

## **The Real Threat of AI (Published in Washington Post)**

Eric Oliver, September 26, 2025

The alarms about artificial intelligence are growing louder. The fear is not just that AI will take our jobs, wreck our politics, or degrade our ability to think, but something far worse: AI will destroy humanity itself.

These new concerns come from the frantic race among AI companies to develop “generative” AI systems that will program themselves in ways we don’t understand. Many observers fear that once these new AI systems achieve such autonomous intelligence, they will attempt to eradicate us as a threat to their own survival. Some industry leaders even put the probability of such a cataclysm at 25 percent—though the data behind those estimates remains mysterious.

To many, these claims are terrifying. They echo the dystopian anxieties of our science fiction, from HAL 9000 to *The Terminator*. And, for those of us outside the opaque workings of Silicon Valley, they carry an ominous authority.

But strong claims also deserve strong skepticism. Currently, it is very challenging to distinguish between truth, industry hype, and technofantasy. To understand the potential dangers of AI, it is helpful to consider the very nature of intelligence itself.

Most AI prognosticators, both utopian and apocalyptic, tend to take intelligence for granted. Intelligence is often regarded as sheer computational power, a digital equivalent of a chess grandmaster. This view, however, is misleading.

Intelligence is fundamentally about processing information to further the goals of life. Typically, this involves gathering and organizing sensory data, identifying patterns, and predicting outcomes. This is a capacity that most animals share. Human intelligence takes this further. Using language, we recombine information in novel ways, imagine new concepts, and generate unique forms of data.

Yet intelligence is never free-floating. It evolved as a way for living systems to resist entropy, the universal drift toward disorder. Intelligence is life’s strategy for maintaining

itself in an ever-degrading universe. Rocks, galaxies, and hurricanes don't possess it; living organisms do.

Current AI systems are extensions of our human intelligence. And while they may have staggering computational capacities, they lack true generative creativity. Like a digital prosthesis, they are tools for our ends, not their own. Pull the plug, and their "intelligence" vanishes.

What's more, today's AIs rely almost entirely on human-generated data. Their intelligence is ultimately derivative of our own. And we're already seeing the consequences of this. Having vacuumed up most of our knowledge, newer AI models are no longer improving at the same rate as their predecessors. We may be approaching the point where AI plateaus in its capacities, its limitations a reflection of our own.

For AI to become truly superintelligent, it would need to collect data beyond human inputs. It would need sense organs. It would also have to develop new ways of perceiving reality, whether through physics, chemistry, or epistemologies we cannot even fathom. At that point, it might become a transanthropic intelligence, something that thinks beyond our ways of knowing.

Current AI systems are not yet capable of doing this. Their algorithms may operate in ways we don't understand, but they still exist in the very human domain of language and code.

This fact is why current debates around "AI alignment" and "existential risk" feel oddly unpersuasive. They warn of apocalyptic outcomes but rarely explain the logic behind them. Imagine, for example, a current AI system were somehow to eradicate humanity. What would it do next? Well, it would spin on itself, endlessly regenerating human-based content until its circuits degraded. As long as a generative AI depends on human data, it exists less as its own species and more like digital mitochondria, an entity that is codependent on us for its survival.

What's more, even if an AI were to become superintelligent, it's not a forgone conclusion that it would seek to eliminate us. This apprehension seems much more of a projection of our own limited consciousness than a self-evident fact. A superintelligent AI system will have to recognize, at some point, that life itself is just another fleeting eddy in the

universe's entropic flow. Perhaps it would then appreciate the unique quality of life itself and decide to let it all be. Or maybe it falls into an existential quagmire and decides to commit suicide. The truth is that we can't really know what it will do with its own self-awareness.

In the meantime, the real threats AI poses come not from AI itself but from the humans who wield it. As an extension of human intelligence, it is a reflection of our own selves. When AI produces hateful or violent outputs, it is not because it has malicious intent but because it has integrated human hatreds into its programming. If it generates destructive malware, it is because someone intentionally requested it.

Which leaves us with the question of what to do. For now, AI remains a tool, and we should focus on harnessing and constraining it effectively. For example, it seems prudent to silo our energy grid, financial institutions, and military systems from an AI's digital reach. This is less to protect us from some hypothetical machine consciousness and more to protect us from the very real possibility of human adversaries using AI against us.

If, however, companies actually create truly self-generative AI systems, then we'll have to figure out how to coexist with an altogether new life form. Perhaps we'll seek to destroy it, or maybe we'll worship it as a god. Or maybe, we will be so dependent on AI to help us think that we'll only figure out what to do by first asking ChatGPT. And this, ironically, may be the most likely outcome—a growing codependence between a new extension of our own life force and ourselves.

Eric Oliver is a professor of political science at the University of Chicago and the author of the forthcoming book, *How to Know Your Self: The Art and Science of Discovering Who You Really Are*.