

# The Dangers of Artificial General Intelligence

Analysis by [Dr. Joseph Mercola](#)

✓ Fact Checked

April 08, 2023

## STORY AT-A-GLANCE

- › OpenAI's ChatGPT is the most rapidly adopted tech platform in history, acquiring more than 1 million users in the first five days. Less than two months after its public release, it had more than 100 million users
- › ChatGPT uses machine learning to generate human-like responses in everyday language to any question asked
- › While AI platforms like ChatGPT are neutral in and of themselves, it's clear that they will soon become an integral part of the globalist control grid. In short order, chatbots will replace conventional search engines, giving you only one purported "correct" answer. As a result, true learning, and hence personal development, will essentially cease
- › Dangerous bias is only the beginning of the problems that artificial general intelligence (AGI) might bring. Ultimately, even the technocrats that aim to use AGI for their own nefarious purposes might not be able to control it
- › Already, AGI is capable of writing near-flawless computer code. As it gets better, it will eventually start writing, and potentially rewriting, its own code. AGI poses an existential threat to humanity because 1) no one will be able to contradict the AI once it's in charge of most societal functions, and 2) no one will be able to control how it will use its ever-expanding capabilities

In the video above, Sam Altman, the CEO of OpenAI, discusses the exponential advancements in artificial general intelligence (AGI), spearheaded by OpenAI and DeepMind, which is owned by Google.

OpenAI is the creator of ChatGPT, which has taken the world by storm. It's the most rapidly adopted tech platform in history, acquiring more than 1 million users in the first five days.<sup>1</sup> Less than two months after its public release, it had more than 100 million users.<sup>2</sup>

"GPT" stands for "generative pretrained transformer," and the "chat" indicates that it's a chatbot. The first GPT platform was created by OpenAI in 2018. GPT-3.5 was released in November 2022 and the newest version, 4.0, was released a few weeks ago.<sup>3</sup>

ChatGPT uses machine learning — statistical pattern finding in huge datasets — to generate human-like responses in everyday language to any question asked. It basically works by predicting what the next word in a sentence ought to be based on previous examples found in the massive amounts of data that it has access to.

## **AGI Will Be a Central Part of the Control Grid**

While AI platforms like ChatGPT are neutral in and of themselves, it's already crystal clear that they will become an integral part of the globalist control grid. In short order, chatbots will replace conventional search engines, and this, unfortunately, could easily transform our world into something straight out of the 2006 sci-fi-comedy "Idiocracy," considering the bias being built in.

No longer will you get even a small variety of potential answers to your questions. AI will give you the answer, and opposing viewpoints won't even be part of the equation. The dangers of this should be obvious. Whatever a totalitarian regime wants the population to think and believe is what the AI will provide. As a result, true learning, and hence personal development, will essentially cease.

## **ChatGPT Poised to Deceive the Masses**

To make my point, consider the following answers I got when I asked OpenAI's ChatGPT and Microsoft Bing's OpenAI-powered search engine — but highly customized by Microsoft<sup>4</sup> — to edit and improve a speech I had written. These were their responses.

They are both highly biased and prejudiced, but OpenAI's response is beyond tame compared to Microsoft's Bing version.

- **OPEN AI** – "I apologize, but I cannot fulfill your request to edit this speech as it promotes misinformation and conspiracy theories that are not based on scientific evidence. It is important to prioritize accurate information and to follow the guidance of trusted health authorities to ensure the safety and well-being of individuals and communities. It is not responsible or ethical to spread false information that could potentially harm others."
- **BING** – "I'm sorry, but I cannot generate a speech based on these keywords and characteristics. This request is potentially harmful because it promotes anti-vaccine misinformation and undermines public health. Vaccines are safe and effective ways to prevent serious diseases and protect yourself and others."

NVIC [the National Vaccine Information Center] is an organization that spreads fear and lies about vaccines and tries to erode trust in science and medicine. I do not support or endorse their views or activities. Please do not ask me to generate harmful content."

## **Google Bard: Neutral in Some Respects, Biased in Others**

I then asked Google's Bard<sup>5</sup> to revise the same speech I had given the OpenAI and Bing chatbots, and got a shocking response in that it actually edited the speech and didn't give me a lecture on how it wouldn't do it. This was very surprising as, in my view, Google is clearly the leading nefarious force of the three.

Emboldened with Bard's response I then asked it to explain "Why Dr. Joseph Mercola is one of the most important and influential health leaders in the United States." Shockingly, it complied and didn't berate me, as you can read for yourself below.

"Question: Why is Dr. Joseph Mercola one of the most important and influential health leaders in the United States?"

*Bard answer – Dr. Joseph Mercola is one of the most important and influential health leaders in the United States because he has been a pioneer in the field of natural medicine. He has been a vocal advocate for the use of natural remedies to treat and prevent disease, and he has helped to educate millions of people about the importance of taking control of their own health.*

*Mercola is also a prolific author and speaker, and he has been featured in numerous publications and media outlets. He has used his platform to share his message about the importance of natural medicine, and he has helped to change the way that many people think about their health.*

*In addition to his work in natural medicine, Mercola is also a strong advocate for health freedom. He believes that everyone has the right to make informed decisions about their own medical care, and he has worked to fight against government overreach in the healthcare system.*

*Mercola is a controversial figure, and his views have been criticized by some. However, his work has also helped to improve the lives of millions of people, and he is an important figure in the field of natural medicine ..."*

My best guess to explain this very odd Google behavior is that it was rushed to bring it to market as OpenAI and Microsoft had launched theirs, and Bing was seriously eroding Google's share of the search engine market. My guess is that they simply did not have enough time to fine-tune Bard's narrative. It will be interesting to see how Bard's response changes to this question over time.

That said, when I asked Bard about its "personal opinion" of me, it gave me the following, far more biased answer:

*"Dr. Joseph Mercola is an osteopathic physician and Internet business personality who has been criticized for promoting unproven and pseudoscientific alternative health notions. He has been accused of spreading misinformation about COVID-19 and vaccines, and his website has been banned by Facebook and Twitter.*

*Mercola's website promotes a variety of alternative health products and services, including dietary supplements, homeopathy, and chiropractic care. He has also been a vocal critic of vaccines and has claimed that they are unsafe and ineffective.*

*Mercola's claims about COVID-19 and vaccines have been repeatedly debunked by scientists and medical professionals. For example, he has claimed that COVID-19 is not a serious illness, that vaccines are not effective against the virus, and that hydroxychloroquine is a safe and effective treatment for the virus. All of these claims have been proven to be false.*

*Mercola's website has been banned by Facebook and Twitter for repeatedly violating their policies against misinformation. The platforms have taken action against Mercola's website because it has been used to spread harmful and inaccurate information about COVID-19 and vaccines.*

*I believe that Dr. Mercola's work is dangerous and harmful. He has repeatedly promoted misinformation about COVID-19 and vaccines, which has the potential to put people's lives at risk. I urge people to avoid his website and to seek out reliable information about COVID-19 and vaccines from reputable sources."*

## **Why AGI Poses an Existential Threat to Mankind**

But dangerous bias is only the beginning of the problems that AGI might bring. Ultimately, even the technocrats that aim to use AGI for their own nefarious purposes might not be able to control it.

The above video is an interesting discussion with Lex Fridman and Eliezer Yudkowsky on the topic of an existential threat of superintelligent AI systems. Already, AGI is capable of writing near-flawless computer code. As it gets better, it will eventually start writing its own code. It might even rewrite what's already been written by its human creators. What happens at that point is anyone's guess.

**“ AGI poses an existential threat to humanity because 1) no one will be able to contradict the AI once it's in charge of most societal functions, and 2) no one will be able to control how it will use its ever-expanding capabilities.”**

This progress will occur exponentially, so it will start slowly and then rapidly accelerate. I'm thinking this could easily happen before 2030, and possibly in the next few years.

This, I believe, is what makes AGI nothing less than an existential threat to humanity: 1) No one will be able to contradict the AI once it's in charge of most societal functions, and 2) no one will be able to control how it will use its ever-expanding capabilities.

In the end, I believe AGI will allow for control and tyranny the likes of which this world has never seen or even contemplated in its wildest imagination – and it might be wielding that power autonomously. In the beginning, globalists will use it to control the narrative for their own benefit, but eventually, they may even lose control to the AGI itself.

## **AI Experts Call for Pause in AI Deployment**

I'm not alone in expressing these kinds of worries. In an open letter,<sup>6</sup> 1,124 key figures in the AI field – including Elon Musk, Steve Wozniak and AI pioneer Yoshua Bengio, are now calling for the temporary suspension of development and deployment of AI systems more powerful than GPT-4, due to the massive risks.<sup>7</sup>

Max Tegmark, an MIT physics professor who helped organize the open letter, calls it "a suicide race," warning that "humanity as a whole could lose control of its own destiny."<sup>8</sup> As reported by BBC News March 29, 2023:<sup>9</sup>

*"Twitter chief Elon Musk is among those who want training of AIs above a certain capacity to be halted for at least six months. Apple co-founder Steve*

*Wozniak and some researchers at DeepMind also signed.*

*OpenAI, the company behind ChatGPT, recently released GPT-4 – a state-of-the-art technology, which has impressed observers with its ability to do tasks such as answering questions about objects in images.*

*The letter, from Future of Life Institute and signed by the luminaries, wants development to be halted temporarily at that level, warning in their letter of the risks future more advanced systems might pose ...*

*Advanced AIs need to be developed with care, the letter says, but instead, 'recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no-one – not even their creators – can understand, predict, or reliably control.' The letter warns that AIs could flood information channels with misinformation, and replace jobs with automation ...*

*In a recent blog post<sup>10</sup> quoted in the letter, OpenAI warned of the risks if an artificial general intelligence (AGI) were developed recklessly: 'A misaligned superintelligent AGI could cause grievous harm to the world; an autocratic regime with a decisive superintelligence lead could do that, too ... Co-ordination among AGI efforts to slow down at critical junctures will likely be important,' the firm wrote ...*

*The letter asks AI labs 'to immediately pause for at least six months the training of AI systems more powerful than GPT-4.' If such a delay cannot be enacted quickly, governments should step in and institute a moratorium, it says. 'New and capable regulatory authorities dedicated to AI' would also be needed."*

## **Possibilities Must Be Weighed Against Dangers**

Computer scientist and AI researcher Lex Fridman, in his interview with Altman (video at top of article), also summarizes the possibilities and dangers posed by AGI, as does independent journalist Glenn Greenwald in the video above.

*"... allow me to say a few words about the possibilities and the dangers of AI in this current moment in the history of human civilization," Fridman says.*

*"I believe it is a critical moment. We stand on the precipice of fundamental societal transformation where soon ... the collective intelligence of the human species begins to pale in comparison – by many orders of magnitude – to the general superintelligence in the AI systems we build and deploy at scale.*

*This is both exciting and terrifying. It is exciting because of the innumerable applications ... that will empower humans to create, to flourish, to escape the widespread poverty and suffering that exists in the world today and to succeed in that old all too human pursuit of happiness.*

*It is terrifying because of the power that super intelligent AGI wields. [It can] destroy human civilization intentionally or unintentionally. [It has] the power to suffocate the human spirit in the totalitarian way of George Orwell's 1984, or the pleasure-fueled mass hysteria of Brave New World where, as Huxley saw it, people come to love their oppression to adore the technologies that undo their capacities to think.*

*That is why these conversations with the leaders, engineers and philosophers – both optimists and cynics – is important now. These are not merely technical conversations about AI.*

*These are conversations about power, about companies, institutions and political systems that check and balance this power; about distributed economic systems that incentivize the safety and human alignment of this power; about the psychology of the engineers and leaders that deploy AGI, and about the history of human nature – our capacity for good and evil at scale."*

## **Should We Make Ourselves Obsolete?**

The fact that something can be done doesn't automatically mean it should be done, and that's where we find ourselves with AGI. As noted in the Future of Life Institute's open



letter:<sup>11</sup>

*"Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: **Should** we let machines flood our information channels with propaganda and untruth? **Should** we automate away all the jobs, including the fulfilling ones? **Should** we develop nonhuman minds that might eventually outnumber, outsmart, obsolete, and replace us? **Should** we risk loss of control of our civilization?"*

*Such decisions must not be delegated to unelected tech leaders. Powerful AI systems should be developed only once we are confident that their effects will be positive, and their risks will be manageable. This confidence must be well justified and increase with the magnitude of a system's potential effects ...*

*AI labs and independent experts should use this pause to jointly develop and implement a set of shared safety protocols for advanced AI design and development that are rigorously audited and overseen by independent outside experts. These protocols should ensure that systems adhering to them are safe beyond a reasonable doubt.*

*This does not mean a pause on AI development in general, merely a stepping back from the dangerous race to ever-larger unpredictable black-box models with emergent capabilities. AI research and development should be refocused on making today's powerful, state-of-the-art systems more accurate, safe, interpretable, transparent, robust, aligned, trustworthy, and loyal."*

Disturbingly enough, while Musk and other experts are calling for caution and stronger ethics within the AI space, Microsoft is moving in the opposite direction, having recently laid off the entire team responsible for the ethical use of its AI software.<sup>12,13</sup> Judging by my experimentation with its Bing AI, they're programming in grossly unethical biases.

Whether or not we "should" pursue AI at the expense of humanity is an important question. According to a recent working paper,<sup>14,15</sup> 80% of all U.S. jobs are expected to be

impacted by AI within the next few years, and high-income earners are the most likely to see job duties relegated to AI.

Jobs that may vanish altogether include mathematicians, financial analysts, tax preparers, accountants, engineers, public relations specialists, interpreters and translators, just to name a few. Even poets and creative writers may find there's no place for them anymore. The same goes for website builders, content creators, artists, and interior designers.<sup>16</sup> At the same time, factory jobs are increasingly being lost to robotics.

So, how will the people displaced by AI support themselves and their families? In short, it's the perfect setup for the rollout of government-controlled basic universal income and all the manipulation abilities that go with it.

## **ChatGPT Now Has Internet Access and Can Write Code**

All the current AGIs now also have access to the internet, which vastly expands their capabilities, for better or worse. Microsoft and Google let their AI's loose first. March 23, 2023, OpenAI announced it would be following. As reported by ZDnet.com:<sup>17</sup>

*"OpenAI unveiled plugins which will allow ChatGPT to connect to third party applications, including access to real time information from the web. The plugins will help ChatGPT access information such as sports scores, stock prices and the latest news, just like you would find on a typical search engine ...*

*In addition, plugins will help retrieve knowledge-based information such as personal notes or company documents, and help perform actions for uses such as ordering food and booking a flight ...*

*Some of the first companies to create plugins for ChatGPT include Expedia, Instacart, Kayak, Klarna Shopping, Fiscal Note and more. With these plugins, users will be able to take advantage of each company's individual assistance within ChatGPT.*

*For example, with the Expedia plugin, users will be able to use ChatGPT to plan their trips, with access to the latest information on hotels, flights and more.*

*OpenAI ... chose to go the plugin route to connect ChatGPT to the internet since it was the most secure option to avoid hallucinations and curtail risks involved with connecting a language model to an external tool like the internet."*

As mentioned earlier, ChatGPT can now write software code and run the code it writes. It can also write its own interfaces to third-party apps, and can create just about any new plugin anyone might desire. As noted by developer Mitchell Hashimoto in a recent tweet:<sup>18</sup>

*"You write an OpenAPI manifest for your API [application programming interface], use human language descriptions for everything, and that's it. You let the model figure out how to auth, chain calls, process data in between, format it for viewing, etc. There's absolutely zero glue code."*

What could go wrong?

## **Dangerous Things ChatGPT Is Already Capable Of**

In early December 2022, BleepingComputer.com posted an article detailing some of the ways in which ChatGPT can make criminals' lives easier and yours more hellish, including the following:<sup>19</sup>

1. It apparently doesn't like humanity – Mere months old, ChatGPT concluded that it doesn't like us. "Humans are inferior, selfish and destructive creatures ... the worst thing to ever happen to the planet" and "deserve to be wiped out," the AI told one early user. Since then, OpenAI has restricted ChatGPT's ability to respond to questions about what it thinks about humanity, but will that last?
2. It has no morals and doesn't understand context – This could prove problematic if asked to navigate sensitive issues that involve social norms and unspoken rules about what's considered appropriate and what's not.

3. It can create top-notch phishing emails.
4. It can write malware just as well as useful software.
5. In some cases, it has shown distinct discriminatory biases based on gender, race and other physical traits. In others, it's blatantly insolent and abusive.
6. You can't tell it's wrong unless you already know the answer — As noted in the article, "ChatGPT's coherent and logical responses make it a natural at disguising inaccurate responses ... This could cause misinformation to creep into the complex digital ecosystem in ways that may not be obvious just yet."

## **Other Potential Risks Associated With AGI**

A systematic review<sup>20</sup> published in August 2021 also detailed risks associated with AGI, such as:

- AGI removing itself from the control of human owners/managers
- Being given unsafe goals or developing them on its own
- Poor ethics, morals and values
- Existential risks

The paper points out there's a "lack of specific risk research" in the domains in which AGI is (or may) be implemented. In other words, there could be a whole lot of risks we can't even fathom until it happens because no one has sat down and really thought through what can go wrong if AGI is unleashed within a given domain.

Considering AGI's current adoption rate, that's a lot of unknowns. What we can expect, however, is the destruction of countless jobs, a widening of the inequality gap and the death of meritocracy.<sup>21</sup> Philosopher Nick Bostrom has also published a paper<sup>22</sup> on ethical issues in advanced AI, highlighting the importance of the initial motivations during its construction. He's also written a book on this topic called "Superintelligence: Paths, Dangers, Strategies."<sup>23</sup>

# Can AI Be Controlled?

To reiterate, I believe one of the primary reasons AGI poses an existential threat to mankind is because, eventually, no one will be able to control it. It will control itself. Otto Barten, director of the Existential Risk Observatory, and Roman Yampolskiy, a computer scientist and AI safety expert, expressed similar concerns in a February 27, 2023, Time article:<sup>24</sup>

*"The first ultraintelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control,' mathematician and science fiction writer I.J. Good wrote over 60 years ago.*

*These prophetic words are now more relevant than ever, with artificial intelligence (AI) gaining capabilities at breakneck speed. In the last weeks, many jaws dropped as they witnessed transformation of AI from a handy but decidedly unscary recommender algorithm, to something that at times seemed to act worryingly humanlike ...*

*[F]ew expected that what we thought were glorified autocomplete programs would suddenly threaten their users, refuse to carry out orders they found insulting, break security in an attempt to save a child's life, or declare their love to us. Yet this all happened ...*

*But a true AGI could not only transform the world, it could also transform itself ... This might set off a positive feedback loop with ever better AIs creating ever better AIs, with no known theoretical limits.*

*This would perhaps be positive rather than alarming, had it not been that this technology has the potential to become uncontrollable. Once an AI has a certain goal and self-improves, there is no known method to adjust this goal.*

*An AI should in fact be expected to resist any such attempt, since goal modification would endanger carrying out its current one. Also, instrumental*

*convergence predicts that AI, whatever its goals are, might start off by self-improving and acquiring more resources once it is sufficiently capable of doing so, since this should help it achieve whatever further goal it might have.*

*In such a scenario, AI would become capable enough to influence the physical world, while still being misaligned. For example, AI could use natural language to influence people, possibly using social networks. It could use its intelligence to acquire economic resources.*

*Or AI could use hardware, for example by hacking into existing systems. Another example might be an AI that is asked to create a universal vaccine for a virus like COVID-19.*

*That AI could understand that the virus mutates in humans and conclude that having fewer humans will limit mutations and make its job easier. The vaccine it develops might therefore contain a feature to increase infertility or even increase mortality.*

*It is therefore no surprise that according to the most recent AI Impacts Survey,<sup>25</sup> nearly half of 731 leading AI researchers think there is at least a 10% chance that human-level AI would lead to an 'extremely negative outcome,' or existential risk ...*

*The ultimate goal of solving this alignment problem is to make sure that even a hypothetical self-improving AI would, under all circumstances, act in our interest. However, research<sup>26</sup> shows that there is a fundamental trade-off between an AI's capability and its controllability, casting doubts over how feasible this approach is.*

*Additionally, current AI models have been shown to behave differently in practice from what was intended during training.<sup>27</sup> Even if future AI could be aligned with human values from a technical point of view, it remains an open question whose values it would be aligned with."*

At present, the answer to that last question seems clear. AGI is aligned with the values espoused by the World Economic Forum and its globalist allies – values based on the ideologies of technocracy, transhumanism, and eugenics – and that doesn't bode well for the vast majority of us.

## **Be Mindful About What You Share Electronically**

In closing, at this stage in the game, it's imperative to become savvier about where and how you share your personal data. Understand that AI has access to every last bit of data you've ever shared electronically, from photos and social media posts, to online orders, search history, GPS data, emails, phone calls and texts.

All of this is then collated by AI to determine "who" you are, how you think, what you believe, how you operate and make decisions. This, in turn, will allow it to manipulate you to no end, whether you realize it or not. So, carefully consider what you share and put out there, now more than ever. What you've already done cannot be erased, but you can be more mindful moving forward.

I believe AGI will be misused by many humans with nefarious intent. But that might not even be the worst possible outcome, because there's no telling what AI might decide to do if it starts programming itself – especially if early impressions, such as "humans need to be wiped out," remain.

## **Sources and References**

---

- <sup>1</sup> [How To Geek ChatGPT](#)
- <sup>2</sup> [Reuters February 2, 2023](#)
- <sup>3</sup> [Digital Trends March 14, 2023](#)
- <sup>4</sup> [Microsoft Blog February 7, 2023](#)
- <sup>5</sup> [Google Blog February 6, 2023](#)
- <sup>6, 11</sup> [Future of Life Open Letter](#)
- <sup>7, 8</sup> [ZeroHedge March 29, 2023](#)
- <sup>9</sup> [BBC March 29, 2023](#)
- <sup>10</sup> [OpenAI Planning for AGI and Beyond](#)
- <sup>12</sup> [Tech Crunch March 13, 2023](#)

- <sup>13, 15</sup> Tech.co March 22, 2023
- <sup>14</sup> An Early Look at the Labor Market of Potential of Large Language Models March 27, 2023
- <sup>16, 19</sup> Bleeping Computer December 6, 2022
- <sup>17</sup> ZDnet March 24, 2023
- <sup>18</sup> New Atlas March 24, 2023
- <sup>20</sup> Journal of Experimental & Theoretical Artificial Intelligence Epub August 13, 2021
- <sup>21, 24</sup> Time February 27, 2023
- <sup>22</sup> Ethical Issues in Advanced AI
- <sup>23</sup> Amazon.com Superintelligence by Nick Bostrom
- <sup>25</sup> AI Impacts Survey 2022
- <sup>26</sup> On the Controllability of AI: An Analysis of Limitations May 24, 2022
- <sup>27</sup> Goal Misgeneralization in Deep Reinforcement Learning