



[richardheinberg.com](http://richardheinberg.com)

*MuseLetter #387 / June 2025 by Richard Heinberg*

## AI Utopia, AI Apocalypse, and AI Reality

Recent articles and books about artificial intelligence (AI) offer images of the future that align like iron filings around two magnetic poles—*utopia* and *apocalypse*.

On one hand, AI is said to be leading us toward a perfect future of ease, health, and broadened understanding. We, aided by our machines and their large language models (LLMs), will know virtually everything and make all the right choices to usher in a permanent era of enlightenment and plenty. On the other hand, AI is poised to thrust us into a future of unemployment, environmental destruction, and delusion. Our machines will gobble scarce resources while churning out disinformation and making deadly weapons that [AI agents](#) will use to wipe us out once we're of no further use to them.

Utopia and apocalypse have long exerted powerful pulls on human imagination and behavior. (My first book, published in 1989 and updated in 1995, was *Memories and Visions of Paradise: Exploring the Universal Myth of a Lost Golden Age*; it examined the history and meaning of the utopian archetype.) New technologies tend to energize these two polar attractors in our collective psyche because toolmaking and language are humanity's two [superpowers](#), which have enabled our species to take over the world, while also bringing us to a point of [existential peril](#). New technologies increase some people's power over nature and other people, producing benefits that, mentally extrapolated forward in time, encourage expectations of a grand future. But new technologies also come with costs (resource depletion, pollution, increased economic inequality, accidents, and misuse) that evoke fears of an ultimate reckoning. Language supercharges our toolmaking talent by enabling us to learn from others; it is also the vehicle for formulating and expressing our hopes and fears. AI, because it is both technological and linguistic, and because it is being adopted at a frantic pace and so disruptively, is especially prone to triggering the utopia/apocalypse reflex.

We humans have been ambivalent about technology at least since our adoption of writing. Tools enable us to steal fire from the gods, like the mythical Prometheus, whom the gods punished with eternal torment; they are the wings of Icarus, who flies too close to the sun and falls to his death. AI promises to make technology autonomously intelligent, thus calling to mind still another cautionary tale, "[The Sorcerer's Apprentice](#)."

What could go right—or wrong? After summarizing both the utopian and

apocalyptic visions for AI, I'll explore two questions: first, how do these extreme visions help or mislead us in our attempts to understand AI? And second, whom do these visions serve? As we'll see, there are some early hints of AI's ultimate limits, which suggest a future that doesn't align well with many of the highest hopes or deepest fears for the new technology.

## AI Utopia

As a writer, I generally don't deliberately use AI. Nevertheless, in researching this article I couldn't resist asking Google's free AI Overview, "What is the utopian vision for AI?" This came back a fraction of a second later:

"The utopian vision for AI envisions a future where AI seamlessly integrates into human life, boosting productivity, innovation, and overall well-being. It's a world where AI solves complex problems like climate change and disease, and helps humanity achieve new heights."

Google Overview's first sentence needs editing to remove verbal redundancy (vision, envisions), but AI does succeed in cobbling together a serviceable summary of its promoters' dreams.

The same message is on display in longer form in the article "[Visions of AI Utopia](#)" by Future Sight Echo, who informs us that AI will soften the impacts of economic inequality by delivering resources more efficiently and "in a way that is dynamic and able to adapt instantly to new information and circumstances." Increased efficiency will also reduce humanity's impact on the environment by minimizing energy requirements and waste of all kinds.

But that's only the start. Education, creativity, health and longevity, translation and cultural understanding, companionship and care, governance and legal representation—all will be revolutionized by AI.

There is abundant evidence that people with money share these hopes for AI. The hottest stocks on Wall Street (notably [Nvidia](#)) are AI-related, as are many of the corporations that contribute significantly to the NPR station I listen to in Northern California, thereby gaining naming rights at the top of the hour.

Capital is being shoveled in the general direction of AI so rapidly (roughly [\\$300 billion](#) just this year, in the US alone) that, if its advertised potential is even half believable, we should all rest assured that most human problems will soon vanish.

Or will they?

## AI Apocalypse

Strangely, when I initially asked Google's AI, "What is the vision for AI apocalypse?", its response was, "An AI Overview is not available for this search." Maybe I didn't word my question well. Or perhaps AI sensed my hostility. Full disclosure: I've [gone on record](#) calling for AI to be banned immediately. (Later, AI Overview was more cooperative, offering a lengthy summary of "common themes in the vision of an AI apocalypse.") My reason for proposing an AI ban is that AI gives us humans more power, via language

and technology, than we already have; and that, collectively, we already have way too much power vis-à-vis the rest of nature. We're overwhelming ecosystems through resource extraction and waste dumping to such a degree that, if current trends continue, wild nature may disappear [by the end of the century](#). Further, the most powerful humans are increasingly [overwhelming everyone else](#), both economically and militarily. Exerting our power more intelligently probably won't help, because we're already [too smart for our own good](#). The last thing we should be doing is to cut language off from biology so that it can exist entirely in a simulated techno-universe.

Let's be specific. What, exactly, could go wrong because of AI? For starters, AI could make some already bad things worse—in both nature and society.

There are many ways in which humanity is already destabilizing planetary environmental systems; climate change is the way that's most often discussed. Through its [massive energy demand](#), AI could accelerate climate change by generating more carbon emissions. According to the [International Energy Agency](#), “Driven by AI use, the US economy is set to consume more electricity in 2030 for processing data than for manufacturing all energy-intensive goods combined, including aluminum, steel, cement and chemicals.” The world also faces worsening water shortages; AI needs [vast amounts](#). Nature is already reeling from humanity's accelerating rates of resource extraction and depletion. AI requires millions of tons of copper, steel, cement, and other raw materials, and suppliers are targeting [Indigenous lands](#) for new mines.

We already have plenty of social problems, too, headlined by worsening economic inequality. AI could [widen the divide between rich and poor](#) by replacing lower-skilled workers with machines while greatly increasing the wealth of those who control the technology. Many people worry that corporations have gained too much political influence; AI could [accelerate](#) this trend by making the gathering and processing of massive amounts of data on literally everyone cheaper and easier, and by facilitating the consolidation of monopolies. Unemployment is always a problem in capitalist societies, but AI threatens quickly to throw millions of white-collar workers off payrolls: [Anthropic's](#) CEO Dario Amodei [predicts](#) that AI could eliminate half of entry-level white-collar jobs within five years, while Bill Gates [forecasts](#) that only three job fields will survive AI—energy, biology, and AI system programming.

However, the most horrific visions for AI go beyond just making bad things worse. The title of a recent episode of *The Bulwark Podcast*, “[Will Sam Altman and His AI Kill Us All?](#)”, states the worst-case scenario bluntly. But how, exactly, could AI kill us all? One way is by automating military decisions while making weapons cheaper and more lethal (a recent [Brookings commentary](#) was titled, “How Unchecked AI Could Trigger a Nuclear War”). Veering toward dystopian sci-fi, some AI philosophers opine that the technology, once it's significantly smarter than people, might come to view biological humans as pointless wasters of resources that machines could use more efficiently. At that point, AI could pursue multiple pathways to [terminate humanity](#).

## AI Reality

I don't know the details of how AI will unfold in the months and years to come. But the same could be said for AI industry leaders. They certainly understand the technology better than I do, but their AI forecasts may miss a crucial factor. You see, I've trained myself over the years to look for [limits](#) in resources, energy, materials, and social systems. Most people who work in the fields of finance and technology tend to ignore limits, or even to [believe](#) that there are none. This leads them to absurdities, such as Elon Musk's expectation of [colonizing Mars](#). Earth is finite, humans will be confined to this planet forever, and therefore lots of things we can imagine doing just won't happen. I would argue that discussions about AI's promise and peril need a dose of limits awareness.

Arvind Narayanan and Sayash Kapoor, in an essay titled "[AI Is Normal Technology](#)," offer some of that awareness. They argue that AI development will be constrained by the speed of human organizational and institutional change and by "hard limits to the speed of knowledge acquisition because of the social costs of experimentation." However, the authors do not take the position that, because of these limits, AI will have only minor impacts on society; they see it as an amplifier of systemic risks.

In addition to the social limits Narayanan and Kapoor discuss, there will also (as mentioned above) be environmental limits to the energy, water, and materials that AI needs, a subject explored at a [recent conference](#).

Finally, there's a crucial limit to AI development that's inherent in the technology itself. Large language models need vast amounts of high-quality data. However, as more information workers are replaced by AI, or start using AI to help generate content (both trends are accelerating), [more of the data available to AI will be AI-generated](#) rather than being produced by experienced researchers who are constantly checking it against the real world. Which means AI could become trapped in a cycle of declining information quality. Tech insiders call this "AI model collapse," and there's no realistic plan to stop it. AI itself can't help.

In his article "[Some Signs of AI Model Collapse Begin to Reveal Themselves](#)," Steven J. Vaughan-Nichols argues that this is already happening. There have been widely reported instances of AI inadvertently generating [fake scientific research documents](#). The *Chicago Sun-Times* recently published a "Best of Summer" feature that included [forthcoming novels that don't exist](#). And the Trump administration's widely heralded "Make America Healthy Again" report included citations (evidently AI-generated) for [non-existent studies](#). Most of us have come to expect that new technologies will have bugs that engineers will gradually remove or work around, resulting in improved performance. With AI, errors and hallucination problems may just get worse, in a cascading crescendo.

Just as there are limits to [fossil-fueled](#) utopia, [nuclear](#) utopia, and [perpetual-growth capitalist](#) utopia, there are limits to AI utopia. By the same token, [limits](#) may prevent AI from becoming an all-powerful grim reaper.

What will be the real future of AI? Here's a broad-brush prediction (details are currently unavailable due to my failure to upgrade my crystal ball's

operating system). Over the next few years, corporations and governments will continue quickly to invest in AI, driven by its ability to cut labor costs. We will become systemically dependent on the technology. AI will reshape society—employment, daily life, knowledge production, education, and wealth distribution. Then, speeding up as it goes, AI will degenerate into a hallucinating, blithering cacophony of little voices spewing nonsense. Real companies, institutions, and households will suffer as a result. Then, we'll either figure out how to live without AI, or confine it to relatively limited tasks and data sets. America got a small foretaste of this future recently, when Musk-led DOGE fired tens of thousands of federal workers with the expectation of replacing many of them with AI—without knowing whether AI could do their jobs (oops: thousands are being [rehired](#)).

A messy neither-this-nor-that future is not what you'd expect if you spend time reading documents like "[AI 2027](#)," five industry insiders' detailed speculative narrative of the imminent AI future, which allows readers to choose the story's ending. Option A, "slowdown," leads to a future in which AI is merely an obedient, super-competent helper; while in option B, "race," humanity is extinguished by an AI-deployed bioweapon because people take up land that could be better used for more data centers. Again, we see the persistent, binary utopia-or-apocalypse stereotype, here presented with impressive (though misleading) specificity.

At the start of this article, I attributed AI utopia/apocalypse discourse to a deep-seated tic in our collective human unconscious. But there's probably more going on here. In her recent book [Empire of AI](#), tech journalist Karen Hao traces polarized AI visions back to the founding of OpenAI by Sam Altman and Elon Musk. Both were, by turns, dreamers and doomers. Their consistent message: *we* (i.e., Altman, Musk, and their peers) are the only ones who can be trusted to shepherd the process of AI development, including its regulation, because we're the only ones who understand the technology. Hao makes the point that messages about both the promise and the peril of AI are often crafted by powerful people seeking to consolidate their control over the AI industry.

Utopia and apocalypse feature prominently in the rhetoric of all cults. It's no surprise, but still a bit of a revelation, therefore, to hear Hao conclude in a podcast interview that [AI is a cult](#) (if it walks, quacks, and swims like a cult . . .). And we are all being swept up in it.

So, how should we think about AI in a non-cultish way? In his article, "[We Need to Stop Pretending AI Is Intelligent](#)," Guillaume Thierry, a professor of cognitive neuroscience, writes, "We must stop giving AI human traits." Machines, even apparently smart ones, are not humans—full stop. Treating them as if they *are* human will bring dehumanizing results for real, flesh-and-blood people.

The collapse of civilization won't be AI generated. That's because environmental-social decline was already happening without any help from LLMs. AI is merely adding a novel factor in humanity's larger reckoning with limits. In the short run, the technology will further concentrate wealth. "Like empires of old," writes Karen Hao, "the new empires of AI are amassing extraordinary riches across space and time at great expense to everyone else." In the longer run, AI will deplete scarce resources faster.

If AI is unlikely to be the bringer of destruction, it's just as unlikely to deliver heaven on Earth. Just last week I heard from a writer friend who used AI to improve her book proposal. The next day, I went to my doctor for a checkup, and he used AI to survey my vital signs and symptoms; I may experience better health maintenance as a result. That same day, I read a just-published [Apple research paper](#) that concludes LLMs cannot reason reliably. Clearly, AI can offer tangible benefits within some fields of human pursuit. But we are fooling ourselves if we assume that AI can do our thinking for us. If we can't build an equitable, sustainable society on our own, it's pointless to hope that a machine that can't think straight will do it for us.

I'm not currently in the job market and therefore can afford to sit on the sidelines and cast judgment on AI. For many others, economic survival depends on adopting the new technology. Finding a personal modus vivendi with new tools that may have dangerous and destructive side effects on society is somewhat analogous to charting a sane and survivable daily path in [a nation succumbing to authoritarian rule](#). We all want to avoid complicity in awful outcomes, while no one wants to be targeted or denied opportunity. Rhetorically connecting AI with dictatorial power makes sense: one of the most likely uses of the new technology will be for mass surveillance.

Maybe the best advice for people concerned about AI would be analogous to [advice that democracy advocates are giving](#) to people worried about the destruction of the social-governmental scaffolding that has long supported Americans' freedoms and rights: identify your circles of concern, influence, and control; scrutinize your sources of information and tangibly support those with the most accuracy and courage, and the least bias; and forge communitarian bonds with real people.

AI seems to present a spectacular new slate of opportunities and threats. But, in essence, much of what was true before AI remains so now. Human greed and desire for greater control over nature and other people may lead toward paths of short-term gain. But, if you want a good life when all's said and done, learn to live well within limits. Live with honesty, modesty, and generosity. AI can't help you with that.