

POLS 016 AI, ETHICS, AND POLITICS

Fall 2026

Yi-Hsuan Huang [*pronounced as e-shuan hwang*]
yhuang6@swarthmore.edu

Class Meetings: T/TH 2:40 - 3:55 pm
Old Science Hall (Trotter) 303

Office Hours @ Old Science Hall 306

T 4:30-5:30 pm

TH 4:30-5:30 pm (Appointment-Only)

Make an [Appointment](#)

Office Hours [Zoom Link](#)

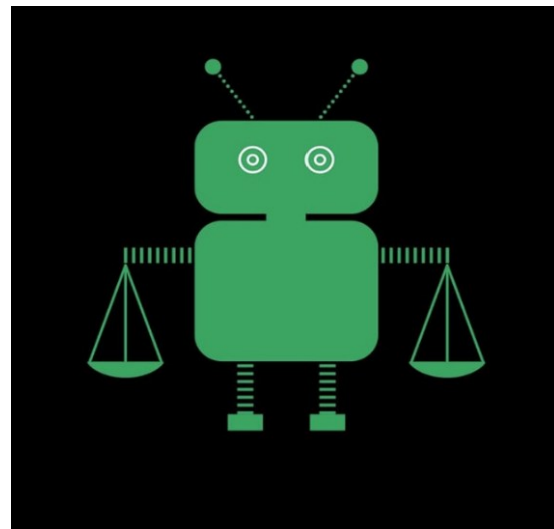
This book tells a story of humanity increasingly lost in its own reflected image, and captive to mechanical mirrors of our own making...We are the source of danger to ourselves from AI, and this is a good thing—it means we hold the power to resist, and the power to heal.

—Vallor, *The AI Mirror* (2024)

Course Description & Objectives

Artificial intelligence (AI) now plays a crucial role in our lives. This class surveys the ethical issues derived from the application of data and algorithm-driven technologies from the angle of political theory, exploring four interrelated topics. “Humanity and Machine” investigates the distinctions between humans and machines and asks what features of humans are irreplaceable. “Moral Agency and Responsibility” asks whether AI can possess certain moral character and be held accountable for ethical mistakes. “Autonomy Under Threats” analyzes ways that AI can undermine our individual autonomy and explores potential solutions. Finally, “Environmental Tolls, Biases, and Regulation” critically evaluates the costs that AI carries.

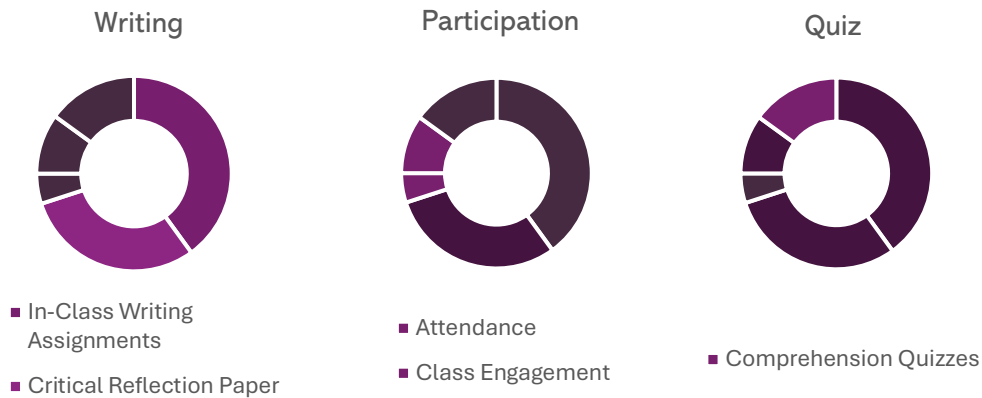
oppression in the use of AI and discusses potential solutions.



(Source: Coeckelbergh, *Robot Ethics*)

This class provides students with the following knowledge and skills:

1. **Familiarity** with the major ethical debates related to the application of AI technologies.
2. **Comprehension** of the critical assessment of AI science from the angle of political theory.
3. **Articulation**, in speaking and writing, of the theoretical significance of AI ethics.
4. **Critical Analysis** of prominent ethical dilemmas of AI technology using learned concepts as an effective speaker, writer, and thinker.
5. **Responsibility** of using AI as an informed user.



Assignments & Evaluations

Attendance	5%	
Class Engagement	10%	
Six Comprehension Quizzes	12%	(can drop two lowest grade)
Two In-Class Issue Analysis Writing	20%*2	(approx. 700 words each)
One Critical Reflection Paper	33%	(5-6 pages, double-spaced)

 **Note:** Final grades are non-negotiable.

Attendance (5%) Regular attendance is required. Each unexcused absence will result in 1% penalty. More than **three** unexcused absences will result in a zero in attendance. Students need to provide documentation (e.g., doctor's notes or proof of medical appointment) to be excused. Please check the Autonomy Token section for attendance flexibility.

Three in-class reviews are part of the attendance grade. Failure to submit any of them will result in 1% grade deduction.

Class Engagement (10%) Discussion is central to this course and will take up 40%-50% of our time. Articulating theoretical concepts is a critical part of the learning process. Conversation in this class is a collective work that helps everyone to clarify key concepts and generate new ideas. Therefore, students are expected to contribute actively and consistently by answering and raising questions, asking for clarifications, and posing their viewpoints.

- Grades are based on three elements: (i) **F**requency of engagement; (ii) **A**ctive listening and responsiveness to others; (iii) **Q**uality of the contributions.

- Students will be asked to compose a series of **Active Reading Logs** throughout the semester. This will be part of the assessment of (iii).
- Behaviors negatively affecting your grade include unapproved phone use, multitasking or distraction, repeated lateness, talking over others, and so on.
- You will receive a midterm engagement review mid-semester. If you have questions about your engagement grade, please reach out before **Oct 22**. This is to ensure sufficient time for improvement. After this date, questions about participation will not be answered.

Comprehension Quizzes (12%) There will be six short quizzes throughout the semester. Quizzes usually include 8-10 multiple-choice and/or 1-2 short-answer questions. You can use up to two autonomy tokens to drop the lowest quiz grade (except for Quiz 6). There will be a comprehensive bonus quiz at the end of the semester, which will earn you a maximum of 1 point to the quiz average.

To guide your reading and the preparation of quizzes, **Reading Questions** will be available on Moodle in advance.

Missed quizzes **cannot** be made up and will result in zeros. If you have an excused absence, the quiz will be *excluded*. If you use an autonomy token to excuse yourself on a quiz date, the quiz will also be dropped automatically, but it will count toward the two-drop limit.

In-Class Issue Analysis Writing (20%*2) You will complete two issue analysis writing assignments. Each assignment will ask you to analyze an ethical dilemma of AI from the perspectives of particular thinkers and provide your own critical responses.

These assignments are **open-note** and are due in class. A few prompts will be given for your selection. Class materials are allowed, but they have to be in physical copies.

Critical Response Paper (33%) This assignment asks you to produce a five-page, double-spaced argumentative essay (approx. 1500 words) that analyzes one of the ethical challenges of AI and proposes and justifies your own solution.

Suggested topics will be provided. You are also allowed to propose your own topic upon approval.

This assignment has three components:

- Focus Statement (3%): You will submit a brief focus statement (max: two pages, double-spaced), which (i) identifies the problem, (ii) explains why the problem requires enhanced attention, (iii) outlines your suggestions, (iv) justifies your suggestions, and (v) provides a list of references. I will either approve or request revisions.
- In-class drafting & peer review (5%): We will sit together in class to produce your first draft. You will be asked to discuss your focus statements with peers and get feedback. This draft will be submitted at the end of class, which will serve as a record of your thought process and inform your final revision.
- Critical Response Paper (25%): Grading is based on (i) command of literature; (ii) clarity and strength of the argument; (iii) originality of the topic; (iv) organization and structure; (v) academic writing ability; (vi) ability to improve and build on feedback.

For all writing assignments, specific guidelines and rubrics will be provided beforehand.

Late Assignment Policy

For unexcused absence, the make-up in-class assignment is subject to a 5% default deduction. You are solely responsible for arranging accommodation if needed. Last-minute requests will not be granted. Unexcused late submissions for every twenty-four hours will result in a deduction of 5%.

Autonomy Tokens!

Autonomy tokens are created to give students autonomy and flexibility in the learning process. Everyone has **three** tokens that can be used in the following three ways:

1. Drop up to two lowest quiz grades with two tokens (except for Quiz 6).
2. Excuse up to three absences without documentation (one token=one absence).
3. Extend the deadline of a take-home assignment to up to 72 hours (one token=24 hours).

How to cash a token:

- Email me about the absence. No explanation is needed. Tokens **will be automatically applied** to any absence requests that **do not** have documentation.
- Indicate the need for an extension when submitting the assignments.
- If you have more than one remaining token after submitting the final essay, I will automatically drop the lowest quiz grade.
- You **should** use tokens for anything that does not have a strong mandatory nature (e.g., travel plans, club events, assignment clashes with other classes, part-time jobs, internships, unless pre-approved by the department or the college).

You can earn **one additional token** by attending a talk related to our course topic (need to demonstrate registration or photo proof) and writing a reflection note of 300 words. Note that you might be asked to share your observations in class.

Accommodation

Accommodation beyond the use of tokens needs to have documented reasons and should be clearly communicated. You should reach out to discuss and confirm the specific request before *each* assignment. Any chronic situation with limited/without documentation should also be communicated clearly to gain approval.

Classroom and Communication

Here are some basic rules to foster mutual respect:

- Laptops or tablets can only be used as note- & exam-takers. Phones are not allowed. Violations will result in penalties in class engagement grades.
- If we have to meet on Zoom, please be in a quiet place, have your camera on, and refrain from multitasking. You should not be walking, driving, or talking to friends simply because we are online.

- Please be mindful of my time, and do not expect last-minute answers or accommodation. Students are expected to do simple research before reaching out. You are welcome to give me a nudge if I do not reply within twenty-four hours.
- Please observe [email etiquette](#).
- Please communicate early and closely with me if you need any assistance. I am committed to supporting your learning process. However, bear in mind that any accommodation I offer needs to be fair for everyone.
- Syllabus may be updated as we move along. I will try to give at least 48 hours' notice for changes that affect your preparation. Please stay tuned. I appreciate your patience in advance.

Office Hours

Office hours are Tuesdays & Thursdays (Thursdays are appointment-only) 4:30-5:30 pm in Old Science Hall (Trotter) 306. Appointments take priority.

- Email me to set up a time outside of these time frames.
- [Zoom](#) is reserved for the time I am not on campus.

Discuss *Democratically*: Practice Mutual Respect in Conversations

Many class topics touch on moral dilemmas and ethical issues that do not have absolute answers. Be ready to encounter viewpoints that challenge your original view. Students are welcomed and encouraged to engage critically with an idea, but should do so in a respectful manner. In class discussion, students are expected to:

- Always assume good rather than malicious intent of others.
- Ask themselves what it means for our society that an idea that seems to be anathema to them is true to another person.
- Ask clarification questions when they do not understand a point of your interlocutor.
- Be ready to move forward on a topic even if no agreement is reached.
- Stay engaged, and use verbal and body language to demonstrate active listening.

Plagiarism and Generative AI Policy

This class has a zero-tolerance policy for plagiarism (See the policy of "[Academic Misconduct](#)"). AI software can only be used as an editing tool unless otherwise specified. Using AI to produce any substantial component of an assignment is considered a violation of academic conduct and will be reported and result in a zero.

Required Materials

All required reading will be provided on Moodle or available online in the Swarthmore Library.

Important Dates!

Sep 3	Semester-Goal Setting Due (In-Class)
Oct 6	In-Class Issue Analysis Writing I
Oct 20	Midterm Reflection Due (In-Class)
Oct 22	Final Date to Discuss Engagement Grade
Nov 19	In-Class Issue Analysis Writing II
Dec 5	Focus Statement Due 10:00 pm
Dec 8	First Draft Writing Due (In-Class)
Dec 16	Critical Response Paper Due 11:59 pm

Schedule

Date	Required Readings	Event
------	-------------------	-------

Sep 1 Introduction
OECD (2019) Artificial Intelligence in Society

Sep 3 Bhargava (2008) Ch1; 2 What is Political Theory?
Coeckelbergh (2020) AI Ethics Ch3 *Semester Goal-Setting*

Sep 8 Winner (1980) Do Artifacts Have Politics?

Humanity & Machine

AI is meant to imitate human features. To what extent is AI similar or dissimilar to humans? Can AI replace humans? Are there elements of humans that cannot or *should not* be replaced?

Sep 10 Vallor (2024) AI Mirror Introduction, Ch2
Chiang (2023) ChatGPT *Is* a Blurry JPEG of the Web
Chiang (2024) Why AI cannot make art?

Sep 15 Danaher (2019) Automation and Utopia Ch1; Ch3 p. 53-56, 80-86; Ch4 p. 87-99, 125 (Table 4.2) *Mock Quiz: 9/3-9/15*

Sep 17 Nozick, The Experience Machine
Danaher (2019) Automation and Utopia Ch7 p. 214-23-, 234-236, 251-257, 263-269 *Quiz 1: 9/8-9/17*

Sep 22 Brennan (2016) Against Democracy, Ch. 1, p. 1-19
Hildago (2019) "Augmented Democracy" (Webpage)

- Sep 24 Fossen and Anderson (2014) What's the point of voting advice applications?
Warren (1992) Democratic Theory and Self-Transformation, excerpt. p. 16-22

Moral Agency & Responsibility

AI is now used to analyze resumes or recommend court decisions. Many of these decisions involve moral considerations. Should AI be trusted to make moral decisions? Who is responsible when machines make mistakes?

- Sep 29 Anderson and Anderson (2007) Machine Ethics: Creating an Ethical Intelligent Agent
- Oct 1 Coeckelbergh (2010) Moral appearances: emotions, robots and human morality **Quiz 2: 9/17-10/1**
- Oct 6 *In-Class Issue Analysis I*
- Oct 8 **No Class Meeting:** Vallor (2024) AI Mirror Ch4
- Oct 13 & 15 Fall Break
- Oct 20 Nissenbaum (1996) Accountability in a Computerized Society *Midterm Reflection*
- Oct 22 Goetz (2022) Mind the Gap: Autonomous Systems, the Responsibility Gap, and Moral Entanglement **Quiz 3: 10/8-10/22**

Autonomy Under Threat: Dangers and Opportunities

The wide application of AI raises a range of questions about surveillance, manipulation, and opacity. As members of liberal society, how do we protect our autonomy against these threats? Are we doomed or are there meaningful way-outs?

- Oct 27 Zerilli et al. (2021) A Citizen Guide to Artificial Intelligence Ch7
Veliz (2021) The Value of Privacy
- Oct 29 Spiekermann et al. (2021) Big Data Justice: A Case for Regulating the Global Information Commons
- Nov 3 Carugati (2020) "A Council of Citizens Should Regulate Algorithms"
Lazar (2024) Legitimacy, Authority, and Democratic Duties of Explanation

- | | |
|--------|--|
| Nov 5 | Himmelreich (2021) Against “Democratizing AI,” excerpt. p. 1342-1344 |
| | Danaher (2016) The Threat of Algocracy: Reality, Resistance and Accommodation, excerpt. p. 245-249; 255-266 |
| Nov 10 | Capraro, V., et al. (2024) The Impact of Generative AI on Socioeconomic Inequalities and Policy Making
Presentation Preparation: Policy Project |

Environmental Tolls, Biases, and Regulation

- | | |
|--------|--|
| Nov 12 | Presentation: Policy Project
Leonelli (2025) Environmental Intelligence |
| Nov 17 | Benjamin (2019) Race after technology Introduction |
| Nov 19 | <i>In-Class Issue Analysis II</i> |
| Nov 24 | Noble (2018) Algorithms of Oppression Ch2 & Ch6 |
| Dec 1 | Barocas et al. (2023) Fairness and Machine Learning Ch4 |
| Dec 3 | Rushoff (2019) Team Human (excerpt TBD) |
| Dec 8 | LDOC: <i>In-Class Drafting & Peer Review</i> |

- | | | |
|-------|---|---|
| Dec 3 | Rushoff (2019) Team Human (excerpt TBD) | Quiz 6: 11/24-12/3;
Bonus Quiz;
Comprehensive |
|-------|---|---|