Conner Jordan
CST 300 – Ethics Analysis Paper
February 12th, 2023

AI and Ethics – The Meteoric Rise of Hyper-Intelligent Bots and the Ethical Implications

A Brief History

The idea of artificial intelligence replacing humans has been a controversial topic of discussion for decades. While this notion has existed for a long time, it has only very recently become extremely relevant with the rise of chatbots and other popular AI. The most current development in the field of AI has been the release of ChatGPT, a linguistics chat bot that can perform many advanced functions and calculations. ChatGPT was released in November 2022 by OpenAI, a research company geared towards the development of user-friendly AI (Bednarski, 2023). The application immediately distinguished itself from other chat bots of its kind with its high-level understanding of conversational context, and its ability to develop complex programs and written text with very basic user inputs (Bednarski, 2023).

By all accounts, AI development will only continue to exponentially increase in the coming years. GPT-4, an upcoming AI application from OpenAI, hints at a multi-modal interface of text, images, and sounds. Further, it reportedly draws from a set of parameters that is orders of magnitude higher than even the most recent version of ChatGPT, allowing it to gain even more insight into large data queries (Montti, 2023). While this is undoubtedly an exciting development, it raises several important ethical questions that should be carefully considered as AI continues to develop daily. Students have already been caught cheating with ChatGPT, using the AI during exams and as a tool to create seemingly original essays and other literary

works. Concerns have also been raised about improper use of these types of AI in the field of cyber security. From Tsarfati et al. in 2023:

> "ChatGPT could easily be used to create polymorphic malware. This malware's advanced capabilities can easily evade security products and make mitigation cumbersome with very little effort or investment by the adversary."

Additional fears have been raised about the utilization of AI in the field of mental health. In just the last few weeks, in early January, there has been much controversy surrounding the use of AI as a tool to communicate with individuals suffering from mental health conditions. Specifically, the mental health website Koko came under fire after it was discovered that many therapists in their online support groups were simply human assisted AI chat bots (Advisory Board, 2023). Further, many are concerned with the fact that ChatGPT and other similar AI can make very convincing and well-articulated arguments that have no foundation or basis in fact. The prospect of AI development can therefore be either extremely exciting, or highly distressing depending on which stakeholders are considered.

The Stakeholders:

The most important question to consider when analyzing stakeholders on opposite ends of the spectrum is how and why they will benefit from the use of modern AI. No matter how you feel about AI use ethically, it's become virtually impossible to argue that AI isn't outperforming humans across several different metrics. The body of evidence has simply become too overwhelming, and AI has proven to be extremely efficient at problem solving and task automation.

It is probably easiest to first consider the viewpoint of those who would advocate for more prominent usage of AI in the professional sector, as the benefits are easy to contemplate. These stakeholders are comprised of large corporations, engineers, and many industry and healthcare professionals who can reap the benefits of AI development firsthand. As task automation becomes more streamlined and AI continues to advance, many industries will see growth that will undoubtedly translate to higher profits, better analysis, and more efficient processes (Davidson et al., 2022). Stakeholders holding this viewpoint recognize that AI presents certain ethical dilemmas, but believe the positives far outweigh the negatives.

For an entirely separate but vocal group of individuals, however, modern AI is developing too quickly and is presenting a plethora of ethical problems that have become too glaring to ignore. These individuals are largely comprised of advocates of human rights and mental health, and a large majority of content creators.

Stakeholder #1:

First, let's consider the advocates of unhindered AI development. These stakeholders believe that machine learning has the potential to track individual health patterns, diagnose and treat various diseases and illnesses, as well as address a range of societal issues, including infrastructure and poverty. They highlight the fact that hindering AI progression would only stunt societal growth and would ultimately be a negative thing for humanity. They stand behind a claim of value, with many advocates pointing to the trend of AI that outperforms its human counterparts in countless sectors. Stakeholders in this demographic stand to lose out on the many societal benefits that AI can provide.

Stakeholder # 2:

On the other end of the ethical conundrum are those who question if it is right to entrust decision making to machines that are not capable of empathy, intuition, or moral reasoning. In direct contention with the first stakeholders, they worry about the consequences of prioritizing economic and strategic goals over societal and human values. A claim of policy is often asserted that, given the potential negative impacts of AI on society, we should not continue to develop and deploy AI systems at such a rapid pace. This claim is often based on concerns about bias, discrimination, privacy, job displacement and other ethical dilemmas. Individuals who are dubious of modern AI developments believe we stand to lose an essential piece of our humanity, in a clouding of individual identity, as we learn to become more reliant on artificial intelligence and machines.

Question:

Should enhanced restrictions be placed on new AI releases so that they cannot be abused by the public? Or should we continue to let the technology progress in its current state without interference?

Argument #1:

Those arguing for the use of AI do not shy away from the obvious fact that this technology will bring its fair share of negative outcomes. People will lose their jobs due to this new technology. In fact, it's highly likely that this will occur in very large numbers throughout several industries. Many will face hardship and uncertainty as a direct result of AI growth. There is no escaping this fact - but hindering new developments will only hamper societal growth. The negative effects of AI are simply a necessary evil, when one considers the greater potential

benefits it can provide to civilization. Consider this ethical dilemma under the framework of ethical utilitarianism. Utilitarianism is the ethical theory that asserts that actions should be evaluated based on their ability to promote the most happiness for the largest number of people. Since its inception in the 18th century, utilitarianism has been a wildly popular framework with which to consider difficult ethical decisions. Under this framework, you need not ignore the obvious issues with unfettered AI development, you must simply consider the greater good in the face of these relatively minor issues. Although AI may make some individuals uncomfortable, it nonetheless needs to be pursued so that society can flourish. Whatever moral uncertainties one may have with AI, it will ultimately generate the best outcome for the largest number of people when compared to a world with regulated AI progression.

Consider the development of the steam engine in the late 18th century and compare it to the progression of modern AI. Both AI and the steam engine represent major technological breakthroughs that have had significant impacts on society and the economy. Like the steam engine, AI has the potential to transform industries and create new ones, leading to significant economic growth and job creation. Both technologies have raised concerns about their potential impact on society and the environment and have been the subject of debate and regulation. However, in hindsight, placing restrictive regulations on steam engine technology because of occupational and environmental concerns would likely have done major damage to the progression of worldwide infrastructure and transport. Instead, the steam engine quickly became a key technology of the Industrial Revolution, powering factories, mines, and transportation systems around the world.

Argument #2:

Consider AI growth under the framework of care ethics - in this context it is important to consider the impact that AI may have on human relationships and social connections. The importance of cultivating relationships and connections with others should be emphasized, and technology should be developed in a way that supports human growth and well-being. Left unchecked, AI could lead to the erosion of social connections and human interaction, as individuals become increasingly reliant on faceless, featureless technology to meet their needs. (Amditis , 2017) A claim of policy is asserted, with many arguing that politicians and policy makers should create strategies designed to allow AI to grow to a certain extent, but also be mindful of the potential negative outcomes and put safeguards in place to mitigate these effects (Lawton & Wigmore, 2023). Many support an alternative course of action, such as investing in other forms of technology, or placing greater emphasis on human-centered design and ethical decision-making in AI development.

Making the same consideration as the first stakeholder under a care-based framework, the steam engine and its development was a completely different process that did not have the same implications that AI promises to have for mental health and the human experience in general. AI has developed at a much quicker pace than the steam engine ever did. While the steam engine took several decades to become widely adopted, AI has seen exponential growth in recent years and months, and is already having significant impacts on society. Further, the steam engine was primarily a physical technology that transformed the way goods and people were transported and powered, while AI is primarily a digital technology that is transforming

the way we process and analyze information. The consequences promise to be far reaching and more profound than the development of the steam engine.

Personal Stance:

I feel that AI is here to stay and that it should be embraced fully. It is true that many important questions are raised by the detractors of unregulated AI progression. In fact, I believe there is a lot more to learn about AI, and that our understanding will not necessarily progress in tandem with AI's evolution. While this does of course pose a dilemma, I choose to believe that the net positives far outweigh the negatives. A utilitarian approach to this ethical setback is the best and most logical solution. To answer the question of whether AI development needs to be more heavily regulated, I firmly believe this would be a bad move and counter to the advancement of society. While we undoubtedly have yet to face many of the growing pains that will come about as this technology advances, attempting to put a proverbial cap on the growth of man-made innovation would run counter to the overall progress of humanity.

References

Advisory Board. (2023, January 10). The ethics of using AI chatbots in mental healthcare. *Advisor.* https://www.advisory.com/daily-briefing/2023/01/10/ai-mental-health

Amditis, J. (2017, March 22). What are some ethical considerations that chatbots raise, and how should they be addressed? *Medium*; NJ Mobile News Lab.

Bednarski, D. (2023, February 16). What is openai? - its history and how it is changing the world: Taskade. *Taskade Blog*. https://www.taskade.com/blog/openai-chatgpt-history/

Davidson, A.-L., Boch, A., Nandaba, B. K., Rougui, M., & Naffi, N. (2022, January 2). AI-powered chatbots, designed ethically, can support high-quality university teaching. *The Conversation*. https://theconversation.com/ai-powered-chatbots-designed-ethically-can-support-high-quality-university-teaching-172719

Lawton, G., & Wigmore, I. (2023, January 18). What are Ai Ethics (AI code of ethics)?: Definition from TechTarget. *WhatIs.com*. https://www.techtarget.com/whatis/definition/AI-code-of-ethics

Montti, R. (2023, January 21). GPT-4 is coming: A look into the future of ai. *Search Engine Journal.* https://www.searchenginejournal.com/openai-gpt-4/476759/#close

Tsarfati, O., & Shimony, E. (2023, January 10). Chatting Our Way Into Creating a Polymorphic Malware. *CyberArk.* https://www.cyberark.com/resources/threat-research-blog/chatting-our-way-into-creating-a-polymorphic-malware