



## Artificial Intelligence (A.I.): A Threat to Civilization and Humanity?

### Description

*How worried should we be as a civilization about artificial intelligence, assuming we aspire to continue to exist?*

*I recently sat for a podcast with Nicolas Creed and The Daily Bell editor Joe Jarvis to discuss the existential threat, or lack thereof, posed by unchecked AI.*

Joe, playing devil's advocate, was bullish on AI as a net positive for humanity. Nicolas and I were less optimistic. We all agreed that the defining factor will be the manner in which it is developed — by whom, for what purposes, and with what precautions, if any.

We on Team Skeptic are now joined by a bevy of experienced AI professionals. One such figure, for instance, recently literally called for the bombing of AI data centers that provide the inputs for AI “cognition.”

Via Futurism:

“One of the world’s loudest artificial intelligence critics has issued a stark call to not only put a pause on AI but to militantly put an end to it — before it ends us instead...

Machine learning researcher Eliezer Yudkowsky, who has for more than two decades been warning about the dystopian future that will come when we achieve Artificial General Intelligence (AGI), is once again ringing the alarm bells.”

(For reference, “artificial general intelligence,” or AGI, is popularly defined as “the representation of generalized human cognitive abilities in software so that, faced with an unfamiliar task, the AGI system could find a solution. The intention of an AGI system is to perform any task that a human being is capable of.”)

Continuing:

“Yudkowsky said that while he lauds the signatories of the Future of Life Institute’s recent open

letter — which include SpaceX CEO Elon Musk, Apple co-founder Steve Wozniak, and onetime presidential candidate Andrew Yang — calling for a six-month pause on AI advancement to take stock, he himself didn't sign it because it doesn't go far enough."

The warning letter signed by Elon Musk and other notable public figures that Yudkowsky alludes to reads, in part:

"Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. **Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable.**"

Other AI heavyweights have echoed similar sentiments, including the "godfather of artificial intelligence," Geoffrey Hinton, who cited a "minor risk" AI would be humanity's undoing.

Returning to Yudkowsky's call to literal arms to stop AI's ascendance, he raises the essential problem, which I have raised elsewhere, of **creating an intelligence that outstrips humanity's cognitive limits**. Without effective guardrails in place to prevent it from becoming either negligent of human welfare or outright hostile to human life, we are at a serious disadvantage:

"It's not that you can't, in principle, survive creating something much smarter than you," he mused, "it's that it would require precision and preparation and new scientific insights, and probably not having AI systems composed of giant inscrutable arrays of fractional numbers."

Much like biomedical researchers using gain-of-function research to soup up viruses, **the AI engineers currently at work creating ever more intelligent AI know not what they do**. They are meddling, recklessly and needlessly, with forces they do not understand when the prudent course of action would require prior study.

As recently reported on elsewhere, AI very recently developed what philosophers and biologists call "theory of mind." This means that it has the newfound lifelike capacity to frame itself in the state of mind of another person or thing and then to act strategically accordingly.

It may not be prudent to cosign calls for kinetic bombing of information warehouses, but these developments certainly should serve as a cause for pause, to grapple with the wide-ranging implications of this technology.

**Ben Bartee** is an independent Bangkok-based American journalist with opposable thumbs.

*Featured image is from TDB*

## Category

1. Main
2. Science-Tech-AI-Medical & Gen. Research

## Date Created

04/18/2023