Nickrod Basiri
48 Vaughan Mills Road
Vaughan, Ontario, L4H1C8
nickrodb@my.yorku.ca
March 26, 2025

Dear Professor Talebzadeh,

I am pleased to submit my technical report titled “Ensuring the Ethical Development of Open-Source AI Technologies.” This report explores the ethical implications of open-source AI. In it, I discuss the pros and cons of open sourcing, using the recent launch of the DeepSeek-R1 model as an example of the potential benefits of open-source. I believe open-source software is a great way to foster collaboration and innovation, allowing everyone to benefit from the technology. I also go over potential risks with the practice, urging the safe and ethical use of AI in the future.

Should you have any questions, comments, or concerns, feel free to reach out. I look forward to hearing your feedback. Thank you for your time.

Sincerely,
Nickrod Basiri
220506614
ENG 2003 - Effective Engineering Communication.

Ensuring the Ethical Development of Open-Source AI Technologies

Nickrod Basiri
Lassonde School of Engineering

March 26th, 2025

Executive Summary

This report will address the grand challenge of the ethical development of artificial intelligence. It will focus on the growing importance of open-source AI technology, discussing both the pros and cons of having AI frameworks be open-source. As AI becomes increasingly more advanced as well as more and more a part of everyday life, it becomes substantially more important to focus on the ethical implications of its development. Open-source AI frameworks are a potential solution to allow everyone equal access to the backbone of the technology, leading to greater innovation. This however also introduces risks, such as security vulnerabilities and potential misuse.

Table of Contents

Introduction

Artificial intelligence is very quickly becoming a part of everyday life. Transforming various major sectors such as healthcare, finance, and engineering, AI has led to immense room for increased optimization and efficiency. As these systems become more advanced, they are integrated into the fabric of everyday life, leading to concerns over the ethical development and use of the technology. Today, the debate over open-source AI is prominent. With AI frameworks such as Deepseek making waves, it becomes more and more controversial whether AI development should continue on an open-source path, or stay behind the closed doors of various major tech corporations. While open-sourcing AI will allow for a much broader reach, allowing developers across the globe to contribute to its innovation, it also opens the door to potential security risks and malicious practices.

By nature, artificial intelligence has the potential to impact major industries and groups of people. Self driving trucks can eliminate the need for truck drivers. AI software engineers will decrease the need for junior human software engineers. Some believe that open sourcing will only increase the rate at which this occurs, eventually leading to millions of lost jobs. While encouraging people to develop for good, there are inherent drawbacks to allowing anyone to have access to the technology. There are also ethical risks, with the potential for misuse of the technology. Deep Fakes and cybersecurity attacks are examples of malpractice that can be done with AI.

This report will focus on the challenge of ensuring that AI development remains ethical, especially with the rise of open-source frameworks. It will explore various expert opinions, both advocating and speaking against open-source AI.

Background

Artificial Intelligence has moved from a theoretical concept to something that over 86% of students use in their studies according to a survey from the Digital Education Council [1], showing how widespread the adoption of AI has become. Peter Norvig, Director of Research at Google, advocated for open-source AI, saying “the challenge now is to make sure everyone benefits from this technology. It’s important that machine learning be researched openly, and spread via open publications and open source code, so we can all share the rewards” [2]. In 2024, a Chinese startup released DeepSeek, an open-source AI model that was on par and in some cases better than the current market leader, ChatGPT. This sparked massive discourse on the open sourcing of AI, with people taking both sides.

Main Topics

DeepSeek-R1

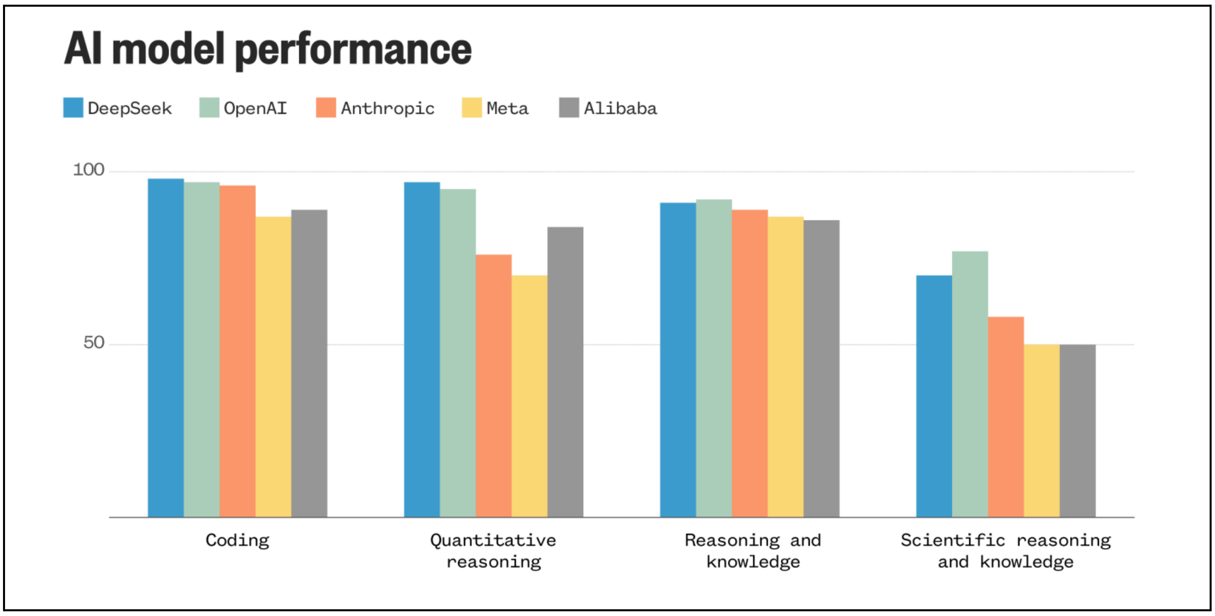

Figure 1. DeepSeek Performance Compared to Other Models [3]

DeepSeek-R1, released in January 2025, is an open-source AI model that matches and sometimes outperforms Open-AI's GPT-4, the most popular AI model at the time. "By releasing models with open weights and transparent code, DeepSeek contributes to a paradigm where AI isn't locked behind paywalls and proprietary systems" [4]. Max Chong Li, a professor at Columbia University, uses DeepSeek as an example of a step in the right direction. With it being completely open-source, anyone can contribute and build using the existing framework, something unable to be done with existing models. Moreover, it also displayed that AI models can be made significantly cheaper, and more efficiently than previously known [4]. Figure 1 compares DeepSeek's performance across various metrics, showing that it is on par and even sometimes ahead of the biggest and most popular AI models [3].

Collaboration and Innovation

The arguments for open-source AI are many. Open source has been beneficial to various other technologies for years. "Making software free has long helped developers make their code stronger. It has allowed them to prove the trustworthiness of their work, harness vast amounts of volunteer labour and, in some cases, make money by selling tech support to those who use it" [5]. The article by the economist highlights the greatest benefit of open-source AI, being collaboration and innovation. Making the software free and publicly accessible allows anyone, anywhere, to build and strengthen the existing technology. The community built around open-source software allows for things that may have never been done to be created by hundreds of people. Volunteer work done by open-source contributors is invaluable to various technologies, and many advancements that are majorly used today. The low barrier to entry of open-source allows for a substantially easier time getting started, which will encourage more people to contribute. Additionally, the fact that anyone can see what is being done helps keep people accountable. Open source enhances transparency and trust. Accessible code allows users to point out bugs, and discuss potential inherent biases present in the algorithms. This leads to more reliable and ethical systems, as there are thousands of developers ensuring things remain safe.

Security Risks

While there are many great benefits, open-source AI has its risks. Seger et al. discuss these risks, expressing the potential for misuse by malicious actors, with intent to cause harm rather than good. “Open-sourcing is helpful in addressing some risks, but could—overall—exacerbate the extreme risks that highly capable AI models may pose” [6]. With AI granting so many opportunities, there are many opportunities for wrongdoing. Deepfake technology is an example of this, allowing someone to create an image or video of another person. This can be used to fake incriminating and defamatory information on a person, with it being hard to prove who the malicious actor is. The act of open-sourcing also increases attackers' knowledge of exploits, something they would not be able to see in software that is not open-source [6].

Discussion

Benefits to Open-Source Artificial Intelligence

The most obvious advantage of open-source, as displayed by DeepSeek, is the opportunity for innovation and collaboration. The fact that it is open-source, as well as completely free and available to everyone is a fantastic step in the right direction. Rather than being limited to the small team in China, DeepSeek can now be worked on and improved by anyone around the world, while also allowing for anyone to benefit. Startups and smaller firms will be able to benefit greatly from DeepSeek, with it being free to use. It avoids the barrier of a paywall and allows for smaller groups to use AI tools to their benefit. This approach encourages a broader range of people to use and adopt AI tools.

The fact that DeepSeek is free also has the benefit of fostering competition between the giants in the AI sphere. Open-AI’s ChatGPT has faced little real competition until the release of DeepSeek-R1. Competition in business leads to innovation out of the desire to keep the greatest percentage of the market share. With DeepSeek rivalling GPT, while also being completely free, will ideally lead to greater innovation from Open-AI, leading to a net positive for the sector.

Potential Risks with Open-Source AI

Despite these benefits, there are also significant risks with open-sourcing advanced AI models. A primary risk is the potential for misconduct. Open-source AI could enable potential bad actors to exploit the models for whatever they would like. As previously discussed, Deepfake technology is a prime example of this. It becomes increasingly easier to deepfake someone's face onto any image or video with malicious intent. It would become substantially easier could these bad actors had access to the technology in an open-source context.

Additionally, Seger et al discuss the increase in knowledge available to attackers. With the code being publicly available, attackers can more easily find and abuse vulnerabilities in the system. It would become substantially easier for these people to exploit the technology, leading to more harm done. There is an inherent lack of security with open-source, being that anyone can see what is going on behind the scenes, even those who may use it maliciously. Closed models such as GPT allow for monitoring of what they're being used for, blocking potential negative inquiries. This would be significantly harder to stop with an open-source model, as anyone can run the software locally. They could simply change and bypass any protocols stopping malicious practices.

Finding a Balance

To move forward, there must be a balance found between the two options. A solution would be to increase the legal scrutiny of these AI models. Having laws that bog down the process of development decreases the capacity of developers to create something innovative, as too many restrictions can hinder the rate at which people can build and advance the technology. However, it also prevents malicious actors from having easy access to the software. As AI advances, there must be laws and regulations put into place that ensure AI models are developed and used ethically. Without the appropriate legal frameworks, the security and ethical risks present with open sourcing may outweigh the benefits. A safe solution must be put into place.

Open-source communities have long helped shape a lot of the software people use. It has been a driving force behind a plethora of technologies that are used by many people around the world. From web browsers to operating systems, the impact of open-source software is clear and incredibly important. While there are potential risks, many open-source communities have codes of conduct, as well as guidelines that ensure ethical development. These codes help make sure that contributors to projects are following good practices and ethical standards, while still being able to innovate. These actions are necessary to make sure that the future of AI is safe. Embedding these practices into AI development will avoid potential harm while encouraging innovation and accountability.

Conclusion

As artificial intelligence continues to evolve and become an important part of various industries, as well as the day-to-day lives of many, the debate over open-source AI is becoming increasingly more relevant. Open-source, demonstrated by models like DeepSeek-R1, offers massive benefits to the industry, while also fostering innovation from developers around the world. Removing paywalls and decreasing the barrier of entry leads to a greater reach, allowing more developers to contribute, while encouraging competition to innovate further.

While the benefits of open-source are clear, it must be implemented securely. The potential for bad actors is high, which is why the necessary rules and regulations must be put in place to stop them. Both lawmakers and the open-source community need to ensure that there is as little room for wrongdoing as possible.

To move forward responsibly, it is essential to find the balance between the benefits and the risks presented. By taking these precautions, we can ensure that AI continues to benefit society. The future of AI is promising, but it must be done carefully, ensuring everyone can benefit.

References

[1] R. Kelly, "Survey: 86% of students already use AI in their studies," *Campus Technology*, Aug. 28, 2024. [Online]. Available: https://campustechnology.com/articles/2024/08/28/survey-86-of-students-already-use-ai-in-their-studies.aspx. [Accessed: Mar. 26, 2025].

[2] M. Marshall, "50 grand challenges for the 21st century," *BBC Future*, Mar. 31, 2017. [Online]. Available: https://www.bbc.com/future/article/20170331-50-grand-challenges-for-the-21st-century. [Accessed: Mar. 26, 2025].

[3] J. Saah, "DeepSeek AI comparison: OpenAI ChatGPT, Google Gemini, Meta Llama," *NBC News*, [Online]. Available: https://www.nbcnews.com/data-graphics/deepseek-ai-comparison-openai-chatgpt-google-gemini-meta-llama-rcna189568. [Accessed: Mar. 26, 2025].

[4] M. Li, "DeepSeek's lesson: The future of AI is decentralized and open-source," *Forbes*, Feb. 28, 2025. [Online]. Available: https://www.forbes.com/sites/digital-assets/2025/02/28/deepseeks-lesson-the-future-of-ai-is-decentralized-and-open-source/. [Accessed: Mar. 26, 2025].

[5] "Why open-source AI models are good for the world," *The Economist*, Nov. 7, 2024. [Online]. Available: https://www.economist.com/leaders/2024/11/07/why-open-source-ai-models-are-good-for-the-world. [Accessed: Mar. 26, 2025].

[6] E. Seger et al., "Open-sourcing highly capable foundation models: An evaluation of risks, benefits, and alternative methods for pursuing open-source objectives," *arXiv preprint arXiv:2311.09227*, 2023. [Online]. Available: https://arxiv.org/abs/2311.09227. [Accessed: Mar. 26, 2025].

Revision Summary

While I did not submit on time for the create / assess portions of the assignment, I still asked two peers to read my report and give me feedback. Here is my revision summary based on the feedback given.

Suggestion: “The introduction is good but I feel like you should probably mention briefly some of the ethical issues with AI.

- I agree with this, and have added a brief explanation of them to my intro.

Suggestion: “I think you should elaborate on the 86% statistic like why is that important.”

- Added a bit about how it shows how widespread the adoption of AI has become.

Suggestion: “Maybe mention what governments already have protections for AI and things like that”

- I don’t necessarily agree with this one, as a large majority of governments have not yet put anything into place.
- America, the dominant leader in AI with Open-AI, has done very little. I don’t feel like it’s helpful to mention a random country that has, if there is any.

Suggestion: “You could have more figures or graphs to help your points”

- I don’t agree with this one. The one graph I have demonstrated DeepSeek’s disruption of the AI space, but I don’t see how another one could be used anywhere to improve a point made.

My peers also pointed out various grammar mistakes or just wording that sounded weird, (for example, I put a - between open source each time, which I shouldn’t have in certain cases) which I fixed.