

DSAN 5450

Data Ethics and Policy

Tuesdays 3:30-6pm, Walsh 498

Jeff Jacobs
jj1088@georgetown.edu

Spring 2027, Georgetown University

Welcome to the Spring 2027 version of **DSAN 5450: Data Ethics and Policy** at Georgetown University! Please note that the most up-to-date version of this syllabus will always be available at <https://jjacobs.me/dsan5450>.

The course meets on **Tuesdays from 3:30-6pm in the Walsh Building, Room 498**

Course Staff

- Prof. Jeff Jacobs, jj1088@georgetown.edu
 - Office hours (Click to schedule): [Tuesdays, 3:30-6:30pm](#)
- TA Ellie Byrd, eb1481@georgetown.edu
- TA Christy Hsu, th1010@georgetown.edu
- TA Wendy Hu, lh1078@georgetown.edu

Course Description

This graduate-level course will train students to navigate the landscape of ethical issues which arise at each step of the data science process, with an eye towards developing policy recommendations for governments and organizations seeking expert advice on how to tackle these issues from a regulatory perspective. Students will explore and critically evaluate a range of data-related issues in contemporary society, such as responsible data collection, algorithmic bias, privacy, transparency, accountability, democratic participation in data usage and data-driven decisions, and the ethical implications of emerging technologies like artificial intelligence and machine learning (self-driving cars, ChatGPT, crowd-sourced training data, etc.).

Beginning with a set of historical case studies—instances in which scientists, engineers, and policymakers have been forced to re-evaluate their ethical intuitions in light of technological developments (nuclear power, use of social media platforms to organize protests and influence political outcomes, deployment of facial recognition software and predictive AI by police and military forces)—the course then introduces a set of general ethical frameworks (consequentialism, deontological ethics, and virtue ethics), challenging students to consider their relative strengths and weaknesses for addressing modern technological-ethical dilemmas faced by businesses, healthcare organizations, governments, and academic institutions. After a final portion of the

course linking these ethical frameworks with practical regulatory and policy considerations, students will write and present a policy whitepaper analyzing a data-ethical issue of particular interest to them, integrating ethical perspectives, regulatory principles, and domain knowledge into a recommendation of best practices for the relevant agency, firm, or institution.

The course will thus equip students with a robust ethical “toolbox” for conscientiously gathering, interpreting, and extracting meaning from data throughout their careers as data scientists, while respecting privacy, fairness, transparency, democratic accountability, and other social concerns. *Prerequisites: None. 3 credits.*

Course Overview

The course revolves around three “pillars”, which we’ll examine individually before bringing them together for your final projects at the end of the class: high-level ethical issues in data science, general ethical frameworks, and public policy applications.

Data Science

A portion of the course will focus on introductions to cutting-edge technologies like self-driving cars, ChatGPT, facial detection algorithms, and various applications of AI to police and military technologies. For this portion, we’ll draw fairly often from the contents of the following books:

- Perez (2019): *Invisible Women: Data Bias in a World Designed for Men*.
- Catherine D’Ignazio and Lauren F. Klein (2020). *Data Feminism*. Cambridge, MA: MIT Press. [Free, open-source!]
- Cathy O’Neil (2016). *Weapons of Math Destruction*. New York, NY: Crown Books.

Since there are plenty of in-depth resources available to you (e.g., other Georgetown courses!) for learning the *technical details* of these technologies, our goal in this course will be to learn just the particular *aspects* of each technology which are most relevant to the ethical and policy issues they present.

For example, we will look at Neural Network-based Machine Learning algorithms, but we will focus specifically on how the **performance** of these algorithms on a given task depends crucially on the existence of effective **training data** for that task. The breakthroughs in Artificial Intelligence which have had an immense impact on society over the past few decades, for example, have **not** come about because of new algorithms (neural networks, for example, have been around since the 1950s). Rather, they have come about because of the massive, exponential increase in the amount of **data** available to **train** these already-existing algorithms: for example, data scraped from across the entire web, or from millions of scanned books, or from Wikipedia’s massive collection of articles. This means, therefore, that these algorithms simply **encode** pre-existing human biases into algorithmically-derived “rules”, thus motivating the next pillar of the course: Ethics!

Ethics

For the ethics-focused portion of the course, we'll be reading selections from the following textbook:

- Lewis Vaughn and Louis P. Pojman (2021). *The Moral Life: An Introductory Reader in Ethics and Literature*. Oxford, UK: Oxford University Press. [PDF]

From the vast array of readings contained in this collection, we'll look at both "standard" ethical readings from e.g. Jeremy Bentham and Immanuel Kant plus readings from literary sources like Ursula Le Guin and Ambrose Bierce.

Public Policy

For the final piece of the course we will take the **technological developments** discussed the first portion, analyze them using the **ethical frameworks** discussed in the second portion, and come to **conclusions** as to what types of things lawmakers, governments, and civil society organizations (NGOs, for example, and Think Tanks) can do **in practice** to address the ethical issues raised by these technologies. This means that, specifically, the recommended final project for the course will be a **Policy Whitepaper**, where you will choose a particular institution and make a **recommendation** to them in terms of how they can use their power (for example, the power to pass laws) to most effectively address an ethical issue that you believe is important.

For this portion of the class we'll have to draw on a wide range of different readings, depending on what particular subdomains of public policy are most interesting to you all, but as a general textbook on ethics in data science which does focus a good amount on **policy** specifically, we will look at:

- Anne L. Washington (2023). *Ethical Data Science: Prediction in the Public Interest*. New York, NY: Oxford University Press.

Now that you have an overview of the trajectory of the course, the following section contains the particulars of what we'll be reading and working on each week!

Schedule

The following is a rough map of what we will work through together throughout the semester; given that **everyone learns at a different pace**, my aim is to leave us with a good amount of **flexibility** in terms of how much time we spend on each topic: if I find that it takes me longer than a week to convey a certain topic in sufficient depth, for example, then I view it as a strength rather than a weakness of the course that we can then rearrange the calendar below by adding an extra week on that particular topic! Similarly, if it seems like I am spending too much time on a topic, to the point that students seem bored or impatient to move onto the next topic, we can move a topic intended for the next week to the current week!

Unit	Week	Date	Topic
Unit 1: Ethical Frameworks	1	Jan 14	Introduction to the Course
	2	Jan 21	Machine Learning, Training Data, and Bias
Unit 2: Fairness in AI	3	Jan 28	Ethical Frameworks
	4	Feb 4	Rights, Policies, and Fairness in AI
	5	Feb 11	Context-Free Fairness
	6	Feb 18	Context-Sensitive Fairness
Midterm	7	Feb 25	In-Class Midterm: Data Ethics, Fairness, Privacy
		Mar 6	No Class (<i>Spring Break</i>)
Unit 3: Privacy Policies and Data Externalities	8	Mar 11	Privacy Policies, Privacy Regulation, and You

Recommended Readings

We start this unit by *framing* privacy policies as contracts, as a simplifying assumption. Across judicial systems, however, there are nuances in terms of how different courts might interpret different ‘pieces’ of a privacy policy as legally-binding. These nuances are covered in incredibly-helpful depth in Hartzog and Solove (2026)!

9	Mar 18	Modeling Data Policies: From Prisoners Dilemmas to Invisible Hand Games
---	--------	---

Recommended Readings

1. On the *legal* background for data externalities: MacCarthy (2015) (Especially pgs. 423-470)
2. The specific *economic model* of data externalities we’ll look at: Acemoglu et al. (2022) (only Sections 1-3 and Appendix B-3)

Unit 4: Policy Evaluation and Inverse Fairness	10	Mar 25	Welfare Economics and Policy Evaluation
---	----	--------	---

Recommended Readings

On the three basic ‘policy structures’ we’ll use to capture what kinds of policy interventions may be necessary: Bowles (2009), pgs. 23-55

11	Apr 1	Fear and Loathing on the Pareto Frontier
----	-------	--

Recommended Readings

Unit	Week	Date	Topic
Unit 5: Discourse Ethics and Identity Formation			On the role of bargaining power in determining policy choice when the Pareto frontier has been reached: Bowles (2009), pgs. 167-183
	12	Apr 8	(Machine) Learning What Policies Value
		Recommended Readings	
		1. <i>Social Welfare Functionals (SWFs)</i>, the tool we're adding to our toolkit this week to evaluate policies along the Pareto frontier at an <i>aggregate</i> level (as opposed to the game-theoretic models from last week, which model how interacting <i>individuals</i> might reach or fail to reach certain Pareto-optimal outcomes), are discussed in exactly the detail you'll need in Adler (2019): Chapter 4 starts to get pretty in-the-weeds, so Chapters 1-3 (which are also shorter) are more than enough! 2. The core paper you'll look at in HW4, which utilizes ML (Causal Random Forests) to derive <i>SWF weights</i> from the algorithmically-determined choices made by Mexico's PROGRESA program, is Björkegren, Blumenstock, and Knight (2025) (An earlier version, Björkegren, Blumenstock, and Knight (2022), has more details on the ML-based estimation)	
	13	Apr 15	SWFLS, Discourse Ethics, and Identity Formation
		Recommended Readings	
		As we move into this unit on 'applications' of the frameworks we've learned throughout the course to various social issues, a crucial paper which we'll unpack first and foremost is Hu (2023)	
	14	Apr 22	AI, Freedom, and the 'Kindly Slavemaster'
		Recommended Readings	
		1. There are many theoretical/historical overviews of liberal vs. republican conceptions of liberty (Pettit 2008; Skinner 1998), and [fewer] applied works exploring how these might relate to AI and data collection (Hamilton 2025). 2. For a book which does <i>both</i> (lays out the tenets of republican liberty while exploring its applications to technology-focused power struggles), my recommendation is Gourevitch (2015)! Given its focus on 19th century labor activism, it may not seem <i>explicitly</i> relevant at first, but read it with the Jacquard loom analogy from Week 1 in mind and the connection should be clear! 3. Chapters 6 through 10 of Acemoglu and Johnson (2023) also focus on how imbalances of economic power impact technological development in general and AI in particular	
		May 10 (Friday)	[Deliverable] Policy Whitepaper

Assignments and Grading

The main assignment in the course will be your **policy whitepaper**, submitted at the end of the semester. However, there will also be a midterm exam and a series of assignments which exist to let you explore each of the modules of the course, in turn.

Assignment	Due Date	% of Grade
HW1: Nuts and Bolts for Fairness in AI	Friday, February 9	9%
HW2: Context-Sensitive Fairness	Friday, February 21	9%
In-Class Midterm	Wednesday, February 25	25%
HW3: Privacy Policies as Incomplete Contracts	Friday, April 12	9%
HW4: Policy Evaluation	Friday, April 26	9%
HW5: Inverse Fairness	TBD	9%
Policy Whitepaper	Friday, May 8	30%

Homework Lateness Policy

After the due date, for each **homework** assignment, you will have a grace period of 24 hours to submit the assignment without a lateness penalty. After this 24 hour grace period, late penalties will be applied up until 66 hours after the due date. Specifically, late penalties will be applied based on the following scale (unless you obtain an excused lateness from one of the instructional staff!):

- **0 to 24 hours** after due date: no penalty
- **24 to 30 hours** after due date: 2.5% penalty
- **30 to 42 hours** after due date: 5% penalty
- **42 to 54 hours** after due date: 10% penalty
- **54 to 66 hours** after due date: 20% penalty
- **More than 66 hours** after due date: Assignment submissions no longer accepted (without instructor approval)

Excused lateness exceptions will be grounded in a policy where you can “pause” the class (due-date-wise, but also just, expectation-of-progress-wise more generally) to take the time you need to handle things happening in your life. Then, you can “resume” work once you feel more centered/once you find yourself back on stable ground. The “[Maslow’s hierarchy](#)” model from psychology (Figure 1) provides a straightforward way to think about this approach.

Learning data ethics is way up there in the blue triangle, so, it’s important that you feel like the first four levels are solidified before you turn your focus back to classwork. And, it’s a win-win, because it means that once you’re back in action you can be more engaged and receptive to new topics, rather than being forced to “simulate” being in the blue triangle while in reality struggling in other spots!

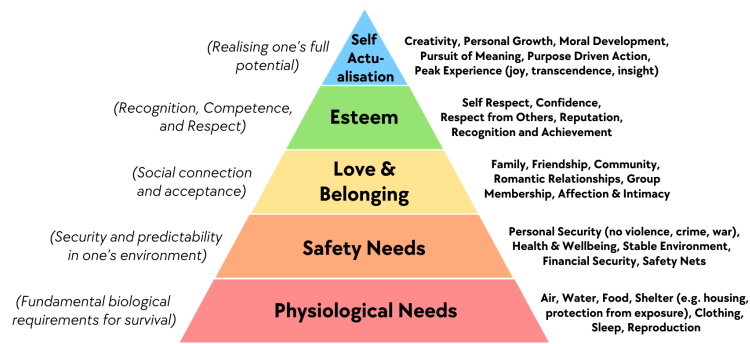


Figure 1: Maslow's Hierarchy of Needs, from [Wikimedia](#)

Generative AI Policy

Generative AI tools such as ChatGPT, Claude, Gemini, GitHub Copilot, and similar technologies are increasingly part of modern data science practice. Students are encouraged to learn how to use these tools **responsibly and critically** as part of their learning in this course.

You may use generative AI to support your learning — for example, to clarify concepts, explore alternative explanations, troubleshoot code, generate additional examples or practice problems, improve code efficiency, or help you understand statistical methods.

However, AI should **support your work, not replace your own thinking and problem-solving**. You may **not use** generative AI to generate complete solutions or answers to homework, quizzes, exams, or other graded assessments and submit those responses as your own. Any work you submit should demonstrate **your own** understanding of the material.

For projects, labs, coding activities, or other assignments where AI use is appropriate, you are responsible for understanding, verifying, and being able to explain any AI-assisted work that you submit. **If generative AI is used for an assignment or activity, you must clearly acknowledge that use and briefly describe how the tool was used.** AI-generated output can be incomplete, misleading, or incorrect, and responsibility for the final work remains with you.

Specific assignments or assessments may include additional instructions regarding acceptable AI use. **AI use is not permitted during quizzes or exams unless explicitly authorized by the instructor.**

When in doubt about whether a particular use of AI is appropriate, please ask.

References

Acemoglu, Daron, and Simon Johnson. 2023. *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity*. PublicAffairs.

- Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar. 2022. “[Too Much Data: Prices and Inefficiencies in Data Markets](#).” *American Economic Journal: Microeconomics* 14 (4): 218–56.
- Adler, Matthew D. 2019. *Measuring Social Welfare: An Introduction*. Oxford University Press.
- Björkegren, Daniel, Joshua E. Blumenstock, and Samsun Knight. 2022. “[\(Machine\) Learning What Policies Value](#).” 2022.
- Björkegren, Daniel, Joshua E. Blumenstock, and Samsun Knight. 2025. “[What Do Policies Value?](#)”
- Bowles, Samuel. 2009. *Microeconomics: Behavior, Institutions, and Evolution*. Princeton University Press.
- Gourevitch, Alex. 2015. *From Slavery to the Cooperative Commonwealth: Labor and Republican Liberty in the Nineteenth Century*. Cambridge: Cambridge University Press.
- Hamilton, Matthew David. 2025. “[Digital Domination: A Case for Republican Liberty in Artificial Intelligence](#).” September 30, 2025.
- Hartzog, Woodrow, and Daniel J. Solove. 2026. “[Privacy as Contract?](#)” SSRN Scholarly Paper. Rochester, NY. January 16, 2026.
- Hu, Lily. 2023. “[What Is ‘Race’ in Algorithmic Discrimination on the Basis of Race?](#)” *Journal of Moral Philosophy* 21 (1–2): 1–26.
- MacCarthy, Mark. 2015. “[New Directions in Privacy: Disclosure, Unfairness and Externalities](#).”
- Perez, Caroline Criado. 2019. *Invisible Women: Data Bias in a World Designed for Men*. Abrams.
- Pettit, Philip. 2008. “[Republican Freedom: Three Axioms, Four Theorems](#).” *Republicanism and Political Theory* 104: 45–56.
- Skinner, Quentin. 1998. *Liberty Before Liberalism*. Cambridge University Press.