

EXPLORING THE IMPACT OF ChatGPT ON ACADEMIC WRITING: A SCIENTOMETRIC ANALYSIS

Ecaterina COMAN

Lecturer, Ph.D., „Transilvania” University of Brasov (Romania),

E-mail: ecaterina.coman@unitbv.ro

Abstract: *The use of ChatGPT in academic writing is a hot topic at the moment, being seen as a threat to academic integrity. It is used both by students in completing assignments and by researchers in writing articles. The aim of this paper is to explore the impact of ChatGPT on academic writing, by analysing the existing literature related to this topic, and to identify which are the research directions in this field. To achieve the purpose of this paper, a scientometric analysis was performed, using the Web of Science database. The results reveal three research directions: 1) The implications of using ChatGPT in academic writing: benefits and limitations; 2) AI-generated content detection methods and tools used to prevent cheating, and 3) Establishing generally accepted ethical standards for using ChatGPT in academic writing. The analysis of the relevant articles included in these clusters highlights the impact of ChatGPT on academic writing, both its potential benefits and technical limitations, but also the ethical challenges it raises. This study calls for a collective effort to establish generally accepted ethical standards and discuss evaluation methods, which should be adapted as AI technology continues to advance.*

Keywords: ChatGPT, Academic Writing, AI-generated content, ethical standards, scientometric analysis

1. Introduction

ChatGPT is a chatbot based on the GPT (Generative Pretrained Transformer) model, that is trained to understand natural language and generate human-understanding texts based on prompts (Cheng, et al., 2023). OpenAI first published GPT in 2018 (Radford, 2018), and since then, its learning parameters, as well as its neural network, have expanded, enabling it to generate texts of progressively higher quality (Floridi and Chiriatti, 2020). ChatGPT is a variant of the GPT-3 artificial intelligence language model, and it was officially launched by OpenAI on November 30, 2022 (<https://openai.com/>). Although there have been concerns regarding the potential misuse of GPT-3 to create 'fake news' and manipulate public opinion (Floridi and Chiriatti, 2020), the launch of Chatbot, which is very easy to use, has raised a series of concerns in the academic community, being a hot topic now. Large Language Models (LLMs) are considered a potential menace to academic integrity at all levels, on the one hand because they help students to cheat on assignments, and on the other hand, they facilitate plagiarism in scientific papers (Williams and Fadda, 2023). Primarily because people are always trying to find easier and more efficient ways to accomplish certain tasks, and in the case of academic writing - to write articles and various projects. But how should teachers and publishers react in this case? The point is that there are still no generally valid rules governing the use of ChatGPT by students and researchers. Banning its use is not a solution because the world is evolving, but it is important to be aware that there are certain limits or even threats that could affect the quality of scientific production in the coming years. The aim of this paper is to analyse the existing literature related to the application of ChatGPT in academic writing, to identify which are the research directions in this field, and to highlight the advantages and disadvantages of its use by students and researchers.

2. Materials and Methods

To achieve the purpose of this paper, a scientometric analysis was performed, using the Web of Science (WoS) database. As ChatGPT was launched at the end of last year, it was not necessary to set a specific time period regarding the publication date of the articles, with

all identified articles being published in 2023. Articles containing the terms "ChatGPT" and "academic writing" in the title or abstract were searched. These terms were used to ensure that the resulting articles would refer to the application of ChatGPT in academic writing and not in other domains. The search was carried out on September 20, 2023 and only 56 works were identified, of which only 52 articles were included in the analysis, being excluded from the analysis letters. The raw data was downloaded as *plain text files* from the Web of Science database. The results were analysed using *VOSviewer software* version 1.6.16 through which scientific mapping can be performed to analyse the titles and abstracts of scientific papers. A term co-occurrence map based on text data was created, the chosen counting method was full counting. Subsequently, a term occurrence threshold was applied such that a term must appear in at least 5 different articles to be considered for inclusion in the semantic map. In order to create the semantic map, the software assesses the relationships between terms using the association strength measure and recommends the optimal number of terms to be included in the map. In the case of this research, of 1461 terms identified, 78 are used more than 5 times. For each of 78 terms, a relevance score was calculated and only 60% of them were selected (47 terms). In addition to this analysis, a short review of the literature was carried out, the most relevant articles from each cluster generated by VOSviewer being analysed.

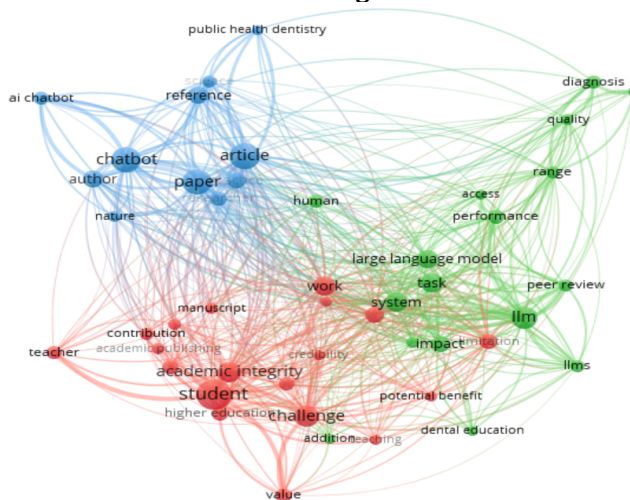
3. Results and Discussions

Three research directions were identified in the studies undertaken related to application of ChatGPT in academic writing, which correspond to the three clusters identified through the scientometric analysis, as can be seen in Figure no. 1. Although there are three distinct clusters, the terms are very interconnected, being many articles that address several aspects related to this topic. The topic's size on the map corresponds to the number of papers associated with that topic. Terms are that frequently appear together in articles are positioned closer to each other.

By analysing the key terms identified, as well as the articles that are associated with each cluster, we can state that the main research directions are:

1. The implications of using ChatGPT in academic writing: benefits and limitations.
2. AI-generated content detection methods and tools used to prevent cheating.
3. Establishing generally accepted ethical standards for using ChatGPT in academic writing.

Figure 1: Static image of the Network Visualization of ChatGPT applications in academic writing



In the following, the relevant articles for each cluster are analysed, identifying the most important aspects specific to each research direction.

4. The implications of using ChatGPT in academic writing: benefits and limitations.

The potential applications of ChatGPT in academia are vast. It is very effective in summarizing information (Xie et al, 2023; Emenike M.E. and Emenike B.U., 2023), feature that can be helpful in the learning process, because AI generates an objective and unbiased summary (Dergaa et al., 2023). However, it may limit the length of certain sections of a paper, without taking into account that certain paragraphs would deserve more detail than others (Lingard, 2023). It is very good at giving structure to complex ideas and can connect them with other ideas. ChatGPT excels as a brainstorming tool for tasks such as generating titles, creating outlines (Lingard, 2023), it can be used to generate a list of potential research questions related to a specific topic (Dergaa et al., 2023), to improve the review process, to enhance metadata, correct grammatical mistakes (Lund et al, 2023) and polish academic papers and presentations to improve readability and language (Cheng et al., 2023). Therefore, Generative AI technologies could improve academic writing and enhance research productivity (Dergaa et al., 2023). But any invention also has negative implications, and ChatGPT as well.

From a qualitative perspective, researchers note that the information provided by ChatGPT is often incorrect (Day, 2023; Williams and Fadda, 2023), even if the results are well written and seem plausible (Howell and Potgieter, 2023). In fact, ChatCPT bases its answers on information from the Internet, which can contain a significant amount of biased, toxic, or erroneous information (Emenike, M.E. and Emenike, B.U., 2023) and, unfortunately, it cannot judge whether the information on which its answers are based is correct or not. Moreover, AIs responses can be influenced by the biases of their developers (Dergaa et al, 2023). Research also shows that ChatGPT fabricates bibliographic resources, generating fictitious citations and references (Day, 2023; Williams and Fadda, 2023; Xie et al, 2023; Ariyaratne et al., 2023). Hence, incorporating AI generated information in research may lead to inaccurate or biased information being inadvertently incorporated into academic papers, which may cause unintentional plagiarism and misattribution of ideas and thus the quality of future scientific production may be affected. Hence, if you are looking for information on a topic that you do not master and you cannot determine whether the information provided is correct, then all the information generated by AI must be double-checked (Lingard, 2023), as in the case of Wikipedia.

As for students using ChatGPT to cheat in exams, that is not a good idea because its performance is low. Williams and Fadda (2023) evaluated ChatGPT's ability to correctly answer various exam-type questions in Carbohydrate Chemistry and Glycobiology. They found that the free version of ChatGPT provided 70% correct answers to the multiple-choice questions, while ChatGPT Plus (the subscription version) only answered 55% of the questions correctly. In terms of true/false questions, both versions of the Chatbot had major difficulties. Instead, it answers well, in a pleasant manner, to descriptive questions about general knowledge. But if several students ask the same question, they may have very similar answers. However, the integration of AI-based technologies into academic learning could help improve the quality of academic outcomes (Hung and Chen, 2023), if it is used “as a learning companion” (Emenike M.E. and Emenike B.U., 2023) to help students better understand concepts, structure information or ideas, counterarguments, etc.

Various limitations derive from the way the prompts for ChatGPT were formulated, as it is essential to define the requirements effectively (Williams and Fadda, 2023). Giray (2023) highlights the common pitfalls of writing prompts, including ambiguity, reinforcement of bias, lack of context and unintended consequences. If the prompt lacks specificity, the resulted output will also be a generalized overview. To have the most accurate results that will be in

line with research objectives, it is necessary to introduce in the prompt distinct parameters and clear guidelines, to eliminate biased language, and to incorporate relevant contextual cues.

5. AI-generated content detection methods and tools used to prevent cheating.

Identifying AI-generated content presents challenges for academic integrity, especially as AI-generated content is increasingly sophisticated. Elkhatat et al (2023) evaluated different paragraphs with AI content detection tools crafted by OpenAI, Writer, Copyleaks, GPTZero, and CrossPlag and found out that these tools showed higher accuracy in recognizing content created by GPT 3.5 compared to GPT 4. Therefore, continuous refinement of AI-generated content detection tools is necessary, especially that in certain cases the instruments presented inconsistencies when human-written paragraphs were analysed. Cingillioglu (2023) states that this has become a challenge and proposes the SVM (support vector machine) algorithm that records 100% accuracy to identify human-generated essays. Even if at this moment the tools for detecting the content generated by AI are good, we must bear in mind that with the development of tools to detect content generated by artificial intelligence, methods to avoid point detection are also being developed, therefore the tools must be continuously reevaluated and improved.

Without access to AI content detection tools, Waltzer et al (2023) claim that high school teachers can tell that certain papers are written with AI assistance in only 70% of cases. To distinguish with the naked eye the work written by an AI from that written by a human, we should look for patterns or anomalies in the language; check the references; verify the accuracy of information; check the grammar and spelling - humans make grammatical errors, unlike robots; analyse if the text is tailored to a specific context or not; verify the originality – if students use similar prompts for a specific task, the content would be very similar (Cotton et al, 2023).

Another way to distinguish between human-written and AI-generated works involves the extraction of linguistic features. Corizzo and Leal-Arenas (2023) use five categories of linguistic features: Text, Repetitiveness, Emotional Semantics, Readability, and Part-of-Speech. Humans use longer sentences and paragraphs than AI, and more punctuation. AI-generated content is characterized by “monotonous writing and lack of narrative and linguistic diversity”. The authors specify that, from a semantic point of view, texts created by people are more subjective, semantic, and emotional. Based on the same principle, Desaire et al (2023) created a model with 20 features based on which it can be determined with 99% accuracy if the author is human or AI. These features are of four types: the complexity of the paragraph; punctuation marks; diversity in sentence length and popular word or numbers. The authors state that human tend to use a more equivocal language, often using words such as: “however”, “but”, “because”, “although”.

Therefore, we can say that there are different methods of identifying if students and researchers misused the chatbot for their papers, but what does „misuse” mean? It is imperative to establish generally accepted ethical standards for using ChatGPT in academic writing, as well as for applying GPT technologies by practitioners in different fields (Cheng et al., 2023). Also, it is necessary to provide training and support for teachers on how to use different methods to detect and prevent cheating (Cotton et al, 2023), as well as on how to guide students in using ChatGPT: how to fact-check the information provided by AI (Day, 2023), how to appropriately incorporate citations and references, and how to optimize the intellectual outputs generated by Chatbot (Hung and Chen, 2023).

6. Establishing generally accepted ethical standards for using ChatGPT in academic writing.

We have been using technology to assist us in the research and article writing process for many years now. Without software such as SPSS, NVivo, Reference Manager or Grammarly,

the process of writing a scientific article would take much longer (Lingard, 2023). Therefore, ChatGPT can also assist us in the research and writing process, but in order to prevent de misuse of ChatGPT in the academic community it is necessary to develop generally accepted policies and procedures. However, this is difficult to achieve considering that AI technology evolves rapidly, and the academic community is always one step behind. Certain policies and rules for the use of generative AI established today may no longer be adequate tomorrow.

Perkins (2023) states that the decision to classify any particular use of LLMs by students as academic misconduct depends on the academic integrity policies of each university, however, in an increasingly globalized world, it is important to have some generally accepted ethical standards. Unfortunately, the research conducted in this direction does not provide clear suggestions on how standards/rules should be established, but only emphasizes the ethical concerns that the use of AI in academia raises. Cheng et al (2023: 594) claim that using Generative AI in academic writing is ethical if “it does not replace key researcher tasks like interpreting data and drawing scientific conclusions.” Altmäe et al (2023) stated that ChatGPT should be seen as a valuable tool that helps authors not get stuck in the writing process and accelerate it, but not actually replace the author's work. On the other hand, Yan (2023) states that it is necessary to reconceptualize plagiarism, because what the students do now by using ChatGPT in their works is not necessarily called plagiarism.

ChatGPT has received authorship recognition in preprints and peer-reviewed published articles. However, it is important to point out that ChatGPT cannot be held responsible for any information provided (Springer Nature Press), and an author has to assume all the responsibility for the correctness of the information and compliance with ethical standards and copywrite legislation. Considering that it is difficult to detect academic dishonesty, some researchers believe that the use of ChatGPT in academic writing should not be prohibited, but its use must be mentioned in Methods or Acknowledgements sections (Lingard, 2023). Some authors go even further, arguing that to ensure transparency, it is necessary for authors to mention the use of any NLP software, including Grammarly and ProWritingAid (Dergaa et al, 2023). In fact, these proposals are about the honesty of the authors, and in this case other researchers can make informed decisions - whether it is worth citing in their paper an article that used ChatGPT or not. Furthermore, Dergaa et al (2023) state that grey literature, should not currently be regarded as valid sources for referencing, as it encompasses materials and research that have not undergone a rigorous peer review process and may have been generated by AI.

At the same time, researchers raise the issue of student evaluation and devaluation of degrees. Students should no longer be evaluated based on descriptive tasks but based on the ability to solve problems/critical thinking, because otherwise people who have access to ChatGPT will have an advantage over those who do not. Teachers could ask students send a preliminary version of their work for assessment prior to the ultimate submission and to present their works in class and evaluate the way they answer the questions. Although now AI systems are free, in the future it is possible to start charging for the services offered, which could raise equity concerns if we do not change the evaluation method (Emenike M.E. and Emenike B.U., 2023; Cotton et al, 2023).

7. Conclusions

This study provides a comprehensive exploration of the implications and challenges of using ChatGPT and other Generative AIs in academic writing. In the last year the studies related to this topic have focused on the following research directions: the benefits and limits of using ChatGPT in academic writing; identifying and testing different AI-generated content detection methods and tools to prevent cheating; discussing ethics and academic integrity concerns and establishing generally accepted ethical standards for using ChatGPT in academic writing.

ChatGPT can be of great help to users in summarizing, structuring ideas, improving the revision process. It excels as a brainstorming tool for tasks such as generating titles, creating outlines. However, ChatGPT should be used with utmost caution, bearing in mind all its current limitations: inaccurate content with the risk of generating fictional information, limited knowledge, fake references, the potential for bias, plagiarism, lack of originality. As for the tools and methods for detecting AI-generated content to prevent cheating on assignments and plagiarism, they need to be continuously improved, considering the rapid advancement of technology. If at this moment there are many tools that allow the distinction between AI-generated content and human-generated content, new versions of AI have learned to "fool" the detection tools. Researchers and students must also consider the ethical, copyright and transparency concerns, as well as the threat of contributing to information epidemics. There is a need for a collective effort to establish generally accepted ethical standards, scholars advocating for transparent reporting and ethical guidelines. At the same time, there is a need to reconsider traditional assessment methods to ensure equity, to evaluate more the ability of students to support their point of view and critical thinking, and less the ability to perform descriptive tasks.

Ultimately, ChatGPT represents a powerful tool in academic writing, but its integration into the academic environment must be accompanied by clear ethical standards, accepted by the entire academic community, and should be combined with careful evaluation methods, adapted according to technological advancement.

References:

1. Altmäe, S., Sola-Leyva, A. and Salumets, A. (2023). Artificial intelligence in scientific writing: a friend or a foe?. *Reproductive BioMedicine Online*, 47(1), 3-9, <https://doi.org/10.1016/j.rbmo.2023.04.009>.
2. Ariyaratne, S., Iyengar, K. P., Nischal, N., Chitti Babu, N. and Botchu, R. (2023). A comparison of ChatGPT-generated articles with human-written articles. *Skeletal Radiology*, 1-4.
3. Author guidelines, Springer Nature Press. Guidance on the use of Large Language Models (LLM) e.g. ChatGPT; 2023. <https://www.springer.com/journal/10584/updates/24013930>
4. Cheng, S. W., Chang, C. W., Chang, W. J., Wang, H. W., Liang, C. S., Kishimoto, T., ... and Su, K. P. (2023). The now and future of ChatGPT and GPT in psychiatry. *Psychiatry and Clinical Neurosciences*. <https://doi.org/10.1111/pcn.13588>, available online at <https://onlinelibrary.wiley.com/doi/full/10.1111/pcn.13588>, accessed on 10.09.2023
5. Cingillioglu, I. (2023). Detecting AI-generated essays: the ChatGPT challenge. *The International Journal of Information and Learning Technology*, 40(3), 259-268. <https://doi.org/10.1108/IJILT-03-2023-0043>
6. Corizzo, R. and Leal-Arenas, S. (2023). One-Class Learning for AI-Generated Essay Detection. *Applied Sciences*, 13(13), 7901.
7. Cotton, D. R., Cotton, P. A. and Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1-12. <https://doi.org/10.1080/14703297.2023.2190148>
8. Day, T. (2023). A preliminary investigation of fake peer-reviewed citations and references generated by ChatGPT. *The Professional Geographer*, 1-4.
9. Dergaa, I., Chamari, K., Zmijewski, P. and Saad, H. B. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40(2), 615-622.
10. Desaire, H., Chua, A. E., Isom, M., Jarosova, R. and Hua, D. (2023). Distinguishing academic science writing from humans or ChatGPT with over 99% accuracy using off-the-shelf machine learning tools. *Cell Reports Physical Science*.

11. Elkhayat, A. M., Elsaid, K. and Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity*, 19(1), 17.
12. Emenike, M. E. and Emenike, B. U. (2023). Was This Title Generated by ChatGPT? Considerations for Artificial Intelligence Text-Generation Software Programs for Chemists and Chemistry Educators. *Journal of Chemical Education*, 100(4), 1413-1418.
13. Floridi L. and Chiriatti M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds Mach.* 30: 681–694, available at <https://link.springer.com/article/10.1007/s11023-020-09548-1>
14. Giray, L. (2023). Prompt Engineering with ChatGPT: A Guide for Academic Writers. *Annals of Biomedical Engineering*, 1-5.
15. Howell, B. E. and Potgieter, P. H. (2023). What do telecommunications policy academics have to fear from GPT-3?. *Telecommunications Policy*, 102576.
16. Hung, J. and Chen, J. (2023). The benefits, risks and regulation of using ChatGPT in Chinese academia: A content analysis. *Social Sciences*, 12(7), 380.
17. Lingard, L. (2023). Writing with ChatGPT: An illustration of its capacity, limitations and implications for academic writers. *Perspectives on Medical Education*, 12(1), 261.
18. Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S. and Wang, Z. (2023). 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*, 74(5), 570-581.
19. Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2), 07.
20. Radford A, Narasimhan K, Salimans T. and Sutskever I. (2018) Improving language understanding by generative pre-training. Available at <https://www.mikecaptain.com/resources/pdf/GPT-1.pdf>. accessed on 10.09.2023
21. Waltzer, T., Cox, R. L. and Heyman, G. D. (2023). Testing the Ability of Teachers and Students to Differentiate between Essays Generated by ChatGPT and High School Students. *Human Behaviour and Emerging Technologies*, 2023.
22. Williams, D. O. and Fadda, E. (2023). Can ChatGPT pass Glycobiology? *Glycobiology*, 33(8), 606-614.
23. Xie, Y., Seth, I., Rozen, W. M. and Hunter-Smith, D. J. (2023). Evaluation of the Artificial Intelligence Chatbot on Breast Reconstruction and Its Efficacy in Surgical Research: A Case Study. *Aesthetic Plastic Surgery*, 1-10.
24. Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 1-25.
25. <https://openai.com/blog/chatgpt>. accessed on 11.08.2023.