

The fight over a 'dangerous' ideology shaping AI debate

August 27 2023



Tesla's Elon Musk is among those claiming that AI could make humanity extinct, while standing to benefit by arguing only their products can save us.

Silicon Valley's favorite philosophy, longtermism, has helped to frame the debate on artificial intelligence around the idea of human extinction.

But increasingly vocal critics are warning that the philosophy is dangerous, and the obsession with extinction distracts from real problems associated with AI like data theft and biased algorithms.

Author Emile Torres, a former longtermist turned critic of the movement, told AFP that the philosophy rested on the kind of principles used in the past to justify mass murder and genocide.

Yet the movement and linked ideologies like transhumanism and effective altruism hold huge sway in universities from Oxford to Stanford and throughout the tech sector.

Venture capitalists like Peter Thiel and Marc Andreessen have invested in life-extension companies and other pet projects linked to the movement.

Elon Musk and OpenAI's Sam Altman have signed open letters warning that AI could make humanity extinct—though they stand to benefit by arguing only their products can save us.

Ultimately critics say this fringe movement is holding far too much influence over public debates over the future of humanity.

'Really dangerous'

Longtermists believe we are dutybound to try to produce the best outcomes for the greatest number of humans.

This is no different to many 19th century liberals, but longtermists have a much longer timeline in mind.

They look to the far future and see trillions upon trillions of humans floating through space, colonizing new worlds.

They argue that we owe the same duty to each of these future humans as we do to anyone alive today.

And because there are so many of them, they carry much more weight than today's specimens.

This kind of thinking makes the ideology "really dangerous", said Torres, author of "Human Extinction: A History of the Science and Ethics of Annihilation".

"Any time you have a utopian vision of the future marked by near infinite amounts of value, and you combine that with a sort of utilitarian mode of moral thinking where the ends can justify the means, it's going to be dangerous," said Torres.

If a superintelligent machine could be about to spring to life with the potential to destroy humanity, longtermists are bound to oppose it no matter the consequences.

When asked in March by a user of Twitter, the platform now known as X, how many people could die to stop this happening, longtermist ideologue Eliezer Yudkowsky replied that there only needed to be enough people "to form a viable reproductive population".

"So long as that's true, there's still a chance of reaching the stars someday," he wrote, though he later deleted the message.

Eugenics claims

Longtermism grew out of work done by Swedish philosopher Nick Bostrom in the 1990s and 2000s around existential risk and transhumanism—the idea that humans can be augmented by technology.

Academic Timnit Gebru has pointed out that transhumanism was linked to eugenics from the start.

British biologist Julian Huxley, who coined the term transhumanism, was also president of the British Eugenics Society in the 1950s and 1960s.

"Longtermism is eugenics under a different name," Gebru wrote on X last year.

Bostrom has long faced accusations of supporting eugenics after he listed as an existential risk "dysgenic pressures", essentially less-intelligent people procreating faster than their smarter peers.

The philosopher, who runs the Future of Humanity Institute at the University of Oxford, apologized in January after admitting he had written racist posts on an internet forum in the 1990s.

"Do I support eugenics? No, not as the term is commonly understood," he wrote in his apology, pointing out it had been used to justify "some of the most horrific atrocities of the last century".

'More sensational'

Despite these troubles, longtermists like Yudkowsky, a high school dropout known for writing Harry Potter fan-fiction and promoting polyamory, continue to be feted.

Altman has credited him with getting OpenAI funded and suggested in February he deserved a Nobel peace prize.

But Gebru, Torres and many others are trying to refocus on harms like theft of artists' work, bias and concentration of wealth in the hands of a few corporations.

Torres, who uses the pronoun they, said while there were true believers like Yudkowsky, much of the debate around extinction was motivated by profit.

"Talking about human extinction, about a genuine apocalyptic event in which everybody dies, is just so much more sensational and captivating than Kenyan workers getting paid \$1.32 an hour, or artists and writers being exploited," they said.

jxb/er-elc

© 2023 AFP

Citation: The fight over a 'dangerous' ideology shaping AI debate (2023, August 27) retrieved 3 May 2024 from <https://techxplore.com/news/2023-08-dangerous-ideology-ai-debate.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.