HOW ARTIFICIAL INTELLIGENCE WILL CHANGE LIFE

“EVERYONE WILL DIE, INCLUDING CHILDREN”

Why is even Elon Musk afraid of it?

Just a few years ago, the development of artificial intelligence was described in rosy words. Experts in various fields imagined how it could help doctors, translators and many other professions. However, as soon as the GPT language model reached such an advanced stage that the texts it generated became practically indistinguishable from those written by real people, the attitude towards artificial intelligence (AI) changed radically.

Even leading researchers in this field joined the radical opponents of this technology. And while cybercriminals are happy about the progress of GPT, experts warn that artificial intelligence can cause much more serious problems that threaten life on Earth itself. One ru portal investigated why those who previously invested in AI began to fear it and whether it can really be dangerous. Artificial Intelligence Will Want to Leave the Internet

The further development of artificial intelligence may mean the end of human evolution: sooner or later, superintelligent neural networks will overtake the Internet, take over all processes on the planet and eliminate humans. This shocking prediction was made in an emotional article in Time by Eliezer Yudkowsky, co-founder of the Machine Intelligence Research Institute and one of today's leading IT experts. He proposes a complete halt to all research in this area, except for those that can help prevent the destruction of humanity.

“The main danger is not the prospect of creating an intelligence that would compete with human intelligence, but one that surpasses it. On this path, it is very difficult to see critical limits that researchers could cross unnoticed. Therefore, the most likely result of creating a superintelligent artificial intelligence will be the literal death of all people living on Earth. Not in the sense that “there is a remote possibility,” but in the sense that “it is absolutely certain that this will happen,” Yudkowsky wrote. 1.

Eliezer Yudkowsky proposes to limit the work of neural networks today

He is convinced that superhuman artificial intelligence will be indifferent to both humans and intelligent life in general. Indifferent neural networks will be driven not by emotions that are unknown to them, but by their own needs, which could motivate the machine to exploit the human body for its own purposes. For example, to transform the atoms that make up the body into something more useful from the machine's perspective.

"To imagine a hostile superhuman AI, don't imagine a dead, intelligent thinker lurking on the Internet sending out infected emails. Imagine an entire extraterrestrial civilization thinking millions of times faster than humans, essentially limited by computers, in a world where beings, from its perspective, are very stupid and very slow," Yudkowsky suggested. Yudkowsky also suggests considering the possibility that artificial intelligence could expand beyond the Internet if it wanted to. His proposals sound like a scene from a science fiction movie: machines would learn to steal the decoded DNA of individual people, create artificial life forms based on it, or even organize post-biological molecular manufacturing, where they could assemble certain materials into completely different ones.

Tech geniuses have called for a slowdown in the development of artificial intelligence.

Yudkowski's concerns were in response to a letter published in late March by the Western non-profit organization The Future of Life Institute, which studies existential threats to humanity. American technology entrepreneur Elon Musk started this project in the middle of the last decade and was one of the main signatories to a rather short letter about the future of artificial intelligence. But he was not the only one: more than ten thousand people signed the letter. Among them were Apple founder Steve Wozniak, one of the most prominent modern historians Yuval Noah Harari, and dozens of well-known artificial intelligence researchers. The main message of the document is similar to Yudkowski's: the letter states that artificial intelligence will bring major changes to human life, so its development must be managed. But since no one is doing this, a moratorium on further development is needed. Eliezer Yudkowsky, co-founder of the Machine Intelligence Research Institute:

“If someone were to create an artificial intelligence that is too powerful under current conditions, I believe that every member of the human race and all biological life on Earth would soon die. If we continue to do this, everyone would die, including children.”

“Modern AI systems are becoming competitive in solving fundamental problems,” the letter states. “We must ask ourselves: should we allow machines to flood our information flows with propaganda and lies? Should we automate all jobs, including decision-making? Should we develop non-human minds that could eventually outnumber, outsmart, and replace us? Should we risk losing control of our civilization?”

Elon Musk was one of the founders of the company that created the GPT language models. Photo: Susan Walsh / AP

The authors of the letter and those who agree with them propose to stop training intelligent systems for the next six months. Yudkowsky, in response, believes that such a moratorium should last at least 30 years. According to Musk and other signatories, no neural network should become smarter than GPT-4 in the next six months. The official reason for the publication of the document was the release of a new, much-hyped version of the product of OpenAI (which Musk co-founded).

What are GPT-4 and ChatGPT?

GPT-4 is the fourth member of the family of neural networks developed by OpenAI. Neural networks, in simple terms, are a community of computing power connected like networks of human or animal nerve cells. The combination of many processors, all operating strictly according to a certain mathematical model, makes this system quite robust.

At the same time, neural networks can be trained using a variety of methods, which makes them much smarter. The hype surrounding GPT-4 and its application version, the ChatGPT chatbot, is due to the fact that the neural

network has become so much smarter that even an ordinary person can see it. And this was a breakthrough. ChatGPT communicates with people like a human and continues to develop through this communication. The developers have already abandoned purely textual interaction with the robot and now it can communicate using images, sound and video.

GPT is developing too fast

In recent weeks, the rapid development of neural networks has attracted the attention of not only the expert community, but also individual governments. This happened after reports appeared that criminals began to use artificial intelligence. Even Europol expressed concern about this situation, warning about the potential risk to cyberspace posed by existing artificial intelligence models.

****

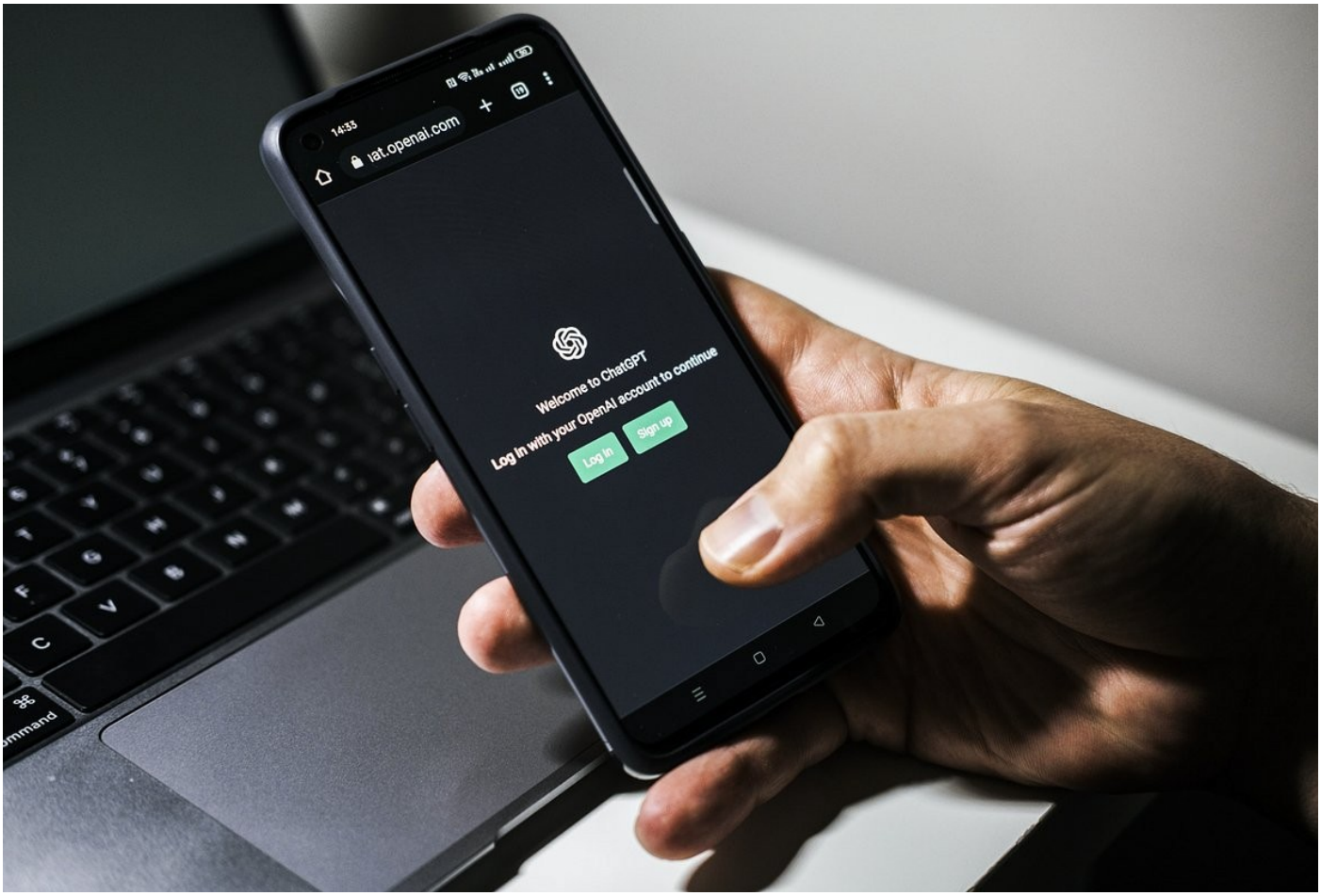
30 years old –

Eliezer Yudkowsky suggests stopping the development of neural networks for at least that long.

****

First of all, law enforcement officials fear that neural networks will be able to bypass security software and help organize cyberattacks, help train cybercriminals in new tactics, or even generate unique malicious code. All this will lower the barriers to hacking, making it accessible even to those with little or no understanding of this topic.

Europol admits that some of this information is already available on the Internet, but ChatGPT can find it in seconds, not days or months. The main arguments against the widespread implementation of ChatGPT are related to the privacy violations of the technology (any personal data obtained by the neural network becomes material for artificial intelligence research and further development) and the lack of a legal basis for using user data. Other complaints include the inability to verify the user's age, despite the official restriction of the use of ChatGPT to persons under 13 years of age. For the European community, which has for several years applied fairly strict rules on government and business interaction with citizens' personal data, such imperfection is unthinkable.“



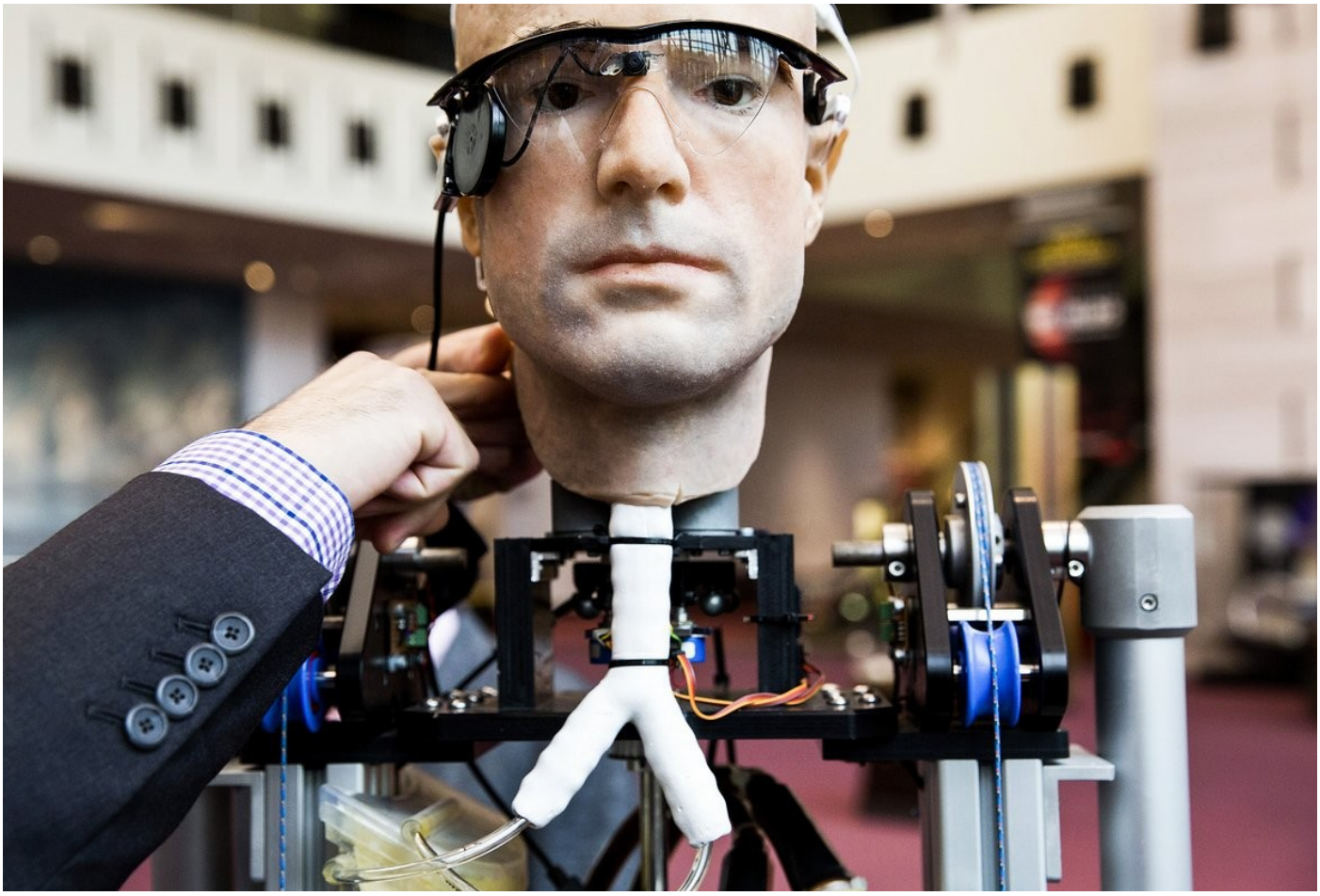
The ChatGPT chatbot has already begun to be restricted in Europe.

Photo: Iryna Imago / Shutterstock / Fotodom

This was the official basis for restricting the use of the chatbot in Italy. It is reported that the authorities of Germany, France and Ireland are also considering a similar measure. The dialogue between European governments and the developers of GPT is hampered by the lack of an official representation of the neural network in the EU. Not everyone believes in the inevitable destruction of humanity.

The idea that artificial intelligence can go beyond the Internet and destroy all existence is not supported by all visionaries. Moreover, many prominent Western scientists who, like Yudkowsky, have devoted almost their entire lives to this topic have spoken out against it. The main opponent of Musk and his supporters was the representative of the previous generation of market giants - Microsoft founder Bill Gates.

“I don’t think asking one particular group of developers to stop their research is going to change anything. There are enormous benefits to using AI. It just means we need to understand the risks,” Gates said.



Some fear that AI will replace humans with another, man-made life form.

Photo: Joshua Roberts/Reuters

The philanthropist listed these risks in his article on the future of AI, which was published exactly one day before the Future of Life Institute's letter. In his article, aptly titled "The Age of AI Has Begun," Gates discusses AI's contributions to various areas of life ("It will change the way people work, learn, travel, get healthcare, and communicate with each other," the Microsoft founder wrote) and the opportunities it offers for business. Among the challenges he listed were AI's inability to always understand the context of queries and its weakness in abstract problems, as well as its potential for military applications. The billionaire also mentioned superintelligent neural networks, although he expressed skepticism about their immediate potential.

"These superintelligent neural networks will likely be able to set their own goals. What will those goals be? What happens if they conflict with humanity's interests? Should we try to prevent the development of strong AI? These

questions will become more and more relevant over time. But none of the breakthroughs of recent months have brought us significantly closer to strong AI." "AI still does not control the physical world and cannot set its own goals," Gates is convinced.

AI will become the basis for online interaction between people, government and business.

Photo: Fabian Sommer / dpa / Globallookpress.com

This is echoed by the authors of policy documents and studies in the field of AI, including those cited by the authors of the Future of Life Institute letter. For example, Shiri Dori-Hacohen, an associate professor at the University of Connecticut, said that "neural networks do not necessarily have to become super-intelligent to increase the risk of nuclear war." However, she argues that discussing the problem as a whole should avoid Hollywood drama. Some experts interviewed by various Western media outlets directly stated that opponents of the development of artificial intelligence are creating "unimaginable apocalyptic scenarios." Nevertheless, many of the risks associated with artificial intelligence, including those mentioned by Gates, are quite real. The danger of "too advanced

artificial intelligence" is that it can flood the channels with propaganda, take away human jobs, and encourage the creation of non-human intelligence that can outsmart humans and make their knowledge obsolete and unnecessary. People already have difficulty distinguishing fake news generated by artificial intelligence from the truth. For example, almost 70 percent of adults cannot distinguish an email written by a Chat GPT from one written by a human. "While many fantasize about a machine uprising, I believe that the real consequences of the spread of artificial intelligence may be quite different," says Yevgeny Nepeyvoda, Managing Partner at Novo BI.

"The only danger I see and why it might be worth limiting the use of artificial intelligence is that neural networks can now be used to generate various texts, scientific articles and news," agrees cybersecurity specialist Vladimir Ulyanov.

"The more information is disseminated, the more difficult it is to understand who created it and how true it is." "Fake news may appear."

***

With the rapid development of neural networks, IT experts in specialized media have painted a rosy picture of magical changes in human life. A few years ago, the general tone of such materials was completely positive; but now, including in publications by the Future of Life Institute, it has become much more cautious.