SCIENTIFIC OASIS

Journal of Soft Computing and Decision Analytics

Journal homepage: www.jscda-journal.org
ISSN: 3009-3481



ChatGPT and OpenAI: A Comprehensive Bibliometric Review

Abdullah Önden[1,*], Mohammed Alnour[2]

[1] Department of International Trade and Finance, Yalova University, Turkey
[2] Department of Economics, Institute of Social Sciences, Erciyes University, Turkey

ARTICLE INFO

Article history:
Received 1 June 2023
Received in revised form 2 September 2023
Accepted 29 September 2023
Available online 29 September 2023

Keywords: OpenAI, ChatGPT; Bibliometric Analysis; VOSviewer; Web of Sciences.

ABSTRACT

The introduction of OpenAI ChatGPT has sparked a lively discussion among researchers in academia, investors in business sector, and governments in policymaking circles. Numerous concerns have been expressed by Commentators from various scientific domains. Therefore, this study is an attempt to review the existing research pertaining to OpenAI's ChatGPT using bibliometric analysis. The study reviewed 210 research collected from Web of Sciences covering all disciplines. The analysis was based on (i) co-occurrence (keywords), (ii) citations and (iii) co-authorship. We found, keywords such as "ChatGPT, artificial intelligence, OpenAI, AI, education, chatbots, natural language processing, medicine, and management" are the most frequently used in OpenAI ChatGPT. Moreover, we identified that areas including cryptocurrencies, healthcare, future nursing, information value, non-human authors, algorithmic bias, clinical otolaryngology, clinical decision-making, large language models, generative pre-trained transfer are not covered deeply investigated using OpenAI ChatGPT. Finally, we found around 155 authors from 109 institutions located in 28 countries led by United Sated and Germany are actively engaged in producing research using OpenAI ChatGPT.

1. Introduction

ChatGPT is a sophisticated language model that employs deep learning techniques to produce responses in natural language that closely resemble human answers [1]. It stands as one of the most extensive publicly accessible language models and is a member of the generative pre-training transformer (GPT) model family developed by OpenAI. ChatGPT can capture the nuances and complexities of human language, allowing it to provide suitable and contextually relevant responses to a wide range of queries[2,3].

The introduction of OpenAI ChatGPT has sparked much interest and discussion among academic researchers in almost all disciplines. Numerous concerns have been expressed by Commentators from various scientific domains, notably in relation to the moral ramifications of using such techniques in academic publications and scientific writing [4]. For instance, the educational landscape has undergone significant change since the advent of AI technology. Artificial Intelligence

* Corresponding author.
E-mail address: abdullah.onden@yalova.edu.tr

https://doi.org/10.31181/jscda11202324

© The Author(s) 2023| Creative Commons Attribution 4.0 International License

has fundamentally changed how teachers and students interact, learn, and share knowledge. AI has fundamentally changed how teachers and students interact, learn, and share knowledge. Example of such Artificial Intelligence is OpenAI's ChatGPT, which is becoming more and more well-liked in the educational space for its capacity to produce responses that are human-like and condense difficult knowledge [5].

The possible uses of ChatGPT in the medical field include everything from helping experts with clinical and laboratory diagnosis to discover issues of research, providing updates on new developments to medical students, doctors, nurses, and all healthcare practitioners [6]. Moreover, the creation of virtual assistants to support patients in managing their health represents another crucial avenue for ChatGPT's utilization within the medical domain [2].

In the field of entrepreneurship, the exploration of AI is a relatively recent development within the broader field of entrepreneurship research. Scholars have begun to investigate its applications in areas such as venture-level assessment, the introduction of new products, and the decision-making processes of entrepreneurs. Notably, tools based on deep learning have introduced novel capabilities, enabling users to scrutinize the language patterns of prominent entrepreneurs in a manner customized to their specific needs. Emerging tools like ChatGPT achieve this by amalgamating a vast body of publicly available information with human feedback during the training phase [7, 8]. OpenAI ChatGPT is regarded as a powerful technology with potential applications in social media, including customer service automation, content development, chatbots, sentiment analysis, and data collection [9, 10].

In economics, ChatGPT implications can offer valuable insights. It is evident that the most critical aspect of these AI-powered services lies in their impact on labor forces and the labor market. Consequently, there exists a substantial body of literature within the field of labor economics dedicated to topics such as automation, robotics, technological advancements, and innovations that influence the labor market. Given that ChatGPT and similar AI-powered services can automate a multitude of tasks previously undertaken by human workers, their deployment has the potential to disrupt specific market segments significantly. The introduction of ChatGPT is anticipated to exert a substantial influence on the labor market dynamics, potentially rendering certain job roles obsolete while simultaneously giving rise to new employment opportunities. However, it is essential to acknowledge that this transitional phase may pose challenges for workers who may need additional training or skill upgrades to maintain their competitiveness in the labor market [11, 12].

Based on the above considerations, this derive motive behind the current research is to provide an overview of prior work on the application of OpenAI ChatGPT in various academic domains through bibliometric analysis and comprehensive literature evaluation. We employed visualization tools to evaluate contributions to marketing research and to visually convey the results of bibliometric methodologies. The academic community and decision-making circles may find this article to be helpful as it adds to the body of knowledge on the application of artificial intelligence.

Many researchers have started providing several future agendas and implementation scenarios in their research conducts. The increasing interest in OpenAI ChatGPT offers new insights in academics which possibly enhance the efficiency in research activities. However, the term OpenAI ChatGPT is still relatively new and there is a need to examine the evolution of OpenAI research across different disciplines to reveal the research trends, focus, and limitations of this research topic using bibliometric analysis. By mining/auditing trails of extensive literature databases, bibliometric analyses aim to decrease research bias and evaluate research systematically. The three key elements in this study that will be determined using a bibliometric analytical technique are co-authorship, co-occurrence, and citation.

Bibliometric analysis was employed to create visual representations depicting the interconnections among core ideas. Consequently, visual mapping aids researchers in uncovering the context of a particular research domain, the interrelationships among pivotal notions, and potential forthcoming directions, facilitating a comprehensive examination of the scrutinized studies. This, in turn, enables the identification of prominent research themes that authors emphasized when delving into the realm of OpenAI ChatGPT. In order to identify the key players in this discipline, significant subtopics, and areas for future research, we looked through published articles. Between 2022 and 2023, we found that there will be an increase in the amount of existing literature on this topic. We use network analysis to map the research clusters as well.

The next up section outlines the methodology used for this review. Th third section accommodates the result and discussion, while the last section shows concluding remarks.

2. Methodology

The combination of bibliometric analysis with citation analysis provides a powerful method for examining the shapes and characteristics of prior literature from a variety of academic disciplines. This methodology can also help in identifying any dominant schools of thinking within research fields. When examining written materials like journals, books, and online resources, bibliometric analysis adheres to an objective philosophical framework and uses a quantitative research approach. According to [13], the citation and co-citation studies are primarily focused on detecting emerging themes within particular fields of study and evaluating the effect of several publications and different schools of thoughts. Previous study has highlighted the ability to determine the substance and progress of an academic area beyond citation tallying and aggregation. This evaluation considers the achievements of individual journals and authors who have benefited the scientific community through collaboration. Bibliographic analysis with co-citation is useful in studying the tendencies and details of what has occurred in the literature, facilitates the examination, organize, and articulate the prior research in a specific field [14].

The Bibliometric analysis is conducted with the help of the modern version of the VOSviewer software. This software is powerful in constructing and visualizing bibliometric networks. Items are used in mapping building, visualization, and analysis. These items can be writers, terms, or publications that reflect the subjects of interest. A link can connect any two items, such as a link in a publication's bibliography, a link in a researcher's co-authorship, or a link in the co-occurrence of a term or concept [15]. Any link's strength can be described using a positive numerical number that indicates a high value when the connection is stronger or vice versa. The strength of each given link can highlight the number of publications that at least two publications share as references or the number of publications that at least two scholars share as co-authorship links. By connecting items and linkages, any network can be created.

Previous studies can be tracked, counted, and analyzed using bibliographic analysis. It includes a list of the scholars' published works, the best journals, the technique used, and the conclusions drawn. With the aid of metadata, any area of research can be condensed. The use of bibliometric methodologies, which necessitate a substantial amount of bibliographic data, has allowed for the analysis of a wide range of topics, journals, and nations [5].

A network is made up of elements and the connections that connect them, which can be further subdivided into groups. Because groups are a set of things included in a map, it is vital to determine whether an object can only be a part of a given group. This part briefly highlights the outcomes from the bibliometric analysis for 210 research items using data collected from the Web of Sciences database (Figure 1 and 2). The survey of literature in Web of Sciences database included all the fields in Scopus database using "OpenAI ChatGPT" as main keywords. Figs 1 and 2 outline the utilization of

OpenAI ChatGPT in different research fields. Medicine General and Education are by far the most fields using artificial intelligence followed by Engineering, Computer Sciences Information Systems, and Surgery.



Fig 1. OpenAI ChatGPT Usage Across Different Research Fields

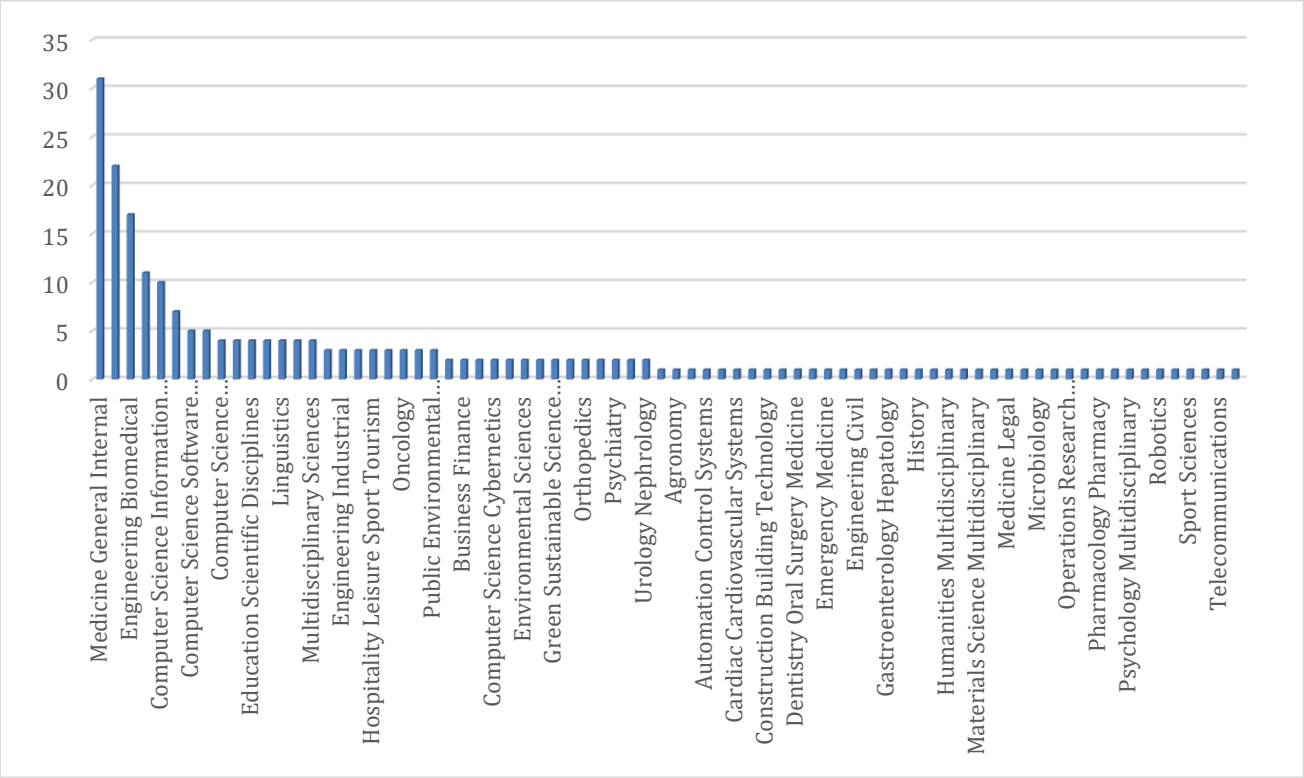


Fig 2. OpenAI ChatGPT Usage Across Different Research Fields

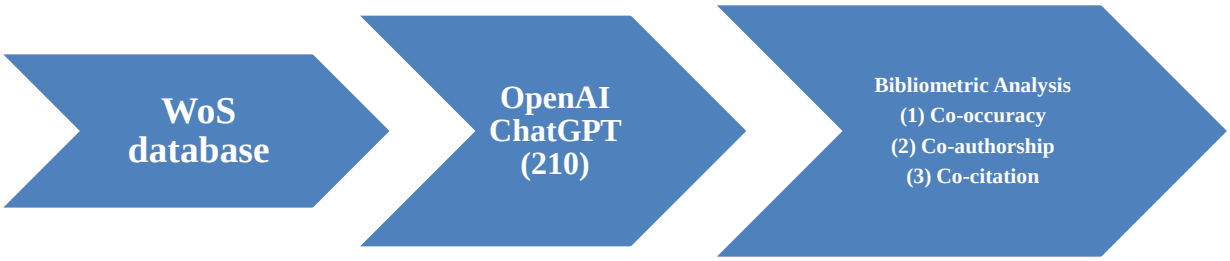


Fig 3. Analysis Flowchart

3. Results

As shown in Figure 3 this study reviews 210 studies related to "OpenAI ChatGPT" collected from Web of Sciences. The search resulted in Article (117), Early Access (41), Review Article (19), Letter (16), Editorial Material (13), Proceeding Paper (4). The use of "keywords analysis" in bibliometric analyses can help to direct collection growth, define institutional scholarship strengths, identify potential schools of thought in a discipline, and identify trends in citation and co-citation. All of these benefits can help researchers spend less time evaluating the impact of their work and the level of collaboration among scholars, which will lead to more effective policymaking [16]. In bibliometric research, citation and co-citation analyses are used as an analytical technique to examine a portion or the totality of a scholarly discipline for written source documents (such as academic journal papers and books), Figure 4 [17].

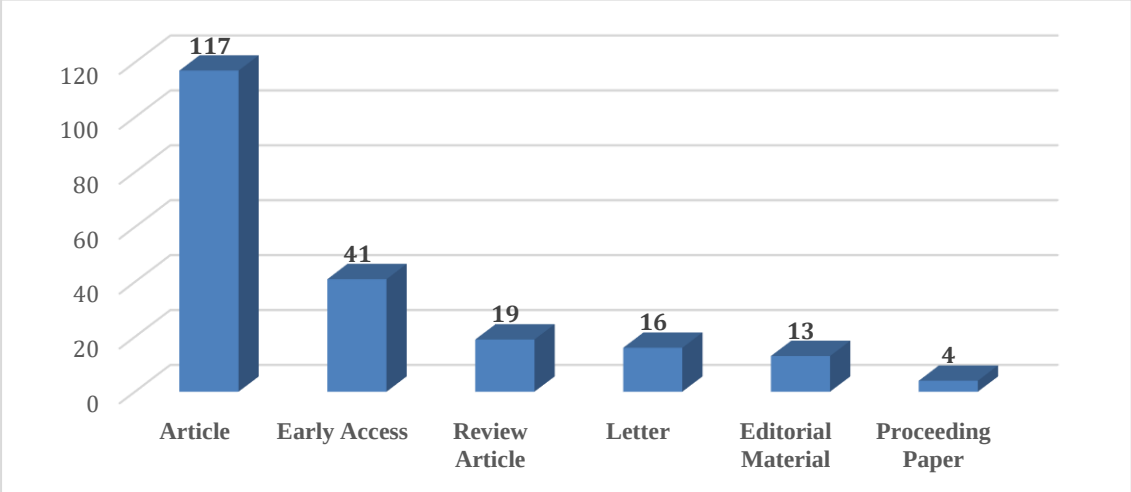


Fig 4. OpenAI ChatGPT related research by documents

3.1. Co-accuracy

Figure 5 highlights the words that occur in the same article at the same time. It is important to remember that the abstract only considers keywords that the author has listed. The main objective is to use "OpenAI ChatGPT" to determine the study fields' most popular terms used by authors. The same figure displays the most relevant keywords set off, and the nodes between keywords. The larger the keyword and the node, so does weight; the smaller the distance between nodes, the stronger the relationship among them. The author used a co-occurrence analysis kind of phrase with at least one co-occurrence to acquire 157 keywords using the full counting method.

Based on the size of the circle, Figure 5 is used to determine which phrases are often used. The figure's combined mapping and grouping provides a high-level picture of the ChatGPT research

subjects' organizational structure. distinct colors are used to represent each group, each of which represents its significance, affinity, and relatedness in a distinct way. This shows that "the density of elements increases with the number of adjacent elements and decreases with the distance between them and the point of interest." The terms that appear the most frequently include "ChatGPT," "artificial intelligence," "OpenAI," "AI," education," "chatbots," "natural language processing," "medicine," and "management," as shown by the co-occurrence analysis (co-occurrence keywords) in Figure 5. The visualization overlay in Figure 6 can show how much a keyword's popularity has increased over time, the more yellow the cluster, the more recent the research. The majority of the keywords that are highlighted in yellow in Figure 6 show how the development of keywords is expanding.

Figures 6 and 7 show the density of related terms and the number of nodes that are close to one another. These results can be used as a starting point for more research to develop other ChatGPT-related keywords in the research disciplines. A saturation level represented by a yellow node indicates the prevalence of frequently occurring phrases. In the case where the keyword "ChatGPT" is most commonly found, the yellow node signifies the region with the most extensive research. Different topics, such as "cryptocurrencies, healthcare, future nursing, information value, non-human authors, algorithmic bias, clinical otolaryngology, clinical decision-making, large language models, generative pre-trained transfer" are covered in the green nodes showing that a subject has not been thoroughly studied.

The size of the yellow nodes, which indicates the prevalence of recurring concerns, will prompt the editor to evaluate whether this subject remains pertinent to the discipline mentioned by the reviewer. The review's results may fall short, particularly in terms of substance, if the scholarly interests of each reviewer are not sufficiently aligned. Including this topic provides editors with evaluative data that can aid in refining and specializing in the journal's focus. Furthermore, the editors select Editors and Reviewers whose scientific expertise aligns with the matter depicted in Figure 7, using this density visualization as a reference.

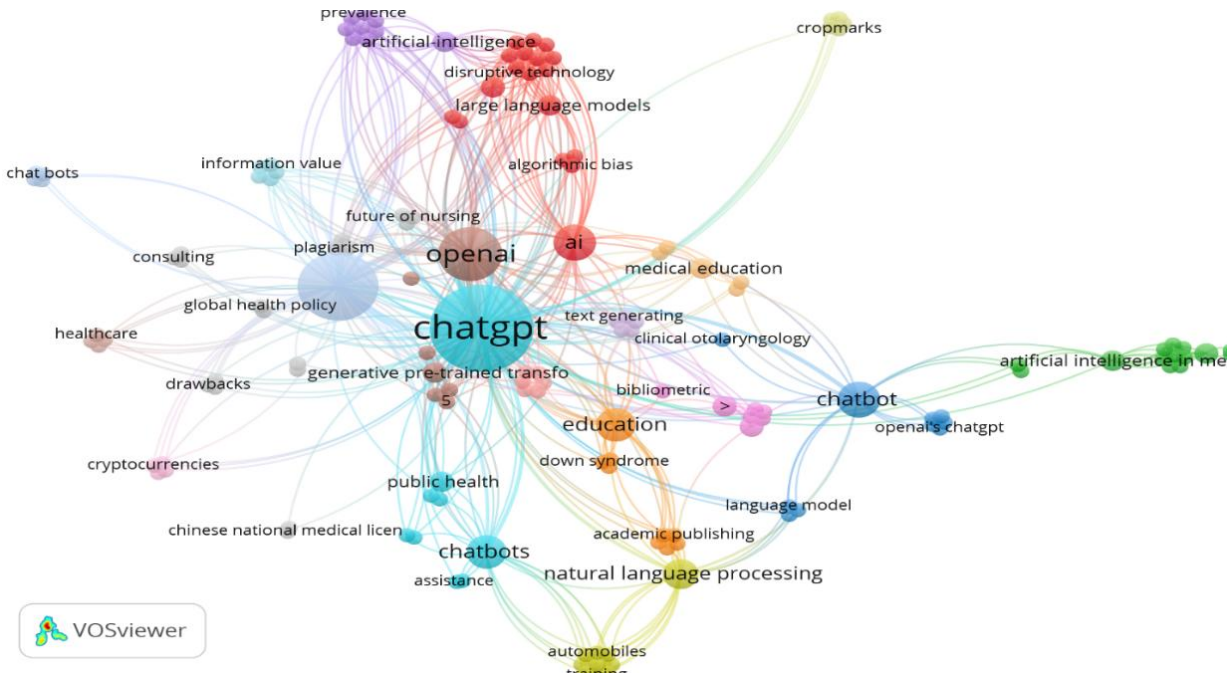


Fig 5. Co-occurrence (Keywords)

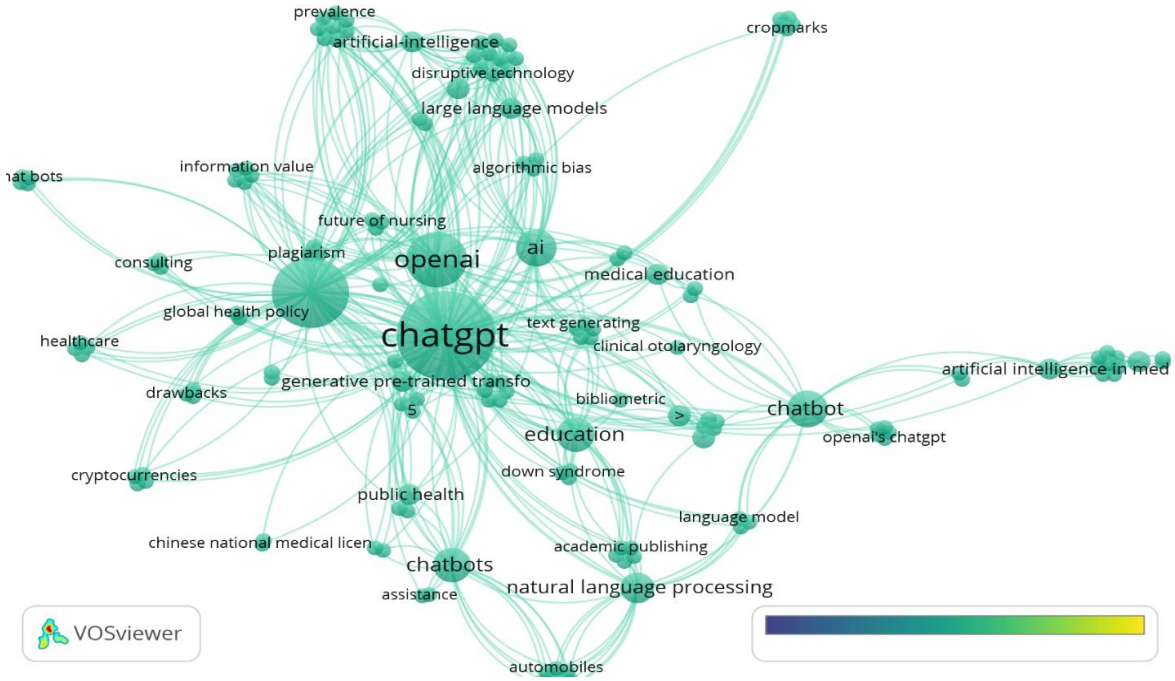


Fig 6. Overlay visualization co-occurrence (Keywords)

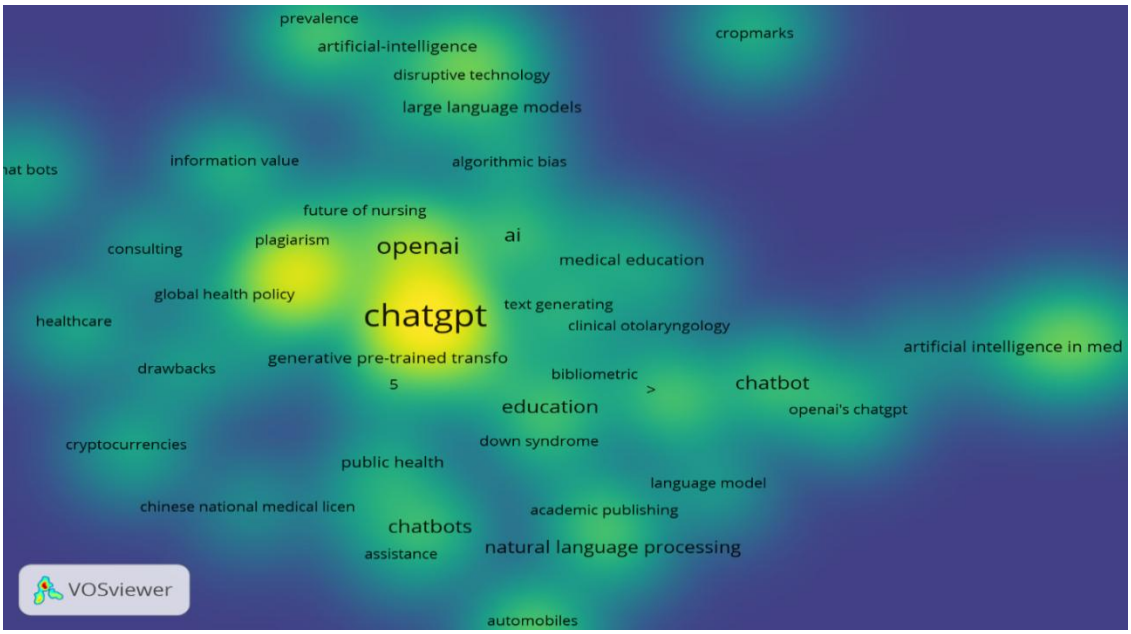


Fig 7. Density visualization co-occurrence (Keywords)

3.2. Citations

This analysis is used to discover which papers have the biggest influence on ChatGPT in scientific research with a minimum of 1 citation. The number of articles to be selected are 29. According to Figure 8, [18] is the most influential articles with 61 number of citations. This study was conducted in response to the call for case reports utilizing ChatGPT's assistance in the Journal of Medical Science Turing Test. The authors have submitted two cases. The first case explores osteoporosis associated with homocystinuria, while the second case delves into late-onset Pompe disease (LOPD), a rare metabolic disorder. The authors employed ChatGPT to elucidate the pathogenesis of these

conditions. Furthermore, they meticulously documented both the favorable, unfavorable, and somewhat concerning facets of the chatbot's performance in their report.

The second most impactful articles is the one of [19] with 29 citations. The authors conducted an analysis of the effect of ChatGPT on the area of public health, alongside an exploration of the benefits and drawbacks associated with its utilization. Third, the study [20] show also a magnificent influence with 13 total number of citations. In this study, the author employed OpenAI ChatGPT as an interviewee instead of a real person. The primary focus of an interview was on the influence of OpenAI's GPT in the realms of higher education and academic publishing. In particular, numerous inquiries were directed towards the effects of OpenAI's ChatGPT and other AI-driven machine learning models within the domains of hospitality, tourism, and education. ChatGPT highlighted that its utilization can aid educators in delegating routine tasks, such as grading, thereby allowing them to concentrate on more intellectual endeavors.

Additionally, students have the potential to employ ChatGPT for brainstorming and idea generation. However, ChatGPT acknowledged the potential risk of diminishing critical thinking skills in students if they become overly reliant on it, along with the issue of educational inequalities that may arise. Regarding academic work, ChatGPT made it clear that it cannot serve as a substitute for human creativity and intellect, as it lacks the capacity to generate outputs characterized by originality and novelty. The last most influential study is the one conducted by [21] with 11 total number of citation. The authors outline the role of ChatGPT in surgical practice.

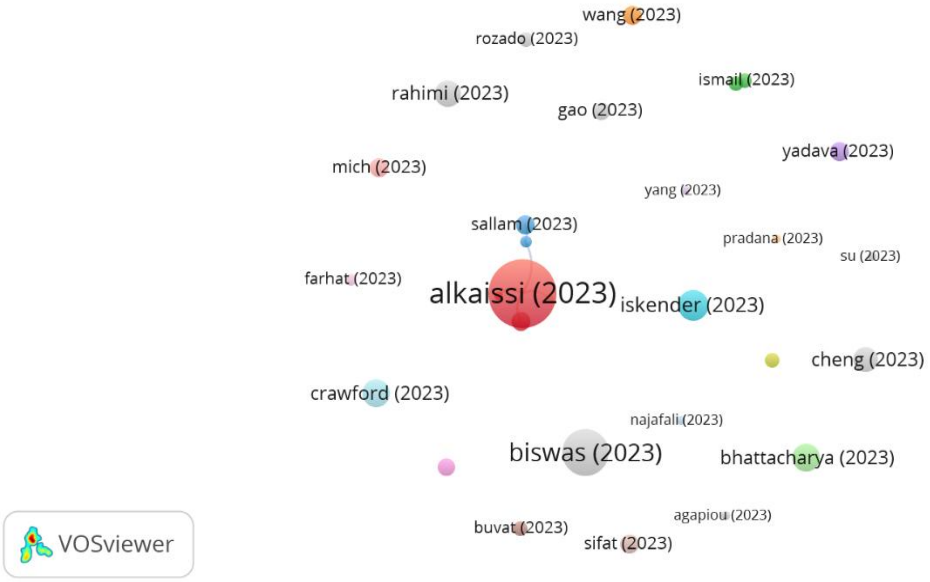


Fig 8. Most cited studies on OpenAI ChatGPT

3.3. Co-authors

Shedding light on the productive authors and effective organizations and research institutions along with countries in the OpenAI ChatGPT, with minimum number of 1 document, Figures 9, 10 and 11 demonstrate that nearly 155 authors contributed to research fields in OpenAI ChatGPT belonging to 109 organizations/institutions/universities located in 28 countries led by the United States and Germany.

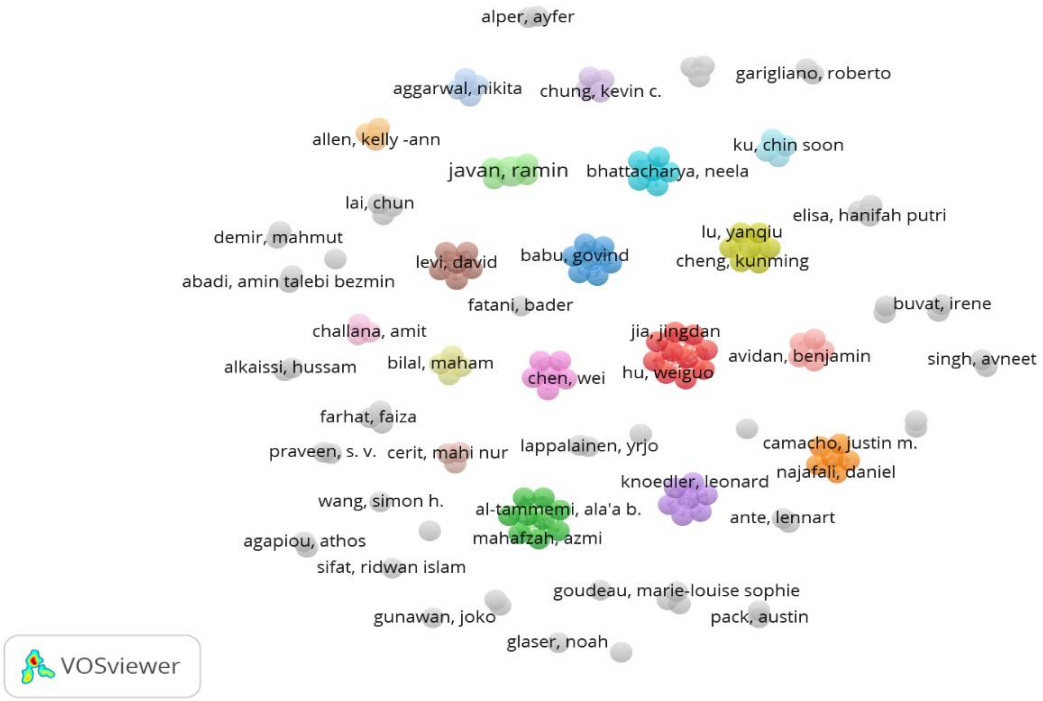


Fig 9. Co-authorship (authors)

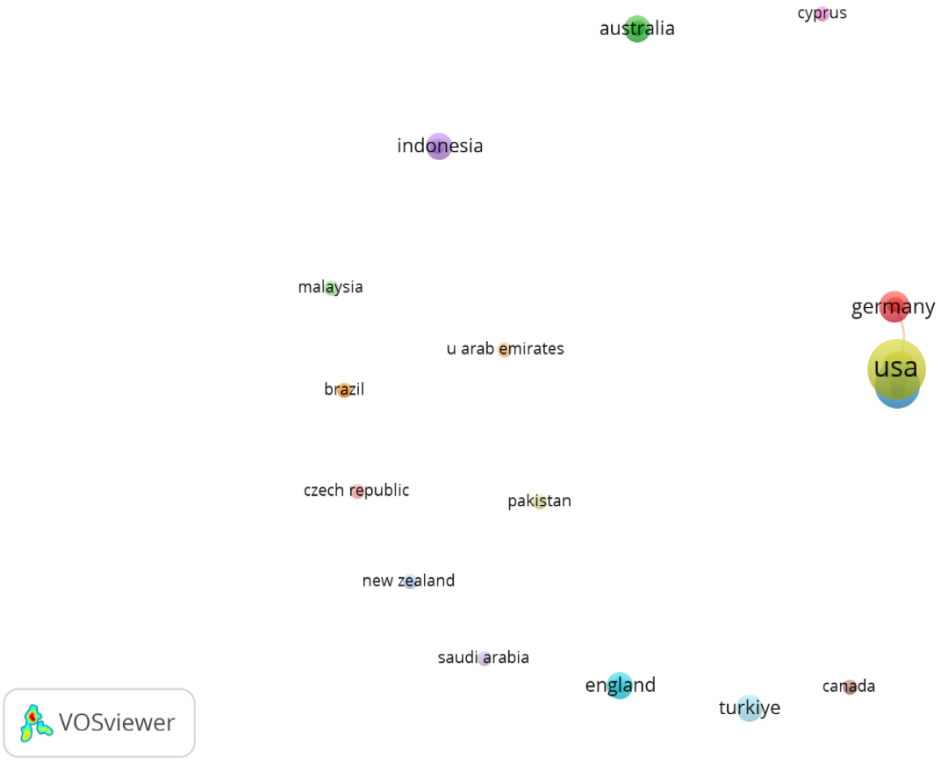


Fig 10. Co-authorship (countries)

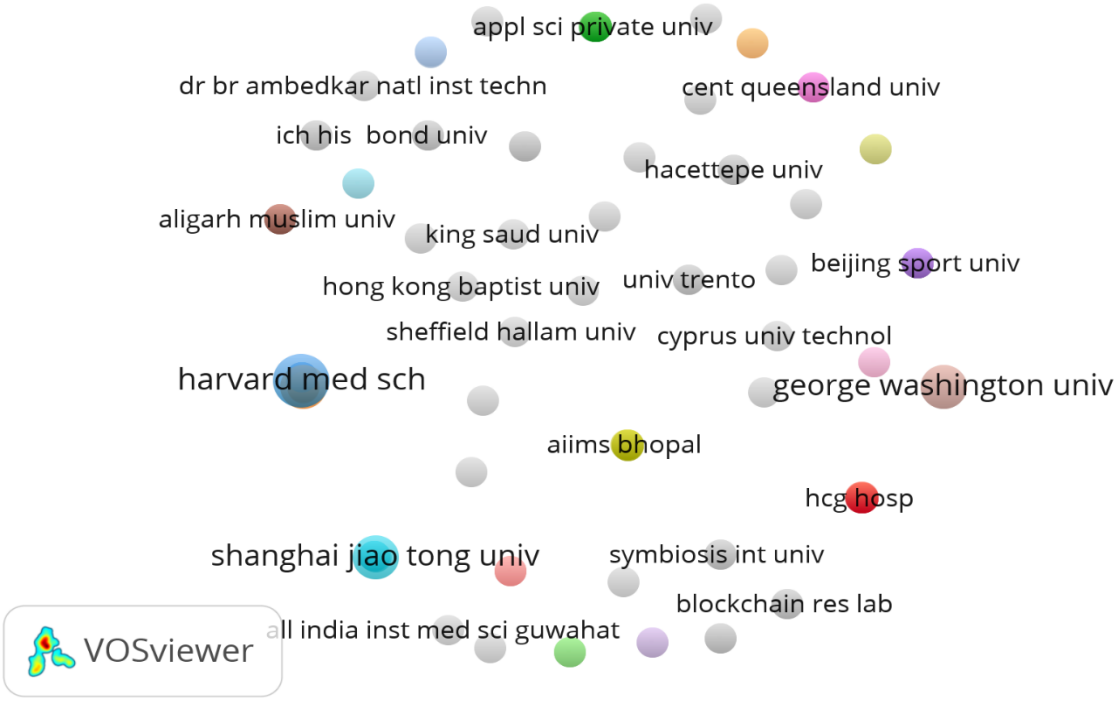


Fig 11. Co-authorship (organizations)

4. Conclusions

In this study, three essential factors, namely co-authorship, co-occurrence, and citation, were identified through the application of a bibliometric analytical approach. It's worth noting that research related to the ChatGPT theme within the scientific community has only started emerging as of 2023 and was sourced from the Web of Science database resulted in 210 studies from all disciplines. Medicine, education, and engineering are the most fields of research used OpenAI ChatGPT since its emergence.

Our results show that the most frequently include "ChatGPT," "artificial intelligence," "OpenAI," "AI," education," "chatbots," "natural language processing," "medicine," and "management,". We also found that subject like cryptocurrencies, healthcare, future nursing, information value, non-human authors, algorithmic bias, clinical otolaryngology, clinical decision-making, large language models, generative pre-trained transfer are covered not deeply investigated using OpenAI ChatGPT. Though, future studies may consider them as area of research utilizing OpenAI ChatGPT.

Regarding the third factor, we found the most influential research used OpenAI ChatGPT showed 61 total number of citations. Additionally, we found around 155 authors from 109 institutions located in 28 countries led by United Sated and Germany are actively engaged in producing research using OpenAI ChatGPT. Our study, therefore, provided insightful information for policymakers and researchers identifying the future research areas and active institutions and productive collaborative researchers for further research collaboration.

Funding

This research was not funded by any grant.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "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*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642.

[2] Dave, T., Athaluri, S. A., & Singh, S. (2023). ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Frontiers in Artificial Intelligence*, *6*, 1169595. https://doi.org/10.3389/frai.2023.1169595.

[3] Zhang, C., Zhang, C., Li, C., Qiao, Y., Zheng, S., Dam, S. K., ... & Hong, C. S. (2023). One small step for generative ai, one giant leap for agi: A complete survey on chatgpt in aigc era. *arXiv preprint arXiv:2304.06488*. https://doi.org/10.48550/arXiv.2304.06488.

[4] Farhat, F., Sohail, S. S., & Madsen, D. Ø. (2023). How trustworthy is ChatGPT? The case of bibliometric analyses. *Cogent Engineering*, *10*(1), 2222988. https://doi.org/10.1080/23311916.2023.2222988.

[5] Pradana, M., Elisa, H. P., & Syarifuddin, S. (2023). Discussing ChatGPT in education: A literature review and bibliometric analysis. *Cogent Education*, *10*(2), 2243134. https://doi.org/10.1080/2331186X.2023.2243134.

[6] Aydın, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT generated literature review: Digital twin in healthcare. *Available at SSRN 4308687*. https://doi.org/10.2139/ssrn.4308687.

[7] Short, C. E., & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights*, *19*, e00388. https://doi.org/10.1016/j.jbvi.2023.e00388.

[8] Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). ChatGPT has aced the test of understanding in college economics: Now what?. *The American Economist,* 68(2), 233–245. https://doi.org/10.1177/05694345231169654.

[9] Biswas, S. (2023). The Function of chat GPT in social media: According to chat GPT. *Available at SSRN 4405389*. https://doi.org/10.2139/ssrn.4405389.

[10] Onden, A., & Kiygi-Calli, M. (2018). The effects of brands' posts on social media: conceptual framework and application method. Research Journal of Business and Management, 5(3), 238-250.

[11] Zarifhonarvar, A. (2023). Economics of chatgpt: A labor market view on the occupational impact of artificial intelligence. *Available at SSRN 4350925*. https://doi.org/10.2139/ssrn.4350925.

[12] Ali, H., & Aysan, A. F. (2023). What will ChatGPT Revolutionize in Financial Industry?. *Available at SSRN 4403372*. https://doi.org/10.2139/ssrn.4403372.

[13] Faruk, M., Rahman, M., & Hasan, S. (2021). How digital marketing evolved over time: A bibliometric analysis on scopus database. *Heliyon*, *7*(12), e08603, https://doi.org/10.1016/j.heliyon.2021.e08603.

[14] Ferreira, M. P., Santos, J. C., de Almeida, M. I. R., & Reis, N. R. (2014). Mergers & acquisitions research: A bibliometric study of top strategy and international business journals, 1980–2010. *Journal of Business Research*, *67*(12), 2550-2558. https://doi.org/10.1016/j.jbusres.2014.03.015.

[15] Lobonț, O. R., Țăran, A. M., & Costea, F. (2020). E-Government Research Still Matter? A Bibliometric Analysis. *Annals of the University Dunarea de Jos of Galati: Fascicle: I, Economics & Applied Informatics*, *26*(2), 58-63. doi: https://doi.org/10.35219/eai15840409106.

[16] Lewis, D. M., & Alpi, K. M. (2017). Bibliometric network analysis and visualization for serials librarians: an introduction to Sci2. *Serials Review*, *43*(3-4), 239-245. https://doi.org/10.1080/00987913.2017.1368057.

[17] Shafique, M. (2013). Thinking inside the box? Intellectual structure of the knowledge base of innovation research (1988–2008). *Strategic management journal*, *34*(1), 62-93. https://doi.org/10.1002/smj.2002.

[18] Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*, *15*(2). https://doi.org/10.7759/cureus.35179.

[19] Biswas, S. S. (2023). Role of chat gpt in public health. *Annals of biomedical engineering*, *51*(5), 868-869. https://doi.org/10.1007/s10439-023-03172-7.

[20] Iskender, A. (2023). Holy or unholy? Interview with open AI's ChatGPT. *European Journal of Tourism Research*, *34*, 3414-3414. https://doi.org/10.54055/ejtr.v34i.3169.

[21] Bhattacharya, K., Bhattacharya, A. S., Bhattacharya, N., Yagnik, V. D., Garg, P., & Kumar, S. (2023). ChatGPT in surgical practice—a New Kid on the Block. *Indian Journal of Surgery*, 0123456789, 1–4. https://doi.org/10.1007/s12262-023-03727-x.