Ethics, Equity, & Justice x AI Discussion Topics, Resources, and Queries

1. Copyright & Consent

Resources
- **How AI steals artists voices and style:**
  - Viral AI-Generated Drake Song 'Heart on My Sleeve' Removed from Spotify, YouTube ([Vice](#)) or "Invasive Diffusion: How one unwilling illustrator found herself turned into an AI model" ([Waxy](#))
  - The AI artist who used Bad Bunny's voice — and shot to fame ([RestofWorld](#))
  - The Specter of AI-Generated 'Leaked Songs' Is Tearing the Harry Styles Fandom Apart ([404](#))
- "Inside the secret list of websites that make AI like ChatGPT sound smart" [WaPo](#)
- The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work ([NYT](#))
  - "Getty Images suing the makers of popular AI art tool for allegedly stealing photos" ([CNN](#)) or "Authors file a lawsuit against OpenAI for unlawfully 'ingesting' their books" ([Guardian](#))
- How Tech Giants Cut Corners to Harvest Data for A.I. ([New York Times](#))
- The tricky truth about how generative AI uses your data ([Vox](#))

Queries
- Is it ethical to train an AI on copyrighted creative work, without permission or attribution?
- Is it ethical to allow people to generate new work in the styles of artists and writers without compensating them?
  - Is it ethical to charge money for that service, built on the work of others?
- How will society cope with the fact that posting a single photo/clip of yourself online allows your entire likeness to be mimicked by AI?

2. Bias & Bigotry

Resources
- **"This is how AI sees the World" [Washington Post](#)**
  - "Humans are Biased Generative AI is Worse" [Bloomberg](#)
  - These robots were trained on AI. They became racist and sexist. [WaPo](#)
  - OpenAI's DALL-E 2 is incredible at turning text into images. It also highlights the problem of AI bias — and the need to change incentives in the industry. ([Vox](#))
    - AI art looks way too European ([Vox](#))
  - AI language models are rife with different political biases ([MIT Tech](#))
- **Rooting Out Anti-Muslim Bias in Popular Language Model GPT-3 ([Stanford](#))**
  - Large language models associate Muslims with violence ([Nature](#))
  - Asking ChatGPT about Israel/Palestine ([Instagram](#))
- ChatGPT Replicates Gender Bias in Recommendation Letters ([SciAm](#))

- Activity: Stable Diffusion Bias [Hugging Face](#) highlights stereotype amplification problems
- Programs to detect AI discriminate against non-native English speakers, shows study ([Guardian](#))
- Google apologizes for 'missing the mark' after Gemini generated racially diverse Nazis ([Verge](#))
- These Women Tried to Warn Us About AI ([Rolling Stone](#))
- Previous:
  - Amazon scraps secret AI recruiting tool that showed bias against women ([Reuters](#))
  - "MIT apologizes, permanently pulls offline huge dataset that taught AI systems to use racist, misogynistic slurs" [The Register](#)

Queries
- Generative AI systems are western- and English-centric. They regularly reproduce racism, sexism, and bigotry. Since LLMs exacerbate inequality, do we have a moral duty to reject this new tech?
- Much like with hallucination issues, issues of racism, misogyny, bigotry can't be solved. Nevertheless, organizations, businesses, and governments are using AI. How should we regulate these systems so that we mitigate the harms?

3. Ethics

Resources
- "Inside the secret list of websites that make AI like ChatGPT sound smart" [WaPo](#)
  - "Revealed: The Authors Whose Pirated Books Are Powering Generative AI ([Atlantic](#)) and ""OpenAI now tries to hide that ChatGPT was trained on copyrighted books, including J.K. Rowling's Harry Potter series" ([BI](#))
  - The tricky truth about how generative AI uses your data ([Vox](#))
- Exclusive: OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic ([Time](#))
  - "The Hidden Workforce That Helped Filter Violence and Abuse Out of ChatGPT" ([The Journal](#) podcast)
  - "The Exploited Labor Behind Artificial Intelligence" [Noema](#)
  - The global labor force of outsourced and contract workers are early adopters of generative AI — and the most at risk. ([RestofWorld](#))
  - Behind the AI boom, an army of overseas workers in 'digital sweatshops' ([WaPo](#))

Queries
- How do we acknowledge the work of these "clickworkers" (as they are often called), when we use generative AI applications?
- Generative AI apps use our inputs to help train their outputs. Is this ethical?

4. Misinformation

Resources
- Generative AI applications often output misformation.

- - - What Can You Do When A.I. Lies About You? ([NYT](#))
    - ChatGPT is making up fake Guardian articles. Here's how we're responding ([Guardian](#))
  - Disinformation Researchers Raise Alarms About A.I. Chatbots ([NYT](#))
    - AI-Generated Propaganda Is Just as Persuasive as the Real Thing, Worrying Study Finds ([Vice](#))
  - Google's and Microsoft's AI Chatbots Refuse to Say Who Won the 2020 US Election ([Wired](#))

Queries
- What do we do when AI "hallucinations" have real world consequences?
  - In the article above, an AI application labeled a Dutch politician a "terrorist." Presumably this could happen to people without them even knowing that the AI did this. What recourse should this politician have?
- What happens when the stream of misinformation and disinformation is nearly constant thanks to the ease of creation afforded by generative AI?
  - To what extent will this change how we think about media we encounter online? Do we need more media literacy instruction in schools to help? What about citizens who are not in school?

5. Surveillance, Privacy, & Deepfakes

Resources
- OpenAI Worries About What Its Chatbot Will Say About People's Faces ([NYT](#))
- "Wrongfully Accused by an Algorithm" [NYT](#) or "Eight Months Pregnant and Arrested After False Facial Recognition Match" [NYT](#)
  - Here Are the Stadiums That Are Keeping Track of Your Face ([Slate](#))
  - GPT-powered deepfakes are a 'powder keg' [FastCo](#)
- Digital Culture Is Literally Reshaping Women's Faces" ([Wired](#))
- Fake Joe Biden robocall tells New Hampshire Democrats not to vote Tuesday ([NBC](#))
- Celebrities Already Live in AI Hell ([Intelligencer](#))
- AI fake nudes are booming. It's ruining real teens' lives. ([WaPo](#))

Queries
-

6. Environment

Resources
- The Secret Water Footprint of AI Technology [The Markup](#)
  - [Paper](#): "Making AI Less 'Thirsty': Uncovering and Addressing the Secret Water Footprint of AI Models"
  - "Artificial intelligence technology behind ChatGPT was built in Iowa — with a lot of water" [AP](#) or "A.I. tools fueled a 34% spike in Microsoft's water consumption, and one city with its data centers is concerned about the effect on residential supply" ([Fortune](#))
  - AI Is Taking Water From the Desert ([Atlantic](#))

- "How much electricity does AI consume?" ([Verge](#))
- "The Generative AI Race Has a Dirty Secret" ([Wired](#))
- "AI Is Hurting the Climate in a Number of Non-Obvious Ways" ([The Markup](#))
- "ChatGPT: tackle the growing carbon footprint of generative AI" [Nature](#)
  - The mounting human and environmental costs of generative AI ([ArsTech](#))
- A.I. Frenzy Complicates Efforts to Keep Power-Hungry Data Sites Green ([NYT](#))
  - Thirsty Data Centers Are Making Hot Summers Even Scarier ([Bloomberg](#))
- "The ugly truth behind ChatGPT: AI is guzzling resources at planet-eating rates" Good summary from the [Guardian](#)

Queries
- How do we balance our need for developing AI with our need for a healthy planet?
- Which communities will be most impacted by the energy consumption of the AI boom?

7. Jobs/Work

Resources
- AI's "tech tsunami" is coming for call center workers ([RestofWorld](#))
- ChatGPT took their jobs. Now they walk dogs and fix air conditioners. ([WaPo](#))
- "The Impact of AI on Developer Productivity: Evidence from GitHub Copilot" [Study](#) and other professionals too ([study](#))
- "Amazon's AI Cameras Are Punishing Drivers for Mistakes They Didn't Make" [Vice](#)
- Co-Writing with Opinionated Language Models Affects Users' Views ([arxiv](#))
- "Hollywood is replacing artists with AI. Its future is bleak." ([Vox](#))
- Instead of being replaced by robots, office workers will soon be pressured to act more like robots themselves. ([Atlantic](#))

Queries
- How will AI change how we work? Will our work output become more like generative AI outputs?
- How will work, jobs and productivity change in an age of high AI use?

—--

Extra Queries

Equity
Will all students having access to Generative AI level the playing field (e.g tutors and parents make less of an impact when everyone has an AI)? Or will privileged students have access to human tutors while less privileged students have to use AI models?
Given some teachers are using and teaching students about AI and some aren't, how will this change our students' AI literacy and AI-enhanced skill development depending on which teachers they get in the registration process?
Citizenship

By uncritically adopting a new technology, are we graduating a generation of tech-solutionists or tech-utopians that believe technology is neutral, inevitable and good?

Mental Health

The last round of disruptive technology (social media) has (net) been bad for the health and humanity of our students. Well-founded research has shown an increase in loneliness in our students. It seems likely Generative AI will amplify these issues. How do we ensure this time it's different?

Students will have an endless stream of personalized content created by humans & AI and delivered by AI algorithms. What media literacy instruction do we need to add to our curriculum? Presumably, students will rely on and trust AI for schoolwork, what happens when they start using it for their social lives as well (creating deepfakes of their friends–or teachers!) but also discussing with AI sensitive issues usually reserved for counselors, doctors, and parents.