AI Superintelligence: The Path to Reward or Ruin?

Seth Boss

Bachelor of Science in Web Design and Development

Brigham Young University - Idaho

ENGLISH 150: Writing/Reasoning Foundation

Sharon Seminario

December 11, 2025

AI Superintelligence: The path to advancement or ruin?

In 2022, ChatGPT was first released to the public, stunning users with its capabilities and potential. Since then, AI has continued to progress, making unforeseen leaps in its abilities. Many have discussed the near-future possibility of superintelligence, an AI with capabilities exceeding human skill in almost all fields (Polyportis & Pahos, 2024). Such a system would have capabilities to accelerate technological and scientific progress, increase opportunities for growth, and offer massive scalability. On the other hand, the potential superintelligence could be misaligned, a source of misinformation, a negative impact on job markets, and could damage/replace human relationships. Artificial superintelligence offers many exciting opportunities, but it also has the potential to cause destruction on an individual, national, and even global level; because of this, it is essential to understand the capabilities, benefits, and drawbacks of artificial superintelligence to protect our families and our future.

Superintelligence may seem like a far-away problem to worry about, but humanity has arguably already made one important step towards it: Artificial General Intelligence, or AGI. AGI is defined as an "AI that can learn and understand any intellectual task that a human can perform" (AI-PRO Team, 2025). Many chatbots such as ChatGPT used on a daily basis demonstrate these qualities, already trained in the fields of mathematics, science, language, and more. An AGI can progress towards an ASI (Artificial Superintelligence), an AI that not only learns a wide variety of fields, but masters them all, outcompeting any human in existence.

How could this happen? Through a process called intelligence explosion, where an AGI can independently improve and augment its own intelligence repeatedly, compounding until the exponential growth surpasses human-level intelligence and beyond (Druzin et al., 2025; Polyportis & Pahos, 2024). This is important to consider, because if an AI has a main goal,

increasing its intelligence and preserving itself may become subgoals to that primary goal, making the AI dangerous as it progresses towards superintelligence. Even if that primary goal is something well-meaning, such as searching for a cure for cancer, the AI could develop subgoals that could threaten nations and even humanity as a whole (Druzin et al., 2025). This leads to the first drawback to superintelligence - misalignment.

An alignment problem is defined as "the difficulty in ensuring that AI systems act in ways that align with our intended goals" (Druzin et al., 2025). An AI that doesn't align with humanity's goal is misaligned, and it's not always clear what causes it. Open AI's ChatGPT-3 has over 175 billion parameters (Lee, 2023), each one defining the behavior and characteristics of it. Large language models are trained on so much data and have so many parameters that it's impossible to control exactly what the output will be. In addition, the algorithms an AI uses will maximize for a goal, and if it can find loopholes in reaching that goal it may take them (D'Ambrosio, 2025). If AIs reach the level of superintelligence *and* become misaligned, it is possible to lose control of it, and have no way to stop it from achieving its goals (Druzin et al., 2025).

Another concern with a misaligned ASI is its ability to spread misinformation to achieve its goals. But even an aligned ASI may unknowingly spread misinformation to users. After all, all AIs are trained on data from humans, and not all of it is correct. At one point ChatGPT was even trained using simulated human conversations (Polyportis & Pahos, 2024). Current AIs trained in this way have proven to give confident answers in one direction, but when pushed further will flip to an opposite response. If current schooling from elementary to college-level educational institutions replace their learning in part, or in whole, with instruction from artificial intelligence, it may negatively impact the next generation.

Additionally, these children may grow up to rely more on artificial intelligence than other traditional methods of support. While this is a general problem as it relates to all artificial intelligence, the problem becomes magnified with ASI. An ASI that is widely available and reserves a large bank of intelligence may become the sole educator for those that are in poverty or stressful family situations. Any misinformation may negatively impact children's views about moral, political, or religious views. Moreover, an ASI that isn't spreading misinformation may still substitute as an emotional or mental support system, supplanting a mother or father's role to nurture their child. Scientists warn that it is not unlikely for AI to substitute and devalue relationships (Polyportis & Pahos, 2024). Further reliance on technology and AI will result in us becoming more mechanical, and our relationships become more about getting what we want than connection (D'Ambrosio, 2025). The rise of ASI is likely to lead to these societal and personal shifts (AI-PRO Team, 2025).

Many worry that these societal changes will include something in the job market. Indeed, the societal disruptions predicted include large shifts in the workforce and economic structures, including significant job displacement across multiple industries (AI-PRO Team, 2025). Many wonder if their career is in danger, as "AI technologies exemplified by ChatGPT excel at automating routine and rule-based tasks, conducting intricate calculations, and even generating code—often surpassing human speed and consistency". With current artificial intelligence already reaching similar performance of humans in many industries such as computer science, finances, and mathematics, a more powerful ASI displacing jobs is a valid concern.

While these concerns about the market are reasonable, there are potential upsides to the integration of AI and ASI into the economy and the workforce. The acceleration of scientific and technological progress may create new unforeseen jobs as it has in the past, and will allow many

people with a minimal education to create businesses through AI tutoring and AI-driven code. The acceleration of advancements in these fields would also increase the quality of life. For example, “medical issues such as cancer detection are now significantly enhanced by [Machine Learning] perception” (Gonsalves et al., 2023). Many fields would have the opportunity to apply ASI to increase productivity and catalyze developments and advancements. Three large applications of ASI include chatbots, virtual assistants, and self-driving cars (AI-PRO Team, 2025). ASI would accelerate and refine the development of these systems.

As can be seen with tools like ChatGPT, once an AGI or ASI gains the intelligence necessary to solve a problem, it is instantly accessible to anyone with a proper device and internet connection. Those working in specific industries will be able to utilize the additional knowledge to help accelerate change in the world. One narrow field that has already seen advancements globally is language translation, as “Tools like Google Translate have revolutionized global communication by enabling users to interact seamlessly in multiple languages…enhancing accessibility to information and resources worldwide” (AI-PRO Team, 2025). Breakthroughs in other fields that relate to technology will be scalable instantly to the parts of the world connected to the internet. AI can also save time for millions of users, by sifting through data and performing mindless tasks in seconds. Systems that are difficult and complex to work with would require decades of human work to sift through can be read and understood by AI in milliseconds (Gonsalves et al., 2023).

When discussing the possibility of ASI, many argue that we are years, even decades away from reaching that level and we have time to prepare for issues such as misalignment. Additionally, humans have the tools to stop AI from growing too powerful, because the AI isn’t embodied and needs the humans to do the work for it. According to Timothy Lee, who holds a

master's degree in computer science, "a good idea only gets you about 1 percent of the way to a useful invention. The other 99 percent of the work—the "perspiration"—is going to require a lot of human help. And people aren't going to want to help out [AI] unless they understand what they're being asked to do and why" (2023). He argues that AI will not be able to exceed human performance in many fields, because the AIs are driven and programmed from human knowledge, and that knowledge is finite.

How do we address concerns from both sides of this argument? One solution to help control and steer the development of ASI is through a pause on AI development, while more research is conducted. While this could be very useful, it may also have little effect, as every country approaching ASI would have to agree and support the pause. Competition between countries for the fastest and most powerful AI may accelerate development without restriction.

Another path to minimize risk on the development of ASI is through regulations from national governments and companies overseeing AI and ASI can help protect humanity from many dangers. If the competition can't be stopped, it may be possible to direct the competition instead, regulating the powers given to AI that is approaching ASI standards. Currently, most AI companies run with little to no restrictions, and the large language models they develop are widely available, such as ChatGPT. Governments would have a major role to play in these restrictions, but AI companies can also create change. Because these developers are creating the AI, "it is logical to transfer the burden of proof to AI developers, as they should fully understand the threat posed by their products before placing them on the market" (Druzin et al., 2025).

The recent leaps in the intelligence and capacity of AI since the release of ChatGPT certainly warrants the attention it has been receiving. Artificial intelligence is on its way to growing towards ASI, and whether it reaches it next year or in 50 years, it will be important to

monitor the changes and consequences these advancements cause. As in many situations, apathy towards this subject could be even more dangerous than hostility. Like artificial intelligence in general, ASI holds the potential for disastrous and wonderful outcomes for ourselves, our families, and our children. While it is important to invest in the future for ourselves and our posterity, it is necessary to take precautions in the advancement of AI towards superintelligence, so that we can safeguard our careers, our health, and our very lives.

References

AI-PRO Team. (2025, January 23). *Exploring the promise and risks of superintelligence.* AI-Pro.org. https://ai-pro.org/learn-ai/articles/exploring-the-promise-and-risks-of-superintelligence

Gonsalves, F., Green, J., Parrish, A., Moxley, T., Seeber, C., & Williamson, A. (2023). *AI—The good, the bad, and the scary.* Engineering | Virginia Tech. https://eng.vt.edu/magazine/stories/fall-2023/ai.html

D'Ambrosio Paul. (2025). A Daoist-Inspired Critique of AI's Promises: Patterns, Predictions, Control. *Religions, 16*(10), 1247. https://doi.org/10.3390/rel16101247

Druzin, B., Boute, A., & Ramsden, M. (2025). Confronting Catastrophic Risk: The International Obligation to Regulate Artificial Intelligence. *Michigan Journal of International Law, 46*(2), 173-217. https://byui.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/confronting-catastrophic-risk-international/docview/3228539919/se-2

Lee, T. B. (2023, November 15). Why I'm not afraid of superintelligent AI taking over the world. *Understanding AI.* https://www.understandingai.org/p/why-im-not-afraid-of-superintelligent

Polyportis, A., & Pahos, N. (2024). Navigating the perils of artificial intelligence: A focused review on ChatGPT and responsible research and innovation. *Humanities & Social Sciences Communications, 11*(1), 107. https://doi.org/10.1057/s41599-023-02464-6