GEOG 384: Principles of the Geospatial Web

Tuesdays and Thursdays, 11:35-12:55, Burnside 511

Instructor, Prof. Renee Sieber, 411 Burnside renee.sieber@mcgill.ca
Sieber Office Hours Tuesday, Thursdays, 1:30-2:30pm

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

TA: Elliot Vosburgh, Burnside 428, elliot.vosburgh@mcgill.ca

Introduction
Most often when students think of geographic information systems (GIS), one company and its software suite comes to mind that provides a one stop shop for data management, analysis and creation of point and click, ready made web app showcases - an closed ecosystem if you will. This old mindset of GIS as a mapmaking tool, a complete yet expensive, computer-bound software package requiring considerable computational power does not represent the bounty of options users now have. Indeed, most geospatial analysis and mapping more often now is in a software "stack" of interoperable and "mashed up" web 2.0 applications, some 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, these functions are being replaced by "smarter" algorithms with the skillset having shifted to coding by non-geographers, engineers and computer scientists. We may soon see the day where all you need is to create a prompt and seconds later your data is sourced, analysis is performed and an application appears that lets you interact with the result. We are not, however, quite there yet.

This smarter software is aided with generative AI and geofoundation models (colloquially known as GPTs or generative pre-transformers). GPTs promise to write perfect computer code, craft explainable documentation and reason over geographic space (function multi-modally e.g., by integrating coordinates and semantics, automatically processing remotely sensed imagery). GPTs propose to basically cut humans out of the loop, although they are also prone to hallucination and geospatial inaccuracies (e.g., asking early MS CoPilot what is the distance from Canada to the US?).

The nature of geographic data has changed as well. Large and streaming datasets, called big data and scraped from sources such as social media and mobile sensors, are commonplace. This includes facial pattern recognition from autonomous vehicles and CCTV cameras. As long as we have a phone, we are almost constantly generating data with location in it. Entire industries have sprung up to harness the power of interpreting patterns revealed by alphanumeric and image 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, produced in a single 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 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. GPTs are the newest addition, enabling the mass harvesting of content, regardless of intellectual property, and generation of text, audio and imagery from text prompts. Thus, 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, epidemic control), wildlife, climate science and natural resources (e.g., with sensors), criminology (e.g., predictive mapping of likely offenders, redlining), warfare and surveillance (e.g., swarming drones, doorbell cameras), psychology (e.g., therapy AI) and transportation (e.g., Uber, Lyft, as well as docked/dockless vehicles). The Geoweb can be used in retail target marketing, political mobilization (e.g., disinformation), tourism development, journalism and the humanities. The Geoweb also is fully linked to social networking platforms and mobile apps. For instance, how often do you use a physical map to navigate? Probably only when its more convenient like in a pamphlet or kiosk.

The course combines theory and practice of the Geoweb. Students won’t learn specific skills, but ‘learn how to learn’. Geoweb software changes constantly so students need to learn about resources and strategies to 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 all digital earth platforms (e.g., Google Maps and OpenStreetMap)
- 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

- Understanding issues related to real time streaming data (e.g., changes in sampling)
- Working with live-location from mobile devices
- 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
- Implementing GPTs for writing and interpreting code

Finally, and time willing, we will hear from experts and host some workshops outside class time to assist in learning. Experts will come from the geospatial industry or are practical users (e.g., in cities). Among possible topics are GeoAI, the Document Object Model, and Scalable Vector Graphics. Stay tuned for times and locations.

Course Prerequisites
GEOG 201 and COMP 202/GEOG 333 are required (COMP 202 can be taken in parallel). OR permission of the instructor.

Required Books and Other Reading Material
During the course, students will read a variety of articles, working papers and other material. These will be posted on the course website (neogeoweb.ca). Students must learn Javascript. The reason why is because JavaScript remains the de facto language for web development. Moreover, this course assumes no prior web dev experience and minimal coding experience so learning JavaScript, which most students in the course don't know, levels the playing field.

This will be done via W3Schools. We have tried many platforms in the past. There is no one perfect platform for learning code online. Many sites that were free have now moved to a fremium model, which have placed practice tests and certificates behind paywalls. If you require additional help in learning material such as geoJson, markup languages, web scraping, or Javascript, you may wish to purchase books. There are other excellent online resources as well—freecodecamp or javascript.info. This course assumes you will pick up HTML and CSS; the afore-mentioned sites are excellent resources for those as well.

We encourage students to seek assistance from GPTs like MS CoPilot. Keep in mind that GPTs can offer some preliminary code generation but be aware that you need to evaluate the quality of the code. Additionally, when you seek help for APIs like Mapbox, it will generate the exact same code available on the Mapbox site so sometimes it's best to return to the source.

Evaluation

Title	Weight	Description	Due Date	Considerations and Late Penalties
Participation	10% (5% in-class	Participation in class discussions, asking	NA	Grades for in class participation may

	and 5% online)	relevant questions, constructively engaging other students' responses. Same rubrics apply online.		be adjusted for classes missed for valid reasons.
Quizzes	20% (2* 10%)	Short questions that require coding and debugging.	Announced 1 week prior to day the quiz is held	If 1 quiz missed then other quiz will count for both.
Assignments (4)	30% (each 15%-25% of that 30%)	Group-based app development	One approximately every 2 weeks.	Late assignments receive a 10% penalty per day.
Final Exam	40% (2 parts *20% each)	Part 1: short answer questions. Part 2: coding, which resembles quiz & practice questions.	In department but scheduled during the final exam period.	Missed final exams are handled by instructor and Service Point.

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 our Discord. Similar to in-class participation, students should not only contribute original content but also react to their colleagues' content.

Assignments: Assignments allow the student to apply lecture material and/or programming to real world cases of geospatial representation on the web. There are four assignments in the course. All assignments are completed 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 1: The temptation in all assignments is to let the experts in the team handle the "hard" technical details. RESIST THIS TEMPTATION because it will hurt you in the quizzes and final lab exam, which are individual based. As you know from working with GIS or RS, knowing the tech ONLY comes from doing (and occasionally, creatively failing) the tech.

Note 2: With the advent of GPTs, the temptation is to have transformers to write all of the code for you. RESIST THIS TEMPTATION to do 'vibe coding' because it will hurt you in the quizzes and final lab exam, which are individual based. Moreover, learning the logic of coding (like the logic of a natural language) requires practice, for example, how to debug software.

Quizzes: Approximately half of the assignments require the use of JavaScript, which the quizzes are based on. There will be two one-hour quizzes during the semester. The dates

of quizzes will be announced the week prior to the day the quiz is held. You will be tested on JavaScript material up to that point.

Learning JavaScript: During the course you need to demonstrate your knowledge of JavaScript. To do this, you must email a screenshot from W3Schools by the instructor by the beginning of Week 6. This screenshot will demonstrate that you have completed the required modules. It is possible to waive out of the coding requirement if you know JavaScript. 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. The final exam is *four* hours long with a one-half hour break in between.

Obligatory Statements

Assessments in this course are governed by the Policy on Assessment of Student Learning (PASL), which provides a set of common principles to guide the assessment of students' learning. Also see Faculty of Science-specific rules on the implementation of PASL.

Legally mandated academic accommodations are handled by Student Accessibility and Achievement. For more information see https://www.mcgill.ca/access-achieve/

In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French written work that is to be graded. This does not apply to courses in which acquiring proficiency in a language is one of the objectives." (Approved by Senate on 21 January 2009)

Conformément à la Charte des droits de l'étudiant de l'Université McGill, chaque étudiant a le droit de soumettre en français ou en anglais tout travail écrit devant être noté, sauf dans le cas des cours dont l'un des objets est la maîtrise d'une langue. (Énoncé approuvé par le Sénat le 21 janvier 2009)

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" (Approved by Senate on 29 January 2003) (See McGill's guide to academic honesty for more information).

In the event of extraordinary circumstances beyond the University's control, the content and/or assessment tasks in this course are subject to change and students will be advised of the change.

ADDITIONAL STATEMENTS

We endeavor to provide an inclusive learning environment in this course. If you are experiencing barriers to learning in this course, do not hesitate to discuss them with a member of the teaching team and/or the Office for Student Accessibility and Achievement, 514-398-6009.

We value intellectual freedom and we also value creating a safe and respectful classroom environment. If you do not feel safe in the classroom by virtue of comments related to your race, gender, sexual orientation, physical ability or for any other reason, please come speak with the course instructors.

Student well-being is a priority for the University. All of our health and wellness resources have been integrated into a single Student Wellness Hub, your one-stop shop for everything related to your physical and mental health. If you need to access services or get more information, visit the Virtual Hub at mcgill.ca/wellness-hub or drop by the Brown Student Services Building (downtown) or Centennial Centre (Macdonald Campus). Within your faculty, you can also connect with your Local Wellness Advisor (to make an appointment, visit mcgill.ca/lwa).

McGill has policies on sustainability, paper use and other initiatives to promote a culture of sustainability at McGill. (See the Office of Sustainability.)

Last, due to transformations of technologies like foundation models, the syllabus may change slightly.

Syllabus

Introduction

Overview of Geoweb and goals of the course.

In-class exercise: Work with existing Geoweb platform, Mapbox GL JS.
Set up to the Discord for online participation
Sign up at OpenAI and/or copilot
Sign up to W3 Schools Spaces
Answer a few questions about your comfort level with various concepts
Get set up on the course server
Sign up to the Discord channel
Sign up for Mapbox account
Create a story map https://www.mapbox.com/solutions/interactive-storytelling/

Topic 1: History and evolution of geospatial data handling and visualization.

Definitions and types of digital earths (e.g., Google Maps). Georeferent systems, tiling, and projections. Differences between data structuration in GIS compared to the Geoweb. Introduction to concepts of markup languages (XML, HTML) and list-based ways of storing data.

In-class exercise: use geojson.io Create your first dataset in geojson, with one point, line, polygon, extrusion, photo.

Assignment 1

Topic 2: Everything You See Online is 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: Work with Unstructured Content

Conduct web scraping. Tag content with XML. Create a web page. Explore what happens when you prompt a GPT to scrape web content.

Assignment 2

Topic 3: Fundamentals of Coding in an Era of GPTs

- Variable Typing
- Data Structures
- Control Logic
- Functions
- User Input, Prompts, Alerts and Event Coding
- Debugging (e.g., use of console.log)
- The foundation of foundation models
- The opportunities & limits of GPTs for coding

In-class exercise: Write sample code with control logic. OR use copilot.

Topic 4: Democratization of Data and Algorithms: Volunteered Geographic Information (VGI) and GeoAI Ethics

Concepts in VGI (e.g., social sensing and crowdsourcing). Motivations for volunteers to contribute. Common methods to assessing accuracy of VGI. Legal issues underlying VGI (e.g., copyright and intellectual property {IP} of using non expert-generated content).

Fundamentals of GeoAI Societal impacts (e.g., privacy, surveillance, IP, environmental impacts); Ethics of Geofoundation Models (i.e., GeoGPTs).

In-class exercise: Create and edit your own VGI.

Assignment 3

Topic 5: 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: Envision a stream of data

Assignment 4

Topic 6: 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.

TBD: Final exams