Research Inventy: International Journal of Engineering And Science
Vol.13, Issue 7 (July 2023), PP 102-105
Issn (e): 2278-4721, Issn (p):2319-6483, www.researchinventy.com

Research on Software Engineering: The Ethical Aspects of ChatGPT

Subhadra Biswal[1], Jharana Paikray[2], Sarmistha Palai[3]
[1,2,3] Dept. of CSE, Einstein Academy of Technology and Management, Bhubaneswar

Abstract:
ChatGPT provides effective, user-friendly information synthesis and analysis through natural language interactions, which can enhance Software Engineering (SE) research techniques. However, ChatGPT may provide ethical issues with regard to data security, privacy, and plagiarism as well as the possibility of producing skewed or potentially harmful data. By focusing on the essential components—motivators, demotivators, and ethical guidelines for utilizing ChatGPT in SE research—this study seeks to close the existing gap in knowledge. In order to accomplish this goal, we reviewed the literature, determined the aforementioned components, and created a taxonomy to show how they relate to one another. Furthermore, a thorough questionnaire-based survey comprising SE scholars was used to experimentally analyze the identified literature-based features (motivators, demotivators, and ethical principles).To develop a cluster-based decision model, we also performed a Cross-Impact Matrix Multiplication Applied to Classification (MICMAC) analysis. By embracing the motivators and addressing the demotivators, these models seek to assist SE researchers in developing ethical strategies for incorporating ChatGPT into their research while adhering to the established standards. With a focus on ethical issues, the study's conclusions will set a standard for using ChatGPT services in SE research.

I. INTRODUCTION

ChatGPT is a cutting-edge language model created by OpenAI [1], designed to generate human-like responses to various prompts. The model employs deep learning algorithms, utilizing the latest techniques in Natural Language Processing (NLP) to generate relevant and coherent responses. GPT, or "Generative Pre-trained Transformer" refers to the model's architecture based on the transformer architecture and pretrained on a vast corpus of textual data [2]. ChatGPTbeen fine-tuned on conversational data, allowing it to generate appropriate and engaging responses in a dialogue context [1], [3]. The model's versatility means that it can be applied to numerous applications, including chatbots, virtual assistants, customer service, and automated content creation. The OpenAI team continues to update and improve the model with the latest data and training techniques, ensuring it remains at the forefront of NLP research and development [4]. ChatGPT has significant potential for use in academic research [5], particularly for performing SE activities [6].

Researchers can utilize ChatGPT to generate realistic and high-quality text for various applications, including language generation, language understanding, dialogue systems, and experts' opinion transcripts [7]. ChatGPT can also be fine-tuned for specific domains or tasks, making it a flexible tool for researchers to create customized language models [8]. In addition, ChatGPT can be used to generate synthetic data for training other models, and its performance can be evaluated against human-generated data. Moreover, ChatGPT can be used for research on social and cultural phenomena related to language use. For example, researchers can use ChatGPT to simulate conversations and interactions between people with different cultural backgrounds or to investigate the impact of linguistic factors such as dialect, jargon, or slang on language understanding and generation [9].

ChatGPT significantly impacts research, particularly in qualitative research using NLP tools. Its ability to generate high-quality responses has made it a valuable tool for language generation, understanding, and dialogue systems [10]. Researchers can leverage ChatGPT to save time and resources, create customized language models, and fine-tune for specific domains or tasks [10]. ChatGPT's simulation capabilities also allow researchers to understand natural language in different contexts and develop more nuanced language models [9], [11]. Overall, ChatGPT has advanced the field of NLP and paved the way for more advanced language models and applications [12]. ChatGPT behaves as a smart, intelligent, and effective tool for SE research [13]–[15]. For instance, the ChatGPT can be used in literature review-based research to extract data by giving specific queries and related text in quotes. Similarly, we noticed that the ChatGPT is also an effective tool for generating the codes, concepts, and categories from transcripts in qualitative research [16].

Considering the effectiveness andusability of ChatGPT in academic research, we conducted this study (1) to explore and understand the motivators (positive factors) and demotivators (negatively influencing factors) across the ethical aspects (principles) of ChatGPT in SE research and (2) to develop Interpretive

Structure Modelling (ISM) and Cross-Impact Matrix Multiplication Applied to Classification (MICMAC) based decision-making models in order to understand the relationships between ethical principles for using ChatGPT in SE research. We believe that the outcomes of this research will benefit the academic research community by providing a body of knowledge and serving as guidelines for considering ChatGPT in SE research.

II. Literature Survey

To identify the motivators, demotivators, and principles associated with the ethical use of ChatGPT in SE research, we conducted a literature survey, examining both peer-reviewed published articles and grey literature [18], [19]. Using the common keywords, we explored the grey literature across general Google search and Google Scholar to investigate peerreviewed literature studies. Furthermore, we employed the snowballing data sampling approach to collect potential literature material related to the study objective [20]. This involved examining reference sections of selected studies (backward snowballing) and citations (forward snowballing), resulting in increasing the sample size by including more relevant studies [20].

The questionnaire survey is an appropriate approach to collect the data from a large and targeted population [21]. In this study, we designed a survey questionnaire to validate the identified motivators, demotivators, and principles for evaluating the ethical implications of ChatGPT in SE research. We divided the questionnaire into two parts. The first part focuses on the demographics of survey participants, while the second part consists of the identified motivators, demotivators, and principles. We used the five-point Likert scale (strongly agree, agree, neutral, disagree, and strongly disagree) to encapsulate the opinions of the targeted population. The second part of the questionnaire also includes an open-ended question, enabling participants to suggest any additional motivators, demotivators, or principles overlooked during the literature survey. To reach the target population, we developed an online questionnaire using Google Forms and sent invitations via personal email, organizational email, and LinkedIn. We employed the snowball sampling approach to collect a representative data sample by encouraging participants to share the questionnaire across their research network. Snowball sampling is efficient, cost-effective, and suitable for large, dispersed target populations [20]. Data collection took place from 15 January to 25 April 2023, returning 121 responses, of which 113 were used for further analysis after removing eight incomplete responses. We used the frequency analysis approach to analyze the collected data, which is appropriate for the descriptive type of data analysis [22]. This approach compares survey variables and computes the agreement level among participants based on the selected Likert scale. Frequency analysis has also been used in other software engineering studies [23], [24].

A Internal Validity

Internal validity is the degree to which the results of observation — namely, the causal relationships — are trustworthy and not influenced by other factors or biases. The potential internal validity threat in this study is the understandability and interpretation of the survey content. The survey respondentsmay have a different understanding of the survey questions, which could bias the responses. To mitigate this threat, we piloted the instrument, seeking feedback from SE researchers to enhance the clarity and readability of the survey content prior to its final distribution.

B. External Validity

External validity is the extent to which the results of a study can be generalized or applied to other situations, populations, or settings. In this study, the questionnaire data were collected from 113 researchers, which may not be representative of the broader SE research community. This could limit the generalizability of the findings. Nonetheless, we gathered 113 valid responses from 19 countries across five different continents. The survey participants had a diverse range of experience, fulfilled various roles in different projects, and worked in research teams of differing sizes (see Figure 3). We agree that the study findings could not be generalized to a larger scale; however, based on the details demographics of the survey participants, the overall results could be generalized to some extent.

C. Construct Validity

Construct validity refers to the degree to which a test or experiment measures what it claims to be measuring. In this study, the constructs such as" motivators,"" demotivators," and" ethical principles" may not have been defined clearly enough, leading to potential misinterpretation. However, we mitigated this threat by defining and elaborating on the mentioned constructs based on the literature survey. The identified "motivators", "demotivators", and "ethical principles" are comprehensively discussed in Section II-A. Moreover, the survey questionnaire was piloted based on the expert's opinion to improve the interpretations of the survey variables (constructs).

D. Conclusion Validity

Conclusion validity is concerned with the relationship between the treatment and the outcome and whether any observed effect in the data is real or not. One possible threat to the conclusion validity is that with only 113 respondents, the statistical power may be insufficient to detect meaningful differences or relationships. However, based on the existing relevant studies and the novelty of the research field, the given sample size is strong enough to draw the study's conclusions. Moreover, we plan to extend this study by widening the pool of potential respondents, extending the data collection period, and using different methods to reach the potential population (see Section VI-B). Finally, all the authors were invited to participate in the brainstorming sessions to collaboratively dissect the primary findings and formulate definitive conclusions.

III. CONCLUSIONS AND FUTURE

PLANS We will now present a summary of the conclusions drawn from the study findings, along with a detailed roadmap outlining potential avenues for future exploration.ChatGPT enhances efficiency in knowledge extraction and collaboration within SE research. Its capacity to produce realistic and contextually appropriate language renders it an attractive tool for use in this research field. However, ethical concerns such as plagiarism, privacy, data security, and the risk of generating biased or harmful data must be addressed. This study explores the motivators, demotivators, and ethical principles associated with using ChatGPT in SE research. We conducted a literature survey, identified 17 ethical principles and their corresponding 14 motivators and 12 demotivators for using ChatGPT in SE research, as detailed in Section III-A. These motivators and demotivators were subsequently mapped to the 17 identified principles. The principles highlight crucial areas that the SE research community must consider in order to conduct ethically responsible research. The associated motivators represent factors that can support adherence to these principles. Conversely, demotivators are factors that may obstruct the consideration of ethical principles when using ChatGPT in SE research.

REFERENCES

[1]. T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell et al., "Language models are few-shot learners," Advances in neural information processing systems, vol. 33, pp. 1877–1901, 2020.

[2]. S. S. Sohail, F. Farhat, Y. Himeur, M. Nadeem, D. Ø. Madsen, Y. Singh, S. Atalla, and W. Mansoor, "The future of gpt: A taxonomy of existing chatgpt research, current challenges, and possible future directions," Current Challenges, and Possible Future Directions (April 8, 2023), 2023.

[3]. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," Advances in neural information processing systems, vol. 30, 2017.

[4]. G. Brockman, A. Eleti, E. Georges, J. Jang, L. Kilpatrick, R. Lim, L. Miller, and M. Pokrass. (2023) Improving language understanding with the gpt-3.5-turbo. openai. [Online]. Available: https://platform. openai.com/docs/guides/gpt-3.5-turbo, Accessedon22ndofMarch2023

[5]. U. Bukar, M. S. Sayeed, S. F. A. Razak, S. Yogarayan, and O. A. Amodu, "Text analysis of chatgpt as a tool for academic progress or exploitation," Available at SSRN 4381394.

[6]. J. White, S. Hays, Q. Fu, J. Spencer-Smith, and D. C. Schmidt, "Chatgpt prompt patterns for improving code quality, refactoring, requirements elicitation, and software design," arXiv preprint arXiv:2303.07839, 2023.

[7]. M. Perkins, "Academic integrity considerations of ai large language models in the post-pandemic era: Chatgpt and beyond," Journal of University Teaching & Learning Practice, vol. 20, no. 2, p. 07, 2023.

[8]. Y. K. Dwivedi, N. Kshetri, L. Hughes, E. L. Slade, A. Jeyaraj, A. K. Kar, A. M. Baabdullah, A. Koohang, V. Raghavan, M. Ahuja et al., "So what if chatgpt wrote it?" multidisciplinary perspectives on opportunities, challenges and implications of generative conversational ai for research, practice and policy," International Journal of Information Management, vol. 71, p. 102642, 2023.

[9]. M. Neumann, M. Rauschenberger, and E.-M. Schon, "We need to talk ¨ about chatgpt": The future of ai and higher education," 2023.

[10]. E. A. van Dis, J. Bollen, W. Zuidema, R. van Rooij, and C. L. Bockting, "Chatgpt: five priorities for research," Nature, vol. 614, no. 7947, pp. 224–226, 2023.

[11]. D. Mhlanga, "Open ai in education, the responsible and ethical use of chatgpt towards lifelong learning," Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning (February 11, 2023), 2023.

[12]. H. Du, S. Teng, H. Chen, J. Ma, X. Wang, C. Gou, B. Li, S. Ma, Q. Miao, X. Na et al., "Chat with chatgpt on intelligent vehicles: An ieeetiv perspective," IEEE Transactions on Intelligent Vehicles, DOI: 10.1109/TIV.2023.3253281, 2023.

[13]. M. M Alshater, "Exploring the role of artificial intelligence in enhancing academic performance: A case study of chatgpt," Available at SSRN, 2022.

[14]. L. Bishop, "A computer wrote this paper: What chatgpt means for education, research, and writing," Research, and Writing (January 26, 2023), 2023.

[15]. B. D. Lund, T. Wang, N. R. Mannuru, B. Nie, S. Shimray, and Z. Wang, "Chatgpt and a new academic reality: Artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing," Journal of the Association for Information Science and Technology, https://doi.org/10.1002/asi.24750.

[16]. B. Mesec, "The language model of artificial inteligencechatgpt-a tool of qualitative analysis of texts," 2023.

[17]. M. Sallam, "The utility of chatgpt as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations," medRxiv, pp. 2023– 02, 2023.

[18]. M. A. Akbar, W. Naveed, A. A. Alsanad, L. Alsuwaidan, A. Alsanad, A. Gumaei, M. Shafiq, and M. T. Riaz, "Requirements change management challenges of global software development: An empirical investigation," ieee access, vol. 8, pp. 203 070–203 085, 2020.

[19]. M. A. Akbar, M. Shameem, A. A. Khan, M. Nadeem, A. Alsanad, and A. Gumaei, "A fuzzy analytical hierarchy process to prioritize the success factors of requirement change management in global software development," Journal of Software: Evolution

and Process, vol. 33, no. 2, p. e2292, 2021.
[20]. C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in Proceedings of the 18th international conference on evaluation and assessment in software engineering, 2014, pp. 1–10.
[21]. M. L. Patten, Questionnaire research: A practical guide. routledge, 2016.
[22]. A. Von Eye, Configural frequency analysis: Methods, models, and applications. Psychology Press, 2003.
[23]. M. Niazi, S. Mahmood, M. Alshayeb, A. M. Qureshi, K. Faisal, and N. Cerpa, "Toward successful project management in global softwaredevelopment," International Journal of Project Management, vol. 34, no. 8, pp. 1553–1567, 2016.