Xinyue Liang
COSC 276 24F
Professor: Soroush Vosoughi
Nov 21, 2024
Word Count: 818

AI Ethics Analysis Report

Part1. Overview

With the current technology and investment focus on AI training and development, raising concern about AI Ethics and Security has emerged. There are four major areas where the concern lies: 1. The black box nature of AI: the inevitable tradeoff between interpretability and training cost 2. Ethical human usage on AI products 3. Social impact: for example educational, fairness between different social groups 4. General security issues: major cyber attacks, data leak concerns. In this report, the main focus will be on the second part: Ethical human usage.

The reason why the second aspect deserves such emphasis is that it is not only including a diverse set of scenarios but also involving the more uncertain part of interaction between stakeholders compared to the other 3. The ethical usage of AI is also directly connected with the other three aspects as well When it comes to the usage of a product or service. As the majority of two main stakeholders are always the service provider and consumers. The provider either gets money or reputation as compensation for providing the service or product while the consumers get convenience or entertainment for the money they spend.

Just like any other products and services, for instance firearms, vehicles and even comics, the way AI was presented and the users' behavior should be taken into consideration when talking about or setting the rules for Ethics of the industry. However, there are two elements that differentiate AI from other products is that it is meant to serve multiple purposes and it can bring large scale social impact. So the

analysis for AI Ethics on human interaction should be started from the three following levels: consumer, professional and enterprise.

Part2. Existing Opinions

The debuts around safe and ethical usage of AI products has long been existing ever since the products have been brought to public attention. On one hand, there are certain views like Google (2024) that it is the provider's responsibility to ensure fair, secure and ethical usage by methods like enforcing detection and providing a “healthy” ecosystem. On the other hand, there is also opinions (Morgan Stanley, 2024) that it is necessary training for users to help them identify AI abuse thus protect their own security

Part3. Personal Opinion

Personally, all of the existing debates leads to one viable solution: forming a balanced committee including all the representatives of stakeholders or government legislation departments and preventing abuse of AI through policies and laws. From the previously mentioned 3 level of scenarios, there were similar previous cases succeed from other industries.

Firstly, on a consumer level. Personally, I view the most common issue of consumer level of AI abuse is that some may use certain AI products such as Deepfake to harm cyber security. Also, on the providers’ scale, we can envision large companies making their AI products accessible for anyone without limitations. We have already seen the successfulness of committee based censorship working on publication products such as comics and movies as illustrated by Chirs (2021). Also, we could see the progressive made in firearm control and quality assurance with the help of a series of legal acts published since 1934 (Gale, 2024). Thus, we might be able to avoid abuse and ensure ethics and safety.

Moreover, as for the enterprise level, I found out the major concern is that companies uses unauthorized data for model training, and missive use of generative AI in creative works such as movies and video games will not only reduce possible job opportunities but also make the productions low

quality. A legislation committee or a government department could also resolve the issue here, just like Ralph (2023) described in his article, protecting the lawful rights of intellectual property but also making adaptations to the according era. For example, adding encryption on unauthorized data to avoid being used for training purposes.

Last but not least, the use of AI in professional fields such as Academic, Military and Medical or Clinical scenarios requires the companies to provide unbiased, precise and correct information while the users are expected to follow the strict integrity rule of their own field. Although punishing laws from a committee or a department might not be able to fully eliminate the existence of abuse, rewarding policies could promote the healthy development and usage of AI to meet up the expectations in professional areas. Similar cases could be seen in the rapid development of Chinese EVs. With the support from government policy, chinese ev manufacturers has obtained 23% of global market in 2023 (Alan).

Part4. Summary

With all the cases above, it could be seen that a third party or government organization could positively influence the future development of AI, regulize AI products and avoid abuse to certain extent. Thus, I personally think forming such an organization and promoting AI specifically legalization could be a viable option for ensuring AI Ethics.

References

Gale. (n.d.). *Gun control*. Gale Open Access. Retrieved November 24, 2024, from https://www.gale.com/open-access/gun-control#:~:text=The%20National%20Firearms%20Act%20(NFA,weapons%20such%20as%20machine%20guns.

Dobson, P. (2024, November 20). *Experts see rapid rise of Chinese EV makers*. EE Times. Retrieved from https://eetimes.com/experts-see-rapid-rise-of-chinese-ev-makers/.

StudioBinder. (n.d.). *Movie censorship in America: A brief history*. StudioBinder Blog. Retrieved November 24, 2024, from https://www.studiobinder.com/blog/movie-censorship-in-america/.

Montague, R. (2023, September 19). *A short history of access to knowledge*. LinkedIn. Retrieved from https://www.linkedin.com/pulse/short-history-access-knowledge-ralph-montague/.

Morgan Stanley. (2024, February 15). *AI and cybersecurity: A new era of defense*. Morgan Stanley Articles. Retrieved from https://www.morganstanley.com/articles/ai-cybersecurity-new-era.

Google. (n.d.). *Advancements in cybersecurity with SAIF*. Google Safety. Retrieved November 24, 2024, from https://safety.google/intl/en_us/cybersecurity-advancements/saif/.