# Apocalyptic AI

## The Convivial Society: Vol. 4, No. 5

[L. M. Sacasas](#)

*Welcome to the* Convivial Society. *A newsletter about technology and culture with both those terms understood quite broadly. AI, of course, is the topic of the moment, and it is the topic of this newsletter. Ordinarily, you would find a rather traditional essay below. What you have here is a sub-genre of the* Convivial Society *which I've taken to labeling "Fragments," a loosely structured list of associated quotations, reflections, and provocations. I'm drawn to this form because it reflects the provisional and associative nature of thinking. It also reflects the way fragments of thought, often surfaced from another time, can gather around a problem to illuminate its contours, disclose its depths, and perhaps even reveal lines of actions. Such fragments, in any case, may be all we have to work with. I also appreciate the fact that the form invites rather than forecloses further thought.*

**1.** Thinking cogently and insightfully about AI is a bit of challenge right now. Or maybe I should be more modest in my claim. I am myself finding it challenging to think cogently and insightfully about AI. Part of the problem is that the term is used rather indiscriminately, so it is hard to pin down what exactly one is talking about with the kind of specificity that sound thinking requires.[1] It's also difficult to fix your thinking on a phenomenon that is rapidly developing. Finally, it is challenging to think about AI because it's hard to distinguish among what is actually happening, sound speculation about what may happen, hype, and [criti-hype](#) (historian Lee Vinsel's term for critical reflection that takes the hype at face value). But you're not reading me for the hottest take on emerging trends, so I'm going to proceed as per usual, deliberately. I'm sure we'll be thinking about AI for the foreseeable

future, and I'll continue share my thoughts insofar as I judge them to be potentially helpful.

In what follows, I'm using the term "AI" in a manner similar to how Kate Crawford uses it in *Atlas of AI*. Crawford observes that "artificial intelligence" is a term that "is both used and rejected in ways that keep its meaning in flux." She notes, too, that "'machine learning' is more commonly used in the technical literature." Consequently, she chooses to "use AI to talk about the massive industrial formation that includes politics, labor, culture, and capital" while using "machine learning" to refer to "a range of technical approaches." Likewise, I am using AI here not to designate an array of specific technical practice and capabilities, but rather the present amorphous techno-cultural idea of AI as it is deployed, debated, feared, and celebrated. This idea of AI is not only as a massive industrial formation including politics, labor, culture, and capital, but also the consummation of a historical development which we ordinarily gloss as modernity.

**2.** Ezra Klein, discussing the disconcertingly high number of researchers work on AI, who appear to believe, rather earnestly, that their work poses a non-trivial risk of causing cataclysmic harm to the human race:

> "I often ask them the same question: If you think calamity so possible, why do this at all? Different people have different things to say, but after a few pushes, I find they often answer from something that sounds like the A.I.'s perspective. Many — not all, but enough that I feel comfortable in this characterization — feel that they have a responsibility to usher this new form of intelligence into the world."

**3.** Robert Oppenheimer testifying before the Atomic Energy Commission in 1954:

> "However it is my judgment in these things that when you see something that is technically sweet, you go ahead and do it and you argue about what to do about it only after you have had your technical success. That

> is the way it was with the atomic bomb. I do not think anybody opposed making it; there were some debates about what to do with it after it was made. I cannot very well imagine if we had known in late 1949 what we got to know by early 1951 that the tone of our report would have been the same. You may ask other people how they feel about that. I am not at all sure they will concur; some will and some will not."

**4.** The late David Noble's _The Religion of Technology: The Divinity of Man and the Spirit of Invention_, first published in 1997, is a book that I turn to often. Noble was adamant about the sense in which readers should understand the phrase "religion of technology." "Modern technology and modern faith are neither complements nor opposites," Noble argued, "nor do they represent succeeding stages of human development. They are merged, and always have been, the technological enterprise being, at the same time, an essentially religious endeavor."

"This is not meant in a merely metaphorical sense," he goes on to explain, "to suggest that technology is similar to religion in that it evokes religious emotions of omnipotence, devotion, and awe, or that it has become a new (secular) religion in and of itself, with its own clerical caste, arcane rituals, and articles of faith." On the contrary, Noble insisted that "it is meant literally and historically, to indicate that modern technology and religion have evolved together and that, as a result, the technological enterprise has been and remains suffused with religious belief."

One of the chapters in Noble's account is about artificial intelligence and artificial life. Here are a few selections from that chapter, beginning with Noble's summary of his findings:

- "At first the effort to design a thinking machine was aimed at merely replicating human thought. But almost at once sights were raised, with the hope of mechanically surpassing human thought by creating a 'super intelligence,' beyond human capabilities. Then the prospect of an

immortal mind able to teach itself new tricks gave rise to the vision of new artificial species which would supersede *Homo sapiens* altogether. Totally freed from the human body, the human person, and the human species, the immortal mind could evolve independently into ever higher forms of artificial life, reunited at last with its origin, the mind of God."

- "[Marvin] Minsky described the human brain as nothing more than a 'meat machine' and regarded the body, that 'bloody mess of organic matter,' as a 'teleoperator for the brain.' Both, he insisted, were eminently replaceable by machinery. What is important about life, Minsky argued, is 'mind,' which he defined in terms of 'structure and subroutines'—that is, programming."

- "[Daniel] Crevier recounts the discussion of such a possibility that began to surface on the AI grapevine in the 1980s, in particular the idea of 'downloading' the mind into a machine, the transfer of the human mind to an 'artificial neural net' through the 'eventual replacement of brain cells by electronic circuits and identical input-output functions.'"

- "If intelligent machines were viewed as vehicles of human transcendence and immortality, they were also understood as having lives of their own and an ultimate destiny beyond human experience. In the eyes of AI visionaries, mind machines represented the next step in evolution, a new species, *Machina sapiens*, which would rival and ultimately supersede *Homo sapiens* as the most intelligent beings in creation. 'I want to make a machine that will be proud of me,' Danny Hillis proclaimed, acknowledging the superiority of his creation. 'I guess I'm not overly perturbed by the prospects that there might be something better than us that might replace us ... We've got a lot of bugs, sorts of bugs of left over history back from when we were animals. And I see no reason to believe that we're the end of the chain and I think better than us is possible.'" [The most striking part of this was the phrase "from when we *were* animals." The past tense is telling.]

- "'The enterprise is a god-like one,' AI enthusiast Pamela McCorduck observed. 'The invention—the finding within—of gods represents our reach for the transcendent … 'Our speculation ends in a supercivilization,' [Hans Moravec] prophesied, 'the synthesis of all solar system life, constantly improving and extending itself, spreading outward from the sun, converting non-life into mind … This process might convert the entire universe into an extended thinking entity … the thinking universe … an eternity of pure cerebration.'"

- "'The manifest destiny of mankind is to pass the torch of life and intelligence on to the computer,' [Rudy] Rucker proclaimed."[2]

**5.** The Enlightenment did not, as it turns out, vanquish Religion, driving it far from the pure realms of Science and Technology. In fact, to the degree that the radical Enlightenment's assault on religious faith was successful, it empowered the religion of technology. To put this another way, the Enlightenment—and, yes, we are painting with broad strokes here—did not do away with the notions of Providence, Heaven, and Grace. Rather, the Enlightenment re-framed these as Progress, Utopia, and Technology respectively. If heaven had been understood as a transcendent goal achieved with the aid of divine grace within the context of the providentially ordered unfolding of human history, it became a Utopian vision, a heaven on earth, achieved by the ministrations Science and Technology within the context of Progress, an inexorable force driving history toward its Utopian consummation.

**6.** It is also important to be a bit more specific, and to classify the religion of technology more precisely as a Christian heresy. It is in Western Christianity that Noble found the roots of the religion of technology, and it is in the context of post-Christian world that it has presently flourished.

It is Christian insofar as its aspirations that are like those nurtured by the Christian faith, such as the conscious persistence of a soul after the death of

the body. Noble cites Daniel Crevier, who referencing the "Judeo-Christian tradition" suggested that "religious beliefs, and particularly the belief in survival after death, are not incompatible with the idea that the mind emerges from physical phenomena." This is noted on the way to explaining that a machine-based material support could be found for the mind, which leads Noble to quip. "Christ was resurrected in a new body; why not a machine?" Reporting on his study of the famed Santa Fe Institute in Santa Fe, New Mexico, anthropologist Stefan Helmreich observed, "Judeo-Christian stories of the creation and maintenance of the world haunted my informants' discussions of why computers might be 'worlds' or 'universes,' .... a tradition that includes stories from the Old and New Testaments (stories of creation and salvation)."

It is a heresy insofar as it departs from traditional Christian teaching regarding the givenness of human nature, the moral dimensions of humanity's brokenness, the gracious agency of God in the salvation of humanity, and the resurrection of the body, to name a few. Having said as much, it would seem that one could perhaps conceive of the religion of technology as an imaginative account of how God might fulfill purposes that were initially revealed in incidental, pre-scientific garb. In other words, we might frame the religion of technology not so much as a Christian heresy, but rather as (post-)Christian fan-fiction, an elaborate imagining of how the hopes articulated by the Christian faith will materialize as a consequence of human ingenuity in the absence of divine action.3

**7.** In a 1979 collection of essays, *Hannah Arendt: The Recovery of the Public World*, edited by Melvyn Hill, you can find the transcript of 1972 conference on Arendt's work in which Arendt herself participated. Below is part of an exchange between the (under-appreciated) philosopher Hans Jonas and Arendt:

> **Jonas**: I share with Hannah Arendt the position that we are not in possession of any ultimates, either by knowledge or by conviction or faith

[...]

However, a part of wisdom is knowledge of ignorance. The Socratic attitude is to know that one does not know. And this realization of our ignorance can be of great practical importance in the exercise of the power of judgment, which is after all related to action in the political sphere, into future action, and far-reaching action.

Our enterprises have an eschatological tendency in them—a built in utopianism, namely, to move towards ultimate situation. Lacking the knowledge of ultimate values—or, of what is ultimately desirable—or, of what is man so that the world can be fitting for man, we should at least abstain from allowing eschatological situations to come about. This alone is a very important practical injunction that we can draw from the insight that only with some conception of ultimates are we entitled to embark on certain things. So that at least as a restraining force the point of view I brought in may be of some relevance.

**Arendt**: With is I would agree.

8. Wendell Berry, *Life Is A Miracle: An essay against modern superstition* (2000):

"What I am against—and without a minute's hesitation or apology—is our slovenly willingness to allow machines and the idea of the machine to prescribe the terms and conditions of the lives of creatures, which we have allowed increasingly for the last two centuries, and are still allowing, at an incalculable cost to other creatures *and to ourselves*. If we state the problem that way, then we can see that the way to correct our error, and so deliver ourselves from our own destructiveness, is to quit using our technological capability as the reference point and standard of our economic life. We will instead have to measure our economy by the health of the ecosystems and human communities where we do our work.

> It is easy for me to imagine that the next great division of the world will be between people who wish to live as creatures and people who wish to live as machines."

**9.** AI is apocalyptic in exactly one narrow sense: it is not causing, but rather revealing the end of a world. We get our word *apocalypse* from a Greek word meaning "to reveal, to disclose, or to uncover." What I am suggesting is that AI, as it is being developed, deployed, and hyped (and criti-hyped), forces us to reckon with the fact that modernity is expiring, and it is expiring precisely to the degree that it no longer serves the interest of and is at various points, particularly in its techno-economic dimensions, openly hostile to the human person. As a second nature, the culture of technological modernity, while undoubtedly improving the lot of humanity in important ways, has become, in other respects, inhospitable to our species. AI can thus be read as a last ditch effort to shore up the old decrepit structures and to double down on the promise of scale, efficiency, rationality, control, and prediction. It can also be read as an effort to extend the logic of late modernity to a point of absurdity. So where we see a proposed or actual application of AI, we might do well to ask how it relates to the end of the world we have called modern.

Another way to think about this is to recognize that modernity derived its cultural power and energy from an unstable ideological compound. The constituent elements of this compound were, on the one hand, a liberal commitment to the individual human person and, on the other, a drive to transcend the perceived deficiencies (later simply the inherent limits) of the human condition. Alternatively, we might also describe the unstable compound as mixture of the promise an unfettered individual will realizing its desires coupled to a system which ultimately demands that human desires be managed, predicted, and channeled to serve the ends of a market economy.

Fears about AI signal the decomposition of the ideological compound. The two constituent elements can no longer be synthesized. The resulting

system demands or threatens the elimination of the human person. But this must not be understood ultimately as the risk of the appearance of a new, alien super-intelligence. Rather, it must be understood as the culmination of a longstanding trajectory. The default eschatology of technological modernity has always been the eclipse of the human person, and this is because its model of the human person was dominated by the image of a disincarnate mind exerting rational control over the material world. Now, as it turns out, those most enthralled by this model of the human being grow increasingly anxious about the prospect of a man-made disincarnate mind overthrowing the human race as it pursues its own rationally optimized goals.

**10.** G. K. Chesterton: "The madman is not the man who has lost his reason. The madman is the man who has lost everything except his reason."

[Share](#)

[1](#)

Ten years ago, we might have been speaking of "Big Data." Five years ago the term of art would've been "Algorithms." For some time, "Social Media" was the object of attention. Now we talk about "AI." In some respects, we have been talking about the same sorts of things all along, which is not to say that there have been no developments or it is all exactly the same. Also, I capitalize the terms and put them in quotation marks to suggest that what I am identifying is not the specific artifact, process, or system, but the object of discourse that encompasses the material artifact, process, or system along with a host of other intertwined concerns and gets refracted through the lens of media analysis.

[2](#)

Curiously, Rudy Rucker's full name is Rudolf von Bitter Rucker, and he happens to be, through his mother, the great-great-great-grandson of the

philosopher G.W.F. Hegel. Hegel's philosophy (with which, to be clear, I profess no deep proficiency) seemed to be lurking in the unspoken background of many of the thinkers Noble explored in this chapter.