

AI in Human Communication
Communication and Media 307-001
Professor: Josh Pasek, Ph.D.
GSI: Haoming Zhou
Fall, 2024

Course Meetings:
MW – 4:00-5:30 PM
G005 Dental Building
jpasek@umich.edu

Professor Drop-in Hours:
Weds – 10:15-11:15AM
5413 North Quad
Zoom (by appt. only)

Drop-in Hours

Professor: Wednesdays 10:15-11:15AM in 5413 North Quad and Zoom or by appointment (email: jpasek@umich.edu).

If you want to meet with Prof. Pasek via Zoom, please sign up **in advance** at:
<https://usemotion.com/meet/josh-pasek/officehours?d=10>

GSI drop-in hours: Tuesdays 1:00-2:00PM (email: hmzhou@umich.edu)

As we venture into the realm of *AI and Human Communication*, we find ourselves standing at the precipice of a new era. The ever-evolving landscape of artificial intelligence (AI) and particularly Generative Artificial Intelligence (also GenAI or GAI) offers unprecedented opportunities and challenges that shape our societies and the way we interact with the world. In an era where AI tools are becoming increasingly integrated into our daily lives, it is crucial to explore their impact on human communication and understand the intricate dynamics that unfold.

Throughout this course, we will embark on an intellectual journey that transcends boundaries and delves into the multifaceted world of AI. Together, we will unravel the enigma of generative AI tools, unraveling their inner workings and uncovering their potential to revolutionize industries, reshape education, and influence various aspects of our lives.

Delving into this realm, we grapple with the ethical considerations that arise when AI intersects with human communication. We will critically examine issues of bias, fairness, and transparency in AI systems, and explore possibilities for mitigating these challenges. Through explorations, discussions, and debates, we examine the social and political implications of AI, attempting to reach a nuanced understanding of its implications on society.

Drawing inspiration from the transformative nature of scientific discoveries, we will assess the current capabilities of AI, contemplate its future trajectory, and try to make sense of its potential impact. We will explore the possibilities of AI-driven advancements

in the workplace, education, and scientific research while considering the potential risks and societal disruptions that accompany these developments.

In this course, you will have the opportunity to actively engage with AI tools, exploring their capabilities and evaluating their potential impact. Through hands-on assignments and collaborative projects, you will develop critical thinking skills, enabling you to analyze and navigate the evolving landscape of AI with (some) confidence. And you will face the tricky social, ethical, and political questions posed by AI's growing suite of capabilities.

By the end of the course, you will emerge with as much of a comprehensive understanding of AI's impact on human communication and society as we can forecast at the moment. You will possess the skills to evaluate the ethical considerations and societal implications of AI applications, empowering you to contribute to the ongoing discourse surrounding AI in your future endeavors. And you will grapple with the ethical and social gravity of a future without a Turing test, where humans may not be the only or even the primary sources of intelligence.

Notably, given the rapidly changing nature of generative AI and the fact that work examining AI across the social sciences is still in its infancy, this course inevitably serves as an experiment. There is no broadly accepted reading list for a class of this sort; nor is there a strong sense, even among experts, about what domains will experience AI's greatest impacts. While I will do my best to bring together a suite of materials that are both timely and relevant, we may find that some seminal works in the field no longer hold water after the developments of the last few years or that arguments that seemed salient in early 2023, when I first assembled some of these materials, read as hopelessly dated by the time we encounter them. Indeed, while the class description you just read was itself aided through the use of Generative AI tools (as are the weekly descriptions), the readings and activities mostly couldn't be, because the training data at the time only went through September of 2021, an eon ago in this rapidly evolving field.

Content Objectives

-
- Develop an understanding of how generative AI tools function and their limitations
 - Explore the societal implications of AI technologies in various domains
 - Analyze the ethical, legal, and social challenges posed by AI in human communication
 - Examine the potential risks and benefits of AI and its impact on employment and education
 - Foster critical thinking skills to evaluate and navigate AI technologies
 - Engage in discussions on policy implications and future developments in AI
-

Course Policies:

Special Accommodations:

Any student who has a need for accommodation based on the impact of a disability, religious practice, physical requirement, or medical need should contact me privately to discuss the specific situation as soon as possible.

Absences:

Sickness. If you are sick, please send me an email as soon as you are aware that you may miss class. In general, if you are sick for an extended period, it will be useful to provide a doctor's note in addition to emailing me.

Religious holidays. Within the first two weeks of the semester, please notify me of any religious holidays for which you will be absent.

Athletic and other university-related absences. If you are travelling to represent the University of Michigan, someone on your team will provide you with the appropriate paperwork to distribute to your instructors.

Other excusable absences. For family emergencies, funerals, and other such absences, please send me an email as soon as you are aware of the potential for missing class.

A Note on Academic Freedom and Controversial Subjects:

Some of the subjects we will be discussing in this class may be highly controversial and sometimes touch on matters of strongly held beliefs. It is both my responsibility as an instructor and your responsibility as students to respect the range of opinions held in the classroom and to recognize that aspects of every topic we will be discussing are open to debate. In addition, some of the materials that we will encounter are coupled with strongly held viewpoints. The viewpoints in these readings are not necessarily correct – indeed, you may see that they sometimes contradict one-another. A reading's inclusion in the syllabus thus does not represent an endorsement of its content rather an assertion that a particular viewpoint is valuable to understand.

Academic Honesty:

A good student-teacher relationship operates on the basis of trust. From that basis, I trust that you will do your utmost to complete course assignments and to be honest with me if for any reason you are unable to fully meet a commitment to the class. I also trust your judgment that any collaboration with peers or use of online tools reflects an honest attempt to present your own work and contributions in the context of the class. That said,

if I encounter evidence that you are taking sole credit for work that is not principally your own, you will fail the course and I will immediately report the incident to the Dean of Student Affairs.

Uses of Generative AI:

Over the course of the term, we will be proposing standards for what constitutes appropriate uses of generative AI in this class and what does not. These may vary over time and across assignments. Appropriate uses of generative AI are NOT the same as academically honest ones. A use of generative AI is appropriate if it is in concordance with standards (we have set as a class) for how and when the tools can and should be used. A use of generative AI is academically honest if any output clearly denotes (either from the terms of an assignment or through the use of citations) what contributions stem from various individuals and from GenAI tools. To be clear, uses of generative AI in this class could simultaneously be academically honest but still inappropriate (e.g., if they are out of bounds for the assignment but well documented) or academically dishonest but appropriate (e.g., material that is taken from GenAI at a time when it would be acceptable to do so, but that is passed off as work that does not use GenAI).

Resources to Know:

Counseling and Psychological Services (CAPS)

Counseling and Psychological Services offers counseling services, educational and preventive initiatives, training programs, outreach and consultation activities, and provide guidance on how to “do something” to fully contribute to a caring healthy community. CAPS can be contacted on their 24-hour crisis line, 734-996-4747 and online at caps.umich.edu.

Sweetland Center for Writing

The Sweetland Center for Writing is a comprehensive center for improving student writing at all levels. They offer one-to-one tutoring for undergraduate students. Find them at www.lsa.umich.edu/sweetland/.

MiTALK

MiTALK offers mental health resources including online screenings for depression and anxiety, skill-building tools to help you manage stress and academic life, and digitally recorded workshops, lectures, and some relaxation exercises. The site is completely free of charge to U-M Students. Find them at mitalk.umich.edu.

Sexual Assault Prevention and Awareness Center (SAPAC)

If you or someone you know has been harassed, assaulted, or stalked, you can receive confidential support and academic advocacy at the Sexual Assault Prevention and Awareness Center (SAPAC). SAPAC can be contacted on their 24-hour crisis line, 734-936-3333 and online at sapac.umich.edu. Alleged violations can be non-confidentially reported to the Office for Institutional Equity (OIE) at institutional.equity@umich.edu. Reports to law enforcement can be made to University of Michigan Police Department at 734-763-3434.

Grading

- Attendance and Participation – 10%
- Weekly reflections – 20%
- Policymaking and Debate 45% (10% written, 5% presentation each)
 - o AI in Class Policies – 15%
 - o AI in University/Workplace Policies – 15%
 - o AI in Society Policies – 15%
- Free-range final assignment – 25%

Assignments

- **Weekly Reflections (20%):** Each week (from weeks 2 through 13), write a 1-page (~250 word) reflection on the readings -OR- have a GenAI tool write a 1-page (~250 word) reflection and write a one paragraph (~150 word) evaluation of what the tool produced. Post this on the discussion on the class Canvas page and write at least 2 sentences of commentary on something someone else posted. These will be graded on a check/check-minus system, with 10 checks needed for full credit (i.e., you can miss 2 weeks). The reflections will be worth 15% of your grade, and the commentary/responses to other students' reflections will be worth 5% of your grade.
- **Policymaking and Debate (45%):** At three times during the semester, you will work with groups in section to develop and refine proposed AI policies at the classroom, university/workplace, and societal levels. You and your group will craft policy language describing the guidelines you think should be in place and a separate document justifying those proposals. In the subsequent week, you will present your proposed rules to the other students in the class, who will question and challenge those rules, and you will be asked to provide support for your proposals. For each of these sessions, your group will receive grades for the written proposal and your presentation/defense. Each written proposal will constitute 5% of your grade, each justification will constitute 5% of your grade, and each presentation/defense will constitute 5% of your grade. Part of this grade will depend on how well you demonstrate that you have thought through the potential issues with your proposal, meaning that a consideration of alternate strategies and asking each other good questions will improve everyone's grades. For each round, everyone in your group will get an identical grade.
- **Free range assignment (25%):** At the end of the class, take one of the topics that we have explored, conduct some additional research on the area, and produce some kind of creative content that will help show (1) why it is an important issue and (2) what members of the public should know about it. You can complete this assignment in any way that you

would like. You are welcome to work in groups on this assignment as long as you clearly denote the contributions of the various group members. You are also welcome to make prolific use of AI tools, again as long as you hew to clear standards for how to use AI sources appropriately and honestly. Some (non-exhaustive) possibilities:

- A traditional essay (boring, but acceptable)
- A formal proposal to the university
- An art piece or exhibit
- A show or movie
- An interactive website
- A board game
- A computer program
- A graphic novel
- Etc.

This should take approximately the same amount of effort as a 10-page paper, so plan accordingly and make sure to get started early! The assignment is due on December 9, our last class day.

Course Schedule

Part 1 – Making Sense of Generative AI: How AI Works

Week 1 (Aug 26, 28): Introduction to AI and Human Communication

As we embark on our journey through the realm of artificial intelligence (AI), we begin by questioning the very nature of intelligence itself. Can machines genuinely think or feel? Or are they merely mimicking human behaviors without understanding? This week, we set the stage by diving deep into foundational readings and debates, probing the very essence of AI and its impact on human communication.

Readings (by Aug 28):

- Turing, A. (1950). Computing Machinery and Intelligence.
- Sejnowski, T.J. (2023). Large Language Models and the Reverse Turing Test.

Week 2 (Sept 4): Understanding Generative AI Tools

[No class Sept 2 – Labor Day]

Generative AI is a formidable force in the AI world, but what exactly is it, and how does it work? This week, we unravel the intricacies of these tools, shedding light on their underlying mechanics and vast capabilities. With insights from industry experts and groundbreaking research, we'll demystify the magic behind AI's ability to generate content.

Readings:

- Wolfram, S. (2023). What is ChatGPT doing and why does it work?

Week 3 (Sept 11): Broad vs. Narrow AI

[No class on Sept 9 due to a family obligation]

AI has evolved from specialized systems performing singular tasks to models capable of tackling multiple, diverse challenges. This week, we'll journey through the history of AI, beginning with early, narrow applications like expert systems and recommendation engines, and moving towards the broader capabilities of modern foundation models. We'll examine how these models are reshaping the landscape of AI, with the potential to operate across various domains. Through engaging discussions and hands-on activities, we'll uncover the implications of this evolution and explore the ethical considerations that come with increasingly versatile AI systems.

- Bommasani et al. (2021). On the Opportunities and Risks of Foundation Models [pages 1-33, 73-92]
- Urban (2022). The AI Revolution – The Road to Superintelligence (parts 1 and 2)

Week 4 (Sept 16, 18): Advancements in Generative AI and Emergent Abilities

This week, we will explore the cutting-edge developments in Generative AI, focusing on both the technological advancements and the surprising emergent abilities of large language models. Through our readings, we will dive into the latest generative models like GANs, GPT, and transformers, understanding their foundational mechanisms and the wide array of tasks they can perform. We'll also examine how large language models have begun to exhibit unexpected capabilities that were not explicitly programmed, challenging our understanding of AI's potential. This discussion will provide a comprehensive view of the current state and future directions of generative AI, highlighting both the opportunities and challenges posed by these powerful tools.

Readings:

- Wei et al. (2022) Emergent Abilities of Large Language Models
- Bengesiet et al. (2023) Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers

Week 5 (Sept 23, 25): Mastering Prompt Engineering: Techniques and Applications

This week, we focus on the art of prompt engineering, a critical skill for enhancing the performance of AI models. Through hands-on activities and case studies, students will explore various strategies for crafting effective prompts that yield high-quality outputs in both text-based and text-to-image generative models. We will delve into the CLEAR framework and other design guidelines to understand how precise and adaptive prompts can significantly improve AI responses. By the end of the week, students will have developed a toolkit of techniques to refine their interactions with AI, making them more effective and creative in their applications

Readings:

- Lo, L.S. (2023) The CLEAR path: A framework for enhancing information literacy through prompt engineering
- Knoth et al. (2024) AI literacy and its implications for prompt engineering strategies
- Liu & Chilton (2022) Design Guidelines for Prompt Engineering Text-to-Image Generative Models

Week 6 (Sept 30, Oct 2): Navigating the Complexities of AI Control and Deception

This week, we tackle the intricate challenges surrounding AI control, security, and ethical use. Our focus will be on understanding how AI models, particularly large language models (LLMs), can be manipulated through prompt engineering to bypass safeguards—an approach known as jailbreaking. We will explore the technical underpinnings of such vulnerabilities, as well as the broader implications for AI governance and security. Through case studies and technical reports, we will discuss how these models can be both a tool for innovation and a potential risk for misuse, emphasizing the balance between AI

capabilities and ethical responsibility. Students will engage in critical analysis and discussions about the dual-use nature of AI technologies, considering both the technical mechanisms and societal impacts of AI deception and control, and will see if they can push the boundaries of what AI can and should be generating.

Readings:

- GPT-4 Technical Report & System Card (2023) [READ pp 1-14 & 41-70]
- Sisonal et al. (2023) ChatGPT: More than a “Weapon of Mass Deception.”
- Liu et al. (2023) Jailbreaking ChatGPT via Prompt Engineering: An Empirical Study

Part 2 – Brace for Impact: AI Across Domains

Week 7 (Oct 7, 9): The Breadth of Applications

AI's influence extends far beyond a few niches—it has permeated or appears poised to permeate almost every sector. This week offers a panoramic view of the applications of AI, from healthcare to entertainment, logistics to finance. Grasping the breadth of AI's reach is essential to appreciate its transformative power.

Readings:

- West & Allen (2018). How artificial intelligence is transforming the world.
- HAI (2023). Generative AI: Perspectives from Stanford HAI.

NOTE: NO CLASS ON OCT 14 OR LAB ON OCT 15; FALL BREAK]

Weeks 8 (Oct 16): AI in the Education

This week, we delve into the transformative potential of artificial intelligence in education. AI is rapidly reshaping how we teach, learn, and manage educational institutions. From personalized learning experiences to AI-driven tools that enhance student engagement and streamline administrative tasks, the possibilities are vast. However, these advancements also bring significant challenges, including concerns about academic integrity, data privacy, and the digital divide. By examining AI's impact on both the classroom and the broader educational landscape, we'll explore how institutions like the University of Michigan are navigating these opportunities and challenges, ensuring that AI serves to enhance education while upholding core academic values.

Readings:

- Bogost, I. (2023). The First Year of AI College Ends in Ruin
- Cotton et al. (2023). Chatting and Cheating: Ensuring academic integrity in the era of ChatGPT
- Chronicle of Higher Education. (2023). How Will AI Change Higher Ed?

- Duraisamy et al. (2023). Generative Artificial Intelligence Advisory Committee Report [skim]

Week 9 (Oct 21, 23): AI in the Workplace

The workplace is evolving, and AI plays a pivotal role in this transformation. This week, we'll delve into AI's impact on the future of work. How does AI reshape industries? What new opportunities arise, and what challenges do we need to confront? Through engaging readings and discussions, we'll forecast the future of the job market in the AI era.

Readings:

- Eloundou et al. (2023). GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models
- Warzel, C. (2023). Here's How AI Will Come for Your Job
- Lowery, A. (2023). How ChatGPT Will Destabilize White Collar Work
- HAI (2023) Artificial Intelligence Index Report 2023. [Skim Chapter 4 The Economy pp. 168-233].

Week 10 (Oct 28, 30): AI in Creative Industries. Art and technology have always intersected, but the advent of AI brings novel dimensions to this relationship. As AI takes on roles in artistic creation and interpretation, what does it mean for human artists, the definition of creativity, and the value of originality? This week, we'll navigate the intersections of AI and art, discussing the controversies and the transformative potential of this union.

Readings:

- Roose, K. (2022). An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy.
- Science Friday (2022). How Will AI Image Generators Affect Artists? <https://www.sciencefriday.com/segments/ai-art/> [listen to the audio]
- Chayka, K. (2023). Is AI Art Stealing from Artists?
- Hutdon, M. (2023). AI learns to write computer code in 'stunning' advance.
- Lee, H-K. (2022). Rethinking Creativity: Creative Industries, AI and Everyday Creativity.

Part 3 – The Pitfalls and the Future

Week 11 (Nov 4, 6): AI Hallucinations and Bias

This week, we confront the darker aspects of AI technology, focusing on two significant challenges: hallucinations and bias. AI hallucinations occur when systems generate plausible but incorrect information, raising concerns about the reliability and safety of AI outputs in critical applications. Simultaneously, we explore the pervasive issue of bias in AI, where data-driven systems inadvertently perpetuate existing societal inequalities or introduce new forms of discrimination. Through a combination of technical insights and

real-world examples, we will analyze the mechanisms behind these problematic outcomes, discuss their ethical implications, and consider strategies for mitigating their impact. This week's discussions aim to equip you with a deeper understanding of the complexities and responsibilities involved in developing and deploying AI technologies.

Week 12 (Nov 11, 13): AI Ethics and the Law.

In Week 12, we will explore the intersection of artificial intelligence, ethics, and the law. As AI technologies rapidly evolve, they bring with them complex legal challenges, particularly around data security, intellectual property, and the transparency of AI systems. We'll delve into the issues surrounding copyright laws and data privacy, and examine the growing concern over AI's "black box" nature, which complicates accountability and raises the need for explainable AI. Through our readings and discussions, we'll unpack the legal frameworks currently in place and the ethical imperatives that drive the call for more transparent, accountable AI systems.

Readings:

- Eitel-Porter, R. (2021). Beyond the Premise: Implementing Ethical AI
- Liao and Vaughan. (2023). AI Transparency in the Age of LLMs
- Appel et al. (2023). Generative AI Has an Intellectual Property Problem
- Alexandrei, D. (2023). Explainable AI, the key to open "black boxes"
- Congressional Research Service (2023). Generative Artificial Intelligence and Copyright Law

Week 13 (Nov 18, 20): AI in public policy and international competition. AI doesn't exist in a vacuum; it plays a significant role in shaping global politics and public policies. This week, we'll explore how nations respond to the rise of AI, the geopolitical implications of AI supremacy, and the strategies governments might adopt in the AI era. Through diverse readings, we'll dissect the intricate dance between AI and global governance.

Readings:

- HAI (2023) Artificial Intelligence Index Report 2023. [Read Chapter 6 Policy and Governance, pp. 263-296].
- Hacker et al. (2023). Regulating ChatGPT and Other Large Generative AI Models.
- Lowrey, A. (2023). Before AI Takes Over, Make Plans to Give Everyone Money
- Anderson, R. (2023). Never Give Artificial Intelligence the Nuclear Codes
- de Oliveira, B.B. (2024). The EU's ambition to influence global standards for Artificial Intelligence amongst regulatory competition with China and the USA. [read section 3: the Regulatory Approaches to AI, pp. 67-91]

Week 14 (Nov 25): The Near Future of AI. In Week 14, we will delve into the immediate future of artificial intelligence, focusing on the next 5 to 10 years. We'll explore the expected advancements in AI technologies, their potential societal impacts, and the urgent questions surrounding AI governance and ethics. Through a combination of expert insights and critical readings, we'll discuss both the opportunities and risks that AI presents as it becomes increasingly integrated into our daily lives.
[NOTE: NO CLASS ON NOV 27; THANKSGIVING BREAK]

Readings:

- Anderson et al. (2018). Artificial Intelligence and the Future of Humans [skim]
- Marcus, G. (2023). Why Are We Letting the AI Crisis Just Happen?
- ARC Evals (2023). Update on ARC's Recent Evaluation Efforts.
- Wong, M. (2023). AI Doomerism is a Decoy.

Week 15 (Dec 2, 4): The Far Future of AI. Week 15 takes us into the far future of AI, where we'll explore speculative but significant scenarios, including the potential for a singularity—where AI surpasses human intelligence—and the risks of AI advancements stalling or leading to self-sabotage. We will consider both utopian and dystopian possibilities, questioning whether AI will help us build a better future or whether it might lead to catastrophic outcomes. The readings this week will challenge us to think deeply about the long-term implications of AI on society and the planet.

- Bubeck et al. (2023). Sparks of Artificial General Intelligence
- Future of Life Institute (2023). Pause Giant AI Experiments: An Open Letter
- Warzel, C. (2023). What Have Humans Just Unleashed?
- Shumailov et al. (2023). AI models collapse when trained on recursively generated data.
- Wong, M. (2023). The Chatbots May Poison Themselves.

Week 15 (Dec 9): Concluding Thoughts. As we culminate our exploration into the vast universe of AI, this week offers an opportunity to reflect, integrate, and showcase. Students will present their final projects, encapsulating their learnings and insights from the course. This collaborative event not only serves as a celebration of knowledge gained but also as a launchpad for future inquiries into the ever-evolving world of AI.