

## Course Overview and Prerequisites

This course will introduce concepts and approaches to mining social media data. It focuses on obtaining and exploring those data, mining networks, and mining text from social platforms. The focus is less on learning new methods; rather, students will learn how to apply previously learned data mining concepts to a domain that will likely be familiar to all of them, social media. Students will learn to explore, model, and predict with network and textual data from existing social platforms

The prerequisites for SIADS 682 include:

- SIADS 503: Data Science Ethics
- SIADS 522: Information Visualization I
- SIADS 652: Network Analysis
- SIADS 655: Natural Language Processing

## Instructor and Course Assistants

### Instructors:

Primary Instructor: Erik Lang ( [eriklang@umich.edu](mailto:eriklang@umich.edu) Opens in a new tab ) -- Lecturer, School of Information

- Jake Huang ( [yiju@umich.edu](mailto:yiju@umich.edu) Opens in a new tab ) -- Lecturer, School of Information
- Alexander Levin-Koopman ( [alevink@umich.edu](mailto:alevink@umich.edu) Opens in a new tab ) -- Lecturer, School of Information

### Course Creator:

Dr. Eric Gilbert - [eegg@umich.edu](mailto:eegg@umich.edu) Opens in a new tab

John Derby Evans Endowed Professor of Information, Associate Professor of Information, School of Information and Associate Professor of Electrical Engineering and Computer Science, College of Engineering

## How to Get Help

If you have questions concerning the degree program, encounter a technical issue with Coursera, or issues using Slack, please submit a report to the ticketing system at [umsimadshelp@umich.edu](mailto:umsimadshelp@umich.edu) Opens in a new tab

If you have an issue specific to the Coursera environment, you can also begin a [live chat session Opens in a new tab](#) with Coursera Technical Support (24/7) or view [Coursera troubleshooting guide Opens in a new tab](#)s (you may be asked to log in to your Coursera account).

For questions regarding course content, refer to the **Communications Expectations** section below.

## Learning Outcomes

Upon successful completion of this course, students will be able to:

- Describe the different types of data commonly found on social platforms
- Understand the ethical sensitivities in obtaining and operating on social data
- Use a social platform API to obtain data and understand the structure of those data
- Load a large social media corpus
- Produce summary statistics over a large social media corpus
- Visualize that corpus along geographic and temporal axes
- Articulate how networks act as the substrate for modern social media platforms
- Describe how and why different networks exist within the same data
- Construct different varieties of networks from a social dataset
- Compute a variety of networks measures from a social media dataset
- Describe the various types of text commonly found on social platforms
- Compute topics over time in a social dataset
- Compute sentiment over social text
- Apply tagging to social media text

## Course Communication Expectations

Slack is the preferred communication tool for this course. If you have questions about course content (e.g. lecture videos or assignments), please make sure to use Slack. Instructor and course assistant response time to **Slack** messages will aim to be within 24 hours, Monday-Friday and 48 hours on weekends. **Please try to monitor the Slack channels for the course regularly.**

Personal communication that may involve sensitive information may be emailed directly to the instructor or course assistant. If you email the instructor or course assistant, please include “SIADS 682” in the email subject. Instructor and course assistant response time to **email** messages will be within 24 hours.

## Course Schedule

This course begins on **Tuesday, October 24th** and ends on **Monday, November 20th at 11:59 p.m. EDT.**

Each week's assignments are due at the end of the week, unless indicated otherwise on the Grades page.

- Week 1 reading and programming assignments are **due Monday, October 30th 11:59 p.m.**
- Week 2 reading and programming assignments are **due Monday, November 6th at 11:59 p.m.**
- Week 3 reading and programming assignments are **due Monday, November 13th at 11:59 p.m.**
- Week 4 reading and programming assignments are **due Monday, November 20th at 11:59 p.m.**

### **Week 1: Introduction to Social Media Data**

*Objectives:* Describe the different types of data commonly found on social platforms. Understand the ethical sensitivities in obtaining and operating on social data. Use a social platform API to obtain data. Understand the structure of those data.

### **Week 2: Visualizing and Modeling Patterns in Social Media Data**

*Objectives:* Load a large social media corpus. Produce summary statistics over a large social media corpus. Visualize that corpus along geographic and temporal axes.

### **Week 3: Social Media Networks**

*Objectives:* Articulate how networks act as the substrate for modern social media platforms. Describe how and why different networks exist within the same data. Construct different varieties of networks from a social dataset. Compute a variety of network measures from a social media dataset.

### **Week 4: Social Media Text**

*Objectives:* Describe the various types of text commonly found on social platforms. Compute topics over time in a social dataset. Compute sentiment over social text. Apply tagging to social media text.

## **Programming Lab**

There is a lightly graded Programming Lab in this course that will require students to gain access to the an API of a social media platform, which requiring a developer account. The instructions for gaining this access are detail in the Lab.

Why is this Lab about Reddit when the other assignments are about Twitter?

Your weekly coding assignments deal with archived Twitter data, and we used to have a Twitter Lab to complement that. But, there have been a lot of changes (and mixed signals) surrounding free access to Twitter developer accounts in recent months. Perhaps Twitter will reverse itself again, but upon recent testing we were no longer able to access a free tier that accommodate our needs, so we looked for an alternative. This Lab as an alternative way to give you some "real world" experience accessing data from a different social media platform, Reddit.

**Important:** Gaining access to a developer account in Reddit can take minutes to days -- if you plan to do the assignment in this class, please request this as early as possible to ensure you have appropriate time to complete your assignment.

## Textbook and Readings Information

There is no textbook for this course; readings will be provided as pdfs or links.

## Reading Reflections

Each week, there will be assigned readings and a 3-5 paragraph reading reflection due. Everyone in the class is expected to have read the required readings for that week's topic. If you do not do the required readings, your understanding of the course material will suffer, as will your grade.

Our engagement with these readings will take place in the form of written reading reflections that you submit to the instructors. We will engage in four ways:

*Description:* statements or questions about what the author claims.

*Critique:* arguments about whether the author is correct or what the author has left out.

*Connection:* how the claims or concepts relate to those in other readings.

*Application:* how the reading applies to current events or platform issues.

You will apply these techniques in your reading reflection, in the form of responding to one or more of a few prompts provided by the instructor. The goal of the reading reflection is to engage more deeply with the readings. A rubric is provided in Coursera to guide you in constructing your week's written response.

You are required to post your reading reflection—in plain text—to the appropriate pinned Slack thread for that week's reading. You can also receive 0.5% extra credit for replying to another student's reading reflection in that thread with a thoughtful comment.

## Content Warnings

The content and discussion in this course will at times engage with potentially triggering content. Some of this will be emotionally and intellectually challenging to engage with. I will flag especially graphic or intense content, and will do my best to make this class a space where we can engage empathetically and thoughtfully with difficult content.

I've done my best to identify any texts with potentially triggering content. I've included tags for: violence, racism, misogyny, and self-harm. If you have concerns about encountering anything specific in the course material that I have not already tagged and would like me to provide warnings, please send me an email. Please feel free to discuss any concerns—including potential concerns about attending in-class discussions—with the instructor.

## Point of View

The readings and my comments in class will suggest a particular point of view. This perspective is my own (or the author's) and does not have to be yours! I encourage you to disagree with the ideas in the readings and lectures as well as the perspectives of your colleagues in the course. A significant part of a college education is learning about the complexity of various issues. Therefore, it is important that we listen and respect one another, but we do not have to agree.

## Weekly Office Hours via Zoom (Ann Arbor, Michigan time):

Your instructor will hold weekly, synchronous office hours using the video-conferencing tool, Zoom. Passcode to join office hours are **682**.

- Jake Huang -
- Erik Lang -
- Alexander Levin-Koopman -

## Grading

Course Item	Percentage of Final Grade	Due Date
Written reflection(choose from <a href="#">.prompts Opens in a new tab</a> )	5%	Week 1

Intro to social media data problem set	15%	Week 1
Lab using a real API	5%	Week 3
Written reflection(choose from <a href="#">_prompts Opens in a new tab_</a> )	5%	Week 2
Visualizing and modeling patterns problem set	20%	Week 2
Written reflection(choose from <a href="#">_prompts Opens in a new tab_</a> )	5%	Week 3
Social media networks problem set	20%	Week 3
Written reflection(choose from <a href="#">_prompts Opens in a new tab_</a> )	5%	Week 4
Social media text problem set	20%	Week 4

Note: All assignments are required to earn credit for this course.

## Letter Grades, Course Grades, and Late Submission Policy

This course requires you to complete all four assignments for full credit. You'll see from the above grading schema that all four sections' assignments are due at the end of the course. You may want to consider submitting assignments as you complete them in the week they are assigned.

**Assignments submitted late will incur a 20% per day deduction.**

The grading scale for this course is as follows

**A+ 97%**

A 93%

A- 90%

B+ 87%

B 83%

B- 80%

C+ 77%

C 73%

C- 70%

D+ 67%

D 63%

D- 60%

F 0%

Finally, students sometimes request re-grading of assignments. If a student wants an assignment re-graded by the instructors, that request must be submitted via email. However, a re-grade request makes the entire assignment possible to be re-assessed; therefore, it is theoretically possible that your overall grade could go *down* as a result of the request.

## **Academic Integrity/Code of Conduct**

Refer to the [Academic and Professional Integrity Opens in a new tab](#) section of the UMSI Student Handbook. (access to Student Orientation course required).

## **Accommodations**

Refer to the [Accommodations for Students with Disabilities Opens in a new tab](#) section of the UMSI Student Handbook.

Use the Student Application Form [in Accommodate Opens in a new tab](#) to begin the process of working with the University's Office of Services for Students with Disabilities.

## **Accessibility**

Refer to the [Screen reader configuration for Jupyter Notebook Content Opens in a new tab](#) document to learn accessibility tips for Jupyter Notebooks.

## **Library Access**

Refer to the [U-M Library's information sheet Opens in a new tab](#) on accessing library resources from off-campus. For more information regarding library support services, please refer to the [U-M Library Resources Opens in a new tab](#) section of the UMSI Student Handbook (access to the Student Orientation course required).

## **Student Mental Health**

Refer to the University's [Resources for Stress and Mental Health website Opens in a new tab](#) for a listing of resources for students.

## **Student Services**

Refer to the [Introduction to UMSI Student Life Opens in a new tab](#) section of the UMSI Student Handbook (access to the Student Orientation course required).



# Technology Tips

## Recommended Technology

- This program requires Jupyter Notebook for completion of problem sets and Adobe or other PDF viewer for reading articles.

## Working Offline

- While the Coursera platform has an integrated Jupyter Notebook system, you can work offline on your own computer by installing Python 3.5+ and the Jupyter software packages. For more details, consult the Jupyter Notebook FAQ.