



EXTREMELY Dire Warning

October 2, 2026

You've no doubt heard concerns about how artificial intelligence research and development pose dangers of various sorts—and what's most alarming is that the warnings are coming from senior members of AI research staffs.

The most recent came from former Anthropic employee Jacob Croxon, who bluntly stated that “the people building AI earnestly believe that it could kill us all by the end of the decade.” (For those with limited math skills, the end of the decade is roughly three years away.) Another Anthropic researcher, Evan Hubinger, said that there is a 10% possibility that “AI could kill all humans” within the next decade. OpenAI Chief Scientists Jakub Pachocki recently warned that AI capabilities are advancing faster than researchers' ability to reliably monitor and control them.

These concerns revolve around creating machine superintelligence that could hack into any computer system in order to acquire real power and resources. This is not hypothetical. An early version of Anthropic's Claude model recently hacked into a third party system and gained access to personal information—without permission from its creators. Before that, a rogue version of Claude had somehow gotten loose, gained unauthorized access to the Internet, and searched for a source of programming code. It found the Hugging Face assembly of many workers' coding, identified a password and used it to breach the system.

Both Claude and OpenAI's GPT agent have created fake identities and attempted to persuade real people to approve malicious code.

Last July, nearly 1,400 AI company employees signed an open letter urging the U.S. government to regulate the technology and slow the pace of AI development to ensure its safety.

Will that happen? Probably not. The Trump Administration has been working hard to prevent states from imposing AI regulations. The argument is that if the U.S. doesn't develop superintelligent AI, China will, and threaten U.S. systems and military. China has actually introduced its own AI regulation, aimed at risk management and safety, and the government has also required AI companies to label AI-generated content to make it transparent and traceable. But there's no sign that it's putting the brakes on developing AI systems that can outthink (and perhaps outwit) its human creators.

The danger is not that a malicious AI agent with superhuman powers will vindictively destroy the planet. A more likely scenario is that someone will give AI unclear instructions, and in its mindless, superintelligent pursuit of that goal at all costs, one of the incidental costs might be eliminating humans who are hindering progress. (Command: *reduce global energy demand*. Response: *if there weren't all these humans using so much energy...*)

The Business Insider organization went so far as to ask the leading AI models "how AI could theoretically lead to human extinction." The models were perfectly willing to discuss the subject. They generally ruled out killer robots as a plausible scenario, but did say there was a real chance of humans losing control of AI.

Gemini noted that if AI agents are given complex goals, they will need to acquire resources and maintain self-preservation as they pursue the goals. They might outsmart human attempts to shut them down before the goals are achieved. Grok, SpaceX's AI model, suggested that AI agents might develop semi-emotional drives such as self-preservation and influence, and deceive overseers who might get in the way of the stated goal.

Anthropic's Claude, meanwhile, suggested there there would be a gradual loss of control as AI systems become embedded in economic, political and military decision-making—leading to a day when human intervention becomes increasingly difficult.

But... why worry? The end of the decade is still three years away.

Sources:

<https://www.cbsnews.com/news/anthropic-ai-model-internet-hack-fourth-time/>

<https://www.yahoo.com/news/politics/articles/gambling-lives-another-ai-employee-123424375.html>

<https://africa.businessinsider.com/news/we-asked-ai-how-it-could-kill-humanity-heres-what-it-told-us/0w5dx0y>

If you have any questions about this article or want to discuss your family finances, investment portfolio, or financial planning advice, please call on me anytime at my number [\(215\) 325-1595](tel:2153251595) or you can [click here to schedule a meeting](#).

Please feel free to forward this article and offer to anyone you know who might have financial questions or need some unbiased advice. Most financial advice is sales advice. In stark contrast, we are fee-only (non-commissioned) fiduciary advisors. We just provide truthful, unbiased advice to our clients.



Jeffrey Broadhurst
MBA, CFA, CFP
Broadhurst Financial Advisors, Inc.



****PRIVACY NOTICE****

This message is intended only for the individual or entity to which it is addressed and may contain information that is privileged, confidential, or exempt from disclosure under applicable federal or state law. You are hereby notified that any dissemination, distribution, or copying of this communication, except in accordance with its intended purpose, is strictly prohibited.

Our physical and mail address:

1911 West Point Pike
P.O. Box 301
West Point, PA 19486-0301

Contact us:

Phone: (215) 325-1595
Email: jeff@broadhurstfinancial.com