

# Data and Society, Spring 2025 [DRAFT]

## Class info

I 320S (Unique Number: 28220)  
MW 2:00 - 3:30pm. SZB 2.418

## Instructor

Hanlin Li, PhD, she/her, Assistant Professor  
Email: [lihanlin@utexas.edu](mailto:lihanlin@utexas.edu) Office: UTA 5.444  
Office hour: TBD and by appointment

## Course Description

What role does data play in our society? What infrastructure and systems does data traverse? And what technologies and practices are associated with data? Who does an increasingly data-driven society affect the most? In this course, students will answer these questions through discussion, readings, and projects. Students will explore common data collection, management, and sharing practices around information technology and emerging technologies such as AI. Students will gain hands-on experiences with collecting and managing user data in ethical and responsible manners. Students will design data-driven systems that are centered around user consent, transparency, and social responsibilities.

## Required Materials

All course readings will be available via the course Canvas site.

## Long Description

What role does data play in our society? What infrastructure and systems does data traverse? And what technologies and practices are associated with data? Who does an increasingly data-driven society affect the most? In this course, students will answer these questions through discussion, readings, and projects.

This course explores the intersection of data collection, technology, and society, focusing on the ethical, social, and legal issues that arise from large-scale data collection practices. Students will critically examine how data is gathered, analyzed, and used by governments, corporations, and other institutions, and its impact on privacy, equity, and power dynamics. Students will delve into topics such as data work, crowdsourcing, surveillance, privacy, algorithmic bias, informed consent, data justice, and the implications of emerging technologies like AI and IoT. The course also emphasizes the role of policy, activism, and design in shaping a more equitable and ethical data landscape.

By the end of the course, students will be equipped to evaluate and critique data practices, understand their societal consequences, and propose solutions that prioritize ethical considerations and social good. This course is ideal for those interested in the fields of data science, public policy, technology ethics, and social justice.

## **Learning outcomes**

Students will learn how to examine what role data plays in our increasingly data-driven society and the pros and cons of different data collection, management, and sharing practices through readings, discussions, and case studies.

Students will gain hands-on experience with data management or systems as UX designers, researchers, and data scientists by completing a class project.

Students will also be exposed to interdisciplinary research on the impact of data-driven technologies and important ethical considerations, e.g. privacy and consent.

## **How will you learn**

This course uses a blended strategy of student-led discussions, mini-lectures, and asynchronous assignments. In addition to attending and participating in discussions and lectures, students will be expected to contribute to Canvas discussion and complete a semester-long project that can take one of the following forms: a datasheet and an evidence-based redesign of an existing data-intensive system.

## **Classroom norms and participation**

This class will be centered around social learning. This means that the instructor will foster a constructive, inclusive discussion space for all students and students are expected to attend all class sessions in person and participate actively in class discussion. Class participation will be graded.

Students should only use laptops for note taking purposes. The instructor may limit technology use to minimize classroom distractions.

If you are a student with a disability, or think you may have a disability, and need accommodations please contact Disability and Access (D&A). You may refer to D&A's website for contact and more information: <http://disability.utexas.edu/>. If you are already registered with D&A, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations.

## **How will you be evaluated**

Students will be evaluated on their completion of assignment and their ability to apply knowledge to several short-term assignments and a design project.

## **Major assignments**

Reading responses every week for 10 weeks. 30%

Students will complete reading assignments each week for ten weeks and submit their reading responses on Canvas the day before class. Each reading response will account for 3%. These

reading responses are meant to help students reflect on what they have learned, what they disagree with the readings, and what lingering questions they have. Reading responses should be of reasonable length (200+ words). Note that there are multiple mandatory readings per week, and you will be asked to submit one response discussing all readings assigned. Your reading responses will be graded for substance and originality. I expect you would spend the first two sentences summarizing the readings and the rest of the response would be to elaborate your reflections and questions. You can relate it to your personal experiences, your own coursework or projects, or other courses you have taken. No late work will be accepted.

Individual assignment - Google crowdsource & Mozilla Common Voice 10%

Students will complete two crowdsourcing tasks, one on Google Crowdsourcing and one on Mozilla Common Voice. These two platforms' purpose is to collect data from volunteers for AI model training. You will perform tasks such as image labeling, audio transcription, and audio recording on these two platforms. You will then write up a report working on the two platforms. Your writeup should be 1500 words long and you can take screenshots of these two platforms to support your argument. I will provide suggested questions for you to answer in your writeup.

Individual assignment - datasheet 10%

Students will produce a datasheet to accentuate critical information about a given dataset. I will provide several datasets as options and you will choose one and read through its existing documentations. You will then perform necessary descriptive analysis and distill critical information about the dataset into a formatted datasheet. You will also include a reflection in your assignment, in response to suggested questions.

Design project - 40%

Students will complete a semester-long design project, focusing on responsible data collection. The project will have multiple milestones, including a project proposal, a midterm presentation, peer reviews, and a final presentation. Detailed instructions will be provided for each milestone.

Quizzes - 10%

Students will complete quizzes or activities during class. The quizzes and activities are designed to test students' understanding of foundational concepts and theories covered in the class.

Participation - 10%

### **Late submission policy**

Late reading responses and late presentations will not be accepted. Other assignments will be subject to 5% grade reduction per day of delay.

### **Required and Recommended class materials**

Readings will include news reports, academic publications, and book chapters. All class materials will be provided via Canvas.

## Weekly Schedule [Subject to change]

Week	Date	Topic	Major assignments
1	Jan 13	Introduction	
	Jan 15	Data and power	
2	Jan 20		
	Jan 22	Crowdsourcing	
3	Jan 27	Crowdsourcing (cont'd)	
	Jan 29	[pending field work]	
4	Feb 3	Data work	
	Feb 5	cont'd	Individual project - crowdsourcing
5	Feb 10	scraping and passive data collection	
	Feb 12	cont'd	
6	Feb 17	consent and privacy	
	Feb 19	cont'd	
7	Feb 24	Research ethics	Project proposal
	Feb 26	cont'd	
8	Mar 3	documentation	
	Mar 5	cont'd	
9	Mar 10	mid term presentation	Mid term presentation slides
	Mar 12	mid term presentation	
10	Mar 17		
	Mar 19		
11	Mar 24	subjectivity and bias	Individual project - datasheet
	Mar 26	cont'd	
12	Mar 31	Labor	
	Apr 2	cont'd	
13	Apr 7	Licensing	Peer review
	Apr 9	cont'd	
14	Apr 14	Guest lecture	
	Apr 16	Guest lecture	
15	Apr 21	Final presentations	Final presentation slides
	Apr 23	Final presentations	
16	Apr 28	drop in discussion and next steps	

