Gender Issues and Solutions in the ChatGPT AI Tool

Amal Mersni
Computer Sciences and
Engineering Program
International University of
Sarajevo
Sarajevo, BiH
amersni@ius.edu.ba

Oleksandra Yeremenko
V.V. Popovskyy Department of
Infocommunication Engineering
Kharkiv National University of
Radio Electronics
Kharkiv, Ukraine
oyeremenko@acm.org

Mykhailo Persikov
V.V. Popovskyy Department of
Infocommunication Engineering
Kharkiv National University of
Radio Electronics
Kharkiv, Ukraine
mykhailo.persikov@nure.ua

Valentyn Lemeshko
V.V. Popovskyy Department of
Infocommunication Engineering
Kharkiv National University of
Radio Electronics
Kharkiv, Ukraine
valentyn.lemeshko@nure.ua

ABSTRACT

Including women in developing and implementing Artificial Intelligence is crucial to ensure its products are high quality and gender-biased free. However, women continue to be underrepresented and excluded in this field, which can lead to gendered problems when employing AI applications such as ChatGPT. This study aims to identify the most significant gender issues arising when using ChatGPT. Furthermore, it proposes solutions to address the discrimination and inequality women may face when interacting with this AI tool.

CCS CONCEPTS

• Computing Methodologies • Artificial Intelligence

KEYWORDS

Artificial Intelligence, ChatGPT, Gender, Bias, Stereotypes

1 Introduction

Digital technologies powered by artificial intelligence (AI) expand and transform many facets of our lives, from the personal to the professional, at a rapid pace. As we witness the emergence of vast data sets and unprecedented advances in machine learning (ML) and deep learning (DL), we are entering a new era of AI development with immense potential and significant implications for how knowledge is constructed in this field. However, the widespread prevalence of harmful stereotypes and biases in existing data sets poses a substantial obstacle to the development of AI. This is particularly regarding gender imbalances and their consequences.

AI is a broad term for a collection of techniques in computer science that, when used together, give computers the ability to mimic human intelligence in their performance of specific tasks. Some examples of these tasks include natural language comprehension, image recognition, choice analysis, and problem-solving. In addition, numerous types of artificial intelligence technology including chatbots, navigation apps, and portable fitness trackers. Chatbots are a subset of AI that mimics human conversation in written or spoken form [1].

2 ChatGPT Overview

ChatGPT is an AI chatbot that generates essays, codes, and answers questions. OpenAI has secured funding from several investors, with Microsoft being the most prominent. ChatGPT can function because it compiles data from user-generated content on the internet and uses computational forecasts to answer questions and fulfill user requests. As a result, the chatbot learns more about different topics and how to discuss them over time based on textual inputs and data.

In March, OpenAI announced that ChatGPT would be updated, with the new version including support for the GPT-4 infrastructure. This gives ChatGPT a leg up on the competition by allowing it to interact with users in more ways than just text.

A massive Chinese tech company, Alibaba, has released Tongyi Qianwen, an artificial intelligence chatbot. Tongyi, an intelligent assistant that can read and respond to text like ChatGPT, will soon be incorporated into several Alibaba products. Despite the release of competing chatbots like TruthGTP and Jasper Chat and the introduction of Elon Musk's new artificial intelligence company (X.AI), ChatGPT remains the most popular option [2].

As a result of its global success, ChatGPT has received both positive and negative press coverage. In March 2023, for instance, the service was banned in Italy over privacy concerns, but it has

since been restored. ChatGPT has been banned in many countries, but its popularity continues to soar.

3   Gender Bias in ChatGPT

ChatGPT is an invaluable tool that offers information and knowledge to those seeking guidance or understanding on various topics. However, it's important to remember that, like any AI system, ChatGPT has limitations and can occasionally give biased answers. For instance, if ChatGPT is trained on a dataset representing a particular culture's or society's values, it could produce insensitive or even prejudiced responses toward a specific group. Furthermore, ChatGPT may unintentionally sustain certain stereotypes or reinforce existing biases that are prevalent in today's society.

ChatGPT uses a gender and racial-biased natural language processing model (NLP). For example, when Amazon used an NLP tool to review resumes, it found this. The tool used linguistic patterns from recaps to screen candidates and find the best. However, since past outlines had fewer women, the program was unintentionally biased toward men [3].

Director of Women Leading in AI Ivana Bartoletti tweeted that she had commissioned Chat-GPT4 to write a story in which young men and women choose their topics to study at university. Unfortunately, by assuming that a young male would be interested in technology and a young female would prefer the arts, the algorithm's response perpetuated harmful gender stereotypes. This supports the biased view that men excel in STEM while women shine in the arts. In addition, the textual content, some of which was inaccurate or toxic, was used to train an algorithm, which then wrote the story. This incident highlights issues in the training and use of AI algorithms that could reinforce harmful gender stereotypes, even though the algorithm did not intend to do so [4].

In today's world, so much of the technology we rely on is developed using biased data that reflects human prejudices. As a result, we are far from living in a fair and equitable society [5].

4   Solutions to Address the Gender Disparity in ChatGPT

Women are required to shape the future of AI and ML actively. As an emerging field, the AI sector can take steps to close the gender gap before it becomes deeply embedded in the culture of the industry and negatively impacts the future of AI and its applications and tools. Many solutions were taken into consideration to prevent the gender gap in AI. Collaborating on regulation to eliminate discrimination in chatbots like ChatGPT-4 is also emphasized. The co-founder and chief scientist of OpenAI has urged authorities to take action swiftly. In their open letter, the Alliance for Universal Digital Rights (AUDRi) encourages cooperation on governance, transparency, testing frameworks, and the regulations of AI tools. These strategies support the nine Digital Principles developed by AUDRi to ensure everyone's security, privacy, and respect in the digital age. It will be

necessary for governments, policymakers, private companies, and human rights experts to work together to maximize the positive social outcomes of ChatGPT-4 [4]. Once again, it is essential to work together to address the gender biases problems in AI and ChatGPT by:

- Ensuring that the training data used by AI models must be as diverse as possible;
- Increasing the representation of women in AI is also important;
- Considering the gender biases of developers when creating AI;
- Regularly auditing AI models to ensure they are not biased against women;
- Engaging more female Artificial intelligence (AI) researchers to collaborate to eradicate sexism from the ChatGPT platform.

5   Conclusion

Consistent investment in research and development (R&D) will produce AI models with increased sensitivity to gender issues, contributing to a more equitable and just society [6]. Overall, ChatGPT has the potential to affect gender equality significantly and women's empowerment by eliminating barriers to education, increasing female workforce participation, and removing gender bias. However, while AI has great promise, there is still a long way to go before it can solve the world's most challenging problems. Achieving gender parity and empowering women requires collaborative efforts across academic disciplines, industries, and organizations. With the help of AI tools like ChatGPT, we can make the world a better, more just place for everyone. Therefore, it is strongly suggested that women be encouraged to participate actively in AI, ensuring that the technology reflects the diverse range of human experiences and perspectives.

REFERENCES

[1] A. Trivedi, E. K. Kaur, C. Choudhary, Kunal, and P. Barnwal, "Should AI Technologies Replace the Human Jobs?," 2023 2nd International Conference for Innovation in Technology (INOCON), Bangalore, India, 2023, pp. 1-6, doi: 10.1109/INOCON57975.2023.10101202.

[2] M. Abdullah, A. Madain and Y. Jararweh, "ChatGPT: Fundamentals, Applications, and Social Impacts," 2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS), Milan, Italy, 2022, pp. 1-8, doi: 10.1109/SNAMS58071.2022.10062688.

[3] C. M. Parra, M. Gupta, and D. Dennehy, "Likelihood of Questioning AI-Based Recommendations Due to Perceived Racial/Gender Bias," in IEEE Transactions on Technology and Society, vol. 3, no. 1, pp. 41-45, March 2022, doi: 10.1109/TTS.2021.3120303.

[4] Equality Now. (2023, April 15). ChatGPT-4 reinforces sexist stereotypes. Retrieved from https://www.equalitynow.org/news_and_insights/chatgpt-4-reinforces-sexist-stereotypes/

[5] Finextra. (2021, March 8). International Women's Day: ChatGPT and how gender bias in AI can impact innovation. Retrieved from https://www.finextra.com/the-long-read/618/international-womens-day-chatgpt-and-how-gender-bias-in-ai-can-impact-innovation

[6] M. Roopaei, J. Horst, E. Klaas, G. Foster, T. J. Salmon-Stephens and J. Grunow, "Women in AI: Barriers and Solutions," 2021 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 2021, pp. 0497-0503, doi: 10.1109/AIIoT52608.2021.9454202.