# Artificial Intelligence Policy

COURSE INFORMATION

*Term:* Spring 2025
*Level:* Intermediate Seminar
*Meet:* W/F, 12:45-2:05
*Room:* Marley 325

*Instructor:* Jack Reilly
*Office:* Eggers 225F
*Appointments:* jacklreilly.github.io
*Office Hours:* Mon & Wed, 2:15-3:15; by appt

DESCRIPTION

This course serves as in introduction to the politics and policy of artificial intelligence. It presumes no particular prior knowledge of artificial intelligence, political science, public policy, or computer science; rather, it will introduce students to what artificial intelligence is as well as to the ethics of AI, its social implications, and the policy choices around AI that currently face governments, corporations, and organizations worldwide. This is a "non-technical" course: assignments will include reading, class discussion and leadership, debates, writing, and learning about the use of AI tools and prompts to generate output. Students will not be expected, however, to program or write any formal code.

STRUCTURE

***Principle - Practice - Politics - Policy***
The course has four sequential units. First, we start with an introduction to the philosophy of artificial intelligence itself. What is "intelligence" and "consciousness" in the first place? What are minds? What are agents and legal persons, and when are agents responsible for their actions? What is artificial intelligence, and what is artificial "general" intelligence? Second, we proceed to discuss the mechanics of current computational approaches to AI, with a particular focus on large language models (such as OpenAI's ChatGPT), including approachable, non-technical overviews of the principles behind machine learning, neural networks, and assorted technologies that have given rise to the current state of the art in AI. Third, we consider the societal impacts of current artificial intelligence systems, including data practices as well as current AI applications, uses, risks, and ethics. Fourth, we turn to policy, examining the decisions in front of governments, companies, and organizations in the United States and worldwide as well as complications in the policy process for producing effective AI legislation.

PREREQUISITES

Introductory coursework in (any) related area(s). No particular knowledge is expected, but the reading level is high and this course may not be suitable for first year students.

## Learning Objectives

1. Explain the core technical, economic, and political forces that shape how modern AI models are developed, deployed, and governed.
2. Critically evaluate and debate alternate perspectives on contemporary AI policy issues, using evidence to construct and defend reasoned arguments.
3. Design and evaluate AI policy proposals by applying principles of data ethics, policy analysis, and policy design to real-world issue domains.

## Materials

BOOKS

*Readings will include selections from:*

- Mitchell, 2019. *Artificial Intelligence: A Guide for Thinking Humans*. Picador.

- Haugeland et al, 2023. Mind Design III: Philosophy, Psychology, and Artificial Intelligence.
- Crawford, 2022. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.* Yale.
- Bullock, et al (editors). *The Oxford Handbook of AI Governance*. Oxford.

*Readings will be available electronically*

COMPUTATION

We will use a variety of AI models in class as end-users, so you should have a computer you can use to access them. No coding will be required in the course.

## Course Requirements

OVERVIEW

This is a seminar-style class.

1. Seminar Preparation *(20%)*
   (a) Attendance, Reading, and Participation
   (b) Discussion Questions
   (c) AI News Submissions
   (d) Discussion Leadership
2. Prepared Work for Class *(50%)*
   (a) Debates *(40%)*
   (b) Short Assignments *(10%)*
3. Core Exam *(10%)*
4. Final Project *(20%)*
   (a) Presentation
   (b) Paper

SEMINAR PREPARATION

**Daily Attendance, Participation, and Reading.** This is an upper level preceptorial-style seminar taught at a high level: it only works if everyone takes part. I have a high opinion of your intellectual capability and a similarly high opinion of the quality of your potential contributions - make sure to show them to me!

**Discussion Questions:** Submit discussion questions or points (broadly construed) to the class by 9 AM on the day of class. Make sure to send messages to the course Teams chat so everyone can read them.

**AI News Submissions:** In class, we'll take formal time in class to discuss AI news from the week. (Especially if something particularly interesting happens . . . which, because this is AI, happens a lot). If you come across interesting AI news - or broad tech industry news - submit it to our Teams chat.

**Discussion Leadership:** Each class has an assigned discussion leader from among the students. While all students are under an obligation to read (or listen, or watch) the material each day

from class, the discussion leader has a special role as the person to starts our discussion for the day rolling, and steps in should our discussion of the day's material falter.

Finally, a note on seminar **etiquette**: course participants should be courteous to the professor and fellow students. Attend class on time, listen to fellow students when they talk, and disagree (or agree) with others arguments professionally . Keep cell phones and other technology silenced and out of sight unless doing something directly relevant to the discussion in the class.

**Preparation Grading Policy:** A seminar like this doesn't work if everyone only seeks to do the bare minimum. Informed class participation is expected and required each day in class; you shouldn't be counting things like *"I submitted a requisite number of news articles already this semester, so I no longer need to submit news articles"*. Participate across all elements of the seminar as and when the spirit moves you, and keep that participation up through the semester, and you'll do fine for this portion of the class.

With that said, it is useful to have some baselines and expectations for everyone to keep in mind. A few that are useful:

- *Attendance*: Everyone has a few bad days. You can miss up to two days of class without excuse or penalty.
- *Discussion Questions*: At least one each day. (Although do note: you should sometimes submit more than one, too!)
- *In-Class Participation*: We should hear your contributions every day. (We'll have systems in class to make sure that everyone gets a turn).
- *News Submissions*: At least four news submissions per semester (hopefully more)!
- *Discussion Leadership*: Split evenly according to the number of students in the class.

CLASS WORK

**Debates**: We will have six formal debates in the class, following a modified Lincoln-Douglas debate format (if you are unfamiliar with these rules, don't worry! We'll cover them.) Each student will have a role in each debate, although not all roles are equal: it is easier to *judge* than it is to *affirm* or *negate*. All students will have equal debate responsibilities across the class.

**Short Assignments:** Small assignments using AI tools will be given through the semester.

CORE EXAM

A short in class quiz on AI basics.

FINAL PROJECT

A final project on an aspect of AI, including both a 10-minute presentation and 5-7 page paper. (There will be options.)

ASSORTED CLASS POLICIES

**Office Hours:** I encourage you to stop by my office hours at any point if you have questions about the course, the readings, school, etc. I have drop-in hours regularly scheduled, as well as additional hours by appointment (schedule by email or at jacklreilly.github.io ).

**E-mail:** Students can generally expect a response to all e-mails within 24 hours, excepting weekends. I am happy to answer any questions over e-mail that require less than a paragraph in response. Questions that require more than a short paragraph in response should be addressed in person.

**Blackboard:** We will organize our class primarily over Blackboard; including reading availability and assignment submission. Make sure you are able to access it effectively.

**Teams:** In addition to Blackboard, we'll have a class discussion channel in MS Teams for discussion question sharing as well as for sharing news articles. Make sure you set up easy access for yourself wherever is convenient.

**Artificial Intelligence and Class Technology Policy:** To be discussed (and decided upon!) in class.

LEARNING OBJECTIVE ASSESSMENT

**Course Assessment Plan**

| Learning Objective | Assessment Measure |
|---|---|
| Explain the core technical, economic, and political forces that shape how modern AI models are developed, deployed, and governed. | Core Exam |
| Critically evaluate and debate diverse perspectives on contemporary AI policy issues, using evidence to construct and defend reasoned arguments. | Debates |
| Design and evaluate AI policy proposals by applying principles of data ethics, policy analysis, and policy design to real-world issue domains. | Final Project |

## Schedule

LIST OF TOPICS *Subject to change*

| W | | Wednesday | Friday |
|---|---|---|---|
| | | UNIT I: AI IN PRINCIPLE AND PRACTICE | |
| 1 | *Jan 13* | Class Introduction | Intelligence, Consciousness, Sentience |
| 2 | *Jan 20* | Emergence | DEBATE I |
| 3 | *Jan 27* | Evaluating Intelligence | AI Embodiment, Agency, & Responsibility |
| 4 | *Feb 3* | DEBATE II | Language & Intelligence |
| 5 | *Feb 10* | Computation & AI | Silicon: Technology & Geopolitics |
| 6 | *Feb 17* | DEBATE III | AI as a Business |
| 7 | *Feb 24* | CORE EXAM | Harms of AI |
| | | UNIT II: AI IN POLITICS AND POLICY | |
| 8 | *Mar 3* | DEBATE IV | Data and AI |
| | | SPRING BREAK | |
| 9 | *Mar 17* | Labor Replacement I | Labor Replacement II |
| 10 | *Mar 24* | DEBATE V | Environmental Factors |
| 11 | *Mar 31* | Democracy & AI | Autocracy & AI |
| 12 | *Apr 7* | DEBATE VI | Geopolitics Redux |
| 13 | *Apr 14* | Governance and AI | Singularities, xRisk, & AGI |
| 14 | *Apr 21* | FINAL PRESENTATIONS | FINAL PRESENTATIONS |
| | | FINALS WEEK | |

All readings may be found on the course blackboard page. Readings will be posted at least one week ahead of time. Each day will have one primary reading that should be read closely, a series of simpler secondary readings (often, news coverage, podcasts, and/or videos) that should be browsed or scanned, and (frequently) further secondary reading for those interested

in diving deeper. The first week's reading may be found below as an example.

SEMINAR 1.1 **Course Introduction**

1. Think: *What does "Artificial Intelligence" mean to you? What are you expecting to learn from this class?"*
2. Read: Syllabus

SEMINAR 1.2 **Intelligence, Consciousness, and Sentience, Oh My!**
Getting clear on terms and possibilities

1. Think: *What is the difference between Intelligence, Consciousness, and Sentience? What about between Autonomous Systems, Artificial Intelligence, and Artificial General Intelligence? How can we become sufficiently precise in our language and thinking?*
2. Read: Mitchell, Melanie. 2023. "Why AI Is Harder than We Think" from *Mind Design III*
3. Listen: *Complexity* podcast, Season 2, Episode 1: "What is Intelligence?"
4. Browse:
   - Lenharo, Mariana. 2023. "Decades-long bet on consciousness ends - and it's philosopher 1, neuroscientist 0" *Nature News*
   - Lenharo, Mariana. 2024. "What should we do if AI becomes conscious? These scientists say it's time for a plan". *Nature News*
   - Marcus, Gary. 2025. "Sam Altman thinks that AGI is basically a solved problem. I don't. Here's why." Blog post.
5. Consult as desired:
   - Altman, Sam. 2025. "Reflections". Blog post.
   - "Consciousness". Stanford Encyclopedia of Philosophy.
   - "Artificial Intelligence". Stanford Encyclopedia of Philosophy.

SEMINAR 2.1 **Emergence**

1. Think: *Is "intelligence" an emergent phenomena? Or something different? What does it mean to be "intelligent"? What about consciousness? When do we start treating a system as more than the sum of its parts, and why?*
2. Read: Boden, "Autonomy and Artificiality". from *The Philosophy of Artificial Life*
3. Browse:
   - Roose, "A Conversation With Bing's Chatbot Left Me Deeply Unsettled" NYTimes.
   - Schultz, "Flocking together: Study shows how animal groups find their way"
   - Krakauer, "What is Complexity?" *Santa Fe Institute Lecture*: `https://www.youtube.com/watch?v=JR93X7xK05o`
4. Consult as desired:
   - Bedau, Mark. "Weak Emergence"
   - Johnson, "Eliminating the mystery from the concept of emergence"

- Hofstadter, "Gödel, Escher, Bach: an Eternal Golden Braid"
- Holland, John. *Emergence*
- Wolfram, *A New Kind of Science*
- Simon, Herb. *Sciences of the Artificial*
- "Emergent Properties", *Stanford Encyclopedia of Philosophy*

SEMINAR 2.2 **Debate I: AI & Common Sense**

SEMINAR 3.1 **Evaluating Minds and Intelligence**

1. Think: *How can we, as fellow humans, assess whether or not a non-human "intelligence" is actually . . . intelligent? What kind of guidelines and rules of thumb should we form? What makes something an autonomous system, an intelligent system, a general intelligence system, or even a superintelligence?*
2. Read: Searle, "Minds, Brains, and Programs" (*aka "The Chinese Room" argument*) from *Mind Design III*
3. Watch: "Artificial Intelligence & Personhood" *Crash Course Philosophy #23*. https://www.youtube.com/watch?v=39EdqUbj92U
4. Listen: *Complexity* podcast, Season 2, Episode 5: "How do we assess intelligence?" https://www.santafe.edu/culture/podcasts/episode-5-how-do-we-assess-intelligence
5. Browse:
   - Mitchell, "The Turing Test and our shifting conceptions of intelligence"
6. Background Reading and Reference:
   - Turing, "Computing Machinery and Intelligence" (*aka "The Turing Test"*)
   - Lycan, "Robots and Minds"
   - Dennett, "True Believers: The Intentional Strategy and Why it Works"
   - Boden, "Escaping from the Chinese Room"
   - "The Chinese Room Argument", *Stanford Encyclopedia of Philosophy*
   - "The Turing Test", *Stanford Encyclopedia of Philosophy*

SEMINAR 3.2 **AI Embodiment, Agency, & Responsibility**

1. Think:
   - *How important is it for intelligence to be "embodied" in a physical agent that interacts with the "real" world, as opposed to digital representations of the world through text, images, video, and the internet? Does embodiment (or lack thereof) change how we think of learning?*
   - *What about responsibility? Are AI agents responsible for their behavior? Should they be, legally? Should the companies that make them? Does whether agents are embodied change how with think about this responsibility?*
2. *Complexity* podcast, Season 2, Episode 4: "Babies vs. Machines" https://www.santafe.edu/culture/podcasts/episode-4-babies-vs-machines"
3. Read:

- Liu and Wu, "A Brief History of Embodied Artificial Intelligence, and its Outlook" https://cacm.acm.org/blogcacm/a-brief-history-of-embodied-artificial-intelligenc
- NHTSA Finds Teslas Deactivated Autopilot Seconds Before Crashes https://www.motortrend.com/news/nhtsa-tesla-autopilot-investigation-shutoff-crash/
- Ziegler, Bart. 2023. When Will Cars be Fully Self Driving? https://www.wsj.com/articles/cars-self-driving-when-c6ae4fdc

4. Browse:
   - The evolving safety and policy challenges of self-driving cars. Brookings https://www.brookings.edu/articles/the-evolving-safety-and-policy-challenges-of-self-dri
5. Background Reading and Reference:
   - Ganesh, "The ironies of autonomy" https://www.nature.com/articles/s41599-020-00646-0
   - Chan et al, "Infrastructure for AI Agents" https://arxiv.org/abs/2501.10114
   - Geistfeld, "A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation"
   - NHTSA, Automated Vehicles for Safety https://www.nhtsa.gov/vehicle-safety/automated-vehicles-safety
   - Google's AI Responsibility Report https://ai.google/static/documents/ai-responsibility pdf
   - Stanford Encyclopedia of Philosophy, "Doing vs Allowing Harm"; especially on the Trolley Problem section
   - Himmelreich, "Never Mind the Trolley: The Ethics of Autonomous Vehicles in Mundane Situations" https://link.springer.com/article/10.1007/s10677-018-9896-4

SEMINAR 4.1 **Debate II: The Turing Test**

SEMINAR 4.2 **Language and Intelligence**

1. Think: *Are thought and language separable? What about intelligence and language? To what extent does intelligence depend upon language acquisition and development, and to what extend does the ability to use language illustrate intelligence?*
2. Listen: *Santa Fe Institute Complexity Podcast,* Episode 2: The Relationship between Language and Thought
3. Read: Mahowald, et al. 2024. "Dissociating language and thought in large language models". *Trends in Cognitive Science.*
4. Browse
   - Mitchell, "The Metaphors of Artificial Intelligence"
   - Mitchell and Krakauer, "The debate over understanding in AI's large language models"
   - *Look at the pictures:* Fedorenko, et al. 2024. "The language network as a natural kind within the broader landscape of the human brain." *Nature Reviews Neuroscience*
5. Reference
   - Chomsky, *Reflections on Language*
   - Rumelhart, "The Architecture of the Mind: A Connectionist Approach" *Mind Design III*
   - Peper, "A general theory of consciousness II: The language problem" *Communicative & Integrative Biology*

SEMINAR 5.1 **Computation & AI**

1. Think: *How do computers work? Can they "simulate" minds? Or can they be minds? Practically, what goes into making a computer, and how do the economics and technology of computation structure both geopolitics as well as the nature of AI itself?*
2. Listen: *Santa Fe Institute Complexity Podcast*, Episode 3: What Kind of Intelligence is an LLM?
3. Watch:
   - CGP Grey, "How Machines Learn" https://www.youtube.com/watch?v=R9OHn5ZF4Uo *(Note: this video has been retitled "How AI, like ChatGPT, learns")*
     - and "How Machines *Really* Learn" https://www.youtube.com/watch?v=wvWpdrfoEv0 *New Title: How AI, like ChatGPT, *really* Learns*
   - 3Blue1Brown, "Large Language Models explained briefly" https://www.youtube.com/watch?v=LPZh9BOjkQs
     - *recommended, but not required:* "Transformers explained visually" https://www.youtube.com/watch?v=wjZofJX0v4M
4. Browse:
   - Hicks et al, "ChatGPT is Bullshit" *Ethics and Information Technology*
   - Wolfram, S. "What is ChatGPT doing?" https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/
   - A few items on evaluating AI Intelligence:
     - Humanity's Last Exam: Website www.lastexam.ai and article: Phan et al, *Humanity's Last Exam*
     - Ullman, "Large Language Models Fail on Trivial Alterations to Theory-of-Mind Tasks" https://arxiv.org/pdf/2302.08399
     - Mitchell, "How do we know how smart AI systems are?" https://www.science.org/doi/10.1126/science.adj5957
5. Reference
   - Only a few videos from 3Blue1Brown's series on Neural Networks are assigned above, but the entire sequence is worth watching for a non-technical overview: https://www.3blue1brown.com/topics/neural-networks
   - Similarly, the *Crash Course: Artificial Intelligence* series provides a high level overview of many current AI principles: https://www.youtube.com/playlist?list=PL8dPuuaLjXtO65LeSb5XQ51par_b
   - Alternately, there is *Crash Course: Computer Science* for more about computers themselves: https://www.youtube.com/playlist?list=PL8dPuuaLjXtNlUrzyH5r6jN9ulIgZBpdo
   - Finally, if you are already statistically inclined, the final episodes of *Crash Course: Statistics* also cover elements of machine learning and big data analysis building off of core statistical concepts
   - Rumelhart, "The Architecture of the Mind: A Connectionist Approach" from *Mind Design III*
   - Churchland and Sejnowski, "The Computational Brain" from *Mind Design III*
   - Cowie and Woodard, "The Mind is Not (Just) a System of Modules Shaped (Just) by Natural Selection" from *Mind Design III*
   - Vaswani et al, "Attention Is All You Need" (Google paper) https://arxiv.org/abs/1706.03762

SEMINAR 5.2 **Silicon: Technology and Geopolitics**

1. Think: *How do the technology demands of AI influence the behavior and relationship of governments and states worldwide? How does the need for advanced silicon fabs and manufacturing, in particular, structure geopolitical relations?*
2. Listen: Sharp Tech Podcast, "Gelsinger Out at Intel, 20 Years of Structural Challenges and Strategic Blindspots, The Board and Whats Next" https://sharptech.fm/member/episode/gelsinger-out-at-intel-20-years-of-structural-challenges-and-strategic-blind
3. Read:
   - "A Chance to Build" https://stratechery.com/2024/a-chance-to-build/
   - "It's the Most Indispensable Machine in the World - and It Depends on This Woman" *Wall Street Journal* https://www.wsj.com/tech/ai/asml-euv-machine-lithography-chips-96
   - "The CHIPS Act: How U.S. Microchip Factories Could Reshape the Economy" *The Councile on Foreign Relations*. https://www.cfr.org/in-brief/chips-act-how-us-microchip-
   - "Micron signs historic deal with feds for $6.1 billion to build chip plants in Central NY" *Syracuse.com* https://www.syracuse.com/business/2024/12/micron-signs-historic-d html
   - *The Economist*, "America takes on China with a giant microchips bill" https://www.economist.com/united-states/2022/07/29/america-takes-on-china-with-a-giant-micro
   - *The Economist*, "The gap between global threats and American power will grow in 2025" https://www.economist.com/the-world-ahead/2024/11/18/the-gap-between-global
4. Reference
   - Miller, 2023. *Chip War: The Fight for the World's Most Critical Technology*. Simon & Schuster.
   - *The Silicon Engine: A Timeline of Semiconductors in Computers*. Computer History Museum. https://www.computerhistory.org/siliconengine/
   - *The Economist* "New fronts will open up in the chip wars in 2025" https://www.economist.com/the-world-ahead/2024/11/20/new-fronts-will-open-up-in-the-chip-war
   - *Video:* Lessons from Intel's First Foundry https://www.youtube.com/watch?v=-Y9LWYmVQu0&t=5s
   - Stratechery: "Intels Death and Potential Revival" https://stratechery.com/2024/intels-death-and-potential-revival/
   - *Broader references on geopolitics, broadly:*
     - Nye, "Soft Power" https://www.jstor.org/stable/1148580
     - "Foreign Aid as Foreign Policy Tool" https://doi.org/10.1093/acrefore/9780190228637.013.332
     - Belt and Road: Beijing Spins a Web of Chinese Infrastructure https://www.wsj.com/articles/beijing-spins-a-web-of-chinese-infrastructure-1484560801 and China Reboots Belt and Road https://www.wsj.com/world/china/china-reboots-its-belt-and-road-initiative-99784632
     - China in Latin America: https://www.cfr.org/backgrounder/china-influence-latin-america-argentina-br
     - Current Taiwan-related events: https://www.economist.com/leaders/2025/02/13/countering-chinas-diplom https://www.economist.com/international/2025/02/09/chinas-stunning-new-campaign-to-turn-the-wo
     - American Military Power: https://www.economist.com/briefing/2025/02/13/americas-military-supremacy

SEMINAR 6.1 **Debate III: Silicon Geopolitics**

SEMINAR 6.2 **Is AI Good Business?**

1. Think: *What monetizable product is it that AI companies are making? Is it cost-effective, given labor, compute, and energy costs? Absent venture capital, how profitable is it likely to be short-, medium-, and long-term?*
2. Read/Listen: Stratechery Daily Update, February 18, 2025: Grok-3, The Nvidia Shortcut, Competitive Implications (Blackboard upload)
3. Read:
   - Paterson, Cal. "Building LLMs is probably not going be a brilliant business" Blog post: https://calpaterson.com/porter.html
   - "Researchers created an open rival to OpenAIs o1 reasoning model for under $50" https://techcrunch.com/2025/02/05/researchers-created-an-open-rival-to-openais-o
   - *The Economist:* "Will the bubble burst for AI in 2025, or will it start to deliver?" https://www.economist.com/the-world-ahead/2024/11/18/will-the-bubble-burst-for-a and and "The very real constraints on artificial intelligence in 2025" https://www.economist.com/the-world-ahead/2024/11/20/the-very-real-constraints-on-artificial
   - Zuckerberg, Mark. "Open Source AI is the Path Forward" https://about.fb.com/news/2024/07/open-source-ai-is-the-path-forward/
   - Leaked Google Memo: "We Have No Moat, And Neither Does OpenAI" https://semianalysis.com/2023/05/04/google-we-have-no-moat-and-neither/
   - Glazer, "Why AI Risks Are Keeping Board Members Up at Night" *Wall Street Journal* https://www.wsj.com/business/c-suite/ai-risk-management-boardroom-b8956c61
   - "Apple Engineers Show How Flimsy AI Reasoning Can Be" Wired. https://www.wired.com/story/apple-ai-llm-reasoning-research/
4. Browse:
   - *Make sure to be familiar with the recent release of Deepseek, an open LLM from a Chinese company. Sources include:*
     - Thompson, Ben. DeepSeek FAQ https://stratechery.com/2025/deepseek-faq/
     - Wade, "American AI Has Reached its Sputnik Moment" https://thehill.com/opinion/technology/5024271-ai-sputnik-moment-china-challenge/ *The Hill*
     - "Chinese AI Is Catching Up, Posing a Problem for Donald Trump" https://www.economist.com/leaders/2025/01/23/chinese-ai-is-catching-up-posing-a-dile
   - Thompson, Ben, "AI's Uneven Arrival" https://stratechery.com/2025/ais-uneven-arrival
   - WSJ, Musk's $97.4 Billion OpenAI Bid Piles Pressure on Sam Altman https://www.wsj.com/tech/ai/musks-97-4-billion-openai-bid-piles-pressure-on-altman-f6749e6c
5. Reference
   - GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models https://arxiv.org/pdf/2410.05229

SEMINAR 7.1 **Core Exam**

SEMINAR 7.2 **The Harms of AI**

1. Think: *What societal harms are there to the rise of current formulations of AI?*
2. Read:

- Acemoglu, "The Harms of AI," *The Oxford Handbook of AI Governance*

3. Browse:
   - "Pause Giant AI Experiments: An Open Letter" https://futureoflife.org/open-letter/pause-giant-ai-experiments/ and "Policymaking in the Pause" https://futureoflife.org/document/policymaking-in-the-pause/
   - Grace et al, 2024. "Thousands of AI Authors on the Future of AI" Preprint. https://arxiv.org/abs/2401.02843

SEMINAR 8.1 **Debate IV: AI Harm Mitigation**

SEMINAR 8.2 **Data & AI**

1. Think: *How does the reliance on data in AI structure the way we think of AI's potential and power? When we think of the assorted challenges, from the technical to the social to the legal and back again, including training set issues, algorithmic bias, long tail problems, copy-write issues, how does the AI industry respond? How should the AI industry respond?*
2. Read:
   - Bender et al, 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* https://dl.acm.org/doi/10.1145/3442188.3445922
   - "How Tech Giants Cut Corners to Harvest Data for A.I." NYTimes. https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligen
   html
3. Listen: Ezra Klein Show, "The Government Knows AGI is Coming""
4. Browse:
   - "Oh Dear, Did Someone Steal Something from OpenAI?" https://www.wsj.com/tech/ai/oh-dear-did-someone-steal-something-from-openai-8e8a267c
   - "Meta torrented over 80 TB of Pirated Books to Train AI, Authors Say" https://arstechnica.com/tech-policy/2025/02/meta-torrented-over-81-7tb-of-pirated-books-
   - "OpenAI Says DeepSeek May Have Improperly Harvested Its Data" NYTimes. https://www.nytimes.com/2025/01/29/technology/openai-deepseek-data-harvest.html
   - "Thomson Reuters wins AI copyright 'fair use' ruling against one-time competitor" Reuters. https://www.reuters.com/legal/thomson-reuters-wins-ai-copyright-fair-use-
   - "Federal Court Rules Work Generated by Artificial Intelligence Alone Is Not Eligible for Copyright Protection" Press Release. https://www.klgates.com/Federal-Court-Rules-Wo

SEMINAR 9.1 **Labor Replacement: Challenges**

1. Think: *What happens as computers and automation replaces humans as the best way to produce economic value?*
2. Watch: CGP Grey, "Humans Need Not Apply" https://www.youtube.com/watch?v=7Pq-S557XQU
3. Read:
   - Frey and Osborne, 2017. "The future of employment: How susceptible are jobs to computerisation?"

- Frey and Osborne, 2023. "Generative Al and the Future of Work: A Reappraisal"

4. Browse:
   - Felten, "How will Language Modelers like ChatGPT Affect Occupations and Industries?" https://arxiv.org/abs/2303.01157
   - Google, "Towards an AI co-scientist" https://research.google/blog/accelerating-scienti
   - Generative AI for Economic Research https://genaiforecon.org/JEL-2024-Dec-LLMsCollabo pdf
   - The Economist, "How AI will divide the best from the rest" https://www.economist.com/finance-and-economics/2025/02/13/how-ai-will-divide-the-best-from-the-rest
   - Pew Research, Workers Experience with Chatbots and their Jobs https://www.pewresearch.org/social-trends/2025/02/25/workers-experience-with-ai-chatbots-in-their-jobs/
   - Thompson, 2015. "A World Without Work". *The Atlantic* https://www.theatlantic.com/magazine/archive/2015/07/world-without-work/395294/

SEMINAR 9.2 **Labor Replacement: Solutions**

1. **Reminder: Assignment 1 (LLMs) is due today!**
2. Think: *How do we re-organize society and the economy in light of AI-sparked labor market disruption? Or do we even need to - will the labor market take care of itself?*
3. Read: Stiefenhofer, 2025. "Artificial General Intelligence and the End of Human Employment: The Need to Renegotiate the Social Contract" https://arxiv.org/abs/2502.07050
4. Browse:
   - *Perhaps AI won't be as destructive to labor markets as we think?*
     - Edwards, "AI could create 78 million more jobs than it eliminates by 2030" *Ars Technica* https://arstechnica.com/ai/2025/01/ai-could-create-78-million-more-jo
     - Pethokoukis, "Another Promising Data Point on the Productivity Impact of AI" *AEIdeas* https://www.aei.org/economics/another-promising-data-point-on-the-pro
   - *Some UBI (Universal Basic Income) Explainers:*
     - "What is Universal Basic Income" Stanford Basic Income Lab. https://basicincome.stanford.edu/about/what-is-ubi/
     - "What Would Happen If We Just Gave People Money?" FiveThirtyEight.com. https://fivethirtyeight.com/features/universal-basic-income/
5. Background/Further Reading:
   - *General:*
     - Acemoglu and Johnson, "Learning from Ricardo and Thompson: Machinery and labor in the early industrial revolution and in the age of AI" https://www.annualreviews.org/content/journals/10.1146/annurev-economics-091823-025129
     - Korinek, "Economic Policy Challenges for the Age of AI"
     - Lobel, 2024. "The Future of Work in the Era of AI"
     - Frey, 2019. *The Technology Trap: Capital, Labor, and Power in the Age of Automation*
   - *Universal Basic Income, in particular:*
     - Bidadanure, "The Political Theory of Universal Basic Income" https://www.annualreviews.org/content/journals/10.1146/annurev-polisci-050317-070954
     - Pettit, "A Republican Right to Basic Income?" https://www.degruyter.com/document/doi/10.2202/1932-0183.1082/html

SEMINAR 10.1 **Debate V: Labor Displacement and Universal Basic Income**

SEMINAR 10.2 **Power & the Environment**

1. Think: *How do we manage the massive energy demands of AI?*
2. Listen: Stratechery, "An Interview with OpenAI CEO Sam Altman" https://stratechery.com/2025/an-interview-with-openai-ceo-sam-altman-about-building-a-consumer-tech-comp *(Not about power and energy, directly, but very interesting nonetheless!)*
3. Read: Stover, 2024. "AI goes nuclear" *The Bulletin of the Atomic Scientists* https://thebulletin.org/2024/12/ai-goes-nuclear/
4. Browse:
   - Chen, 2025. "How much energy will AI really consume? The good, the bad and the unknown" *Nature* https://www.nature.com/articles/d41586-025-00616-z
   - The Economist, "How safe is nuclear energy?" https://www.economist.com/graphic-detail/2022/07/19/how-safe-is-nuclear-energy
   - Mandler, 2024. "Three Mile Island nuclear plant will reopen to power Microsoft data centers" https://www.npr.org/2024/09/20/nx-s1-5120581/three-mile-island-nuclear-po
   - Ars Technica, "Can we make AI less power-hungry? These researchers are working on it." https://arstechnica.com/ai/2025/03/can-we-make-ai-less-power-hungry-these-
   - Erdenesanaa, 2024. "A.I. Could Soon Need as Much Electricity as an Entire Country" *NYTimes* https://www.nytimes.com/2023/10/10/climate/ai-could-soon-need-as-much-el html
   - Goldman Sachs, "Is nuclear energy the answer to AI data centers power consumption?" https://www.goldmansachs.com/insights/articles/is-nuclear-energy-the-answer
   - Bacquero, "The Energy Footprint of Humans and Large Language Models" https://cacm.acm.org/blogcacm/the-energy-footprint-of-humans-and-large-language-models
5. Background/Further Reading: Crawford, 2022. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.* Yale.

SEMINAR 11.1 **Democracy & AI**

1. Think: *How does the use of AI support or undermine shared civic life?*
2. Read: Risse, 2022. "Artificial Intelligence and the Past, Present, and Future of Democracy" *Cambridge Handbook of Responsible Artificial Intelligence*
   https://www.cambridge.org/core/books/cambridge-handbook-of-responsible-artificial-in artificial-intelligence-and-the-past-present-and-future-of-democracy/B6A19E65F15179E
3. Browse:
   - Martin, et al. "Political Machines: Understanding the role of Generative AI in the US 2024 elections and beyond" https://mediaengagement.org/research/generative-ai-electi
4. Background:
   - Jungherr, 2023. "Artificial Intelligence and Democracy: A Conceptual Framework" https://journals.sagepub.com/doi/10.1177/20563051231186353
   - Zhang, "Public Opinion toward Artificial Intelligence" *Oxford Handbook of AI Governance* https://academic.oup.com/edited-volume/41989/chapter-abstract/411053295?redirectedFrom=fulltext&login=false
   - Sifry, 2024. "How AI is Transforming the Way Political Campaigns Work" https://www.thenation.com/article/politics/how-ai-is-transforming-the-way-political-camp

SEMINAR 11.2 **Autocracy & AI**

1. Think: *How does the use of AI support or undermine authoritarian control?*
2. Read: Beraja et al, 2024. "AI-tocracy" *Quarterly Journal of Economics*
3. Browse: *Look at both of the following articles as in conversation with the primary reading*
   - Xu, 2020. "To Repress or to Co-opt? Authoritarian Control in the Age of Digital Surveillance" *American Journal of Political Science*
   - Yang, 2024. "The Limits of AI for Authoritarian Control" Working paper.

SEMINAR 12.1 **Debate VI: Energy**

SEMINAR 12.2 **Geopolitics & the AI Race**

1. Think: *How does the use of AI support or undermine relations between states?*
2. Read: Horowitz et al, "AI, the International Balance of Power, and National Security Strategy". *The Oxford Handbook of AI Governance*
3. Browse:
   - "What Google's Return to defense AI Means" https://www.defenseone.com/business/2025/02/what-googles-return-defense-ai-means/402816/

SEMINAR 13.1 **Governance, Policy, and AI**

1. Think: *Given all that we have read in the course - and particularly the topics of the last few weeks - what approaches to governance should nations take with AI? What are the critical areas of focus and what can wait? What is reality and what is hype? And how will AI change the nature of governance itself?*
2. Read: Zhang, et al. "Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers" *Journal of Artificial Intelligence Research*
3. Browse:
   - Toby Ord, Inference Scaling Reshapes AI Governance: https://www.tobyord.com/writing/inference-scaling-reshapes-ai-governance
   - "Against Democratizing AI" Himmelreich https://link.springer.com/article/10.1007/s00146-021-01357-z

SEMINAR 13.2 **The Future: Singularities, xRisk, AGI, & ASI**

1. Watch: Mitchell, "The Future of Artificial Intelligence" https://www.youtube.com/watch?v=GwHDAfAAKd4
2. Read: Blili-Hamelin et al, 2025. "Stop Treating AGI as the North Star of AI research" https://arxiv.org/pdf/2502.03689
3. Browse:
   - Sutton, 2019. "The Bitter Lesson" https://www.cs.utexas.edu/~eunsol/courses/data/bitter_lesson.pdf

- Wait But Why (Urban), 2015. "The AI Revolution: The Road to Superintelligence"
  - Part 1: https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html
  - Part 2: https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-2.html
- "AI 2027" https://ai-2027.com/

Reference:

- Kurzweil, *The Singularity is Near* and *The Singularity is Nearer*

SEMINAR 14.1 **Final Presentations** ***or workshop day***

SEMINAR 14.2 **Final Presentations**

## Campus Academic Resources & Policies

ACADEMIC INTEGRITY

As a pre-eminent and inclusive student-focused research institution, Syracuse University considers academic integrity at the forefront of learning, serving as a core value and guiding pillar of education. Syracuse University's Academic Integrity Policy provides students with the necessary guidelines to complete academic work with integrity throughout their studies. Students are required to uphold both course-specific and university-wide academic integrity expectations such as crediting your sources, doing your own work, communicating honestly, and supporting academic integrity. The full Syracuse University Academic Integrity Policy can be found by visiting class.syr/edu, selecting, Academic Integrity, and Expectations and Policy.

Upholding Academic Integrity includes the protection of faculty's intellectual property. Students should not upload, distribute, or share instructors' course materials, including presentations, assignments, exams, or other evaluative materials without permission. Using websites that charge fees or require uploading of course material (e.g., Chegg, Course Hero) to obtain exam solutions or assignments completed by others, which are then presented as your own violates academic integrity expectations in this course and may be classified as a Level 3 violation. All academic integrity expectations that apply to in-person assignments, quizzes, and exams also apply online.

Students found in violation of the policy are subject to grade sanctions determined by the course instructor and non-grade sanctions determined by the School or College where the course is offered. Students may not drop or withdraw from courses in which they face a suspected violation. Any established violation in this course may result in course failure regardless of violation level.

DIVERSITY

It is the intent of this course for students from all diverse backgrounds and perspectives to be well served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that students bring to this class be viewed as a resource, strength, and benefit. It is also critical to present materials and activities that are respectful of diversity: gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture. Your suggestions are encouraged and appreciated. Please let your instructor know ways to improve the effectiveness of the course for you personally or for other students or student groups.

**INCLUSION**

Syracuse University values diversity and inclusion; we are committed to a climate of mutual respect and full participation. There may be aspects of the instruction or design of this course that result in barriers to your inclusion and full participation in this course. I invite any student to contact me to discuss strategies and/or accommodations (academic adjustments) that may be essential to your success and to collaborate with the Center for Disability Resources (CDR) in this process.

If you would like to discuss disability-accommodations or register with CDR, please visit Center for Disability Resources. Please call (315) 443-4498 or email disabilityresources@syr.edu for more detailed information.

The CDR is responsible for coordinating disability-related academic accommodations and will work with the student to develop an access plan. Since academic accommodations may require early planning and generally are not provided retroactively, please contact CDR as soon as possible to begin this process.

**DISCRIMINATION AND HARASSMENT**

The University does not discriminate and prohibits harassment or discrimination related to any protected category including creed, ethnicity, citizenship, sexual orientation, national origin, sex, gender, pregnancy, disability, marital status, age, race, color, veteran status, military status, religion, sexual orientation, domestic violence status, genetic information, gender identity, gender expression or perceived gender.

Any complaint of discrimination or harassment related to any of these protected bases should be reported to Sheila Johnson-Willis, the University's Chief Equal Opportunity & Title IX Officer. She is responsible for coordinating compliance efforts under various laws including Titles VI, VII, IX and Section 504 of the Rehabilitation Act. She can be contacted at Equal Opportunity, Inclusion, and Resolution Services, 005 Steele Hall, Syracuse University, Syracuse, NY 13244-1120; by email: titleix@syr.edu; or by telephone: 315-443-0211.

**FAITH AND RELIGIOUS ACCOMMODATIONS**

Syracuse University's Religious Observances Policy recognizes the diversity of faiths represented in the campus community and protects the rights of students, faculty, and staff to observe religious holy days according to their traditions. Under the policy, students are given an opportunity to make up any examination, study, or work requirements that may be missed due to a religious observance, provided they notify their instructors no later than the academic drop deadline. For observances occurring before the drop deadline, notification is required at least two academic days in advance. Students may enter their observances in MySlice under Student Services/Enrollment/My Religious Observances/Add a Notification.

**HEALTH AND WELLNESS**

Mental health and overall well-being are significant predictors of academic success. As such it is essential that during your college experience you develop the skills and resources effectively to navigate stress, anxiety, depression, and other mental health concerns. Please familiarize yourself with the range of resources the Barnes Center provides (https://ese.syr.edu/bewell/) and seek out support for mental health concerns as needed. Counseling services are available 24/7, 365 days, at 315-443-8000.