

# 'Terminator' tech could one day take over humanity, 'Godfather of AI' warns

Some experts have called on Congress to act before AI become unsafe

By [Michael Lee Fox News](#)

A British computer scientist who earned the nickname "the Godfather of AI" warned that the dangers of [artificial intelligence](#) made famous in films like "The Terminator" could become more reality than fiction.

"I think in five years' time, it may well be able to reason better than us," Geoffrey Hinton, a British computer scientist and cognitive psychologist, said during an interview with "60 Minutes," according to a report from Yahoo News.

Hinton, who became well known for his work on the framework for AI, urged caution in the continued development of [AI technology](#), questioning whether humans can fully understand the technology that is currently seeing rapid development.

"I think we're moving into a period when for the first time ever, we have things more intelligent than us," Hinton said.



Children interact with an AI robot at the China Science and Technology Museum in Beijing on Jan. 10, 2023. (Huang Yong/VCG via Getty Images)

Hinton argued that while humans develop the algorithm AI tools use to learn, they have little understanding of how that learning actually takes place. Once the concepts AI begins to learn get more complicated, Hinton said understanding what the technology is thinking is just as difficult as reading a human mind.

"We have a very good idea sort of roughly what it's doing," Hinton said. "But as soon as it gets really complicated, we don't actually know what's going on any more than we know what's going on in your brain."

The computer scientist warned that one way humans could lose control of AI is when the technology begins to write "[their own computer code](#) to modify themselves."

"That's something we need to seriously worry about," Hinton said.

That reality presents dangerous problems, Hinton argued, saying that he doesn't see a way for humans to guarantee that the technology continues to be safe.

### **PENTAGON LOOKING TO DEVELOP 'FLEET' OF AI DRONES, SYSTEMS TO COMBAT CHINA: REPORT**

"We're entering a period of great uncertainty where we're dealing with things we've never done before," he said. "And normally the first time you deal with something totally novel, you get it wrong. And we can't afford to get it wrong with these things."

Such uncertainty could even lead to dangers that just a few years ago were seemingly safely in the [realm of fiction](#), including an AI takeover of humanity. "I'm not saying it will happen. If we could stop them ever wanting to, that would be great. But it's not clear we can stop them ever wanting to," Hinton said.

Christopher Alexander, chief analytics officer of Pioneer Development Group, told Fox News Digital that he shares many of the same concerns as Hinton, including fears over what will happen to human workers who find themselves displaced by AI.

### **CHINA'S AI DOMINANCE SHOULD BE A WAKE-UP CALL FOR US ALL**

"The ability of AI to do the job of a routine worker is going to be the first shockwave, as there is a very real perspective of large numbers of people who are no longer employable," Alexander said, while noting that Hinton wasn't saying that AI would "rule humanity," but could gain capabilities that will "permanently alter" human civilization. "He is correct in noting human beings may become the second most powerful intelligence on the planet."

The danger means that [Congress should act](#) now with regulation, Jon Schweppe, policy director of American Principles Project, told Fox News Digital.

"One wonders: have any of these AI enthusiasts read a book?" Schweppe questioned. "The fears about the dangers of AI are absolutely justified. And just like the fictional characters in these stories, our tech titans appear to be filled with self-assured hubris, certain that nothing will go wrong. We can't afford to take that risk. Congress must enact AI safeguards to protect humanity."

### **WHAT IS ARTIFICIAL INTELLIGENCE (AI)?**

Phil Siegel, founder of the Center for Advanced Preparedness and Threat Response Simulation, agreed that some of Hinton's concerns are "warranted," though he argued that the main concern is not that the technology will "overwhelm humanity."

"I do not think these algorithms are sentient, so the worry they will on their own overwhelm humanity are not the main concern," Siegel said. "However, that may not matter much because bad actors, who are sentient, will absolutely be able to use these systems to enable them to damage humanity and at the least 'program' them to do bad things."



**Cutting-edge applications of artificial intelligence are seen on display at the Artificial Intelligence Pavilion of Zhangjiang Future Park during a state-organized media tour in Shanghai on June 18, 2021. (Andrea Verdelli/Getty Images)**

Regardless of where threats originate, Siegel believes that it is prudent for people to "prepare for unpredictable consequences of AI advancements," not only with regulation, but by using "the current tools in defense, preparing for potential threats by practicing against them, and developing new capabilities to raise the probability we will respond and protect ourselves well."

During his TV interview, Hinton said the "main message" he was hoping to get across is that there is still "enormous uncertainty" about the future of AI development.

"These things do understand, and because they understand we need to think hard about what's next, and we just don't know," Hinton said.

But Samuel Hammond, a senior economist at the Foundation for American Innovation, agreed that AI technology will be able to reason better than humans in the next five years and that there is great uncertainty about "what

happens next," though he pointed to a recent breakthrough in research that could suggest humanity's worst fears could [remain in theaters](#).

"A recent breakthrough in AI interpretability research suggests the risk of AI taking over the world or deceiving its users may be overblown," Hammond told Fox News Digital. "Mechanistic interpretability is neuroscience for the AI brain, only unlike the human brain, we can directly measure everything the AI brain is doing and run precise experiments."

Hammond noted that such research has demonstrated that humans will be able to "read the AI's mind," allowing humans to detect if it is lying and giving developers a chance to control potentially dangerous behaviors.

"The risks from bad actors and broader societal disruptions remain and have no easy solutions," Hammond said. "Institutional adaptation and coevolution may be the most important way to ensure AI leads to a better world, but unfortunately our [government is deeply resistant to reform](#)."

Retrieved December 29, 2023, from [‘Terminator’ tech could one day take over humanity, ‘Godfather of AI’ warns \(foxnews.com\)](#)