Exploring the Engineering Ethicality of ChatGPT

With the launch of ChatGPT on November 2022, its impact on current and future education has become gamified and questionable. Without OpenAI engineers factoring in the need for proper acknowledgment of identified sources on which ChatGPT bases its answers, the use of ChatGPT in engineering and research raises several ethical concerns for its engineers and its users. Issues include plagiarism, bias, misinformation, and undermining the credit of authors with fabricated citations. Due to the large misconception around ChatGPT’s ability to generate immediate responses by recognizing substantial patterns and datasets of internet sources, this paper will evaluate how ChatGPT raises ethical concerns related to plagiarism and misinformation due to the failures of prompt engineers in creating clear and consistent policies that practice generating accurately cited sources in artificial intelligence (AI).

 ChatGPT, Chat Generative Pre-training Transformer, is a language model that operates on machine learning. Trained to digest large text data, it replicates detailed conversational and educational writing styles promptly. With its versatile and convenient features, a study by the Mesopotamian Journal of Computer Science reports that “70% of its users” favored it over traditional citation methods, suggesting that AI-generated writing or chatbots could be the new preferred mode of writing in the future [1]. This popularity extends to college students as Forbes reports that 43% of students have had experience using ChatGPT [2]. With over 100 million users in the first five days of launch and over 10 million queries per day, ChatGPT’s relevance and popularity have surpassed the launch records of popular companies such as Facebook, Instagram, Spotify, and Netflix [3]. However, with its undeniable popularity and widespread usage, users and OpenAI engineers must think about who is accountable for misusage and digesting such information to avoid unintentional plagiarism.

 Prompt engineering deals with a user and an AI interacting in a text conversation to complete a task. AIs like ChatGPT use prompt engineering to output a human-like response based on the user’s prompt. However, one major shortcoming of prompt engineering is that it compresses copious sources of information, resulting in inaccurate or misleading answers to users' questions. This ultimately leads to unethical citing circumstances that need to be considered. ChatGPT generated errors prose concerns in plagiarism as it can generate “incorrect, incomplete, misleading, biased, harmful information, instructions, or content” [1]. Additionally, ChatGPT outputs fabricated citations, presenting them as real, which undermines the credibility of the author but also the ethicality of the engineers who designed it. Because ChatGPT does not have the ability of a human to analyze and vet all outputs, this results in ethical violations.

 Plagiarism in college is a serious academic violation, resulting in receiving a zero in a course or even expulsion. A situation that both students and educators want to avoid, the introduction of ChatGPT has shaken up the way people approach plagiarism and academic dishonesty. Implemented in our education system, educators encounter plagiarism with “over 89% of students admit[ing]” they use it for assignments [5]. Even within USC, the Viterbi Code of Integrity for Engineers outlines a standard expected from its students, with plagiarism as the most common example of academic misconduct. This includes “submission of material authored by another” but presented as your own and “improper acknowledgment of sources” [6]. It underscores that even unintentional plagiarism or inaccurate citations are not acceptable excuses. Thus, ChatGPT's failure to adhere to these standards violates the code of ethics for engineers but also opens a need for new policies disciplining students who abuse this public resource.

To promote the ethical fairness of ChatGPT, engineers and users must collaboratively think about accountability and transparency. The Code of Ethics for Engineers outlined by the National Society of Professional Engineers asserts the need for honest, equal, standardized, and professional behavior that universalizes ethical conduct in the engineering community. This certainly applies to prompt engineers in ChatGPT where engineers must be accountable when presenting information that deceives the readers. Any misrepresentations of a fact or hiding proper citations leads to serious consequences, flagged as plagiarism which may damage their reputation for violating the ethical principles of engineers. In any case, plagiarism, especially intentional, is unethical and leads to professional and criminal charges [4].

The code also asserts that engineers must not share information without consent, especially if it is confidential information, and give credit to work that is not theirs, “recognizing the proprietary interests of others.” Within the plagiarized text such as in fabricated sources of ChatGPT, it is important to obtain proper permissions, especially when dealing with sensitive or confidential material, from the original author. Even summarizing, rewording, or quoting texts without permission of confidential texts is considered unethical and against the Code of Ethics for Engineers. Thus, to avoid damage to one’s reputation, legal consequences, professional consequences, and ethical consequences, prompt engineers should implement greater awareness by either developing a warning or also programming in-text citations from where it pulls its sources. With these initial measures, the engineers will practice giving credit and encourage users to use ChatGPT as a guideline or a suggestive reference that humans should fact-check [4].

Apart from plagiarism, there are additional unethical engineering violations that ChatGPT brings. For example, ChatGPT provides outdated information as it contains a database of information up to 2021. With information and technology changing constantly, engineers should consider updating ChatGPT’s inaccurate or outdated information that is presented as true. This results in misinformation as users may interpret the information as true or still accurate. While ChatGPT does acknowledge itself as a model that contains information up to 2021, it believes that any year beyond that is the future – which is not the case. This is problematic as ChatGPT can perpetuate biased information without appropriate assessment. One solution to reduce this is considering bias mitigation. This includes users and engineers “actively working to identify and address biases” in AI models like ChatGPT through a series of tests and monitoring its consistency [1]. By doing so, prompt engineers can ensure that the content presented by ChatGPT is “fair” and “does not perpetuate existing biases” [1].

Apart from implementing bias mitigation, another solution includes determining accountability and transparency. One of the public professional obligations outlined by the Code of Ethics for Engineers is that “Engineers shall be guided in their relations by the highest standards of honesty and integrity” [4]. This includes admitting errors and not falsifying facts. To ensure accountability and transparency, engineers should acknowledge the faulty design of ChatGPT to ensure that its users understand the shortcomings of the AI. Because it is unrealistic to completely ban or punish the engineers for the AI's shortcomings, engineers must “strive to serve the public interest” and participate in broadening public knowledge [4].

To reduce the dissemination of false information in AI writing, it is quintessential for engineers to offer AI plagiarism checkers. OpenAI has recently launched its OpenAI classifier, which is trained to detect human and AI text. It notes that the classifier “is not fully reliable” and contains limitations, admitting that “wrong prediction[s]” may occur [7]. This resource is focused primarily from an educator standpoint, to serve as a guideline to detect plagiarism. Furthermore, OpenAI is encouraging anyone impacted by issues of credibility or academic dishonesty, there is a feedback form on their website for the engineers to adjust. This shows that the engineers of OpenAI are making an effort to fix their shortcomings and offer alternatives for educators and those who are negatively affected by this resource.

The lack of AI policies leads to unethical and unintentional breaches of the Code of Ethics for Engineers. While it is stated by the National Society of Professional Engineers (NSPE) that “engineers shall avoid all conduct or practice that deceives the public,” this has been a challenge to maintain with automated systems that put public rights in danger [4]. Thus, changes in Congress include altering the Bill of Rights, which now applies to AIs that have the potential to “impact the public’s rights, opportunities, or access to critical resources or serves” [8]. This means that the engineers of ChatGPT must ensure that there is equitable access to the resource – even its critical resources – and ensure “civil rights, liberties, and privacy” [8]. While there is a premium option to obtain ChatGPT’s faster response, upgrades, and priority in peak times, the information presented is still accessible to the public. Along with the premium, there are plugin options to help users and developers filter outdated and fabricated sources to a certain extent. However, there are still limitations within AI writing in terms of plagiarism as it is a relatively new and developing project that is continually adjusted based on feedback. Overall, standardized policies paired with assistive tools could clarify unintentional plagiarism and misinformation.

 While there is no concrete solution at this time to combat plagiarism in AI writing, there are a few promising solutions that ChatGPT and other AI writing software can implement. One idea includes implementing a detection system based on “perplexity and burstiness” [9]. The binary distinguishing system between humans and AI creates errors, however, evaluating the text’s perplexity and burstiness measures how complex and varying a text is. By examining the text’s homogeneity, predictability, and overall structure, educators and students can utilize this tool to avoid plagiarism, especially in colleges where this issue is prevalent.

 The prevalence of plagiarism in education must be recognized given the development of technologies that can both write and even detect it. It is critical that engineers address concerns of plagiarism, bias, misinformation, and compromising the credit of writers with services that give the ability to identify AI plagiarism. This is because AIs like ChatGPT occasionally present misleading information and even create fake material. Although the OpenAI engineers have made an effort by developing AI text recognition software, the issue of improper identity sources recognition, upon which ChatGPT bases its answer, continues to exist. In order to lessen disinformation and raise the ethical standards of prompt engineers who design AIs, it is crucial to address and raise awareness of the unethical types of plagiarism it introduces.

Works Cited

[1] P. Rivas and L. Zhao, “Marketing with ChatGPT: Navigating the Ethical Terrain of GPT-Based Chatbot Technology,” AI, vol. 4, no. 2, pp. 375–384, Apr. 2023, doi: 10.3390/ai4020019.

[2] M. T. Nietzel, “More than half of college students believe using chatgpt to complete assignments is cheating,” Forbes, 21-Mar-2023. [Online]. Available: https://www.forbes.com/sites/michaeltnietzel/2023/03/20/more-than-half-of-college-students-believe-using-chatgpt-to-complete-assignments-is-cheating/?sh=5822268b18f9.

[3] Nerdynav, “91 important CHATGPT Statistics & user numbers in April 2023 (GPT-4, plugins update),” *Nerdy Nav*, 16-Apr-2023. [Online]. Available: https://nerdynav.com/chatgpt-statistics/#:~:text=ChatGPT%20User%20Stats,-Platform&text=ChatGPT%20has%20become%20the%20fastest,in%202%20months%20of%20launch.&text=According%20to%20Similarweb%2C%20the%20ChatGPT,gets%2025%20million%20daily%20visits.

[4] “NSPE Code of Ethics for Engineers,” *Code of Ethics | National Society of Professional Engineers*, 2019. [Online]. Available: <https://www.nspe.org/resources/ethics/code-ethics?gclid=Cj0KCQjw27mhBhC9ARIsAIFsETFC2dcipPtgC5OmhP2iLDiQddhwHtvJon57Nc3Y6BvcyySOn8VB9lEaAsqWEALw_wcB>.

[5] C. Westfall, “Educators battle plagiarism as 89% of students admit to using OpenAI's chatgpt for homework,” *Forbes*, 30-Jan-2023. [Online]. Available: <https://www.forbes.com/sites/chriswestfall/2023/01/28/educators-battle-plagiarism-as-89-of-students-admit-to-using-open-ais-chatgpt-for-homework/?sh=1a92c835750d>.

[6] “Academic integrity - USC Viterbi: School of Engineering,” *USC Viterbi | School of Engineering*, 11-Nov-2022. [Online]. Available: <https://viterbischool.usc.edu/academic-integrity/>.

[7] “New Ai classifier for indicating AI-written text,” *New AI classifier for indicating AI-written text*. [Online]. Available: <https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text>.

[8] “Blueprint for an AI bill of rights - OSTP,” *The White House*, 16-Mar-2023. [Online]. Available: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>.

[9] V. Izzo, “Best plagiarism checkers for AI-generated content,” *WordLift Blog*, 17-Apr-2023. [Online]. Available: <https://wordlift.io/blog/en/best-plagiarism-checkers-for-ai-generated-content/>.