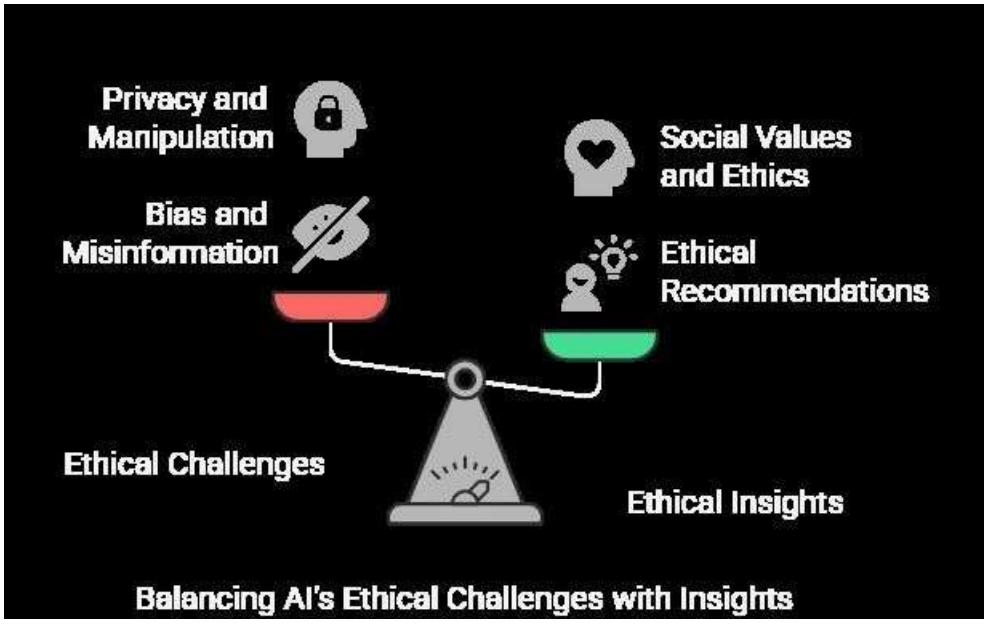
SJIF Impact Factor (2025): 8.688| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 10 | Issue: 5 | May 2025 - Peer Reviewed Journal

ETHICS IN AI CHATBOTS: LESSONS FROM CHATGPT

Ruchi Rai, Afsar, Neeraj Verma, Himanshu Tomar, Ritik Kumar
Computer Science and Engineering, Uttar Pradesh, India
Indraprastha Institute of Management and Technology, Saharanpur, Uttar Pradesh, India

ABSTRACT
The advent of AI chatbots, especially OpenAI's ChatGPT, has shifted the framework of human-computer interaction in virtually every sector. This progress comes with new sociotechnical challenges. This research analyzes the ethical issues of AI chatbots, particularly through the prism of ChatGPT, focusing on bias, misinformation, privacy data, user manipulation, and accountability. It examines relevant literature, case studies, and OpenAI's policies regarding user data and autonomous systems, stressing the development of appropriate artificial intelligence systems that honor social values and human ethics. In its concluding sections, it offers ethical insights and recommendations for the betterment of future chatbots.

KEYWORDS—*AI Ethics, ChatGPT, AI Chatbots, Bias in AI, Misinformation, Privacy and Data Security, AI in Education, Responsible AI, Policy and AI Governance.*



SJIF Impact Factor (2025): 8.688| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 10 | Issue: 5 | May 2025 - Peer Reviewed Journal

1. INTRODUCTION

AI chatbots like ChatGPT can now communicate in a coherent and contextually relevant manner. They aid in customer support, educational services, healthcare, and creative writing. While the functions of these chatbots are tremendously helpful, their growing proficiency brings sobering debates concerning ethical considerations of autonomy, privacy, and trust. With these chatbots interacting with millions of users every day, the burden of ensuring ethical design is monumental. This paper uses ChatGPT to consider broader ethical questions about the design and use of chatbots.

2. LITERATURE REVIEW

The earliest considerations of chatbot ethics centered on misinformation, anthropomorphizing the bots, and biases within the training data. Binns et al. (2018) reviewed some aspects of algorithmic fairness, while Floridi and Cowls (2019) provided rationale on ethical AI practices, highlighting principles of beneficence, non-maleficence, and accountability. More recent literature examines constraints of ChatGPT, including OpenAI's admission of it "hallucinating" facts, data biases, and potential misuse by bad actors. Collectively, these works indicate concern towards the effectiveness of ethics in AI.

3. METHODOLOGY

This paper is based on qualitative research, including peer-reviewed literature, OpenAI's documents, and case examples from the real world. It also integrates policy analysis from ethical frameworks provided by UNESCO, IEEE, the EU, and other leading bodies on AI governance.

4. ETHICAL ISSUES CONCERNING AI CHATBOTS

4.1 Bias and Inequity

The AI model used in ChatGPT receives training on large portions of the internet, which contains explicit information about societal injustices. The model has the potential to foster discrimination as it might repeat negative stereotypes concerning race, gender, religion, or ethnicity.

4.2 Disinformation

ChatGPT seeks to provide factual information, but it can also produce content that is false, misleading, or untrue. This becomes a problem in academic or clinical settings where users rely on the chatbot without employing critical thinking skills.

4.3 Information confidentiality and privacy

A growing number of people are sharing their private information and data with chatbots. If revealed, these data can be used to infer an identity or directly target to change someone's actions. Even though OpenAI has taken steps to protect the anonymity of the inputs provided, the dangers remain.

4.4 Freedom and Excessive Control

AI chatbots have the ability to control actions and make decisions in indirect ways. For example, the way these bots handle certain questions can influence what consumers purchase or change their political beliefs. It changes the user's freedom and draws attention to the fact that there must be clarity in what bots say.

4.5 Responsibility and Accountability Successor:

In a proactive AI design, responsibility is always reserved, either for the developer to blame the deployer, or if the blame goes to the user. Responsibility for the chatbot causing damage needs to lie within one of the three. Constructing a meaningful answerability hierarchy is a must concerning making ethical AI.

5. THE REVOLUTION OF AI: CHATGPT IN THE CLASSROOM

Conversations around ChatGPT in classrooms are polarized. For some students, it is an invaluable assistive tool. But for others, it serves as a source of an essay writer and a problem solver. Although the ChatGPT promtes learning for some students, there are issues like academic dishonesty, loss of individuality, and extreme dependency on technology that also come into play, raising major concerns. A proper balance is what is needed here – effective literacy of AI rounds and etiquette rules around it.

SJIF Impact Factor (2025): 8.688| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 10 | Issue: 5 | May 2025 - Peer Reviewed Journal

6. ETHICAL GUIDELINES AND RECOMMENDATIONS

In light of the issues discussed earlier, the following strategies addressing ethics have been suggested.

1. Enforcing celebration of cognitive diversity by changing and adding underrepresented voices to the training dataset.
2. Adding transparency regarding AI users by communicating that they are indeed talking with a chatbot.
3. Cultivating critical minds by teaching concepts of digital literacy alongside AI.
4. Taking extra safety measures regarding conversations that are personal and not allowing them to be saved in any way.
5. Including specialists in ethics and education, as well as users of the applications in the formative stages of development."

7. CONCLUSION

From learning and communication to productivity, AI chatbots like ChatGPT promise great value. However, these tools may unintentionally hurt users or perpetuate biases without ethical guidelines. Analyzing the gains and gaps of ChatGPT reveals lessons for developing AI that is intelligent but also responsible, just, and human-focused. The unarguable, escorted future of AI ethics does not lie within the technology, but rather in how we intend to build and use it.

8. REFERENCES

1. *Binns, R. et al. (2018). "Fairness in Machine Learning." ACM Computing Surveys.*
2. *Floridi, L. & Cowls, J. (2019). "AI In Society: A Unified Framework Of Five Principles For AI In Society." Harvard Data Science Review.*
3. *AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). Chatgpt as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. Journal of Artificial Intelligence and Technology.*
4. *Aron, J. (2011). How innovative is Apple's new voice assistant. Siri, NewScientist, 212(2836), 24*
5. *IEEE. (2020). Aligned Ethically Designed: A Vision for Autonomous and Intelligent Systems for Humanity.*
6. *Cooper, G. (2023). Examining science education in chatgpt: An exploratory study of generative artificial intelligence. Journal of Science Education and Technology.*
7. *Fauzi, F., Tuhuteru, L., Sampe, F., Ausat, A. M. A., & Hatta, H. R. (2023). Analysing the role of ChatGPT in improving student productivity in higher education. Journal on Education.*
8. *Lally, A., & Fodor, P. (2011). Natural language processing with prolog in the ibm watson system. The Association for Logic Programming (ALP) Newsletter, 9, 2011.*