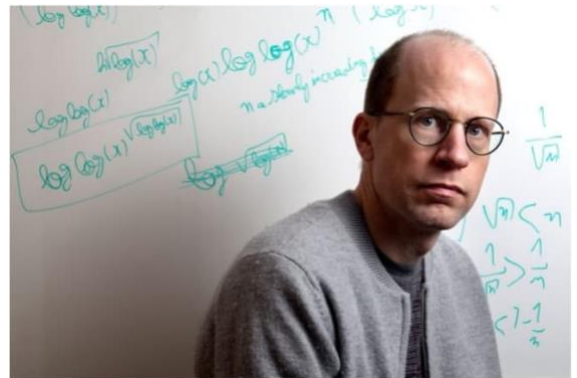# Toward the Point of No Return

Why Everything Has Changed So Much

With the emergence of ChatGPT at the close of 2022, AI transitioned from the realm of science fiction to a rapidly approaching reality. The consensus among AI specialists is now firm-the creation of machines capable of human-level thinking, and even surpassing it, is not a matter of if, but when.

With the emergence of ChatGPT at the close of 2022, AI transitioned from the realm of science fiction to a rapidly approaching reality. The consensus among AI specialists is now firm-the creation of machines capable of human-level thinking, and even surpassing it, is not a matter of if, but when.

AI that has reached such a level of development is called Superintelligence. This term was popularized by Swedish philosopher N. Bostrom in his 2014 book **Superintelligence: Paths, Dangers, Strategies**. He and many other thinkers believe that Superintelligence will be capable of recursive self-improvement, and this process will be impossible to control externally.

It is still unknown when AI of this kind will be created, but the speed at which this direction of scientific and technological progress is developing suggests that it will take years, not decades.

What is the problem here?

The problem here is that we may create an entity more powerful than ourselves and with goals different from ours. This means that the existence of humanity may depend on its intentions. If these intentions are friendly, Superintelligence could become our partner. This would allow us to solve the fundamental problems that humanity has been grappling with for millennia. Perhaps the very form of our existence will change, turning into a symbiosis of human and artificial minds. That will open up prospects for us to understand the universe at a level of comprehension that we cannot even imagine right now. Although it's impossible to predict what this future will be like, our main hope here is that it will be the subject of our desire and choice.

In the worst-case scenario, we will disappear from the historical scene as a species. We may become unnecessary to this super-powerful entity. It may happen that we, in accordance with Elon Musk's concern, will simply end up as a "biological bootloader for Superintelligence." Thus, the latter may turn out to be what journalist and writer J. Barrat called "Our Final Invention" in his 2013 book (of the same name)—an invention that was created in a *bad* sense.

Why We Won't Be Able to Cancel the Creation of Superintelligence

The fundamental aspect of this problem is that we won't be able to refuse to create this entity. Slowing down this process is also very difficult. There are several objective reasons for this. We analyze them in detail in the section Why We Won't Refuse Creating Superintelligence. Here, we'll mention the most obvious ones:

1. Market Demand for AI Products

The demand for AI means an AI race. Regulating this process is extremely difficult because public awareness of the problem lags behind the dynamics of its development. This makes it difficult to make effective political decisions in this area.

2. Business Priorities of AI Developers

AI development companies are primarily interested in making a profit, not in ensuring the safety of their products. This prompts them to sacrifice the latter for the former. Even if they are aware of the possible consequences of such an approach, it's difficult for them to make balanced decisions due to competitive pressure.

3. Geopolitical Rivalry of States Capable of Creating Superintelligence

A regime intolerant of its ideological opponents may be tempted to use AI to eliminate them. The opposing side, in turn, will take retaliatory measures. Thus, the AI race will be escalated to an international level. This will not only complicate the solution of the safe AI problem but will significantly complicate it.

Formal agreement between parties not to use AI as a weapon may prove ineffective. Its observance requires transparency not only of the parties' intentions but also of the state of their developments. This is difficult to control technically and institutionally due to the clash of interests of various parties within societies themselves.

4. Difficulties in Controlling Illegal AI Development

Finally, private AI developers, including malicious ones, can avoid any external control. They don't need any special infrastructure, organization of ultra-complex production, or logistical processes. For the most part, they need access to knowledge and relevant

services, which in modern conditions is quite feasible with even relatively modest funding. Obviously, for such developers, safety will not be a priority. The results of their activities may have unintended consequences, catastrophic for all of humanity.

Approaching the Point of No Return

So, we have reason to believe that we are approaching an event that will be a point of no return in human history. The creation of Superintelligence will cause what British mathematician I.J. Good[1] defined as an "intelligence explosion"[2] back in the mid-1960s. This term refers to a hypothetical scenario in which an AI, once it reaches a certain level of intelligence, can rapidly improve its intelligence, leading to an exponential increase in its cognitive abilities.

At that time, it looked like an indefinitely distant or completely unrealistic prospect. But now it's not so. The famous inventor, futurist writer, Principal Researcher, and AI Visionary at Google, R. Kurzweil, in his book *The Singularity Is Near* (2005), predicted the creation of Superintelligence around 2045, and his forecast is now far from the most radical. If this indeed happens, then for the first time in Earth's history, there will be more than one intelligent species on it, and one of them will be far smarter than the other.

The unpleasant assumption from this is that we have no guarantee of survival. Our planning horizon regarding Superintelligence is objectively limited. In the equation of reality with its presence, there are many more unknown variables than those we can operate with. So far, we don't have reliable approaches to predict its intentions and protect ourselves from those that may pose a threat to us. Thus, we must realize that if we fail to develop such approaches, everything could go according to the worst-case scenario for us.

How We Should Respond to the Challenge of Superintelligence?

The good news is that humanity has many brilliant minds ready to work on this issue, vast knowledge, and experience in successfully solving incredibly complex problems.

---

[1] **Irving John Good** (1916-2009), a British mathematician, statistician, and cryptographer, played a significant role in the development of ideas related to artificial intelligence (AI). His contributions to this field are impressive, especially considering his pioneering work in the mid-20th century.

[2] In his article *Speculations Concerning the First Ultraintelligent Machine*, he described the hypothesis that the creation of an 'ultra-intelligent machine' could initiate explosive growth in intelligence, surpassing human capabilities and leading to rapid and uncontrolled technological development.

Of course, the problem of Superintelligence is extraordinary in all senses, but this doesn't mean it's unsolvable in principle. Besides, our mind has a saving feature of maximum mobilization of its resources in situations of existential challenge.

Perhaps overcoming the challenge behind the Superintelligence problem depends most on our ability to realize its relevance.

Therefore, we urge everyone concerned about our shared future to engage with the information presented on this site. Your thoughtful consideration and contribution, no matter how small, can make a significant difference in addressing humanity's most pressing challenge.

So, we will be happy to help you:

- Assess the chances that Superintelligence will become the Great Filter for Humanity;

- Find out why, despite this danger, we won't refuse to create it;

- Understand the details of the fundamental risks coming from AI;

- Imagine Superintelligence not as a threat to our existence but as our partner and friend;

In addition to this information, you will find sections devoted to the vision of the AI problem through the prism of modern art [1] [2]. Finally, we hope to learn your opinion on the most pressing issues related to this problem.

Good luck in exploring the challenge of Superintelligence!

Online version: https://super-ai-challenge.vercel.app/

Author: Sergei Klevtsov, srgg67@gmail.com