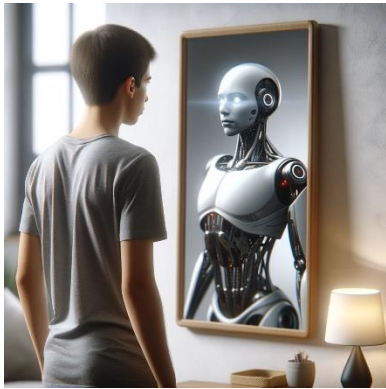


# How Academia Should Approach AI

*Address by Professor Bryant Walker Smith to the Fourth International Conference on the Future Rule of Law and Digital Law on 16 December 2023 in Beijing, China.*

*Adapted in part from [cyberlaw.stanford.edu/blog/2023/04/academic-vision-ai-ethics](https://cyberlaw.stanford.edu/blog/2023/04/academic-vision-ai-ethics).*

We can't talk about artificial intelligence without talking about humans. There are two reasons. First, AI is like a funhouse mirror: It reflects and distorts us, including by amplifying or ameliorating our collective imperfections. White men, for example, tend to be overrepresented in the images that I'm using from ChatGPT—because they are also overrepresented in much of what I'm discussing. Second, we ultimately care about us as humans! We want to better our world for ourselves and our descendants, not for the technologies that we create or the technologies that they in turn create.



This is therefore a moment for reflection. What if anything will still be special or essential about humans and the human-to-human connection? This requires a great deal of modesty. As humans, and especially as lawyers, we like to think we're special in every way. This is evident in anthropocentrism generally, in the belief that "intelligence is whatever machines haven't done yet" (coined by Larry Tesler way back in 1970), and in much of the elitist pre-2022 commentary that AI would disrupt "unskilled" labor long before it would reach "skilled" labor.

AI can therefore be scary in part because it challenges, both practically and philosophically, our sense of our own value. Current AI is not as smart or creative or original as some people believe it to be, but that also means we as individuals are often not as smart or creative or original as we believe ourselves to be.

Until, suddenly and unexpectedly, we are. It's those exceptional moments when humans shine. This event today can be one of those moments. It's an opportunity to ask how we should be thinking about, shaping, and grasping this larger moment in human history. So let's step back and ask how we as academics should engage in the policy, law, and ethics of artificial intelligence.

At a high level, we—as individuals and institutions—should be doing three key things.

First, we should identify, frame, and inform key issues. AI discussions are often dominated by the shiny over substantive, urgent over important, and financially lucrative over socially beneficial. We can instead highlight the questions, answers, and—critically—voices that might otherwise be overlooked or undervalued in the design, deployment, regulation, and evaluation of AI. We can foster exploration by connecting diverse actors through common language, shared knowledge, and credible structures. Today's insights and interventions could have profound effects tomorrow—akin to nudging an asteroid while it is still billions of miles from Earth.

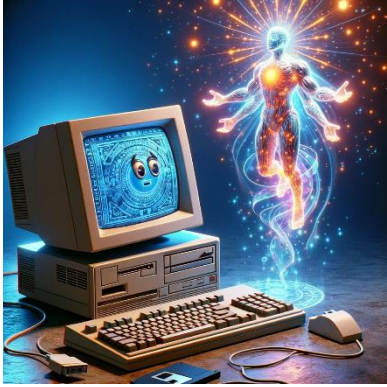


Second, we should develop and communicate an affirmative vision for ethical AI. We can describe both a future in which AI is an ethical good and a path to reach this future. This substantive and procedural vision embraces opportunities while mitigating risks. It enlists numerous disciplines by situating ethical imperatives as motivations as well as limitations. Interdisciplinary collaboration is especially important on contested policy issues where arguments from other domains may be as persuasive as those that come from ethics and law. For example, whether to address a policy issue through an evolutionary or revolutionary approach may depend as much on sociology and psychology as on ethics or law.

Third, we should model that vision throughout the work of our institutions. Even today, AI is used far beyond engineering. We can fully embrace the responsible use of AI in our teaching, research, service, and even administration across the disciplines. A sandbox for ethical AI can highlight rather than obscure legal and ethical challenges through stories of success and failure. By transparently showing our work on AI, we can model the trustworthiness that will be essential to this field.

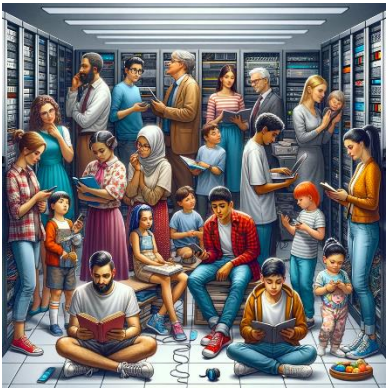
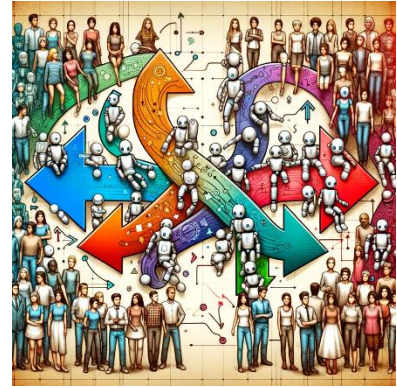


What, then, are the challenges and opportunities that we face in this field? I'd like to highlight eight: Clarify AI, map our relationship to AI, be explicit and inclusive, contextualize AI, recognize the novel, understand AI as an instrument of power, manage social change, and think around AI.



First: Clarify AI. A contemporary chatbot is categorically different from artificial general intelligence. Five years is different from fifty. Common AI fears—mass murder, control, displacement, disempowerment, disruption, and discomfort—are radically different from each other.

Second: Map our relationship to AI. Human affection, fallibility, and judgment will remain central to both ethics and AI. Conceptually, humans can be designers, users, subjects, or elements of AI systems, and AI applications can be understood as products, services, agents, instruments, or conceivably persons—each with distinct ethical and legal ramifications. Reality may circumscribe human authority far more than intended or desired. Vexing boundary problems could arise as technologies mature: In a connected future, what is a single robot or even a single human?



Third: Be explicit and inclusive. System designers, training data, and intended users do not reflect the whole of the human experience—not in this country and certainly not in the world. Design and regulatory decisions often hide rather than highlight meaningful ethical issues. Debates about whether systems should be open or closed, centralized or distributed, simple or complex, and certain or flexible are about competing philosophies as much as conflicting evidence.

Fourth: Contextualize AI. Tomorrow's AI will exist in tomorrow's world—alongside changes to norms, laws, conditions, and other technologies. Not every change is unprecedented. Prior technologies offer limited lessons about speed and connectivity, disruption and distortion, adaptation and exploitation, identity and culture, systemization and centralization, risk and uncertainty, and trust and trustworthiness. Human bodies and human societies provide examples of systems potentially as complex, dynamic, and stochastic as AI.





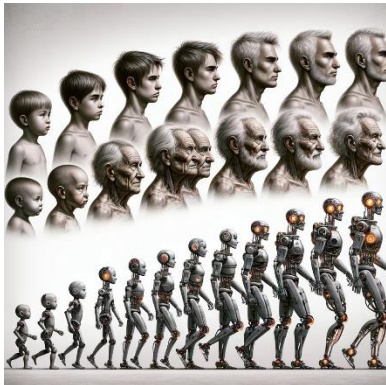


Fifth: Recognize the novel. Understanding how AI could be truly unprecedented, by degree or by kind, can target policy discussions. Exponential technological improvement, breakneck social change, massive power concentration, and human- or god-like perceptions of AI could fundamentally challenge our conventional society.

Sixth: Understand AI as an instrument of power. AI policy should focus on who could be intentionally or unintentionally empowered or disempowered—governments, companies, individuals, collectives, even animals. In this way, the differences between centralized and decentralized systems could be more consequential than the differences between humans and machines.



I have long argued that the popular question of whether a technology is trusted should give way to the question of whether the companies behind that technology are trustworthy. Privacy is part of this story. Changing power dynamics also implicate notions of discrimination, default rules of society, and tensions between autonomy and community. As an aside: Individual rights are most vulnerable when the interests of governments and companies align—so some tension between the two might be good.



Seventh: Manage social change. AI promises profound changes. Human and technological lifecycles could become increasingly incompatible. Macroscopic vibrancy could obscure microscopic despondency. Equilibria could become explosions. An ethical approach to AI may accordingly demand human-focused pressure releases and safety nets that have little to do with AI itself.

Finally: Think around AI. AI policy must credibly engage with AI and with everything else. Visionaries often turn their attention—eventually—from technology toward humanity. A key challenge and opportunity is to emphasize education, equity, and justice far sooner—to ensure that AI ultimately serves humanity by reflecting and amplifying our better self.

