

GEOG 4593/6593:

Geospatial Semantics and Geo-Text Mining

Map of Manhattan Island showing the distribution of 10,000 points. A scale bar at the top indicates distances up to 6 Kilometers. A north arrow is in the top right. A blue rectangle highlights a specific area in the lower part of the island. The map is credited to Esri, HERE, DeLorme, MapmyIndia, © OpenStreetMap contributors, and the GIS user community.



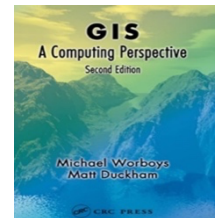
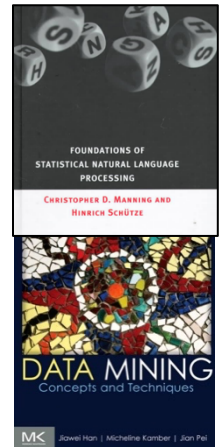
- **Course number:** GEOG 4593/6593
- **Credits:** 3
- **Lectures:** TR, 15:55-17:10, Geography-Geology Building, Room 321
- **Instructor:** Dr. Gengchen Mai
 - **Email:** gengchen.mai25@uga.edu
 - **Office hours:** Thursday, 14:30-15:30, Geography-Geology Building, Room 312
 - **Office:** Geography-Geology Building, Room 312

- **Teaching Assistant:** Chintan Maniyar
 - **Email:** chintanmaniyar@uga.edu
 - **Office hours:** TBD
 - **Office:** Geography-Geology Building, Room 313

Textbook (Optional)

There is no single textbook that systematically introduces geospatial semantics and geo-text mining. Accordingly, we will use book chapters here and there, and some chapters are from the three textbooks as below.

- Manning, C. D., & Schütze, H. (1999). Foundations of statistical natural language processing, Cambridge: MIT press. (refer as “NLP” later)
- Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier. **(DM)**
- Worboys, M. F., & Duckham, M. (2004). GIS: a computing perspective. CRC press. **(GIS)**



Course Website

The website of this course is available on UGA ELC.

Prerequisites

This course does not have a formal prerequisite yet. However, before taking this upper level course, you may consider taking the following courses first:

- GIS background: Geog4470/6470, Geog4300/6300 or equivalent
- GeoAI: GEOG 4/6920, GEOG 4/6921, GEOG 8350, or equivalent
- Basic programming background: CSCI 1301, C1360, Geog4590/6590, or equivalent

Software Environment

- Anaconda (a python-based software suite)
- Jupyter Notebook.

They have been installed on lab computers for you to use. However, students are also expected to download and install them on their personal computers. Both are free packages.

Tentative Course Schedule*

Week	Date	Topics	Reading
Week 1	08/17	- Lecture 1: Course introduction I – Course Logic	NLP Chap. 1
Week 2	08/22 08/24	- Lecture 1: Course introduction II -- The emergence of geo-text data - K2 Geo-Large language model invited talk by Mr. Cheng Deng from Shanghai Jiaotong University, China (10:00 – 11:00 EST) [No Class on 08/24]	DM Chap. 1
Week 3	08/29 08/31	- Lecture 2: Why mine geo-text data? - Lecture 3: Sub Areas and Perspectives in Geospatial Semantics	-
Week 4	09/05 09/07	- Lecture 4: Place name recognition and disambiguation - Lecture 5: Digital Gazetteer	NLP Chap. 7
Week 5	09/12 09/14	- Lecture 6: Geospatial Knowledge Graph I - Lab 1: Geotagged Wikipedia webpage retrieval and place name recognition with Stanford NLP toolkit (Paper #1 Presentation)	KnowWhereGraph (Janowicz 2022)
Week 6	09/19 09/21	- Lecture 6: Geospatial Knowledge Graph II (Paper #2 Presentation) - Lab 1: Continue (Paper #3 Presentation)	GIS Chap. 3
Week 7	09/26 09/28	- Lecture 7: Vector Space Model, TF-IDF, Geographic Information Retrieval - GeoGPT model invited talk by Mr. Yifan Zhang from Chinese University of Geoscience, China (10:00 – 11:00 EST) [No Class on 09/28]	DM Chap. 10
Week 8	10/03 10/05	- Lecture 8: Examining Data From Spatial, temporal, and Semantics Perspective - Lab 2: Analyzing geotagged Wikipedia data (Paper #4 Presentation)	-
Week 9	10/10	- Lecture 9: Topic Modeling for Geo-Text Data	-

	10/12	- Lab 2: Continue (Paper #5 Presentation) (Project Proposal Due)	
Week 10	10/17 10/19	- Lecture 10: Sentiment Analysis and Place Emotions - Lab 3: Topic modeling on the Yelp POI review data (Paper #6 Presentation)	-
Week 11	10/24 10/26	- Lecture 11: Word Embedding - Lab 3: Continue	-
Week 12	10/31 11/02	- Lecture 12: Knowledge Graph Embedding - Lab 4: Sentiment analysis on the neighborhood review data	-
Week 13	11/07 11/09	- Lecture 13: Geographic Question Answering - Lab 4: Continue	-
Week 14	11/14 11/16	- Lecture 14: Visualization for Geo-Text Analysis Results - Lab 5: Language Model Demo	-
Week 15	11/21 11/23	Thanksgiving – No classes	
Week 16	11/28 11/30	- Working on your final project	
Week 17	12/05	- Final Project Presentation	

***Course project paper is due on Dec 10, 2023, 12:00 pm (at noon).**

Course Requirements

1. Lab assignments (40%)

There will be 4 lab assignments in total led by TA. You will learn how to use different NLP and GIS techniques on different geo-text data.

2. Student research paper presentation (10%) and participation (5%)

Each student will read one scientific paper on relevant topic and present it in class (**8 min presentation + 2 min Q&A**). You will receive a paper list early in the semester to choose the paper to present and the date of presentation:

- Undergraduate: each student can pick a general geospatial semantics paper
- Graduate: each student can pick a geo-text mining technical paper

Please submit the **PDF file of your presentation**. To share with the class, please also **post the original article on a designated discussion board**. The due date is your presentation day. For technical papers, your presentation should introduce the research problem, research objective, methodology, and findings/results addressed in the article, as well as your critique if any. The presentation will be graded by the instructor.

3. Invited Talk Summary (5%)

There will be 1-2 invited talks at UGA GeoAI talk series. All students are required to attend at least one of them and submit a **one-page summary (12 font, 1.5 line space) of one invited talk** after it.

4. Course project (40%)

The course project can be about using any geo-text mining techniques on a specific geospatial research problem. The project needs to be conducted by:

- A team of no more than 3 undergraduate students;
- Or one graduate student.

Submission:

- 1) Write a project proposal (**1-2 pages, dual in Week 9**). The instructor will provide feedback on the proposed project idea. **(10%)**
- 2) Finish the proposed project and submit a scientific paper of it (**4-12 pages including references**). Submit the paper in ELC. **(20%)**
- 3) Make a final presentation of the project. **(10%)**

Both project proposal and final paper should use [IJGIS word template](#) or [LaTeX template](#). Please **submit a single PDF for each submission**.

The presentation will be graded by all other students, teaching assistant, and the instructor. The weight for each group is - **all other students (40%), teaching assistant (30%), and the instructor (30%)**. The **maximum and minimum score from students will not be considered**.

Grade weighting scheme

Component	Weight
Lab assignment (4)	40%
Research Paper Presentation (1)	10%
Participation	5%

Invited Talk Summary (1)	5%
Course project proposal	10%
Course project paper	20%
Course project presentation	10%

Final Letter Grades: The final letter grade will be determined according to the scale below.

A : [90,100]	C : [74, 77)
A-: [88,90)	C-: [70, 74)
B+: [86, 88)	D+: [67, 70)
B : [83, 86)	D : [64, 67)
B-: [80, 83)	D-: [60, 64)
C+: [77, 80)	F: <60

Due Dates: All Assignments must be completed on time. Submittal of assignments after due dates is accepted but with a penalty as 10% of the percentage grade for each day they are late (note: anytime passing the due time will be counted as late for one day, and anytime passing the first late day will be counted as late for two days, and so forth). Submissions that are late for more than 5 days will not be accepted. Late in-class assignments will not be accepted. In each case, exceptions are possible only with documentation of a medical or family emergency.

Academic Honesty: Cheating and plagiarism will automatically earn zero (0) points for the assignment or exam. All academic work must meet the standards contained in “A Culture of Honesty.” Each student is responsible to inform themselves about those standards before performing any academic work.