

AI Ethics

UNDER CONSTRUCTION

Spring 2021
MW 9:00-10:15AM
EST
Zoom on Canvas

Instructor: Jason Farr
jaf301@georgetown.edu
OH: Wed 2:00- 4:00PM EST

course description and aims

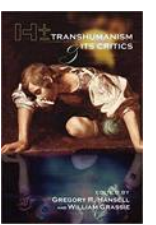
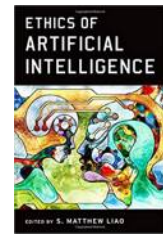
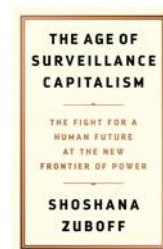
In this course we will set out to better understand the philosophical issues surrounding the use and progression of artificial intelligence. Is the promise of increased economic development, efficiency, convenience, and medical breakthroughs reason to forge ahead at full speed? Or are there issues concerning inequality, racial discrimination, privacy violation, harms to democracy, and personal targeting that might give us pause? What does human life look like once we fully incorporate artificial intelligence into our lives—indeed, into our *bodies* and *minds*? What will it mean to flourish as a human being? What will our social lives look like? Are these lives we want? Are they morally laudable? Are they inevitable?

There are a number of objectives for the course. You will learn new philosophical concepts and perspectives that will enable you to reflect on human life in new ways. You will practice interrogating your assumptions about ethical issues relating to artificial intelligence, and you will practice clarifying your assumptions and clarifying the ethical issues that arise in the course of your reflection. You will also practice reading and writing philosophical texts, charitably reconstructing philosophers' arguments, and crafting careful, well-constructed arguments of our own. And, importantly, we will practice respectful and detailed discussion (face-to-face on zoom and written as part of a thread) of important ethical issues relating to tech and AI.

course requirements and policies

Course Texts*

- [The Age of Surveillance Capitalism](#) ● ● ● ●
Shoshana Zuboff
- [Ethics of Artificial Intelligence](#) ● ● ● ● ● ● ● ●
S. Matthew Liao
- [H+/- Transhumanism and its Critics](#) ● ● ● ● ● ● ● ● ● ● ● ● ● ●
Gregory R. Hansell and William Grassie



*All readings will be posted on Canvas. You may purchase these books at the bookstore or online if you would rather have the physical copies. I'd rather you not have to pay a lot of money for this course!

Technology Resources

- [Canvas](#) - Try to familiarize yourself with our course page, including where to submit assignments, how to view your grades, where to participate in discussions, how to join office hours / class meetings, etc.
- [Zoom](#) - The course will be *entirely online*. Whether you can participate during class hours or not, it's a good idea to familiarize yourself with zoom.
- [Miro](#) - This is a collaborative whiteboard (like Google Jamboard, but better). We'll be using it inside and outside of class, for activities and assignments.

Discussions

You will be expected to respond directly to **three** discussion threads of your choice throughout the semester and you will be expected to respond to **three** posts from your peers. Discussions are a chance to **expand** upon your immediate reactions to course material in a format less formal than your papers.

Papers

There are **two** papers. Both papers will be based on your choice among several prompts that I assign. The first paper will be roughly roughly 1,200 words (about 4 pages), and the second roughly 1,600 words (about 6 pages). I will offer extended office hours on Zoom the weeks before and after a paper is due. The first paper will be written in sections and graded accordingly.

Small Assignments

You will complete **four** small writing assignments throughout the semester. For example, you may be asked to write an introduction to a fake paper.

Miro Board Engagement

You will be expected to engage thoughtfully and creatively with readings on a Miro board. Minimum **10** Miro boards over the course of the semester, and maximum **1** board per week.

Attendance

You may skip **three** classes (or **three** reflections if you are an asynchronous student) without needing a valid excuse. Beyond that, you will incur participation grade reductions.

Extensions & Late Papers

Extensions of an appropriate length will still be happily be granted, but only if you email me more than 24 hours in advance of the deadline. Late papers will incur grade reductions of a third of a letter grade per day past the deadline. Exceptions may be granted on a case by case basis due to the challenges associated with our instructional situation.

Participation

Students who meet synchronously with the class will be graded on genuine and respectful engagement with the material, with other students, and with me. You won't be expected to speak to the whole class every day, but I will look to make sure that you are putting in the effort. Students who conduct their work entirely asynchronously (e.g., those in very different time zones) will complete short reflections in which you engage with the recorded discussions.

Drafts

I do not read drafts of papers because there are just too many students. However, I am *more* than happy to go over an outline of your paper in office hours on Zoom!

Grade Breakdown

- Paper 1 [20%]
- Paper 2 [25%]
- Participation [20%]
- Miro Board Engagement [20%]
- Discussions [5%]
- Short written assignments [10%]









Plagiarism and Academic Integrity

Cheating and plagiarism will still not be tolerated in this course. Please acquaint yourself with Georgetown policies [here](#) and [here](#).

Accommodations for Students with Disabilities

Students with a multitude of different disabilities and learning styles will absolutely be given accommodation when and where possible. Please inform me via email of any issues that arise.

course readings and due dates (*Under Construction)

DATE / THEMES	UNITS	READINGS WE WILL DISCUSS	Additional Materials If you're Interested / Due Dates
TBD	AI ETHICS: INTRODUCTION		Stanford Encyclopedia of Philosophy: Ethics of Artificial Intelligence and Robotics
TBD	WHAT IS AI? MACHINE LEARNING, DEEP LEARNING, NEURAL NETWORKS	NVIDIA Podcast: Deep Learning 101	
TBD	AI AND ETHICAL AGENCY	Peter Railton "Machine Morality: Building or Learning" "Why Tool AIs Want to Be Agent AIs"	
TBD	IDENTITY & MORAL LEARNING	Selections from... John Haugeland "Heidegger: On Being a Person" Hilde Lindemann "What Child is This?"	Film: Her
TBD	SELF-DRIVING CARS	The Use and Abuse of the Trolley Problem: Self Driving Cars, Innocent Threats, and the Distribution of Harm, F.M. Kamm	Writing Assignment #1 Due 
TBD	AI ETHICAL REASONING	Modeling and Reasoning with Preferences and Ethical Priorities in AI Systems, Andrea Loreggia, Nicholas Mattei, Francesca Rossi, and K. Brent Venable	
TBD	AI AND EMBODIED VIRTUE	Moral Machines: From Value Alignment to Embodied Virtue, Wendell Wallach and Shannon Vallor	
TBD	ROBOT RIGHTS?	Designing AI with Rights, Consciousness, Self-Respect, and Freedom, Eric Schwitzgebel, with Mara Garza	
TBD	WRITING / REVIEW		Writing Assignment #2 Due 
TBD	WRITING / REVIEW		Writing Assignment #3 Due 
TBD	SURVEILLANCE CAPITALISM	Shoshana Zuboff "The Discovery of Behavioral Surplus"	
TBD	SURVEILLANCE CAPITALISM	Shoshana Zuboff "The Elaboration of Surveillance Capitalism: Kidnap, Corner, Compete"	
TBD	CONTEXTUAL INTEGRITY	Nissenbaum, Helen, "Contextual Integrity Up and Down the Data Food Chain"	Introduction Due 
TBD	SURVEILLANCE EVERYWHERE	Shoshana Zuboff "The Reality Business"	
TBD	MANIPULATION, SELF-AUTHORSHIP	Susser, Roessler, and Nissenbaum, "Technology, Autonomy, and Manipulation"	Daniel Susser "Information Privacy and Social Self-Authorship" Synthesis Due 
TBD	FACIAL RECOGNITION	Shoshana Zuboff "Rendition: From Experience to Data"	
TBD	INTERNET OF THINGS	Shoshana Zuboff "Rendition From the Depths"	
TBD	PROFILING AND NUDGING	Shoshana Zuboff "Make Them Dance"	Argument Due 
TBD	WRITING / REVIEW		
TBD	WRITING / REVIEW		Objections Due (Full 1st Paper Due) 
TBD	Algorithmic Bias	Gabrielle M. Johnson, "Algorithmic Bias: On the Implicit Biases of Social Technology" On Canvas	
TBD		Mohamed, Png, and Isaac, "Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence" (On Canvas)	
TBD		Neely, Erica, "Augmented Reality, Augmented Ethics" (On Canvas)	
TBD		Ramirez and LaBarge, "Real Moral Problems in the Use of Virtual Reality"	Writing Assignment #4 Due 
TBD		James, "Planning for Mass Unemployment: Precautionary Basic Income"	
TBD	SPRING BREAK		
TBD	TRANSHUMANISM	Hava Tirosh-Samuels, "Engaging Transhumanism" (On Canvas)	Transhumanism Vs. Humanism Debate
TBD	PRO TRANSHUMANISM	Bostrom, Nick "In Defense of Posthuman Dignity" (On Canvas) Marsen, Sky "Playing by the Rules—Or Not? Constructions of Identity in a Posthuman Future" (On Canvas)	Elon Musk Neuralink Presentation TEDx Presentation
TBD	AGAINST TRANSHUMANISM	Dupuy, "Cybernetics is Antihumanism: Advanced Technologies and the Rebellion Against the Human Condition"	AI and Artificially Enhanced Brains
TBD	FINAL PAPER		Final Paper Due 