

# Data Science and AI for Global Leaders

## Fall 2024

**Instructor(s):** Anna Haensch (anna.haensch@tufts.edu)

**Teaching Assistant:** -

**GMAP Technology Support:** Stephanie Depauw, [stephanie.depauw@tufts.edu](mailto:stephanie.depauw@tufts.edu)

**Office Hours:**

**Semester Hour Units:**

**Live Session Dates:** December 13 - 15, 2024

**Course Materials:** Course materials will be made available on Canvas

### Course Information

#### Course Overview

In an era where data permeates every aspect of our world, understanding its potential is more valuable than ever. As a global leader it is important to understand both its capabilities and the potential pitfalls. This course will focus on the full data life-cycle including the generation, collection, processing, analysis, and management of data. Special attention will also be paid to data ethics as it pertains to representativeness and missingness of data. We will also discuss how recent advances in artificial intelligence have impacted the data life-cycle and the opportunities and threats that this poses. The technical aspects of this course will be supplemented by relevant readings and case studies showing data science in action.

While data science is a very technical subject, the focus on this course will be on high level concepts offering optional inroads for deeper technical exploration. No prior mathematical, statistical or programming experience will be assumed.

#### Course Structure

- 4 days of online asynchronous pre-work — i.e., skill preparation, reading, group work, discussion boards
- 3 days (Fri., Sat., Sun. 10:00AM - 1:00PM) of live, online, skill-building instruction; 3 hours of instruction per day

#### Learning Objectives

By the end of this course, you will be able to:

1. Plan new data projects incorporating a full data life-cycle.
2. Critique the social, political, and economic contexts of data acquisition and analysis.
3. Understand the basic underlying technology and terminology around data and AI.

## Content by Session

Pre-Synchronous Work	<b>Day 1: Big Data and the Data Lifecycle</b>		9 Dec. 2024
	<b>Learning Objectives:</b>	Develop a working definition for data and big data. Understand the processes that generate data and why they are important.	
	<b>Required Reading:</b>	boyd et al. (2012), <i>Critical Questions for Big Data</i>	
	<b>Optional Reading:</b>	D'Ignazio et al. (2020), <i>Data Feminism</i> , 4. "What Gets Counted Counts"	
	<b>Day 2: Data Processing and Management</b>		10 Dec. 2024
	<b>Learning Objectives:</b>	Understand the meaning of missingness, representativeness, and reproducibility. Identify the key considerations for proper data management.	
	<b>Required Reading:</b>	D'Ignazio et al. (2020), <i>Data Feminism</i> , 2. "Collect, Analyze, Imagine, Teach"	
	<b>Assignment:</b>	Discussion Board Initial Post (Due 10 Dec. 2024 11:59 PM Eastern)	
<b>Optional Reading:</b>	Wilkinson et al. (2016), <i>The FAIR Guiding Principles for scientific data management and stewardship</i>		
Live Synchronous Sessions	<b>Day 3: Modeling and Learning with Data</b>		11 Dec. 2024
	<b>Learning Objectives:</b>	Understand the basic goals of data modeling. Critique the assumptions that go into creating a data model.	
	<b>Required Reading:</b>	D'Ignazio et al. (2020), <i>Data Feminism</i> , 3. "On Rational, Scientific, Objective Viewpoints from Mythical, Imaginary, Impossible Standpoints"	
	<b>Optional Reading:</b>	Buolamwini et al. (2018), <i>Gender Shades</i>	
	<b>Day 4: Data Visualization</b>		12 Dec. 2024
	<b>Learning Objectives:</b>	Understand the principle of proportional ink, and the fundamentals of good data visualization. Identify the proper use of color and color maps in data visualizations.	
	<b>Required Reading:</b>	Angwin et al. (2016), <i>Machine Bias</i>	
	<b>Assignment:</b>	Discussion Board Response (Due 12 Dec. 2024 11:59 PM Eastern)	
<b>Optional Reading:</b>	Wilke (2019), <i>Fundamentals of Data Visualization</i> , Chapters 1-19		
Live Synchronous Sessions	<b>Day 5: AI and Deep Learning</b>		13 Dec. 2024
	<b>Learning Objectives:</b>	Develop a working definition for AI, machine learning, and deep learning. Identify the necessary input and resulting output of an AI system.	
	<b>Optional Reading:</b>	D'Ignazio et al. (2020), <i>Data Feminism</i> , 6. "The Numbers Don't Speak for Themselves"	
	<b>Day 6: Generative AI and Human Systems</b>		14 Dec. 2024
	<b>Learning Objectives:</b>	Understand the difference between AI and Generative AI. Identify the impacts of AI on the world from social systems to the environment.	
	<b>Optional Reading:</b>	Solaiman et al. (2023), <i>Evaluating the Social Impact of Generative AI Systems in Systems and Society</i>	
	<b>Day 7: Large Language Models</b>		15 Dec. 2024
	<b>Learning Objectives:</b>	Understand the basic technology underlying the large language model. Identify appropriate use cases for LLMs.	
<b>Assignment:</b>	Final Policy Presentation (In-Class 15 Dec. 2024)		
<b>Optional Reading:</b>	Bender et al. (2021), <i>On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?</i> 🗣️		

This schedule is subject to modifications at the instructor's discretion.

## Assignments and Grading

### Assignments

- Discussion Board:
  - Your discussion board posts will consist of two parts:
    - 1) A 250-300 word reflection on the reading will be posted prior to seeing the full discussion board. The reflection **must** explicitly state **1-2 discussion questions** based on the concepts discussed in the readings. This can be a *curiosity question*, where you're interested in finding out more, a *critical question*, where you challenge the author's assumptions or decisions, or an *application question*, where you think about how concepts from the reading would apply to a particular context you are interested in exploring.
    - 2) A response to a post by a peer addressing one of the stated discussion questions. This doesn't need to be an explicit answer to the question, but maybe an idea around it, or some supporting context.
- Participation:
  - During the live 3-day portion of the skill summit students are expected to participate fully in the discussions and in-class exercises. As a prerequisite for participating, students are expected to have completed the required readings prior to the day's meeting. In-class exercises will consist of reflection, discussion, writing, and working through case studies. In some cases in-class work will result in written artifacts to be submitted.
- Final Presentation:
  - In small groups, you will deliver a short (~10 minute) presentation on a data/AI policy that you think would be locally, nationally, or globally useful for our current moment. Groups will be assigned by the instructor prior to the start of the Skill Summit but groups are not expected to meet prior to the live synchronous sessions. Time will be given during each of the synchronous days to begin formulating and clarifying your proposed policy through a series of guided exercises.
    - Assessment Criteria: Clearly identify the stakeholders, the problem your policy is intended to address, the concrete actions of your policy, and data you might need to demonstrate the efficacy of your policy. You will be assessed on your verbal communication, the clarity and supporting evidence of your stated problem, and appropriateness of your proposed policy solution.
    - Due Date: Your group will present your proposed policy during the second half of our final synchronous meeting on Dec. 15 2024.

### Grading

Discussion Boards	30%
Participation	50%
Final Presentation	20%

## Expectations

- **Attendance:** Students must participate virtually (via Zoom) 100% of the time during the live 3-day portion of the skill summit.
- **Zoom:** During synchronous zoom sessions it's required that you turn on your camera and mute your microphone (unless you are speaking). Visual cues are really important in gauging comprehension and facilitating lively conversation.
- **Integrity:** Students must abide by the school's Honor Code, as detailed in [Fletcher Student Handbook](#), pages 32-40.

## Accommodation of Disabilities

- **Accommodations for Students with Disabilities:** Tufts University is committed to providing equal access and support to all qualified students through the provision of reasonable accommodations so that each student may fully participate in the Tufts experience. If you have a disability that requires reasonable accommodations, please contact the [StaarCenter@tufts.edu](mailto:StaarCenter@tufts.edu) or 617-627-4539. Please be aware that accommodations cannot be enacted retroactively, making timeliness a critical aspect of their provision.
- **Student Accessibility and Academic Resources (StAAR) Center:** The StAAR Center offers academic support through one-on-one academic coaching, writing consultations, tutoring, study groups and study strategies and discipline specific workshops, including time management. Services are free to all enrolled students. Students can make an appointment for any of these services at <https://students.tufts.edu/staar-center>. Resources specific to graduate students can be found at: <https://students.tufts.edu/staar-center/academic-support/graduate-student-resources>.

## Diversity Statement

As a school of international affairs, Fletcher prepares students for global careers in complex and diverse contexts. In designing your syllabus, consider embedding diversity along multiple dimensions:

- Pedagogy—diverse ways in which students learn and are evaluated;
- Representation—diverse identities of authors, speakers, case protagonists;
- Content—diversity of topics, concepts and frameworks, geographies, etc.

It is my goal that all students feel empowered to have a successful and affirming experience in this course. I strive to make sure that all voices are heard and all ways of learning and being are respected. I understand that every student has a life beyond this classroom and that life can bring unexpected challenges. If you find that you are having trouble meeting the requirements of the course please speak to me as early as possible.

## Other Resources

- **Student Information System (SIS):** This is where you will find your course rosters and list of advisees. Final course grades must be entered here (not on Canvas).
- **Canvas:** This is the [course platform](#) where you will post your syllabus and other materials. See Canvas [training](#) and guides.
- **Tufts Center for the Enhancement of Learning and Teaching (CELT)**
- **IT Support:** <https://it.tufts.edu>