required to use Udemy (specifically https://www.udemy.com/javascript-for-beginners-learn-javascript-from-scratch/). If you require additional help in learning material such as KML, web scraping, or Javascript, you may wish to purchase books (e.g., *Eloquent JavaScript: A Modern Introduction to Programming*). There are excellent online resources as well—freecodecamp, Mozilla Web Developer, W3Schools and a free version of CodeAcademy (these last two have been used in the course before). This course assumes you will pick up HTML and CSS; the afore-mentioned sites are excellent resources for those as well.

Evaluation
In-class and online participation 10% (5% in-class and 5% online)
In class quizzes – 20%
Assignments – 40%
Final Exam – 30%

In-class and online participation: Student preparation and participation as well as performance during class will be assessed for half (5%) of the participation grade. We are looking for quality of contributions over the quantity of contributions. Class participation will be evaluated based on evidence that students have read assigned readings, done exercises and otherwise prepared for class. Students also will be assessed in their ability to thoughtfully and reflectively build on other students' contributions. The other five percent of the participation grade will be based on online contributions on Twitter, using the hashtag #neogeoweb. Like above, quality of contributions in tweets is preferred over number of tweets. Alternate to Twitter, students may contribute to the course Wiki database of relevant geospatial web knowledge. The wiki can be found at http://neogeoweb.ca/wiki.

Assignments: Assignments allow the student to apply lecture material and/or programming to real world cases of geospatial representation on the web. There are five assignments in the course. All assignments will be done in groups of three to four students. The goal is to balance levels of computational experience so every group possesses a similar level. All assignments are graded as a group regardless of individual contribution.

Note: The temptation in all assignments is to let the experts or intermediates in the team handle the hard technical details. RESIST THIS TEMPTATION because it will hurt you in the final lab exam, which is individual-based. As you know from working with GIS or RS, knowing the tech ONLY comes from doing (and occasionally, creatively failing) the tech.

Quizzes: Approximately half of the assignments require the use of a specific coding language, which is Javascript. Quizzes will be announced the class prior to the day the quiz is held. You will be expected to learn the udemy material up to that point (udemy course schedule TBA). You must email a .js file (TBA) by the instructor by the beginning of Week 6. This file will demonstrate that you have completed the required Udemy module.

It is possible to waive out of the Udemy requirement if you know JS. Contact the instructor for details.

Final Exam: The final exam is comprehensive of all material in the course. It is divided into two equally weighted components—a written exam and a lab practical exam.

Obligatory Statements

1. McGill University values academic integrity. Therefore, all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures (see www.mcgill.ca/students/srr/honest/ for more information). (approved by Senate on 29 January 2003)
2. In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded. (approved by Senate on 21 January 2009 - see also the section in this document on Assignments and evaluation.)

Syllabus
Introduction Overview of Geoweb and goals of the course. In-class exercise: Work with existing Geoweb platform, Mapbox GL JS. Set up Twitter account for online participation Get JSFiddle account Answer a few questions about your comfort level with various concepts Get set up on the course server Purchase and get started on Udemy
Module 1: Digital Earths, the prime mapping platform in the Geoweb. Definitions and types of digital earths. Georeferent systems, tiling, and projections. Differences between data structuration in GIS compared to the Geoweb. First introduction to web hosting. Introduction to concepts of markup languages (XML, HTML) and as used on Digital Earths (KML). In-class exercise: Create your first dataset in kml, with one point, line, polygon, extrusion, photo. Assignment 1: Write a smart city welcome page, using KML, HTML and CSS
Module 2: Treating Everything You See in Cloud as a Data Source Introduction to web architectures, software stacks. Principles of web harvesting/scraping. Legal issues in "repurposing" data, for example, intellectual property and liability. Example from tourism. Structured and unstructured data. In-class exercise: Conduct web scraping. Tag content with XML. Create a web page Assignment 2: Use Google Spreadsheets and XPath to automatically scrape and then map a popular classified advertising site
Module 3: Democratization of Data: Volunteered Geographic Information (VGI), the crowd and Beyond Concepts in VGI (e.g., citizen sensors, crowdsourcing, and neogeography). Motivations for volunteers to contribute. Common methods to assessing accuracy of VGI. Legal issues underlying VGI (e.g., copyright and intellectual property of using citizen-generated content). Emergence of citizen science, crisis mapping and open data. In-class exercise: Create and edit your own VGI with Google Forms. Assignment 3: Create tracker app with the Mapbox on the browser with Javascript and utilizing Google spreadsheets as the Internet database

Module 4: Streaming, real time BIG Geospatial Data

Introduction to concepts of big data and data-intensive science. Brief discussion of geosensors, which are a prime source of geolocated data. Challenges to working with big data, like sampling.

In-class exercise: Collect and manipulate Twitter data

Assignment 4: Create a transportation app with bixi bike or streaming bus data

Module 5: Geo- and Data Visualization

Principles of cartographic/ geo-visualization and the emergence of data visualization. Exploratory data visualization vs explanatory data visualization. Visualization as storytelling.

In-class exercise: Explore the "maps" of D3JS. Tell a story with a Geoweb platform.

Module 6: Geoweb on Mobile Devices

Concepts of geospatial awareness on devices--location based services (LBS), including location intelligence, vehicle tracking, and RFIDs. Examples from mobile commerce. Social issues in LBS (privacy, surveillance)

In-class exercise: Mobile messaging and mapping of geographic data

Assignment 5: Conduct data journalism using D3JS using heroku or jsfiddle

TBD: Final exams