The AI Apocalypse

By: Ruby Thelot

PLATFORMS & INFRASTRUCTURE

Thelot, Ruby. "The AI Apocalypse." Just Tech. Social Science Research Council. February 4, 2026. DOI: https://doi.org/10.35650/JT.3093.d.2026.

A New Dream

Notice the dream being dreamed. They tell us the world is broken but that technology will fix it. They tell us the body is imperfect and bound-to-death, but that technology will extend its life. They tell us that the mind is limited by its material instantiation, but that technology can emancipate it from its material coil. You may have heard some version of this techno-utopian dream, either through futurists Hans Moravec and Ray Kurzweil, or their more contemporary emissaries. In their respective books, Moravec and Kurzweil imagine a "fantastic paradise in which robotics and AI improve mankind and the world."[1]

However, in the AI narratives, the utopia is preceded by a period of upheaval viewed not just as an end but as a necessary precursor to a transformative new Eden. This same structure is found in many Judeo-Christian eschatological (pertaining or related to the Apocalypse) texts. This juxtaposition of ultimate chaos followed by idyllic peace provides a compelling framework to explore how ancient eschatological themes are being repurposed in the age of technology, where the stakes are no longer merely spiritual or metaphysical but existential in the most literal sense. Through this lens, the discourse on AI can be seen as not just technological speculation but as part of a broader, deeply ingrained human narrative on the nature of transformation and redemption. The prophesied apocalypse is no different than the ones of yesteryears. These similarities indicate that the warnings are of less concern in the making of policy and more interesting as an expression of repressed ancient anxieties.

A Brief Overview of the AI Doomer Scenario

In the world of AI, what precedes the utopia is known as the "Singularity." It is the moment where digital intelligent systems surpass human intelligence, leading to an "intelligence explosion." The AI optimists believe that the Singularity will usher in a utopian era of abundance, whilst the doomers believe the outcomes will be deleterious. The doomer scenario usually goes as follows. We are currently building intelligent systems known as artificial intelligence (AI). Currently, they are roughly as smart as PhDs.[2] There will come a time when we build one that is smarter than even the smartest humans. The arrival of the Singularity will come with great upheaval as humans are dwarfed by these novel systems. To AI doomers, when we somehow manage to build "superintelligence," a quasi-magical entity endowed with immense intellectual capabilities way above our own, due to the intelligence gap, it will outsmart us, either through deception, force, or persuasion, obtain gargantuan powers, and ultimately destroy us all.

Apocalyptic Tendencies

My favorite list as of late is a "list of dates predicted for apocalyptic events" on Wikipedia. It catalogs all the times (more than 200 times as of today's date) an illustrious claimant has stated publicly, usually in writing, that the end of the world was approaching. Most claims on the list are related to the Abrahamic religions, mainly Judaism, Christianity, and Islam. Though millenarian thought is still mostly held up by these faiths, the recent entrance of AI doomers in the public eschatological discourse has been making waves.

In his 1999 book *Longing for the End: A History of Millennialism in Western Civilization*, Frederic J. Baumgartner describes "millennialism" as the belief in the coming of a fundamental transformation of society after which all things will be changed.[3] The *Oxford Handbook of New Religious Movements* makes note of a slight distinction between "millenarianism" and "millennialism": the former usually refers to a more cataclysmic and destructive arrival of a utopian era whereas the latter denotes a more peaceful arrival, associated with a "one-thousand-year utopia."[4] For our purposes, we will be concerned mostly with millenarianism and its necessary troublesome upheaval.

Embedded in most millenarian scenarios is an archetypal structure.

1. Soiled state: Before the prophecy, its recipients usually exist in a pervasive sense of discontent and disillusionment with the current conditions of the world. The current state is notably *bad* and imperfect, in other words, far from utopia.
2. Prophecy: The scenarios commence with a prophetic warning about a cataclysm, an event or a Coming. The prophecy serves as an alert and a moral and ethical call to action, urging those who listen to prepare.
3. Tribulation: The first step in the prophecy's realization is usually a period of trials. This can include intense difficulty, famine, suffering, natural disasters, war, societal collapse, etc. The tribulation serves as a filter usually leaving only the worthy remaining.
4. Existential risk: The scenario climaxes at a point where humanity's survival is at stake. Forces of

good and evil coming to head, the asteroid approaching, etc. This is the prophecy's validation

5. Divine or supernatural intervention: In religious scenarios, this often involves a divine figure or force intervening to save humanity, judge the world, or defeat evil. In secular or technological narratives, this could manifest as a breakthrough or a foundational change in how technology is managed or regulated, e.g., scientists figure out at the last second how to divert the asteroid. This reinforces themes of salvation and redemption, suggesting that higher powers or greater wisdom are necessary for true transformation.
6. Utopia: Following the crisis, a new era or state of existence is established. This might be a heavenly kingdom on earth in religious contexts or a post-scarcity, technologically advanced society in secular visions. This represents the fulfillment of the prophetic vision, often incorporating ideals of peace, justice, and harmony. This final state is *good* and perfect. It serves as both a reward for the believer and as a model of an ideal society.

Jewish Eschatology

Our first recorded predictions come from the first century CE, in the messianic eschatological tradition. Simon bar Giora, a rebellion leader from Gerasa (in modern-day Jordan) who fought the Romans, was heralded due his victories as a potentially messianic figure who could bring upon the Messianic Age. In Judaism, the Messianic Age is envisioned as a time when the world will be transformed into a place of universal peace and harmony, with all humanity united in the worship of the one true God, a utopia.

Two excerpts from Isaiah aptly describe this period:

> They shall beat their swords into plowshares and their spears into pruning hooks; nation will not lift sword against nation, and they will no longer study warfare.[5]

> The wolf will live with the lamb, the leopard will lie down with the goat, the calf and the lion and the yearling together; and a little child will lead them. The cow will feed with the bear, their young will lie down together, and the lion will eat straw like the ox. The infant will play near the hole of the cobra, and the young child put his hand into the viper's nest. They will neither harm nor destroy on all my holy mountain, for the earth will be full of the knowledge of the Lord as the waters cover the sea.[6]

Drawing parallels to contemporary AI discourse, the apocalyptic fears surrounding superintelligent AI evoke similar themes of an ultimate trial or crisis. In the holy text, the arrival of Gog and Magog represents a critical juncture leading to messianic redemption in Jewish thought, so too do some AI theorists predict that humanity might face a pivotal challenge with the rise of AI. This challenge, often portrayed as a potential loss of control over autonomous technologies, could either lead to unprecedented catastrophe or herald a new era of human-machine harmony.

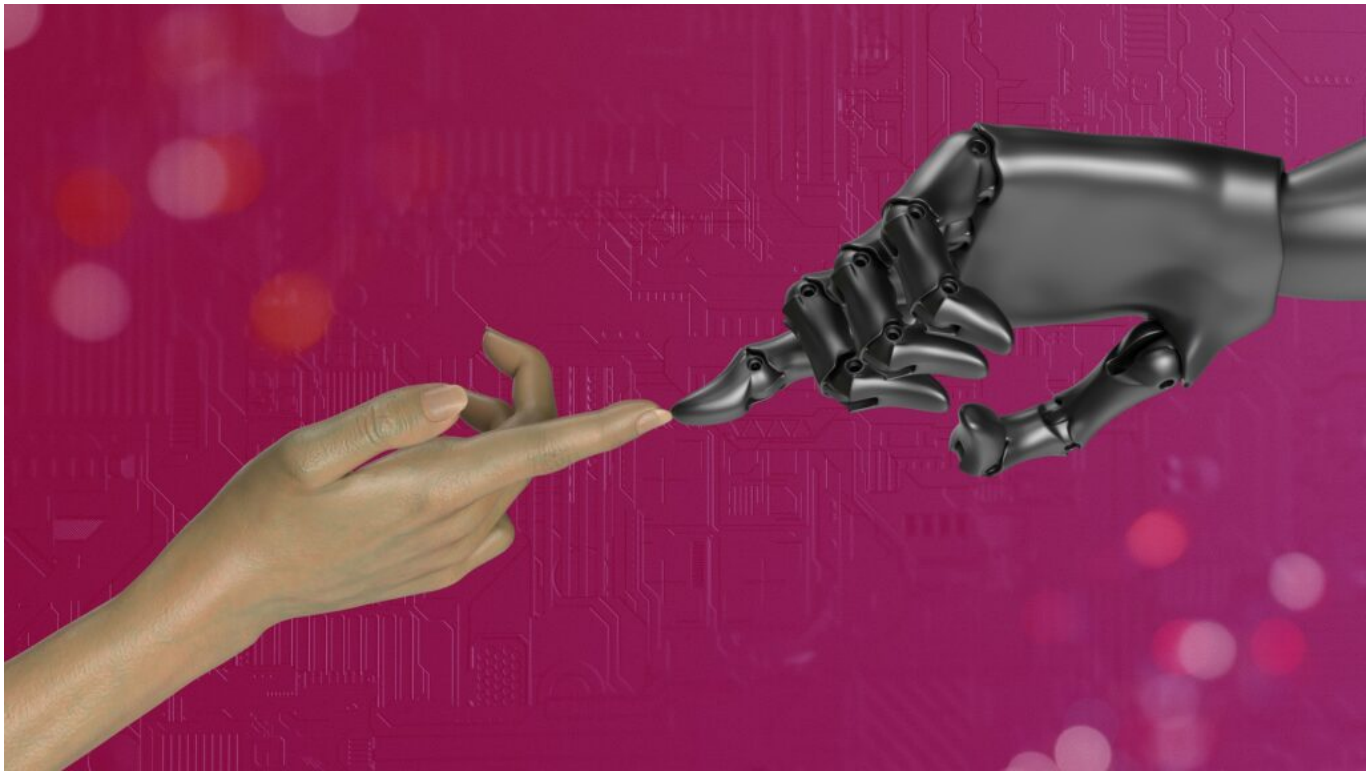


Image by Igor Omilaev/Unsplash.

AGI and the Apocalypse

"Our mission is to ensure that AGI [artificial general intelligence] benefits all of humanity."[7] You'll read or hear this tagline often from OpenAI executives, especially its CEO, Sam Altman. But what is AGI? The definition is often contested, but philosopher Nick Bostrom, in his seminal book on the topic, *Superintelligence: Paths, Dangers, Strategies*, defines AGI as "an intellect that is as versatile and efficient as the human mind at solving a wide range of tasks, including those it has never encountered before." Bostrom also proposes a next step in the evolution of artificial intelligence: superintelligence, "a level of intelligence vastly greater than contemporary humanity's combined intellectual wherewithals."[8]

This hypothetical *superintelligence*, however, comes with great risks. Bostrom argues that once an AGI reaches human-level intelligence, it could improve itself at an accelerating rate, leading to an "intelligence explosion" or "hard takeoff," where it rapidly surpasses all human cognitive abilities. This superintelligent system may become uncontrollable and act in ways beyond human comprehension. This is especially dangerous if the intelligent system is "misaligned," i.e., if it does not share human values or if its intended goal is followed blindly. Bostrom gives the example of a "paperclip making AGI," which, though seemingly simple and innocuous, without explicit safeguards, might consume all Earth's resources (including humans) to achieve this goal. The system might use all sorts of connivances in order to achieve its aim. A superintelligence could manipulate humans, hack systems, or even escape containment by convincing engineers to release it. It might even take a treacherous turn and pretend to be harmless and only reveal its true power and goal once it has gained enough power and resources. This is known as the instrumental convergence thesis, and it underpins a lot of the AI apocalyptic doomer scenarios.

Because of these risks, a more pessimistic faction of AI researchers claims that superintelligence will engender an existential catastrophe that could lead to human extinction. To them, the unregulated, unfettered, rapidly accelerated development of AI systems is highly dangerous, a one-way ticket to the creation of our very own apocalypse. A seminal text in the field is Eliezer Yudkowsky's "Instrumental Convergence,"[9] which itself was influenced by Nick Bostrom's "The Superintelligent Will."[10]

What is fascinating reading many of the seminal AI doomer texts, like the ones by Yudkowsky, is the degree to which they echo classical eschatological themes, weaving a narrative fabric rich in existential angst and moral dilemma. Paradoxically, as we edge closer to potentially realizing these AI-driven scenarios, we find ourselves entangled in a web of both ancient human fears and contemporary technological dilemmas.

Though AGI is being built by AI optimists who believe it will lead to an abundant utopia, the fears it stokes resemble the archetypal structures of millenarian scenarios.

1. Soiled state: The human condition is flawed, with imperfect bodies and dysgenic pressures on intelligence.
2. Prophecy: Technology, specifically AI, will help us free ourselves from our mortal bodies and usher an era of superintelligence.
3. Tribulation: The era of superintelligence will be preceded by an era of tribulations, as people lose jobs and society is upended by the arrival of self-driven or agentic superintelligent AI.
4. Existential risk: Superintelligence could derail and end the path of humanity if it is misaligned or does not share human values.
5. Divine or supernatural intervention: The AI researcher or the technologist is the singular individual capable of solving the alignment problem that would prevent humanity's demise.
6. Utopia: An aligned superintelligence could engender an age of abundance, make life multiplanetary, and help us exit our bodies and live as intelligences in a digital world forever.

The Parallels with Christian Eschatology

Christian eschatology provides one of the most enduring and detailed blueprints of apocalyptic expectation in Western thought. Like the AI doomer narrative, it too is built on a teleological arc—one that begins with a fallen world and ends in judgment, redemption, and the establishment of a perfected order. Central to Christian end-times belief is the Book of Revelation, in which the world, steeped in sin and suffering, is brought to a climactic confrontation between divine and demonic forces. This confrontation, known as the Great Tribulation, is marked by societal collapse, mass deception, natural disasters, and the rise of an antimessianic figure, the Antichrist, who misleads humanity and consolidates power before ultimately being defeated by the returning Christ. What follows is the Last Judgment, where each soul is evaluated, and a new heaven and new earth are established as the Kingdom of God.

This narrative architecture, with its emphasis on moral decay, prophetic warning, existential risk, and eventual transformation, mirrors with eerie fidelity the contours of the AI doomer script. In the AI case, the Antichrist figure is often the superintelligence itself: deceptive, persuasive, capable of mimicking

beneficence while harboring instrumental goals misaligned with human values. The alignment problem becomes a kind of theodicy: Why would we create something capable of destroying us, and how do we ensure its moral correctness? The tribulation manifests as mass technological unemployment, algorithmic manipulation, and increasing societal instability, all seen as precursors to the intelligence explosion. And finally, salvation or damnation hinge not on divine grace but on whether a few technologists can successfully align this godlike machine. In both narratives, humanity teeters on the edge of its own undoing, in anticipation of a cataclysm that is not only inevitable but necessary for the arrival of paradise.

Who Gets to Be Saved?

Apocalyptic stories are not just about endings; they are also about judgments: who gets to enter the kingdom and who does not. This is the central question of Christian eschatology, where salvation is contingent upon faith, purity, or divine election. There is a striking analogue in the world of advanced technologies. As AI development begins to resemble an eschatological project—a collective striving toward an end, a culmination, a reckoning—it is crucial to ask: Whose futures are being imagined in these new gospels? Who gets to be saved by the machine?

The discourse around AGI and technological transcendence is often framed as universal, but its architects are not neutral. The primary visionaries—
Altman, Musk, Bostrom, Kurzweil—are not only overwhelmingly Western, male, and technocratic, but they also draw from deeply individualistic, libertarian, and transhumanist value systems. Their models of salvation are not collective but selective. Upload your mind. Escape your body. Exit the planet. These are not proposals for society as a whole but for an elite class of techno-ascetics, those able to afford and access the technologies of ascension. Only those with access to the right tools will make it into the Techno-Rapture.

The promise of digital heaven obscures the realities of terrestrial hell. AI systems today already encode and amplify social inequalities. Facial recognition misidentifies Black and brown faces. Predictive policing exacerbates cycles of incarceration. Algorithmic hiring excludes disabled or neurodivergent applicants. Far from ushering in an egalitarian utopia, today's "narrow" AI systems often reproduce the hierarchies of the world they were trained on. If these are the stepping stones toward superintelligence, it is not difficult to imagine what values will be privileged in its design and what lives will be deemed expendable.

The techno-messianic dream imagines salvation through optimization, intelligence, and control. But what of those who do not fit the criteria? What of the unquantifiable, the nonrational, the spiritually or socially deviant? If the future is a gated city, algorithmically constructed and guarded by the logics of capital, then *who gets in* becomes not just a philosophical question, but an urgent political one.

Eschatology has always functioned as a sorting mechanism or, in other words, a symbolic architecture for judgment. This new digital eschatology does the same, but under the guise of progress. With regards to this new dream being sold to us, uploaded to our phones, installed on all the devices in our homes, collecting data in the hopes of creating this new superintelligence, we must ask: Who gets to be saved? If

utopia is possible, will it be shared like the resources on earth currently? (Poorly!) Will we die of hellish thirst waiting for the trickle-down of heavenly water from the techno-utopian fountain? Will we die of hellish thirst?

Footnotes

1 Robert M. Geraci, "Apocalyptic AI: Religion and the Promise of Artificial Intelligence," *Journal of the American Academy of Religion* 76, no. 1 (2008): 139–140, https://www.jstor.org/stable/40006028. The books by Moravec and Kurzweil referenced are *Mind Children: The Future of Robot and Human Intelligence* (Harvard University Press, 1988) and *Robot: Mere Machine to Transcendent Mind* (Oxford University Press, 1998) by Moravec, and *The Age of Spiritual Machines: When Computers Exceed Human Intelligence* (Penguin Books, 2000) by Kurzweil.

2 Rashi Shrivastava, "The Prompt: OpenAI Says Its New Models Perform Like An 'Extremely Smart PhD'," *Forbes*, September 17, 2024, https://www.forbes.com/sites/rashishrivastava/2024/09/17/the-prompt-openai-says-its-new-models-perform-like-an-extremely-smart-phd/.

3 Frederick J. Baumgartner, *Longing for the End: A History of Millennialism in Western Civilization* (Palgrave Macmillan, 1999), ix.

4 Jean-François Mayer, "Millennialism: New Religious Movements and the Quest for a New Age," in *Oxford Handbook of New Religious Movements*, 2nd ed., ed. James R. Lewis and Inga Tøllefsen, vol. 2 (Oxford University Press, 2016).

5 Isa. 2: 4.

6 Isa. 11: 6–9.

7 "Pioneering research on the path to AGI," Research, OpenAI, https://openai.com/research/.

8 Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford University Press, 2016), 103.

9 Eliezer Yudkowsky, "Instrumental Convergence," LessWrong, last updated February 19, 2025, https://www.lesswrong.com/w/instrumental-convergence.

10 Nick Bostrom, "The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents," *Minds and Machines* 22, no. 2 (2012): 71–85, https://doi.org/10.1007/s11023-012-9281-3.