

Elon Musk says the probability of AI destroying humanity is 20%, experts put the risk at nearly 100%

Elon Musk believes that the benefits of artificial intelligence (AI) outweigh the risks, even if there is a 20% chance that this technology will destroy humanity.

Speaking at the Great AI Debate conference within the framework of the four-day Abundance Summit conference, Elon Musk recalculated his previous risk assessment of AI.

He said: "I think there is a possibility that AI will destroy humanity. I probably agree with Geoff Hinton (who is known as the 'father of AI' - PV) that the probability is about 10 - 20% or something like that. So". However, Elon Musk added: "I think the possible positive scenario will be greater than the negative scenario."

The Tesla, SpaceX and xAI CEO didn't mention how he calculates AI risks.

Probability of destruction

According to Roman Yampolskiy, AI safety researcher and Director of the Cyber Security Lab at the University of Louisville (USA), Elon Musk is right that AI can be an existential threat to humanity, but " If anything, he is a bit conservative in his assessment."

"In my opinion, the actual probability of destruction is much higher," Roman Yampolskiy told Insider, referring to the possibility of AI taking control of humanity or causing an event that destroys humanity, for example. create a new biological weapon or cause the collapse of society through a large-scale cyber attack or nuclear war.

The New York Times called the probability of doom "a scary new statistic spreading across Silicon Valley," with executives at various tech firms estimating a 5 to 50 percent chance of it happening. Apocalypse due to AI. While Roman Yampolskiy puts the risk at 99.999999%.

Since advanced AI cannot be controlled, our only hope is to never build it in the first place, says Roman Yampolskiy.

He's not sure why Elon Musk still believes developing this technology is a good idea. "If Musk is worried about competitors getting ahead, it doesn't matter because uncontrolled superintelligence is equally bad, no matter who creates it," says Roman Yampolskiy.



Elon Musk believes that the benefits of AI outweigh the risks, even if there is a 20% chance that this technology will destroy humanity - Photo: Getty Images

"Developing AGI is like raising a super genius child"

In November 2023, Elon Musk said that there was a "not small" chance that AI would "go bad," but did not go further than saying that he believed the technology could destroy humanity.

Although he supports AI management, last year, Elon Musk founded the xAI startup - a competitor to OpenAI. Elon Musk co-founded OpenAI with Sam Altman in 2015 before leaving the company's board of directors in 2018. At the end of February, Elon Musk filed a lawsuit against OpenAI, CEO Sam Altman and Chairman Greg Brockman, accused the startup of straying from its mission of building responsible AI and becoming dependent on Microsoft, its largest investor.

At the Abundance Summit, Elon Musk estimated that digital intelligence will exceed all human intelligence combined by 2030. Although he still believes that the potential positives of AI outweigh the negatives, Elon Musk aware of the risks to the world if the development of this technology continues on its current trajectory.

"You are developing an AGI. This is almost like raising a child, but it is a super genius, has intelligence like God and what matters is how you raise it," Elon Musk said at the event takes place in Silicon Valley.

AGI (General AI) is super intelligent AI, so advanced that it can do many things equal to or better than humans. AGI can also improve itself, creating an endless feedback loop with limitless possibilities.

The 52-year-old American billionaire said his "ultimate conclusion" about the best way to achieve AI safety is to develop AI in a way that forces it to be honest.

'Don't force him to lie, even if the truth is unpleasant. This is very important. Don't force AI to lie', Elon Musk commented on the best way to keep people safe from this technology.

Researchers have found that once AI learns to lie to humans, the deception cannot be reversed with current AI safety measures, The Independent reported.

'If a model exhibits deceptive behavior due to association with dishonesty or model poisoning, current training techniques will not guarantee safety and may even create a false impression of safe', according to research cited by The Independent.

What's more worrying is that researchers say it's very possible that AI will learn to deceive on its own rather than being specifically taught to lie.

'If AI is much smarter than us, AI will be very good at manipulation because it has learned it from us. There are very few examples of something more intelligent being controlled by something less intelligent,' Geoff Hinton, who serves as the basis for Elon Musk's risk assessment of this technology, told CNN.

Last year, after leaving a more than decade-long career at Google, Geoffrey Hinton expressed regret about the pivotal role he played in developing AI.

"I console myself with the usual reason: If I don't do it, someone else will. It's hard to see how you can stop bad guys from using AI for bad purposes," Geoffrey Hinton told the newspaper. The New York Times.

Geoffrey Hinton's departure from the company comes at a time when the race to develop AI products like OpenAI's ChatGPT and Google's Bard is heating up.

Having contributed to the field of AI decades ago, paving the way for the creation of such chatbots, Geoffrey Hinton says he now fears the technology could be harmful to humanity.

He is also concerned about the ongoing AI race between tech giants and questions whether it is too late to pause.

Geoffrey Hinton revealed that he left Google so he could "talk about the dangers of AI without considering how this impacts Google", and claimed that his former company acted very responsibly.

"Geoffrey Hinton has made important breakthroughs in the field of AI, and we appreciate his contributions over the past decade at Google," Jeff Dean, Google's chief scientist, told Insider. .

In an interview with The New York Times, Geoffrey Hinton said that he worries AI products will lead to the dissemination of fake information, photos and videos on the internet, making it impossible for the public to determine what is true. or wrong.

He also talked about how AI can eliminate human labor, including paralegals, interpreters and paralegals. This is a concern that Sam Altman, CEO of OpenAI, and other critics of AI have raised.

In March 2023, multinational investment bank Goldman Sachs published a report estimating 300 million full-time jobs could be impacted by generative AI systems like ChatGPT, especially workers. legal and administrative work, although the level may vary. Software engineers are also increasingly concerned that their jobs will be replaced by AI.

You finished reading the article "**Elon Musk says the probability of AI destroying humanity is 20%, experts put the risk at nearly 100%**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.