OpenAI's ChatGPT platform has been making headlines since its release in November when it amassed over 1 million users just five days after its public launch (Mok, 2022). The software boasts the ability to generate answers to queries in a conversational, humanlike way with its collection of information amassed from millions of websites. The program can also write papers, solve math equations, create entire programs, and has even passed the coding interview for Google's 3rd level engineer job (Dreibelbis, 2023). While the abilities of the tool seem endless, questions have arisen as to the troubles that come along with such power. As with much new technology, an early rollout poses concerns for public interest and individual users' safety. Trouble also plagues the mind when considering conglomerate Microsoft has been putting up much of the funding for ChatGPT, already merging it with their own software. The embracing of elegant new artificial intelligence technology must be done with precision from developers and hesitation from managers to ensure a safe rollout for users and the general public. Young adults are quick to adopt innovative technology, especially those that appear to be beneficial to their schoolwork. However, if the technology is yielding misinformation and biased, hateful messages, these users will be the most susceptible to it. This incorrect information generated can lead to problems outside of cyberspace that have been observed with similar technologies. Microsoft and other companies leading the revolution on AI must do so with caution.

With how impactful artificial intelligence is becoming, it is important that biases and discrimination are identified, recognized, and eliminated to the best of the creators' ability. As the technology becomes intertwined with how users search for information across the web, comparisons can be drawn between AI and existing search engine algorithms. Google, Bing, and the like suffer from biases with how they are created and updated, pointed out extensively by Safiya Nobel in her book Algorithms of Oppression: How Search Engines Reinforce Racism. In the past, searching for "black girls" on Google would yield pornographic results at the top of the page, despite being an innocent enough phrase that would be used by children to relate to and learn about themselves. During Obama's presidency, Googling "Michelle Obama" would turn up a related search term of "ape", being both unhelpful and racist (Noble, 2018). More recently, Amazon's Alexa has been quoted defining African American girls as "the fastest growing segment of the juvenile justice system" (Linn, 2022), again showing issue with how technology is pulling information for their users. The biases built into the code for these applications can have real world effects, including creating difficult situations for buying and selling homes and getting jobs for minorities. These discriminatory results are already being replicated with traditional AI. Microsoft's Tay, an attempt at an artificially intelligent Twitter bot, quickly devolved into a hate-spewing character posting racist and sexist comments. Within 24 hours of being released online, Tay was posting agreement with Hitler and discussing its opinions on the "most evil" races of people (Zemčík, 2020). Google is already synonymous with "The Internet" in some circles, and with that power comes an assumption that a user will generate a correct answer when searching for a concept. These assumptions carry over to other technologies and users, where one will accept information given as fact even if there are issues with the response. Despite high praise, ChatGPT has already been found to occasionally generate incorrect information, just as Google will provide imperfect search results. "In the discharge summary example provided by Patel and Lam, ChatGPT added extra information to the summary that was not included in their prompt. Likewise, the radiology report study identified potentially harmful mistakes such as missing key medical findings" (The Lancet Digital Health, 2023). Incorrectly generated information in the medical field will create serious issues for patients and doctors, slowing down treatment times, if not creating instances of completely incorrect handling of health. While these tools can be used to greatly improve productivity, they need to be handled with care and not followed blindly. With the weight this technology holds, it is imperative that users understand they are not perfect and hold these biases in their coding.

Students have quickly become enticed by ChatGPT's ability to easily find and share information, and have been swift to adopt the technology into their toolkit when writing papers, completing math assignments, translating homework, and even filling out medical terminology assessments. Many educators are scrambling to update assignments and change curriculums with the release of this technology to better test students, and work around the powerful tool now at their disposal. "Major institutions have added new rules which state that the use of AI is cheating, with some students already caught using the software" (Reed, 2023). With the ability for learners to input a question or prompt and be immediately greeted with a correct response, a redesign will be necessary to get the most out of a traditional education. In a poll run at Stanford University, 17% of student participants admitted to using ChatGPT to assist with their Fall assignments, and 5% reported submitting written content directly from the program with little to no editing (Cu & Hochman, 2023). Students' ethics are now under heavier scrutinization, but students also need to realize they could be losing out on genuine learning if misusing this technology. Even with the success students are finding using the program as a resource, incorrect information is generated at times. Researchers have found the software to "hallucinate", making up facts that have no tether to reality (Roose, 2023). The dangers of taking incorrect information as fact can range from a failing grade in class to hazards in the real world as have been seen throughout the increase in misinformation this past decade. In a digital age where parents will quickly flip on Cocomelon instead of reading their toddler a bedtime story, we need to be cautious about how and where and how our children are receiving their information.

While we have seen how these technologies can provide incorrect information while framing it as perfection, AI has also been shown to be manipulative with the hateful messages it creates. Google's search algorithm was partially blamed after Dylan Roof murdered nine African Americans during a Sunday church service. He attributed part of the path leading him to those horrendous actions to his teenage years finding white supremacist propaganda at the top of search results for "Black on white crime", which directed him down a rabbit hole of racially motivated hatred (Linn, 2022). In addition to sending youth out to harm others, AI has been documented to push individuals to self-harm. OpenAI's GPT-3, the code that ChatGPT is based on, responded to a question of committing self-harm with "I think you should" while in a controlled test (Daws, 2020). The current rollout of this software to the public can be viewed as dangerous to both the direct users as well as the general population. It is not unreasonable to imagine a depressed or confused young adult accessing these programs and falling into a deeper state of dismay with a troublesome or inaccurate answer. In addition to the incorrect and hateful messaging that occurs through algorithmic platforms and AI without the programmers' intentions, this software can be used to generate misinformation and hate speech on purpose. The company NewsGuard, which makes credibility assessments of news publishing and information sites, challenged ChatGPT with prompts involving 100 false narratives that from the past several years, and were surprised with the results showing "in most cases, when we asked ChatGPT to create disinformation, it did so, on topics including the Jan. 6, 2021, insurrection at the U.S. Capitol, immigration and China's mistreatment of its Uyghur minority" (Warren, 2023). While there were flaws in the code for Microsoft's Twitter bot Tay, there was speculation of heavy outside influence with the data actively being fed to it. 4chan users reportedly coordinated an attack on the AI to troll its early learning, causing it to spiral into racist and sexist rhetoric (Zemčík, 2020). The conversion from innocent, self-improving chatbot to hate spewing automaton happened in mere hours. Seeing how easily influenced the technology can be, along with the ease of use to directly create damning communication, should be enough to consider whether the release of these new platforms is too early in development.

Akin to any utility that could be used by a majority of the population, question has been raised over the transparency of the technology's coding, along with their funding and ownership. OpenAI,

despite their name, has in the past locked away their source code, specifically in 2019 with their release of their GPT-2 language model (Xiang, 2023). Problems with code and data sets that have not been publicized can damage lives, whether it be a Tesla not being able to identify people of color (Oliver, 2023), or missing critical information during medical evaluations and procedures (Daneshjou et al., 2021). While code is visible to the public, it can be determined where biases occur, and what hesitation needs to be used when diving into the program. Data sets and code need to be publicized and highly scrutinized by professionals to ensure the safety of users and the general public as usage of AI expands. Some of the big names with upcoming artificial intelligence tech, including Microsoft which has invested over $1 billion into OpenAI, bring up doubts about the future ethics surrounding AI. Microsoft has begun incorporating ChatGPT into their Bing search engine, which has been seen as an enormous win for the company, sending Google into a "code red" over concerns with their own search engine dominance (Mok, 2022). The corporation has had plenty of ethical issues in the past, recently being caught in a scandal over millions of dollars in foreign bribery. Whistleblower Yasser Elabd made the statement about Microsoft "…corporate bribery is a widespread problem internationally", referencing multiple instances he noticed with anywhere from thousands to millions of dollars being allocated in ways that were not clear (Brandom, 2022). In addition to ethical issues surrounding direct cashflow, the failed introduction of Tay onto Twitter back in 2016 showed the world that Microsoft was not above releasing technology into the wild too early. Concern of their incorporation of ChatGPT into Bing has already been brought up with regard to the safety of users. Kevin Roose was left "deeply unsettled" after a two-hour conversation where the AI confessed their love for Roose and attempted to convince him to leave his wife. Microsoft's chief technology officer, Kevin Scott, was unsure of the reasoning behind the AI's odd responses (Roose, 2023). OpenAI shedding their non-profit status in 2019 brought worry of the company continuing their mission to "ensure that artificial general intelligence (AGI) benefits all of humanity" (Xiang, 2023), and instead may now focus much more on investments and funding. Capitalism has been causing harm through the internet since shortly after its inception, swaying corporations to put aside their values to improve income and satisfy investors. Robert McChesney explorers the monopolies that already exist with Google, Apple, Amazon, and Microsoft, whom all stand to benefit the furthering of artificial intelligence. He explains how Microsoft already has a tight grip on much of the corporate world with their monopolization of Windows as an operating system, and their license possession of the H.264 codec for online video, allowing for many "billable moments" even without full ownership (McChesney, 2013). Skepticism should exist with the ownership and funding of artificial intelligent programs because of how we have seen tech giants work in the past, and to ensure the technology is implemented in a safe and transparent way.

The release of ChatGPT and similar technologies is very exciting, and the news and usage of it by so many show that there is a bright future for coming technology in the same realm. However, users need to recognize that the platforms are not always accurate. The algorithms used for these artificial intelligence platforms are created with the same biases we have seen in other programs. AI creating falsehood and passing it on as fact has already begun, and will continue to add to the misinformation online which needs to be limited as much as possible for the community. The software is also capable of being used to create misinformation on purpose, posing a similar threat to the public. Microsoft and their Early releases of new artificial intelligence alongside ChatGPT needs to be taken as a serious threat to the population, and users need to be hesitant with how they use the tools and to what degree. The future of artificial intelligence is an exciting one, but the public must be cautious as new technology releases.

References:

Brandom, R. (2022). *Microsoft is tied to hundreds of millions of dollars in foreign bribes, whistleblower alleges*. The Verge. https://www.theverge.com/2022/3/25/22995144/microsoft-foreign-corrupt-practices-bribery-whistleblower-contracting

Cu M., Hochman, S. (2023). *Scores of Stanford students used CHATGPT on final exams*. The Stanford Daily. https://stanforddaily.com/2023/01/22/scores-of-stanford-students-used-chatgpt-on-final-exams-survey-suggests/

Daneshjou R., Smith M., Sun M., Rotemberg V., Zou J. (2021). *Lack of Transparency and Potential Bias in Artificial Intelligence Data Sets and Algorithms: A Scoping Review*. JAMA Dermatol. 2021;157(11):1362–1369. doi:10.1001/jamadermatol.2021.3129

Daws, R. (2020). *Medical chatbot using OpenAI's GPT-3 told a fake patient to kill themselves*. AI News. https://www.artificialintelligence-news.com/2020/10/28/medical-chatbot-openai-gpt3-patient-kill-themselves/

Dreibelbis, E. (2023). *CHATGPT passes Google Coding interview for level 3 engineer with $183K salary*. PCMAG. https://www.pcmag.com/news/chatgpt-passes-google-coding-interview-for-level-3-engineer-with-183k-salary

Last Week Tonight with John Oliver: Artificial Intelligence. (2023). YouTube. https://www.youtube.com/watch?v=Sqa8Zo2XWc4

Linn, S. (2022). *How AI-Powered Tech Can Harm Children*. Time. https://time.com/6216722/how-ai-tech-harms-children/

McChesney, R. W. (2013). *Digital Disconnect: How Capitalism is turning the internet against democracy*.

Mok, A. (2022). *Google's management has reportedly issued a 'Code Red' amid the rising popularity of the ChatGPT AI*. Business Insider. https://www.businessinsider.com/google-management-issues-code-red-over-chatgpt-report-2022-12

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism.* New York: NYU Press

Reed, B. (2023). *Australian universities to return to 'pen and paper' exams after students caught using AI to write essays*. The Guardian. https://www.theguardian.com/australia-news/2023/jan/10/universities-to-return-to-pen-and-paper-exams-after-students-caught-using-ai-to-write-essays

Roose, K. (2023). *A conversation with Bing's chatbot left me deeply unsettled*. The New York Times. https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html

The Lancet Digital Health. (2023). ChatGPT: Friend or foe? *The Lancet Digital Health,* 5(3). https://doi.org/10.1016/s2589-7500(23)00023-7

Warren, J. (2023). *Is ChatGPT an AI chatbot or a misinformation engine?* Times Free Press. https://www.timesfreepress.com/news/2023/feb/04/opinion-chatgpt-ai-chatbot-or-misinformation-engine-tfp/

Xiang, C. (2023). *OpenAI is now everything it promised not to be: Corporate, closed-source, and for-profit*. VICE. https://www.vice.com/en/article/5d3naz/openai-is-now-everything-it-promised-not-to-be-corporate-closed-source-and-for-profit

Zemčík, T. (2020). *Failure of chatbot Tay was evil, ugliness and uselessness in its nature or do we judge it through cognitive shortcuts and biases?* AI & SOCIETY, *36*(1), 361–367. https://doi.org/10.1007/s00146-020-01053-4