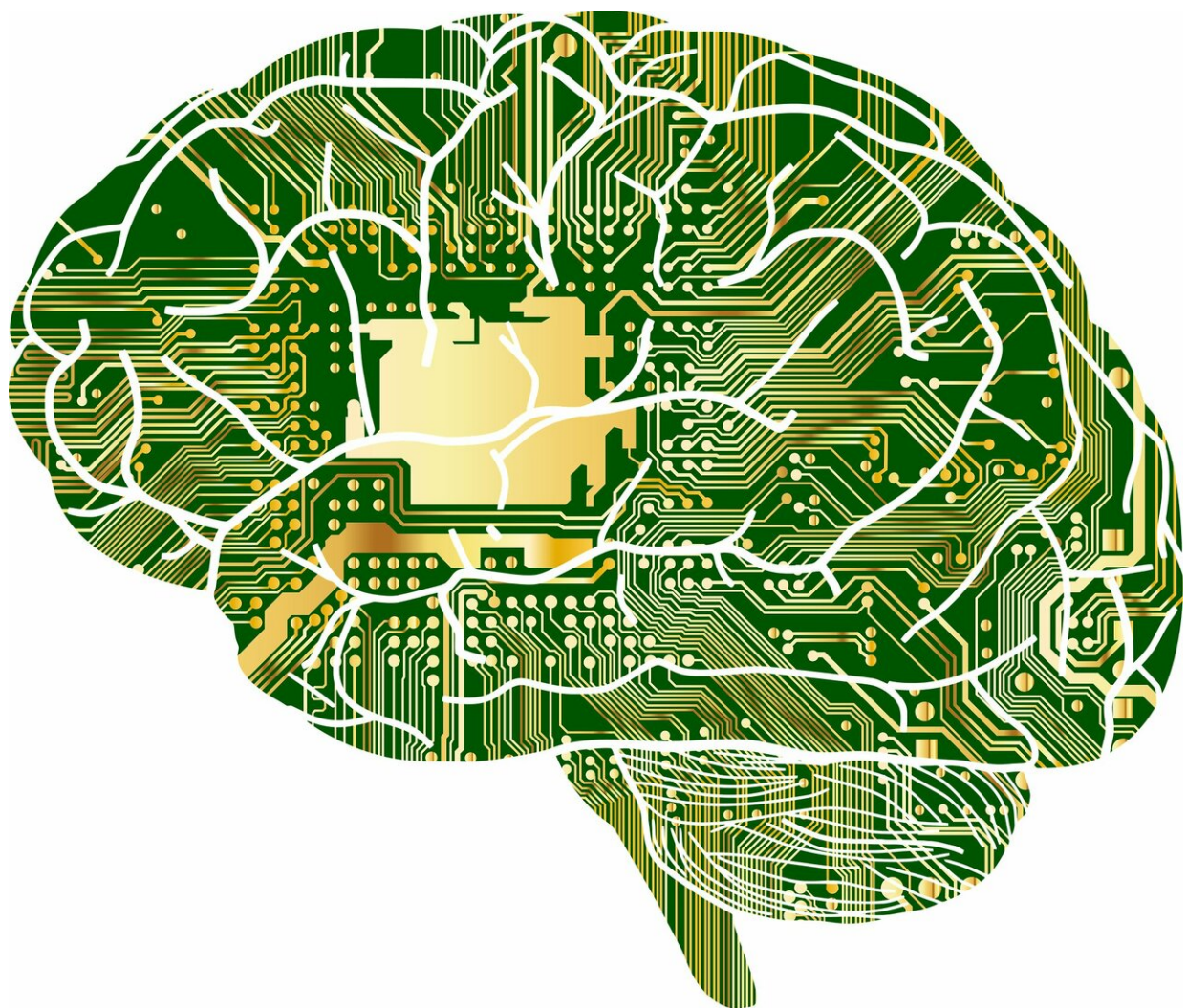


# No, AI probably won't kill us all—and there's more to this fear campaign than meets the eye

June 1 2023, by Michael Timothy Bennett

---



Credit: CC0 Public Domain

Doomsaying is an old occupation. Artificial intelligence (AI) is a complex subject. It's easy to fear what you don't understand. These three truths go some way towards explaining the oversimplification and dramatization plaguing discussions about AI.

Yesterday outlets around the world were plastered with news of yet another [open letter claiming](#) AI poses an existential threat to humankind. This letter, published through the nonprofit Center for AI Safety, has been signed by industry figureheads including Geoffrey Hinton and the chief executives of Google DeepMind, Open AI and Anthropic.

However, I'd argue a healthy dose of skepticism is warranted when considering the AI doomsayer narrative. Upon close inspection, we see there are commercial incentives to manufacture fear in the AI space.

And as a researcher of artificial general intelligence (AGI), it seems to me the framing of AI as an existential threat has more in common with 17th-century philosophy than computer science.

## **Was ChatGPT a 'breaththrough'?**

When ChatGPT was released late last year, people were delighted, entertained and horrified.

But ChatGPT isn't a research breakthrough as much as it is a product. The technology it's based on is several years old. An early version of its underlying model, GPT-3, was released in 2020 with many of the same capabilities. It just wasn't easily accessible online for everyone to play with.

Back in 2020 and 2021, [I](#) and many [others](#) wrote papers discussing the capabilities and shortcomings of GPT-3 and similar models—and the world carried on as always. Forward to today, and ChatGPT has had an

incredible impact on society. What changed?

In March, Microsoft researchers [published a paper](#) claiming GPT-4 showed "sparks of artificial general intelligence." AGI is the subject of a variety of competing definitions, but for the sake of simplicity can be understood as AI with human-level intelligence.

Some immediately interpreted the Microsoft research as saying GPT-4 *is* an AGI. By the definitions of AGI I'm familiar with, this is certainly not true. Nonetheless, it added to the hype and furore, and it was hard not to get caught up in the panic. Scientists are no more immune to [group think](#) than anyone else.

The same day that paper was submitted, The Future of Life Institute [published an open letter](#) calling for a six-month pause on training AI models more powerful than GPT-4, to allow everyone to take stock and plan ahead. Some of the AI luminaries who signed it expressed concern that AGI poses an [existential threat](#) to humans, and that ChatGPT is too close to AGI for comfort.

Soon after, prominent AI safety researcher Eliezer Yudkowsky—who has been commenting on the dangers of superintelligent AI [since well before](#) 2020—took things a step further. [He claimed](#) we were on a path to building a "superhumanly smart AI," in which case "the obvious thing that would happen" is "literally everyone on Earth will die." He even suggested countries need to be willing to risk nuclear war to enforce compliance with AI regulation across borders.

## **I don't consider AI an imminent existential threat**

One aspect of AI safety research is to address potential dangers AGI might present. It's a difficult topic to study because there is little agreement on what intelligence is and how it functions, let alone what a

superintelligence might entail. As such, researchers must rely as much on speculation and philosophical argument as evidence and mathematical proof.

There are two reasons I'm not concerned by ChatGPT and its [byproducts](#)

.

First, it isn't even close to the sort of artificial superintelligence that might conceivably pose a threat to humankind. The models underpinning it are slow learners that require immense volumes of data to construct anything akin to the versatile concepts humans can concoct from only a few examples. In this sense, it's not "intelligent."

Second, many of the more catastrophic AGI scenarios depend on premises I find implausible. For instance, there seems to be a prevailing (but unspoken) assumption that sufficient intelligence amounts to limitless real-world power. If this was true, more scientists would be billionaires.

Cognition, as we understand it in humans, takes place as part of a physical environment (which includes our bodies)—and this environment imposes limitations. The concept of AI as a "software mind" unconstrained by hardware has more in common with 17th-century [dualism](#) (the idea that the mind and body are separable) than with contemporary theories of the mind existing as [part of the physical world](#).

## Why the sudden concern?

Still, doomsaying is old hat, and the events of the last few years probably haven't helped. But there may be more to this story than meets the eye.

Among the prominent figures calling for AI regulation, many work for

or have ties to incumbent AI companies. This technology is useful, and there is money and power at stake—so fearmongering presents an opportunity.

Almost everything involved in building ChatGPT has been published in research anyone can access. OpenAI's competitors can (and have) replicated the process, and it won't be long before free and open-source alternatives flood the market.

This point was made clearly in a memo [purportedly leaked](#) from Google entitled "We have no moat, and neither does OpenAI." A moat is jargon for a way to secure your business against competitors.

Yann LeCun, who leads AI research at Meta, says these models should be open since they will become public infrastructure. He and many others are [unconvinced by the AGI doom](#) narrative.

Notably, [Meta wasn't invited](#) when US President Joe Biden recently met with the leadership of Google DeepMind and OpenAI. That's despite the fact that Meta is almost certainly a leader in AI research; it produced PyTorch, the machine-learning framework OpenAI used to make GPT-3.

At the White House meetings, OpenAI chief executive Sam Altman suggested the US government should issue licenses to those who are trusted to responsibly train AI models. Licenses, as Stability AI chief executive Emad Mostaque [puts it](#), "are a kinda moat."

Companies such as Google, OpenAI and Microsoft have everything to lose by allowing small, independent competitors to flourish. Bringing in licensing and regulation would help cement their position as market leaders, and hamstringing competition before it can emerge.

While regulation is appropriate in some circumstances, regulations that are rushed through will favor incumbents and suffocate small, [free and open-source competition](#).

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: No, AI probably won't kill us all—and there's more to this fear campaign than meets the eye (2023, June 1) retrieved 2 July 2024 from <https://techxplore.com/news/2023-06-ai-wont-alland-campaign-eye.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.