A Systematic Review of Large Language Models (LLMs) in Scientific Literature: Scientometric Analysis and Tree of Science Application *

Una Revisión Sistemática de Modelos Largos de Lenguaje (MLL) en Literatura Científica: Análisis Cienciométrico y Aplicación de Tree of Science

Recibido: marzo 2 de 2024 - Evaluado: abril 2 de 2024 - Aceptado: junio 20 de 2024

Juan Guillermo Saurith Moreno**
Daniel Camilo Blanco-Galán***
Sebastián Mindiola-Garizado****
José Francisco Ruiz-Muñoz *****

Para citar este artículo / To cite this Article
J. G. Saurith-Moreno, D. C. Blanco-Galán, S. Mindiola-Garizado, J. F. Ruiz-Muñoz "A Systematic Review of Large Language Models (LLMs) in Scientific Literature: Scientometric Analysis and Tree of Science Application" Revista de Ingenierías Interfaces, vol. 7, no.1, pp.1-19, 2024.

Abstract
The analysis of scientific literature has had a significant impact on multiple areas of knowledge. Its importance lies in the characterization and extraction of information related to scientific production, guaranteeing the homogeneity of references. Although the use of Large Language Models (LLMs) and algorithmic tools such as ChatGPT has positive aspects, the application of ChatGPT in the review of scientific literature presents challenges due to its lack of precision in the construction of bibliometric analyses. To address this limitation, the Tree of Science (ToS) tool, which provides a more reliable characterization of the articles, has been used. This scientometric review includes an analysis of the research output in the main databases, Scopus and Web of Science on MLLs and their different applications. International collaborations, the main journals of publication and the dynamics of collaboration between authors have been considered as study variables, using the algorithm employed by ToS to identify and explain the main contributions, articles, authors, journals, keywords, and countries highlighted. The findings reveal the novelty of the topic, with academic production concentrated in 2023 and 2024, and a notable influence of researchers from the United States and China. Medical fields lead in this area, with prominent journals and prolific authors in research.

---

*Artículo inédito: "A Systematic Review of Large Language Models (LLMs) in Scientific Literature: Scientometric Analysis and Tree of Science Application".

**Estudiante de Estadística, Universidad Nacional de Colombia – Sede de La Paz, Correo Electrónico: jsaurith@unal.edu.co, https://orcid.org/0009-0004-3885-3663, La Paz-Cesar, Colombia.

***Estudiante de Ingeniería Mecatrónica, Universidad Nacional de Colombia – Sede de La Paz, Correo Electrónico: danblanco@unal.edu.co, https://orcid.org/0009-0007-8578-8691, La Paz-Cesar, Colombia.

**** Estudiante de Mecatrónica, Universidad Nacional de Colombia – Sede de La Paz, Correo Electrónico: smindiola@unal.edu.co, https://orcid.org/0009-0005-0651-2327, La Paz-Cesar, Colombia.

***** Doctorado en Ingeniería – Automática, Magister en Automatización Industrial, Ingeniero Electrónica, Correo Electrónico: jfruizmu@unal.edu.co, https://orcid.org/0000-0002-0895-2036, La Paz-Cesar, Colombia.

Keywords: Large Language Models, ChatGPT, Tree of Science, Systematic Review, Scientometrics, Output Analysis.

Resumen

El análisis a la literatura científica ha tenido una repercusión significativa en múltiples áreas del conocimiento. Su importancia radica en la caracterización y extracción de información relacionada con la producción científica, garantizando la homogeneidad de las referencias. Aunque el uso de Modelos Largos de Lenguaje (MLLs) y herramientas algorítmicas como ChatGPT tiene aspectos positivos, la aplicación de ChatGPT en la revisión de literatura científica presenta desafíos debido a su falta de precisión en la construcción de análisis bibliométricos. Para abordar esta limitación, se ha utilizado la herramienta Tree of Science (ToS), que proporciona una caracterización más confiable de los artículos. Esta revisión cienciométrica incluye un análisis de la producción investigadora en las principales bases de datos, Scopus y Web of Science sobre MLLs y sus diferentes aplicaciones. Se han considerado variables de estudio las colaboraciones internacionales, las principales revistas de publicación y la dinámica de colaboración entre autores, utilizando el algoritmo empleado por ToS para identificar y explicar las principales contribuciones, artículos, autores, revistas, palabras clave y países destacados. Las conclusiones revelan la novedad del tema, con producción académica concentrada en 2023 y 2024, y una notable influencia de investigadores de Estados Unidos y China. Los campos médicos lideran en esta área, con revistas destacadas y autores prolíficos en la investigación..

Palabras clave: Grandes modelos de lenguaje, ChatGPT, Tree of Science, Revisión Sistemática, Cienciometría, Análisis de Producción.

1. Introducción

Reviewing and analyzing bibliographic information is a challenge for researchers, so using computational algorithms decreases to some extent the time spent performing these activities [1]. Tools that use long language models, i.e., systems that perform deep analysis of natural language [2], such as ChatGPT (Chat Generative Pre-Trained Transformer), interpret and can find data, reports, and notes (among many other types of information) that allow speeding up this process [3].

Even though these tools facilitate the bibliographic analysis work of researchers, they have important limitations that should be exposed [4]. Recent studies have shown that citations generated by the ChatGPT 3.5 model are 55% false, and although this rate is reduced to 18% in ChatGPT 4.0, it is still a considerable concern [5]. Furthermore, only 24% of the bibliographic citation management by ChatGPT 4.0 proves to be relevant, indicating a significant gap in the accuracy, handling, and usefulness of bibliometric data by these models.

ChatGPT is not a tool for information analysis, but one based on Natural Language Processing (NLP), so it is not optimized for these tasks. To combat the limitations of this and other LLMs in data analysis, we will focus on the use of the Tree of Science (ToS) tool, which offers a robust and accurate alternative for scientometric and bibliometric analysis [6], [7].

For this study, we employed a methodology centered on an extensive search of articles and their relevant information using keywords in reputable databases such as Scopus and Web of Science (WoS). These articles underwent a thorough filtering and review process to ensure their significance. Subsequently, we created various graphs and visualizations, and based on these, we drew a series of conclusions. Finally, we applied the ToS classificatory algorithm to structure and hierarchically analyze the information, which facilitated the identification of key trends and relationships between the articles [8]. The use of the ToS algorithm is reflected in different areas of knowledge, as can be seen in articles such as that of Alejandro Echeverri [9] or those of [10], [11].

This article continues with the details of the methodological part, providing a comprehensive explanation of the process involved in searching for and identifying the primary scholarly articles relevant to a specific topic for subsequent analysis.

2. Materiales y Métodos

The research methodology involved a thorough review of articles from two prominent databases, Scopus and Web of Science (WoS). Articles with specific keywords in their titles such as "large language models," "LLMs," "AI language models," or "chatgpt" were initially identified. Additionally, articles containing keywords like "literature review," "scientific research," or "systematic review" in their abstracts or keywords were included in the search. The search yielded 108 documents from Scopus and 136 from Web of Science. It's worth noting that within Scopus' search parameters, there is another parameter called "Doctype," which specifies the type of document being used, in this case, articles.

Table I. Search parameters used in Scopus and WoS databases.

Parameter	Web of Science	Scopus
Range	2000-2024	
Date	May 19, 2024	
Document Type	Article	
Words	Title: ("large language models" OR "LLMs" OR "AI language models" OR "chatgpt") AND Title-Abstract-Keywords: ("literature review" OR "scientific research" OR "systematic review")	
Results	136	108
Total (Wos+Scopus)		190

Once we have these searches, the next step is to download the information of each document like Title, Authors, Abstract, Keywords, Crossreferences, and DOI. Figure 1 displays the general process from the search results to the data analysis. The preprocessing is a complex task due to the differences between Scopus and WoS formats. It is necessary to perform text mining and web scraping to extract and normalize the references of both datasets. Finally, we presented the results in two sections, the first one is a traditional scientometric analysis using new techniques [12], and the second, we used the ToS metaphor to identify the main theoretical contributions[13], [14].

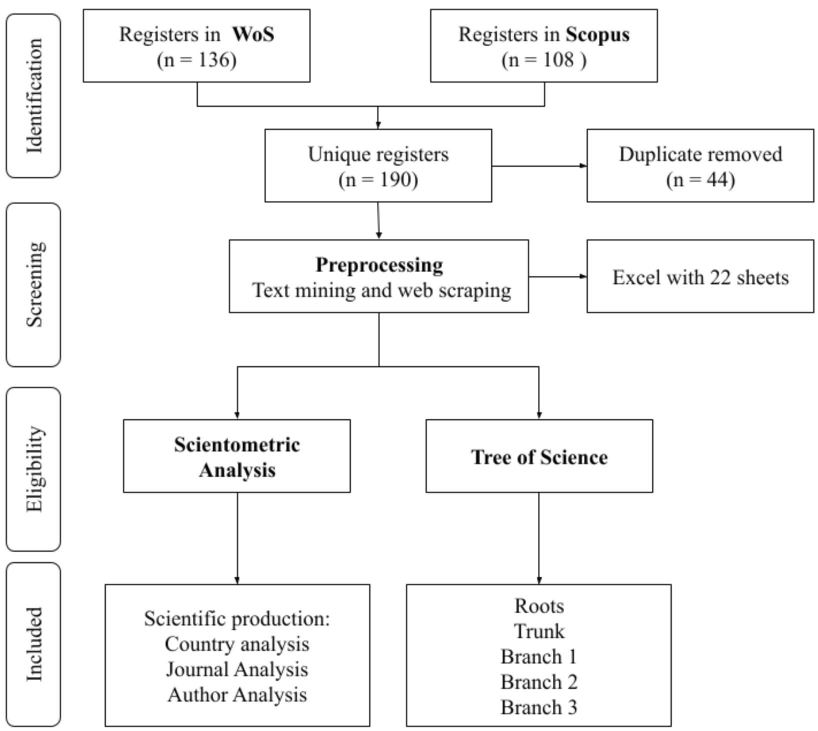

Figure 1. Flowchart of Data Collection and Analysis Process for Scientometric and Tree of Science Studies.

3. Resultados

3.1 Scientometric Analysis

Scientific Production

Understanding the annual analysis of scientific production within a specific topic is fundamental, as it allows for the identification of key moments and trends over time [15], [16]. In Figure 2, it is evident that scientific production recorded in WoS (yellow bar) and Scopus (green bar), about the use of LLMs for scientometric analysis, begins notably from 2023, with some publications also emerging in 2024. This phenomenon is undoubtedly due to the impact of Open AI's GPT-3.5 model, which has revolutionized scientific research and the generation, analysis, and dissemination of scientific articles.

Figure 2. Yearly Comparison of Total Scientific Production and Citations in Scopus and WoS (2023-2024).

It is noteworthy that, compared to last year, the scientific production this year already shows a significant increase in the first five months, indicating a more intensive exploration of the topic.

The considerable number of citations is particularly important given the time elapsed since publication and the total number of articles. Among them, the article by Sallam [17] stands out, having accumulated a total of 478 citations (Scopus + WoS). This article is highly valued in scientific circles for its exhaustive exploration of the role of ChatGPT in scientometric analysis, emphasizing how this technology can reshape scientific research, improve the communication of ideas, and promote equity and diversity in scientific domains.

Country Analysis

Table II provides a detailed breakdown of scientific production and citation metrics across various countries, along with the distribution of publications across journal quartiles (Q1, Q2, Q3, Q4) and those not categorized.

The USA leads significantly with 57 publications, constituting 25.68% of the total output. China follows with 18 publications (8.11%), highlighting its growing contribution in this field. Both India and Saudi Arabia have 10 publications each (4.5%), demonstrating their active participation. Other notable contributors include Italy (9 publications, 4.05%), Australia (8 publications, 3.6%), and Korea (6 publications, 2.7%).

Table II. Country-wise Scientific Production and Citation Analysis with Journal Quartile Distribution.

Country	Production		Citation		Q1	Q2	Q3	Q4	No category
USA	57	25.68%	386	15.3%	39	7	3	1	7
China	18	8.11%	49	1.95%	8	2	5	2	1
India	10	4.5%	66	2.62%	6	1	1	0	2
Saudi Arabia	10	4.5%	144	5.72%	3	5	0	0	2
Italy	9	4.05%	195	7.75%	5	2	1	0	1
Australia	8	3.6%	40	1.59%	3	4	0	0	1
Korea	6	2.7%	34	1.35%	5	1	0	0	0
Brazil	5	2.25%	17	0.68%	2	3	0	0	0
Turkey	5	2.25%	44	1.75%	2	3	0	0	0
United Arab Emirates	5	2.25%	78	3.1%	1	0	0	0	4

One of the most representative studies from the USA examines ChatGPT's capability to generate innovative ideas for systematic reviews on topics related to cosmetic surgery [18]. The results showed that, overall, ChatGPT achieved 55% accuracy, which dropped to 35% when the topics were stratified and focused on specific ideas. Conversely, Asian researchers assessed ChatGPT's potential in scientific research on infectious diseases [19]. The results indicated that this tool is useful and saves researchers time in organizing and structuring their articles.

Figure 3 illustrates the global scientific collaboration network among countries based on their scientific research activities. The network highlights the connections between countries, with the thickness of the lines indicating the strength of collaboration. The size of the nodes represents the degree of connectivity, with larger nodes indicating higher levels of collaboration. The figure is color-coded by community clusters, revealing distinct groups of countries that frequently collaborate. The United States emerges as the central hub with extensive connections to multiple countries, demonstrating its pivotal role in fostering international research partnerships. Other notable hubs include China, the United Kingdom, and Germany, which also exhibit significant collaborative ties. The inset graph on the left shows the number of countries in each community, with seven distinct communities identified. This network analysis underscores the importance of international collaboration in advancing scientific research and highlights key players and emerging collaborative clusters in the global scientific community.

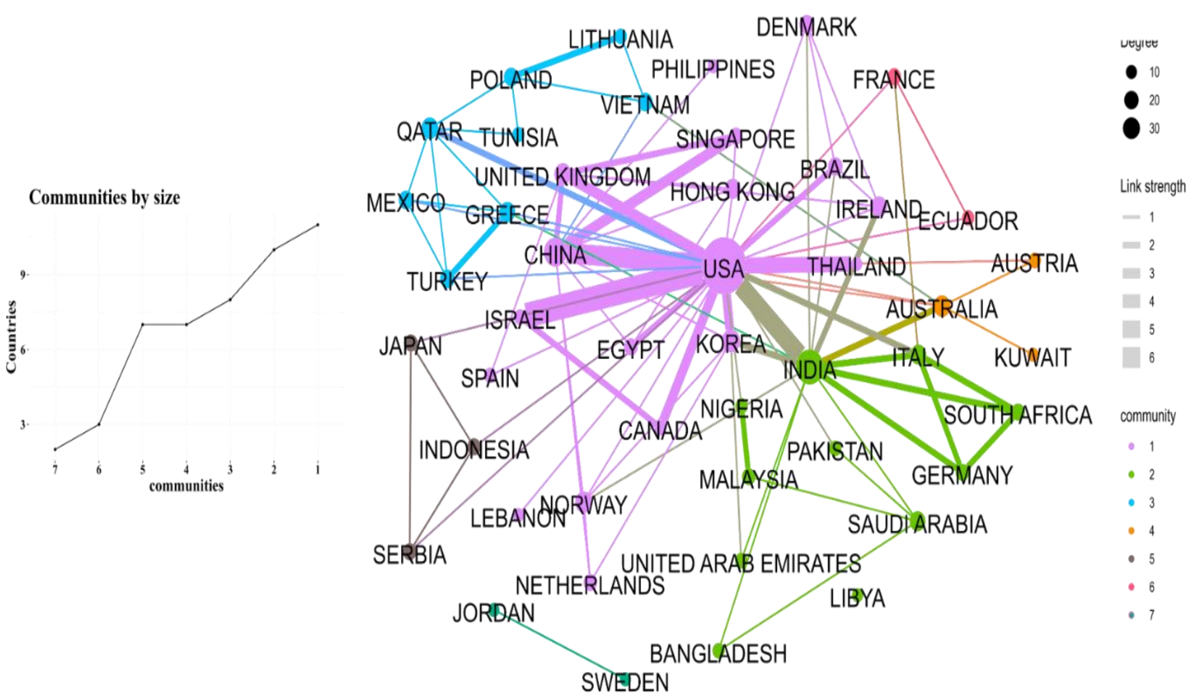

Figure 3. Global Collaboration Network Among Countries in Scientific Research.

One of the collaborations between authors from the USA and Israel is a systematic review on the use of LLMs in breast cancer research [20]. This study highlights the capability of these models to extract textual information and respond to clinical questions. However, it suggests that improvements are needed to address certain inaccuracies in their results. Another strong relationship shown in the figure is between the USA and India. For instance, researchers affiliated with these two countries recently conducted a systematic review to identify the potential uses of ChatGPT [21]. The authors propose that this tool is incredibly useful in fields such as healthcare, marketing, and finance. They also uncover a significant potential for advancements in conversational capabilities with ChatGPT.

Journal Analysis

Table III presents the ten journals with the highest scientific production on LLMs, along with their impact factor, h-index, and Scimago quartile. The journal with the most publications on LLMs is the Cureus Journal of Medical Science, which is not indexed in Scimago but is known to be in Q3. The most cited article in this journal is authored by Professor Temsah O, examining the early presence of ChatGPT in the medical literature, highlighting its potential applications, ethical considerations, and the need for further research in the healthcare field [22].

A recent study investigated whether GPT-4 can match humans in writing introductions for scientific articles. The results showed no significant differences in terms of applicability and content quality between GPT-4 and human-written introductions. Most evaluators preferred the introductions generated by GPT-4, and many could not distinguish between those written by GPT-4 and humans [23]. These findings suggest that GPT-4 can be a useful tool for scientific writing. In conclusion, although Table 3 reflects that there is not yet a large volume of writings on the subject, it is certain that the use of ChatGPT and other generative language models in scientific production and bibliographic analysis will continue to grow and develop in the future.

Tabla III. Top Journals by Scientific Production on Large Language Models (LLMs) with Impact Factor, h-index, and Scimago Quartile Distribution.

Journal	WoS	Scopus	Impact Factor	h-index	Quartile
Cureus Journal Of Medical Science	10	0	1.2	–	Q3
Aesthetic Surgery Journal	4	1	1.27	78	Q1
Educational Process: International Journal	0	3	0.42	8	Q2
Frontiers In Education	3	0	0.63	40	Q2
Journal Of Medical Internet Research	2	3	2.02	197	Q1
Value In Health	3	0	1.51	132	Q1
Contemporary Educational Technology	2	0	0.63	20	Q2
Elife	1	0	3.93	206	Q1
Entrepreneurial Business And Economics Review	2	2	0.54	26	Q1
Healthcare	4	0	2.5	29	Q2

Figure 4 illustrates the citation network among journals, distinguishing three main groups. The first group is notable for studies on the automation of LLMs like ChatGPT for scientific research [24], [25]. The second group addresses the use of ChatGPT in medicine, focusing on the potential analysis of metadata it can offer [19], [26]. The third group centers on the application of ChatGPT in clinical decision-making and the improvement of efficiency in medical practice, as well as exploring its role in scientific research and publication [27]. The diversification of approaches in the application of ChatGPT, as reflected in the citation network among journals, shows a growing interest in its versatility across various scientific domains. This phenomenon demonstrates how journals enhance the use of these models, promoting their adoption in interdisciplinary research and improving scientific efficiency overall.

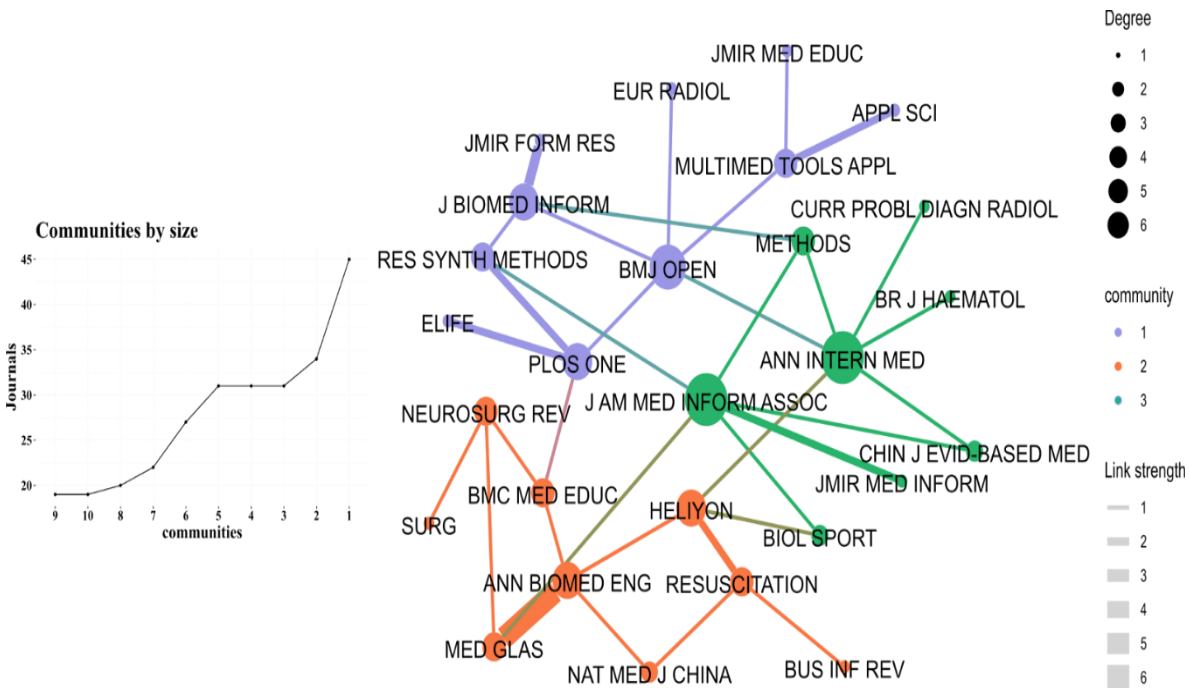

Figure 4. Citation Network Among Journals Highlighting Collaborative Communities.

Author Collaboration Network

Table IV presents the ten most prominent researchers in the use of LLM in systematic reviews of the scientific literature. Among them, Dr. Rohun Gupta stands out for having the highest article production. In his most recent article, he addresses the limitations in systematic searches, highlighting the importance of the user's ethical responsibility when employing these LLM-based tools such as ChatGPT [28]. On the other hand, Professor Girish N. Nadkarni stands out for having the highest H-index. Among his contributions, he addresses the challenges related to inconsistent precision, while significantly recognizing the potential of LLMs in textual information extraction [20].

Table IV. Top Researchers in Large Language Models (LLMs) by Total Articles, Scopus H-Index, and Affiliation.

No	Researcher	Total Articles*	Scopus H-Index	Affiliation
1	Gupta R	4	7	Saint Louis University School Of Medicine,St. Louis, United States
2	Klang E	4	34	Icahn School Of Medicine At Mount Sinai,New York, United States
3	Demirkol M	3	7	Firat Üniversitesi,Elaziğ, Turkey
4	Karakose T	3	16	Dumlupinar Üniversitesi,Kütahya, Turkey
5	Konen E	3	38	Tel Aviv University,Tel Aviv-Yafo, Israel
6	Nadkarni G	3	65	Icahn School Of Medicine At Mount Sinai, New York, United States
7	Sorin V	3	11	Tel Aviv University, Tel Aviv-Yafo, Israel
8	Yirci R	3	12	Kahramanmaras Sütçü Imam Üniversitesi, Kahramanmaras, Turkey
9	Lee S	2	8	University Of Nottingham Ningbo China, Ningbo, China
10	Wang Z	2	6	Shanghai Polytechnic University, Shanghai, China

Figure V illustrates the scientific collaboration network among the researchers mentioned in Table 4. This network is based on Ego Networks or personal networks. The network represented in the figure is divided into four groups or components. The first component is the largest within the network and provides contributions of significant interest. Professor Zheng Wang proposes a framework for the integration of LLMs in systematic reviews and evaluates the concordance between the GPT-4 model and human reviewers [29]. In the second component is Professor Ayanesh Kumar, who recently performed a comparative analysis of the efficacy and accuracy between ChaGPT and human reviewers [30]. In the third component, we observed a great deal of synergy on the part of the investigators. For example, Vera Sorin, Eli Konen, Girish Nadkarni, and Eyal Klang examine LLMs, such as GPT-3.5 and GPT-4, for the generation of legitimate and valid multiple-choice questionnaires (MCQs) in medical examinations [31]. In the last group, Professors Turgut Karakose, Murat Demirkol, and Ramazan Yirci participated in the collaborative network by analyzing the use of LLMs in human-IA collaboration and their usefulness in generating accurate and clear information in scientific research [32].

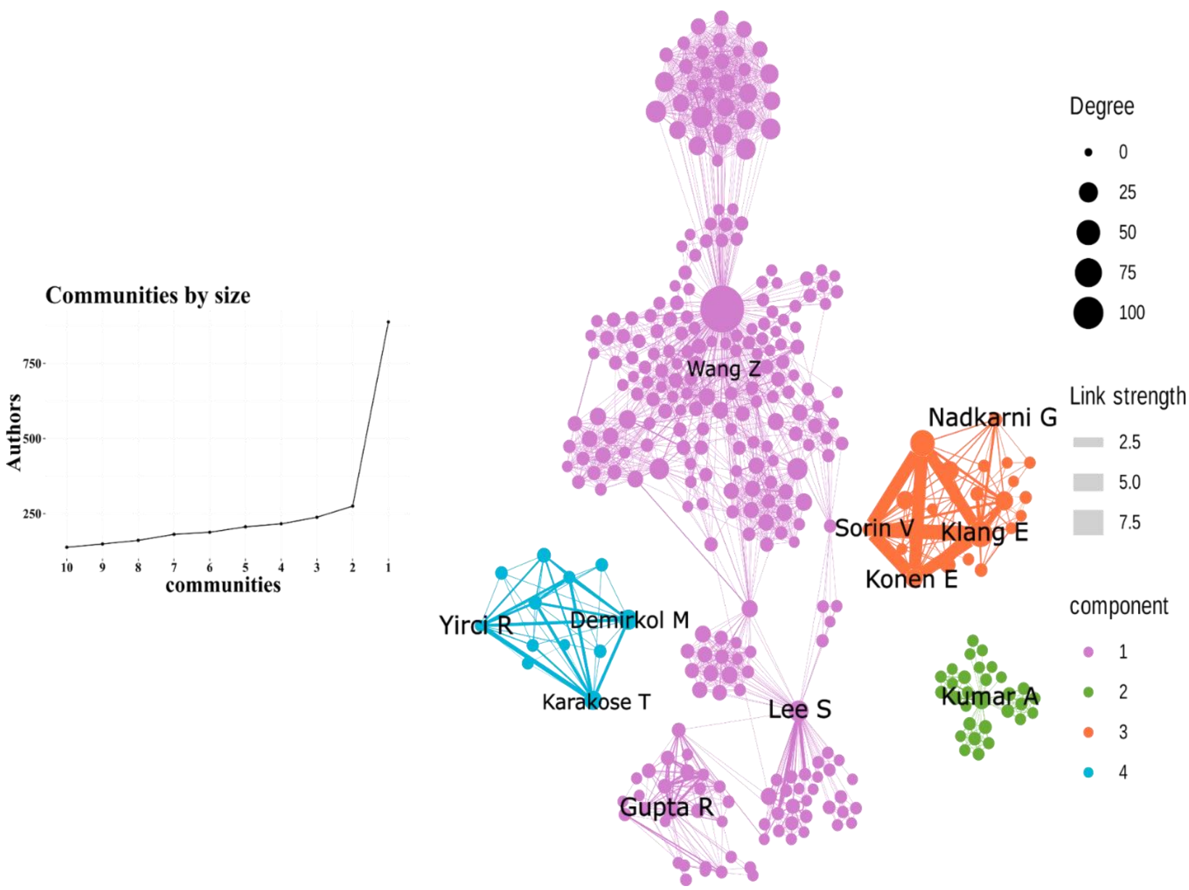

Figure V. Egonetwork Analysis of Top Researchers in Large Language Models (LLMs): Collaborative Communities and Link Strength.

3.2 Tree of Science

Root

The foundational articles in LLMs show a global perspective of the applications and implications of this type of models in different areas, specifically in health and scientific research. These theoretical contributions form the basis for understanding the capabilities and limitations of LLMs. An example is the work of Sallam [14] and Biswas [30] where they show the usefulness of Chatgpt in health, highlighting its impact on education, interaction with patients, diagnosis and clinical decisions.

Regarding advances in writing articles with LLMs, Thorp [31] presents a general overview of the improvement in the efficiency of scientific writing but shows the need to carry out ethical questions about these new methodologies. This is important because it is necessary to integrate policies for the use of AI in the development of scientific advances.

On the other hand, Salvagno [32] highlights the need for continued research to integrate AI in environments where there is a lot of pressure to obtain results.

Trunk

The articles in the trunk make important contributions to the field of LLMs that support this topic. Dergaa [36] and Alexis-Tahta [37] emphasize the innovative potential of tools such as ChatGPT and IA in academic, educational, medical, and administrative contexts. These models can automate a wide range of tasks and also showcase the ability to produce accurate summaries, thereby promoting creativity and innovation in the research process.
In addition, a thorough scientometric review offers a comprehensive overview of the different subdomains that make up the field of LLMs. This detailed analysis pinpoints important areas such as language generation and machine translation, enhancing our understanding of how they are utilized in academic and professional contexts. These findings are detailed in the article referenced by Roumeliotis [38].

The incorporation of technologies like ChatGPT in education and writing presents both challenges and opportunities that require careful consideration. One primary concern highlighted by the authors is the preservation of academic integrity, with an emphasis on maintaining originality of content and preventing plagiarism. However, they also acknowledge and advocate for the inherent benefits of ChatGPT, including its efficacy in content generation and its potential to stimulate creativity. These key points are discussed in the referenced article [39].
The algorithm developed by Blondel et al. [40] was used to identify the branches in the Tree of Science (ToS). The primary branches of research on Large Language Models (LLMs) and literature reviews are represented by the three largest clusters, as shown in Figure VI.

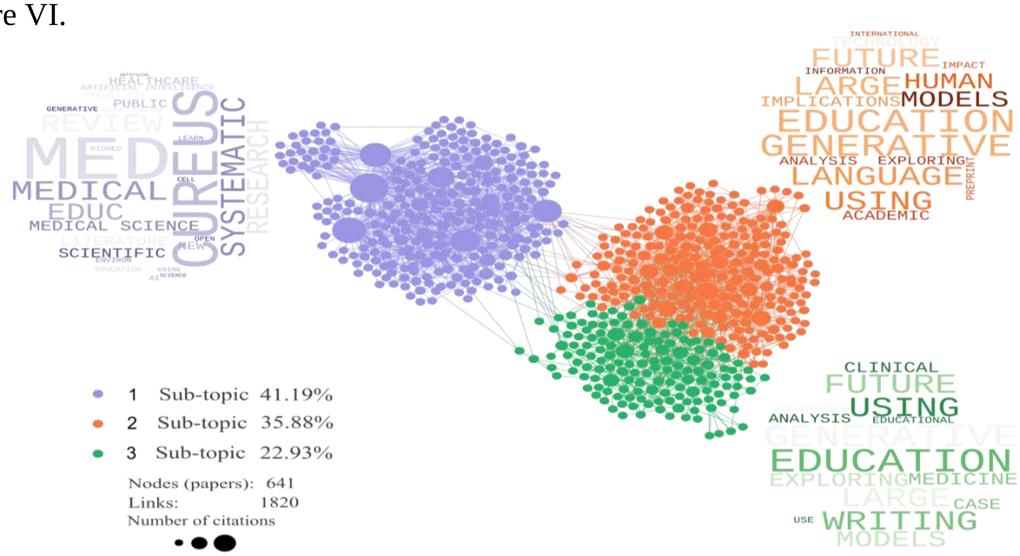

Figure 6. Citation Network Analysis of Sub-Topics in Large Language Model Research.

Branch 1: The role of LLMs, ChatGPT and AI in Medical research

The first branch focuses on the application of artificial intelligence in medical sciences, highlighting the significant role of ChatGPT in current research. For example, a recent study conducted a comprehensive evaluation of this tool, ranging from writing literature reviews to improving medical reports [41]. Likewise, Treviño-Juárez [42] highlights the use of ChatGPT as a tool to mitigate bias in numerous studies, integrating automation processes and ChatGPT-4 to accelerate the synthesis of information in systematic reviews in evidence-based medicine. On the other hand, some studies have analyzed the pros and cons of the use of artificial intelligence in educational settings, addressing ethical challenges, concerns, and implications of this tool, and concluding on the need to encourage initiatives that promote the effective use of AI among educators and students [43]. Undoubtedly, the evolution of LLMs and the revolutionary impact of ChatGPT in recent years in the field of artificial intelligence is a topic that researchers are exploring in depth. All ethical issues are being addressed and the enormous potential for impact in areas such as public health is being recognized. As a result, public policy initiatives have been proposed to regulate AI-generated texts [44].

Branch 2: ChatBots and generative tools in the education sector

The analysis of the second branch delves into the correlation of uses of artificial intelligence, in specific, generative models and natural language processors like ChatGPT and education (AIEd, artificial intelligence in Education), particularly in the dissemination of knowledge to students and its instructional role at the secondary level. Rudolph et al. [45] highlight that the use of ChatGPT enables students to access various databases or knowledge through a single connection/search engine, thus positively influencing the learning process by streamlining information retrieval and text generation. However, at its current developmental stage, it is constrained in performing specific tasks such as language translation or addressing niche-specific queries due to its training algorithm, posing challenges in providing precise and accurate information for students. According to Montenegro-Rueda [46], it is crucial for educators to be well-versed with the tool and its functionalities, as responsible use of text generation tools has been proven to enhance the teaching-learning process.

Branch 3: Benefits, Challenges, and Ethical Considerations of ChatGPT in Medical Education

Branch Three collectively explores the integration of LLMs, particularly ChatGPT, in medical education, providing insights into their benefits and limitations. Artsi et al. [31] and Bečulić et al. [47] demonstrated how LLMs can generate content and support neuro-surgical training, respectively, showcasing practical applications of AI to enhance the learning experience. These studies also highlighted ethical considerations and practical implications. Eysenbach [48] urges the academic community to conduct comprehensive research and address the challenges associated with integrating AI into medical education. Additionally, there is preliminary evidence of the benefits of AI platforms in neurosurgical education [49] and [50] for illness script abilities, but Han et al. [51] emphasize the importance of cautious use of ChatGPT, noting the necessity for human supervision and critical evaluation.

Conclusiones

This scientometric review presents two analyses of LLMs applied to the review of scientific literature. The first analysis involves a scientific mapping of research output, international collaborations, key publishing journals, and the dynamics of scientific collaboration among authors. The second part applies the Tree of Science (ToS) algorithm to identify and explain the major contributions in the form of roots, trunk, and branches.

The primary findings reveal that academic production only exists for the years 2023 and 2024, despite the query spanning from 2000 in both WoS and Scopus. This indicates the novelty of the topic and its significant impact within just two years. It is also important to highlight the role of researchers affiliated with institutions in the United States, which accounts for 25.68% of the total output, followed by China with 8.11%. Furthermore, the medical fields are pioneers in this area, significantly increasing scientific production. For instance, the journals with the most publications are the Cureus Journal of Medical Science and the Aesthetic Surgery Journal. This is further supported by the analysis of researchers, with Dr. Rohun Gupta from SLUCare Academic Pavilion being the most prolific author and Professor Girish N. Nadkarni from Icahn School of Medicine at Mount Sinai having the highest h-index.

References

[1] C. A. Martínez-Pinzón, «Mapeando el Panorama Cienciométrico de los Activos Intangibles: Tendencias, Mediciones y Dinámicas Colaborativas», Interfaces , vol. 6, no. 2, Dec. 2023, Accessed: May 22, 2024. [Online]. Available: https://revistas.unilibre.edu.co/index.php/interfaces/article/view/11190

[2] I. A. Blank, «What are large language models suppodatosed to model?», Trends in Cognitive Sciences, vol. 27, n.o 11, pp. 987-989, nov. 2023, doi: 10.1016/j.tics.2023.08.006.

[3] J. De Winter, «Can ChatGPT be used to predict citation counts, readership, and social media interaction? An exploration among 2222 scientific abstracts», Scientometrics, vol. 129, n.o 4, pp. 2469-2487, abr. 2024, doi: 10.1007/s11192-024-04939-y.

[4] F. E. Sandnes, «Can we identify prominent scholars using ChatGPT?», Scientometrics, vol. 129, n.o 1, pp. 713-718, ene. 2024, doi: 10.1007/s11192-023-04882-4.

[5] W. H. Walters y E. I. Wilder, «Fabrication and errors in the bibliographic citations generated by ChatGPT», Sci Rep, vol. 13, n.o 1, p. 14045, sep. 2023, doi: 10.1038/s41598-023-41032-5.

[6] A. E. Rubio, G. Y. F. Yepes, and L. A. V. Marín, «Gobernanza para el desarrollo y la sostenibilidad de los destinos turísticos: una revisión de la literatura con ToS», Interfaces , vol. 5, no. 1, Sep. 2022, Accessed: May 22, 2024. [Online]. Available: https://revistas.unilibre.edu.co/index.php/interfaces/article/view/9459

[7] P. Duque, M. Cárdenas, y S. Robledo, «Marketing social: evolución y tendencias», entramado, vol. 20, n.o 1, oct. 2023, doi: 10.18041/1900-3803/entramado.1.10378.

[8] A. M. G. A., S. Robledo, y M. Zuluaga, «Topic Modeling: Perspectives From a Literature Review», IEEE Access, vol. 11, pp. 4066-4078, 2023, doi: 10.1109/ACCESS.2022.3232939.

[9] A. Echeverri Rubio, G. Yaneth Flórez Yepes, y L. A. Vargas Marín, «Gobernanza para el desarrollo y la sostenibilidad de los destinos turísticos: una revisión de la literatura con ToS», Revista de Ingenierías Interfaces, vol. 5, n.o 1, pp. 1-22, 2022, Accedido: 23 de mayo de 2024. [Online]. Available: https://dialnet.unirioja.es/servlet/articulo?codigo=8661443

[10] C.-A. Gonzalez-Correa, L.-O. Tapasco-Tapasco, y P.-A. Gomez-Buitrago, «A Method for a Literature Search on Microbiota and Obesity for PhD Biomedical Research Using the Web of Science (WoS) and the Tree of Science (ToS)», ISTL, n.o 99, feb. 2022, doi: 10.29173/istl2679.

[11] D. S. Valencia-Hernandez, S. Robledo, R. Pinilla, N. D. Duque-Méndez, y G. Olivar-Tost, «SAP Algorithm for Citation Analysis: An improvement to Tree of Science», Ing. Inv., vol. 40, n.o 1, pp. 45-49, ene. 2020, doi:10.15446/ing.investig.v40n1.77718.

[12] S. Robledo, J. E. Vasquez, N. D. Duque3-Méndez, y V. Duque-Uribe, «Networking as an entrepreneurial marketing tool: the link between effectuation and word of mouth», JRME, vol. 25, n.o 2, pp. 270-285, mar. 2023, doi: 10.1108/JRME-08-2020-0112.

[13] C. M. Botero, C. B. Milanes, y S. Robledo, «50 years of the Coastal Zone Management Act: The bibliometric influence of the first coastal management law on the world», Marine Policy, vol. 150, p. 105548, abr. 2023, doi: 10.1016/j.marpol.2023.105548.

[14] G. A. Ruiz, S. Robledo Giraldo, y H. Hernando Morales, «Static Code Analysis: A Tree of Science Review», Entre cienc. ing., vol. 17, n.o 34, pp. 9-14, 2023, doi: 10.31908/19098367.2846.

[15] S. Robledo-Giraldo, J. G. Figueroa-Camargo, M. V. Zuluaga-Rojas, S. B. Vélez-Escobar, y P. L. Duque- Hurtado, «Mapping, evolution, and application trends in co-citation analysis: a scientometric approach», rev. investig. desarro. innov, vol. 13, n.o 1, pp. 201-214, feb. 2023, doi: 10.19053/20278306.v13.n1.2023.16070.

[16] D. Díez, J. Díaz-Ospina, S. Robledo, y M. D. P. Rodríguez-Córdoba, «Tendencias teóricas y desafíos en la comunicación de la responsabilidad social corporativa», anagramas rumbos sentidos comun., vol. 20, n.o 40, pp. 146-176, abr. 2022, doi:10.22395/angr.v20n40a7.

[17] M. Sallam, «ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns», Healthcare, vol. 11, n.o 6, p. 887, mar. 2023, doi: 10.3390/healthcare11060887.

[18] R. Gupta et al., «Expanding Cosmetic Plastic Surgery Research With ChatGPT», Aesthetic Surgery Journal, vol. 43, n.o 8, pp. 930-937, jul. 2023, doi: 10.1093/asj/sjad069.

[19] K. Cheng et al., «Potential Use of Artificial Intelligence in Infectious Disease: Take ChatGPT as an Example», Ann Biomed Eng, vol. 51, n.o 6, pp. 1130-1135, jun. 2023, doi: 10.1007/s10439-023-03203-3.

[20] V. Sorin et al., «Utilizing large language models in breast cancer management: systematic review», J Cancer Res Clin Oncol, vol. 150, n.o 3, p. 140, mar. 2024, doi: 10.1007/s00432-024-05678-6.

[21] S. S. Sohail et al., «Decoding ChatGPT: A taxonomy of existing research, current challenges, and possible future directions», Journal of King Saud University – Computer and Information Sciences, vol. 35, n.o 8, p. 101675, sep. 2023, doi: 10.1016/j.jksuci.2023.101675.

[22] O. Temsah et al., «Overview of Early ChatGPT's Presence in Medical Literature: Insights From a Hybrid Literature Review by ChatGPT and Human Experts», Cureus, abr. 2023, doi: 10.7759/cureus.37281.

[23] B. Sikander, J. J. Baker, C. D. Deveci, L. Lund, y J. Rosenberg, «ChatGPT-4 and Human Researchers Are Equal in Writing Scientific Introduction Sections: A Blinded, Randomized, Non-inferiority Controlled Study», Cureus, nov. 2023, doi: 10.7759/cureus.49019.

[24] Z. N. Khlaif et al., «The Potential and Concerns of Using AI in Scientific Research:ChatGPT Performance Evaluation», JMIR Med Educ, vol. 9, p. e47049, sep. 2023, doi: 10.2196/47049.

[25] H. Raja et al., «Automated Category and Trend Analysis of Scientific Articles on Ophthalmology Using Large Language Models: Development and Usability Study», JMIR Form Res, vol. 8, p. e52462, mar. 2024, doi: 10.2196/52462.

[26] H. Bagde, A. Dhopte, M. K. Alam, y R. Basri, «A systematic review and meta-analysis on ChatGPT and its utilization in medical and dental research», Heliyon, vol.9, n.o 12, p. e23050, dic. 2023, doi: 10.1016/j.heliyon.2023.e23050.

[27] N. Schopow, G. Osterhoff, y D. Baur, «Applications of the Natural Language Processing Tool ChatGPT in Clinical Practice: Comparative Study and Augmented Systematic Review», JMIR Med Inform, vol. 11, p. e48933, nov. 2023, doi: 10.2196/48933.

[28] R. Gupta, K. Bagdady, y B. A. Mailey, «Response to: Truth or Lies? The Pitfalls and Limitations of ChatGPT in Systematic Review Creation», Aesthetic Surgery Journal, vol. 43, n.o 8, pp. NP656-NP657, jul. 2023, doi: 10.1093/asj/sjad108.

[29] B. Hasan et al., «Integrating large language models in systematic reviews: a ramework and case study using ROBINS-I for risk of bias assessment», BMJ EBM, p. bmjebm-2023-112597, feb. 2024, doi: 10.1136/bmjebm-2023-112597.

[30] S. Ruksakulpiwat et al., «Assessing the Efficacy of ChatGPT Versus Human Researchers in Identifying Relevant Studies on mHealth Interventions for Improving Medication Adherence in Patients With Ischemic Stroke When Conducting Systematic Reviews: Comparative Analysis», JMIR Mhealth Uhealth, vol. 12, p. e51526, may 2024, doi: 10.2196/51526.

[31] Y. Artsi, V. Sorin, E. Konen, B. S. Glicksberg, G. Nadkarni, y E. Klang, «Large language models for generating medical examinations: systematic review», BMC Med Educ, vol. 24, n.o 1, p. 354, mar. 2024, doi: 10.1186/s12909-024-05239-y.

[32] T. Karakose, M. Demirkol, R. Yirci, H. Polat, T. Y. Ozdemir, y T. Tülübaş, «A Conversation with ChatGPT about Digital Leadership and Technology Integration: Comparative Analysis Based on Human–AI Collaboration», Administrative Sciences, vol. 13, n.o 7, p. 157, jun. 2023, doi: 10.3390/admsci13070157.

[33] S. Biswas, «ChatGPT and the Future of Medical Writing», Radiology, vol. 307, n.o 2, p. e223312, abr. 2023, doi: 10.1148/radiol.223312.

[34] H. H. Thorp, «ChatGPT is fun, but not an author», Science, vol. 379, n.o 6630, pp.313-313, ene. 2023, doi: 10.1126/science.adg7879.

[35] M. Salvagno, F. S. Taccone, y A. G. Gerli, «Can artificial intelligence help for scientific writing?», Crit Care, vol. 27, n.o 1, p. 75, feb. 2023, doi: 10.1186/s13054-023-04380-2.

[36] I. Dergaa, K. Chamari, P. Zmijewski, y H. Ben Saad, «From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing», bs, vol. 40, n.o 2, pp. 615-622, 2023, doi: 10.5114/biolsport.2023.125623.

[37] Alexis-Tahta, Joaquin. 2023. «Inteligencia Artificial en la Administración: una revisión de la literatura utilizando Tree of Science.» Interfaces 6 (1). https://revistas.unilibre.edu.co/index.php/interfaces/article/view/10731.

[38] K. I. Roumeliotis y N. D. Tselikas, «ChatGPT and Open-AI Models: A Preliminary Review», Future Internet, vol. 15, n.o 6, p. 192, may 2023, doi: 10.3390/fi15060192.

[39] M. Imran y N. Almusharraf, «Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature», CONT ED TECHNOLOGY, vol. 15, n.o 4, p. ep464, oct. 2023, doi: 10.30935/cedtech/13605.

[40] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, y E. Lefebvre, «Fast unfolding of communities in large networks», J. Stat. Mech., vol. 2008, n.o 10, p. P10008, oct. 2008, doi: 10.1088/1742-5468/2008/10/P10008.

[41] S. Ruksakulpiwat, A. Kumar, y A. Ajibade, «Using ChatGPT in Medical Research: Current Status and Future Directions», Journal Of Multidisciplinary Healthcare, vol. Volume 16, pp. 1513-1520, may 2023, doi: 10.2147/jmdh.s413470.

[42] A. S. Treviño-Juarez, «Assessing Risk of Bias Using ChatGPT-4 and Cochrane ROB2 Tool», Med.Sci.Educ., abr. 2024, doi: 10.1007/s40670-024-02034-8.

[43] I. Maita, S. Saide, A. M. Putri, y D. Muwardi, «ProsCons of Artificial Intelligence-ChatGPT Adoption in Education Settings: A Literature Review and Future Research Agendas», IEEE Eng. Manag. Rev., pp. 1-12, 2024, doi:10.1109/EMR.2024.3394540.

[44] L. De Angelis et al., «ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health», Front. Public Health, vol. 11, p. 1166120, abr. 2023, doi: 10.3389/fpubh.2023.1166120.

[45] J. Rudolph, S. Tan, and S. Tan, «ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?», JALT, vol. 6, no. 1, pp. 342–363, Jan. 2023.

[46] M. Montenegro-Rueda, J. Fernández-Cerero, J. M. Fernández-Batanero, y E. López-Meneses, «Impact of the Implementation of ChatGPT in Education: A Systematic Review», Computers, vol. 12, n.o 8, p. 153, jul. 2023, doi:10.3390/computers12080153.

[47] H. Bečulić, E. Begagić, R. Skomorac, A. Mašović, E. Selimović, y M. Pojskić, «ChatGPT's contributions to the evolution of neurosurgical practice and education: a systematic review of benefits, concerns and limitations», Medicinski Glasnik, n.o 1, 2024, doi: 10.17392/1661-23.

[48] G. Eysenbach, «The Role of ChatGPT, Generative Language Models, and Artificial Intelligence in Medical Education: A Conversation With ChatGPT and a Call for Papers», JMIR Med Educ, vol. 9, p. e46885, mar. 2023, doi: 10.2196/46885.

[49] U. T. Sevgi, G. Erol, Y. Doğruel, O. F. Sönmez, R. S. Tubbs, y A. Güngor, «The role of an open artificial intelligence platform in modern neurosurgical education: a preliminary study», Neurosurg Rev, vol. 46, n.o 1, p. 86, abr. 2023, doi: 10.1007/s10143-023-01998-2.

[50] Y. Yanagita, D. Yokokawa, F. Fukuzawa, S. Uchida, T. Uehara, y M. Ikusaka, «Expert

assessment of ChatGPT’s ability to generate illness scripts: an evaluative study», BMC Med Educ, vol. 24, n.o 1, p. 536, may 2024, doi: 10.1186/s12909-024-05534-8.

[51] Z. Han, F. Battaglia, A. Udaiyar, A. Fooks, and S. R. Terlecky, «An explorative assessment of ChatGPT as an aid in medical education: Use it with caution», *Med. Teach.*, vol. 46, no. 5, pp. 657–664, May 2024.