International Journal of Combinatorial Optimization Problems and Informatics, 17(2), 2026, 1-13. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2026.v17i2.1281

www.editada.org

Generative AI and the scientific landscape: a bibliometric exploration of its global impact

Edgar Gonzalo Cossio Franco[1], *Juan Humberto Sossa Azuela*[2], *Víctor Manuel Larios Rosillo*[3-1], *Rocio Maciel Arellano*[3-2], *María Esmeralda Arreola Marín*[4], *Mariela Chávez Marcial*[4], *José Iraic Alcántar Alcántar*[4], *Samuel Efrén Viñas Álvarez*[5], *Ramón Alejandro Briseño Martínez*[6], *Carolina Del Valle Soto*[6]

[1] Instituto de Información Estadística y Geográfica de Jalisco, Zapopan, Jalisco, México.
[2] Instituto Politécnico Nacional, Centro de Investigación en Computación, México.
[3-1] Universidad de Guadalajara, CUGDL, [3-2] Universidad de Guadalajara, CUCEA
[4] Tecnológico Nacional de México, Instituto Tecnológico Superior de Ciudad Hidalgo.
[5] Tecnológico Nacional de México, Instituto Tecnológico de Zitácuaro.
[6] Universidad Panamericana. Facultad de Ingeniería. Zapopan, Jalisco, México.

edgar.cossio@iieg.gob.mx, hsossa@cic.ipn.mx, vmlarios@cucea.udg.mx, rmaciel@cucea.udg.mx, marreola@cdhidalgo.tecnm.mx, mchavez@cdhidalgo.tecnm.mx, jiraic@cdhidalgo.tecnm.mx, samuel.va@zitacuaro.tecnm.mx, rbriseno@up.edu.mx, cvalle@up.edu.mx

Abstract. The present comparative bibliometric study (2020-2025) of the Scopus and WoS databases on Generative Artificial Intelligence (GenAI) reveals accelerated growth, concentrating more than 95% of the production and reaching its peak impact in 2025. Thematically, the intersection of communication and technology/education dominates. Geographically, the United States leads production, but Asia-Pacific institutions (Hong Kong) are key. The field of GenAI is a massive trend driven by concentrated collaboration between North America and Asia.
Keywords: Generative Artificial Intelligence, Artificial Intelligence, Bibliometrics, Scopus, WoS.

Article Info
Received Dec 26, 2025
Accepted January 4, 2026

1 Introduction

The context of the rise of Generative Artificial Intelligence (GenAI) in scientific and academic research is marked by exponential growth, rapid interdisciplinary discussion, and a set of ethical and regulatory challenges that directly affect publication integrity (Orchard & Tasiemski, 2023; Ding, Lawson, & Shapira, 2024; Ganjavi et al., 2024).

1. Definition and the technological turning point

GenAI is a branch of artificial intelligence that uses deep learning models to generate novel and creative content (such as text, images, audio, code, or simulations) from the data it has been trained on (Orchard & Tasiemski, 2023; Eke, 2023; Wang, Li, Wang, & Zhao, 2024; Dwivedi & Elluri, 2024).

This technology is considered more creative and flexible than traditional AI techniques, as it can create entirely new content rather than simply reasoning or predicting based on existing rules or data (Wang, Li, Wang, & Zhao, 2024).

The emergence of LLMs as GPTs as pioneers in the field is identified as a canonical point. (Orchard & Tasiemski, 2023; Eke, 2023).

- ChatGPT was launched in November 2022 (Eke, 2023; Khan et al., 2024).
- At that time, it was the fastest-adopted technology with 100 million users (Ganjavi et al., 2024).
- Due to its maturity, this technology is considered the foundation for the next industrial revolution (Orchard & Tasiemski, 2023; Dwivedi & Elluri, 2024).

2. Scientific production and its exponential growth.

GenAI's growth is clearly shown in the production of (scientific/academic) literature with an interest in recent years. (Dwivedi & Elluri, 2024).

- Regarding the publication trend, an initial moment is identified in terms of GenIA and LLM, from 2023, it is not until 2018 that an interesting increase is observed (Dwivedi & Elluri, 2024; Doshi & Kaleel, 2025; Wang, Li, Wang, & Zhao, 2024).
- Regarding bibliometric data, an analysis in Scopus from 2023-2025 showed a substantial increase from 27 to 211, respectively (Doshi & Kaleel, 2025).
- Research topics: Initially, the most prominent topics focused on technical advances and developments in GenAI systems (such as Generative Adversarial Networks or GANs) and their applications to image processing, pattern recognition, and computer vision.
- More recently (2023 onwards), ChatGPT, LLMs, and the application of GenAI to the health and education sectors have become emerging research topics (Dwivedi & Elluri, 2024).

3. Interdisciplinary diffusion and key applications

GenAI is positioning itself as a general-purpose technology that promises to revolutionize science (Ding, Lawson, & Shapira, 2024). Its use has rapidly spread beyond its origins in computer science, influencing diverse academic disciplines (Dwivedi & Elluri, 2024).

The main fields of application of GenAI in research include (based on recent bibliometric analyses):

Discipline	% Implementation (ChatGPT)	Research Applications
Healthcare/Medicine	33.6% (Dominant field)	Diagnostic assistance, improving medical research efficiency, clinical decision support (CDS), and medical education, passing MBA or law exams (Orchard & Tasiemski, 2023; Eke, 2023; Khan et al., 2024; Dwivedi & Elluri, 2024).
Computer Science/IT	18.6%	Code generation (Github Copilot), algorithm optimization (GANs, neural networks), and big data analysis (Orchard & Tasiemski, 2023; Khan et al., 2024; Dwivedi & Elluri, 2024).
Education/Research	17.3%	Idea generation, literature review, improving productivity in text writing and presentation preparation, acting as a virtual tutor, or creating lesson plans (Ding, Lawson, & Shapira, 2024; Eke, 2023; Khan et al., 2024; Dwivedi & Elluri, 2024).
Other domains	Widespread diffusion	Use extends to chemistry, geography, geology, psychology, sociology, engineering (biomedical, mechanical), and law, being used for predictive modeling or hypothesis testing (Ding, Lawson, & Shapira, 2024; Orchard & Tasiemski, 2023; Khan et al., 2024; Dwivedi & Elluri, 2024).

4. Ethical and regulatory challenges in research

The rapid adoption of GenAI has generated significant challenges, especially regarding academic integrity and the scientific publication process (Ding, Lawson, & Shapira, 2024; Eke, 2023).

A. Integrity and authorship

Using GenIA to create documents and present them as your own or original violates academic integrity, honesty, trust, respect, fairness, and responsibility (Eke, 2023).

- There is a fear regarding the possibility of the essay being replaced as an evaluation mechanism, in place of tools such as ChatGPT and their excessive use by students and researchers.
- A latent risk is hallucinations on the part of the GenAI models, given the certainty in the response, perhaps without the knowledge (Ding, Lawson, & Shapira, 2024; Orchard & Tasiemski, 2023; Eke, 2023; Ganjavi et al., 2024).
- Regarding prohibited authorship, there is near-universal consensus that GenAI should not appear as an author in research (explicit prohibition in 96% of publishers and 98% of surveyed journals). The focus is on the fact that AI should not assume responsibility or legal rights over the presented content. (Ganjavi et al., 2024).

B. Regulatory heterogeneity

The absence of standards or regulation regarding the use of GenIA creates confusion in the scientific community, according to Ganjavi et al. (2004):

- Lack of guidelines: Only 24% of the 100 largest academic publishers had specific guidelines on the use of GenAI, although the majority of the top 100 ranked scientific journals (87%) did provide them.
- Discordance in disclosure: Permitted use and disclosure requirements vary substantially between publishers and journals, even being contradictory in some cases. Editorial journals require disclosure of use, but the location (methods section, acknowledgments, cover letter, or new section) and the detail of what must be reported (model name, version, purpose) are not unified.
- Need for standardization: The lack of cohesive and transdisciplinary guidelines, developed through a formal consensus process, is an urgent problem that must be addressed to protect the integrity of scientific production.

5. Geographic landscape of GenAI research

Research on GenAI and LLMs is a global effort, with 80 countries participating in studies related to ChatGPT (Khan et al., 2024).

- Global leadership: The United States and China are the 2 dominant players. The United States has led in GenAI production (39% of global publications). China, although the second-largest producer of GenAI publications, has a much larger base in other areas of AI (Ding, Lawson, & Shapira, 2024).
- Contributing countries: Other countries with significant contributions include China (238 publications in WoS 2013-2024), India (69 publications in Scopus 2023-2025/157 country mentions in ChatGPT), the United Kingdom, Germany, and Korea (Khan et al., 2024; Doshi & Kaleel, 2025; Dwivedi & Elluri, 2024).
- Emerging institutions: High-growth universities in GenAI research include the University of California, Cornell University, and Nanyang Technological University, as well as leading institutions in ChatGPT research such as King Saud University and All India Institute of Medical Sciences (Khan et al., 2024; Doshi & Kaleel, 2025).

When observing the rise of Genetic Engineering in research, what emerges is an accelerated technological adoption driven by LLMs, with visible effects on working methods and the speed of discoveries in areas such as medicine and engineering. Amidst this progress, the need arises to navigate the regulatory framework and address essential risks if the integrity of and trust in scientific output is to be preserved. (Orchard & Tasiemski, 2023; Eke, 2023; Ganjavi et al., 2024; Ding, Lawson, & Shapira, 2024).

A scenario has been identified that resembles a newly opened digital laboratory for all disciplines: it offers tools with unprecedented power, while the security handbook and ethical standards continue to be written on the fly. (Dwivedi & Elluri, 2024).

A. Justification: Why is measuring scientific impact important?

Measuring the scientific impact of GenAI is relevant for several crucial reasons:

Exploring the structure and dissemination of research allows us to understand its impact (Doshi and Kaleel, 2025; Dwivedi and Elluri, 2024) and how knowledge is disseminated (Doshi and Kaleel, 2025), as well as revealing whether GenAI is transforming scientific practices and the global distribution of output (Ding, Lawson, and Shapira, 2024; Doshi and Kaleel, 2025).

Measuring it also guides future policies and strategies, providing key information for initiatives and avoiding inaccurate estimates of their scope, while anticipating the dynamics that affect human resources (Dwivedi and Elluri, 2024; Ding, Lawson, and Shapira, 2024).

Academic integrity comes into play by highlighting the heterogeneity and contradictions among publishers, underscoring the urgent need for coherent and transdisciplinary ethical standards (Ganjavi et al., 2024).

Finally, rigorous evaluation provides reliable data on capabilities, risks, and opportunities, essential for deciding on the appropriate use of each system (Wallach et al., n.d.).

B. Objective of the study: Global bibliometric analysis

The central objective of a global bibliometric study on GenAI and LLMs is:

To analyze the structure, diffusion among institutions, and geographical distribution of research in order to detect emerging trends, measure their global influence, and identify key players (Doshi & Kaleel, 2025; Ding, Lawson, & Shapira, 2024; Khan et al., 2024).

This involves exploring how GenAI is applied in different areas of knowledge (Khan et al., 2024), providing an organized view of the scientific landscape by identifying key institutions, authors, and journals (Doshi & Kaleel, 2025; Khan et al., 2024; Wang, Li, Wang, & Zhao, 2024), and assessing the impact of research output by considering productivity and the concentration or dispersion of scientific performance at national and disciplinary levels (Doshi & Kaleel, 2025; Ding, Lawson, & Shapira, 2024; Khan et al., 2024).

2 Literature Review

The literature review focuses on the rapid rise of GenAI, its economic applications, and its impact on academic integrity, and identifies key insights into its structure, diffusion, and geography through bibliometric analysis.

The work of (Ding, Lawson & Shapira, 2024) shows that the diffusion of GenAI is advancing rapidly and is cross-disciplinary, with smaller teams but greater international collaboration.

(Doshi & Kaleel, 2025) provides a bibliometric analysis (covering 2023 to 2025) that positions the United States as a leader, identifies IEEE Access as a central source, and confirms exponential growth in 2024.

(Dwivedi & Elluri, 2024) highlights dominant themes (from 2013 to 2024): technical advances in GenAI systems (GANs) and applications in computer vision, pointing to ChatGPT, LLMs, health, and education as emerging trends.

Eke (2023) warns of risks to academic integrity and proposes a joint effort to mitigate impacts and prohibit authorship by GenAI.

Ganjavi et al. (2024) reveal heterogeneity in editorial policies, the almost universal prohibition of authorship by authority (96-98%), and the urgent need for standardized guidelines.

Finally, Khan et al. (2024) confirm the multidisciplinary adoption of ChatGPT, with health as the main area (38.6%), and identify leading countries (80) and institutions.

In (Orchard & Tasiemski, 2023), the business potential and economic effects of GenAI are analyzed, concluding that it has reached a tipping point for broad adoption (possibly the next Industrial Revolution).

(Wallach et al., n.d.) proposes a four-level measurement framework based on social science measurement theory to bring rigor and validity to the evaluation of capabilities, impacts, and risks of GenAI systems.

In (Wang, Li, Wang, & Zhao, 2024), the status and trends of GenAI research (2014-2022) are analyzed, highlighting publication growth starting in 2018 and strong collaboration between China and the United States.

3 Methodology

3.1. Definition and configuration phase

- Source selection: The present study is conducted using the Scopus and JCR databases as primary sources to ensure academic quality.
- Search equation definition: Configure the query using the keyword "GenAI".
- Temporal delimitation: Configure the time filter for the period 2020-2025.

3.2. Extraction and data processing phase

- Search execution: Run the search in the selected databases.
- Data export: Download the resulting metadata.
- Pre-processing: Use Excel or Python to clean the data (remove duplicates, normalize author names, verify date consistency).

3.3. Bibliometric analysis phase (indicator calculation)

- Descriptive analysis: Calculate the total number of publications per year to visualize growth.
- Impact analysis: Calculate total citations and determine the h-index of authors or top reviews.
- Relational analysis: Map international collaboration (countries and co-authorship networks).
- Tool usage: Implement VOSviewer for graphic visualization and Python for more complex statistical analysis, if necessary.

3.4. Interpretation and limitations phase

- Drafting results: Interpret what the obtained indicators mean.
- Declaration of method limitations:
- Language bias: The search was conducted only in English.
- Time window: By cutting off at 2025, data for this year would be incomplete.
- Keywords: The term "GenAI" is very popular now but might exclude older, specific technical terms.

Figure 1 shows the methodology followed for the present proposal: Fig. 1. Methodology.

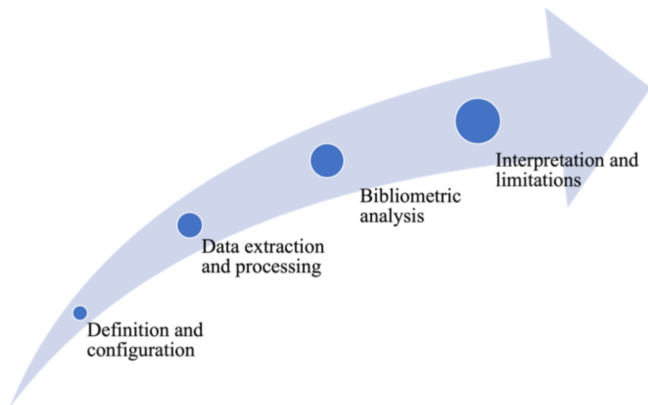

Fig. 1. Methodology.

4 Results

- Databases consulted. For the present work, Scopus and JCR databases were consulted using the term "GenAI" as a search parameter.
  - **Scopus**
    - 4,289 documents found (total)
    - **4,148 documents found (2020-2025)**
    - 96.7% of the results are concentrated in the period 2020-2025.
  - **Web of sciences (WoS)**
    - 2,808 results from Web of Science Core Collection (total)
    - **2,686 results from Web of Science Core Collection (2020-2025)**
    - 95.7% of the results are concentrated in the period 2020-2025.

- Temporal distribution: evolution of publications per year.

  - **Scopus**
    The dataset reflects rapid exponential growth. Despite a decrease in the percentage rate, the jump between 2024 and 2025 was the largest in absolute terms, demonstrating a powerful capacity to scale quickly and sustain massive growth up to 3,000 units.

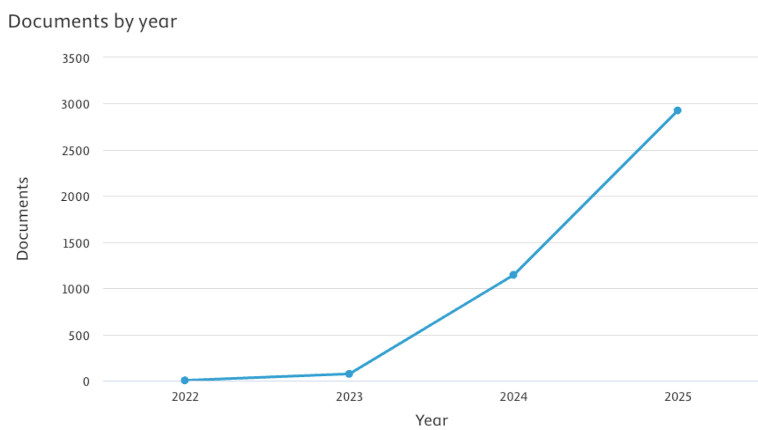

Fig. 2. temporal distribution

  - **Web of sciences (WoS)**
    An accelerated growth pattern is observed in the last three years, characteristic of an emerging technology or research field. From 2020 to 2022, activity was almost null. Growth began modestly in 2023 (with a very low value, barely visible), skyrocketed significantly in 2024 (around 750), and reached its peak in 2025 (with more than 1,800), indicating that the volume of activity doubled drastically in just the last year of the period, confirming an upward and sustained growth trend.

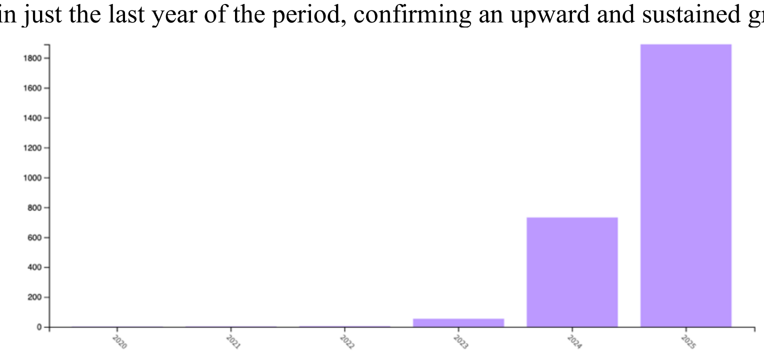

Fig. 3. temporal distribution

- Leading Countries

  - **Scopus**
    The distribution shows a clear dominance of the English language and Western influence, led by the United States, with a figure that doubles the next country, China. The United Kingdom, Australia, and India consolidate the high weight of the Anglosphere, occupying key positions and growing rapidly. Germany and Spain represent the main non-Anglophone contribution, highlighting the marked disparity between the United States and the rest of the countries.

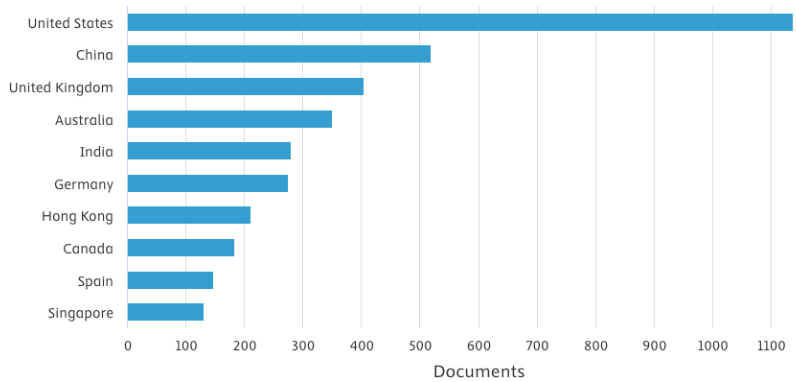

Fig. 4. Documents by country

  - **Web of sciences (WoS)**
    A strong concentration of activity is observed in North America and the Asia-Pacific axis, with the United States being the main driving force.

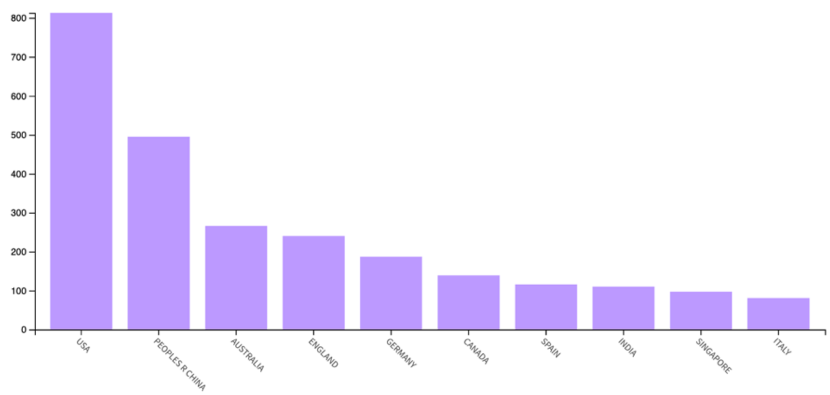

Fig. 5. Documents by country

- Universities

  - **Scopus**
    A marked dominance of Asian and Australian universities is observed, with an outstanding concentration in Hong Kong, which contributes five of the top 10 universities, highlighting The University of Hong Kong as the leader. Nanyang Technological University of Singapore and Monash University of Australia are also in the top positions, while Carnegie Mellon University is the only North American institution.

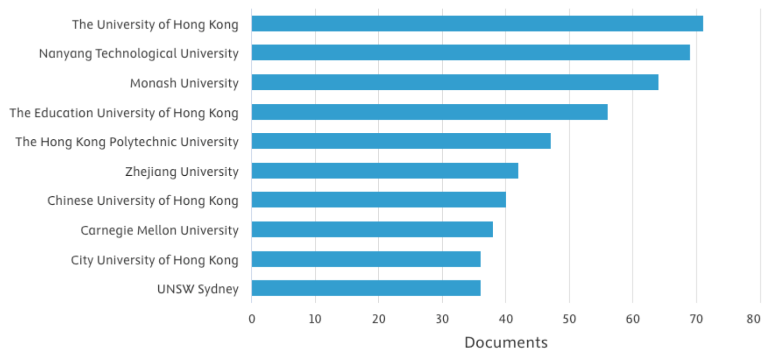

Fig. 6. Documents by university

  - **Web of sciences (WoS)**
    The distribution reveals a high concentration of activity, although without an overwhelming dominance by a single institution. The University of Hong Kong leads with the highest value (around 21), followed closely by several Asian and Australian institutions, such as The Chinese University of Hong Kong and Monash University, all with very similar values (between 16 and 13). The distribution suggests that research is widely distributed among a select group of elite Asia-Pacific universities.

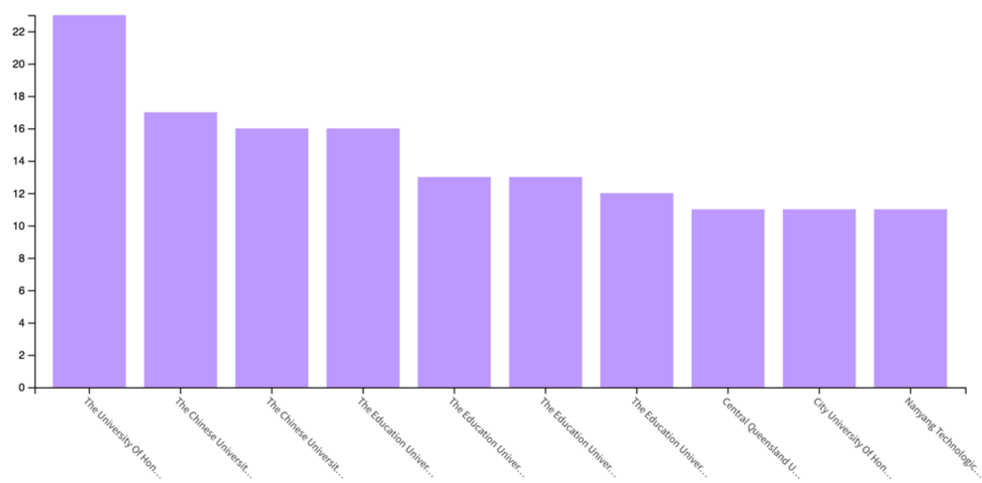

Fig. 7. Documents by university

- Most impacted subject areas.
  - **Scopus**
    Data show that scientific activity is strongly dominated by computer science, leading with the highest value, close to 2,400, and social sciences, following closely with 1,800 units. In this case, the bias is towards technology and the study of society.

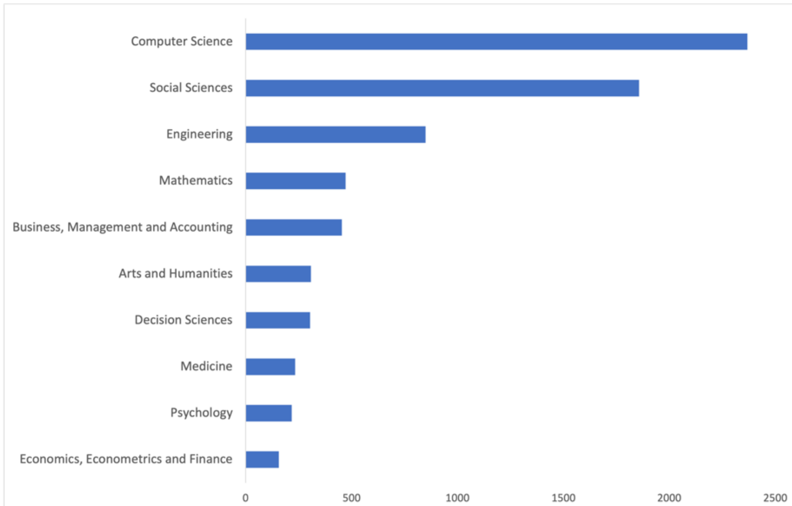

Fig. 8. Documents by subject area

  - **Web of sciences (WoS)**
    Activity is strongly concentrated in two leading disciplines: Computer Science, exceeding 900 units, and the second category, Educational Research, very close to 900. A bias towards the intersection of technology and education is observed.

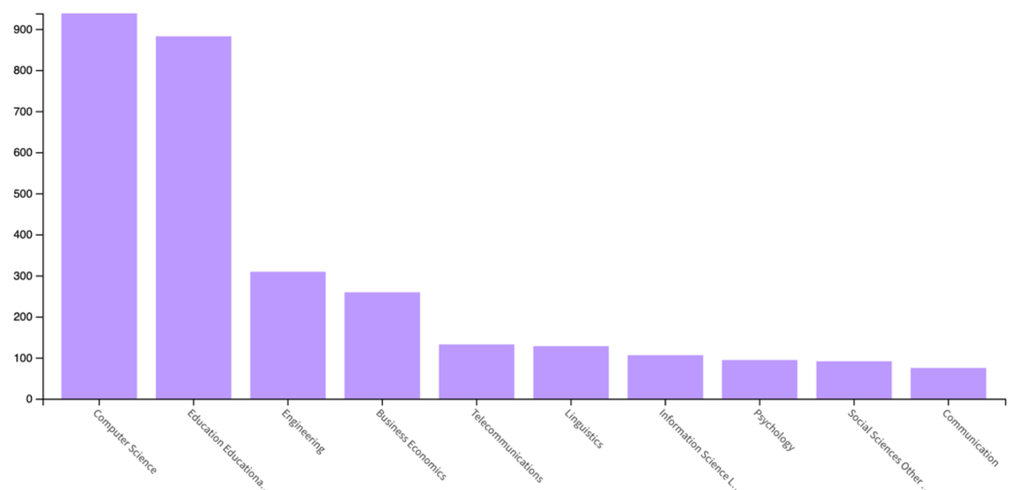

Fig. 9. Research areas

- Citation analysis and impact metrics.
  - **Scopus**
    Figure 11 reveals an explosive and hyper-accelerated growth pattern in the last two years of the period. From 2020 to 2022, activity was almost null. Deployment began in 2023 and reached massive scale in 2024, where document production exceeded the barrier of 800 units. However, the phenomenon skyrocketed dramatically in 2025, where document production reached nearly 2,800 units (a growth of more than 200% compared to 2024) and citations project a record value of more than 28,000, indicating that impact and production have become colossal in the most recent year.

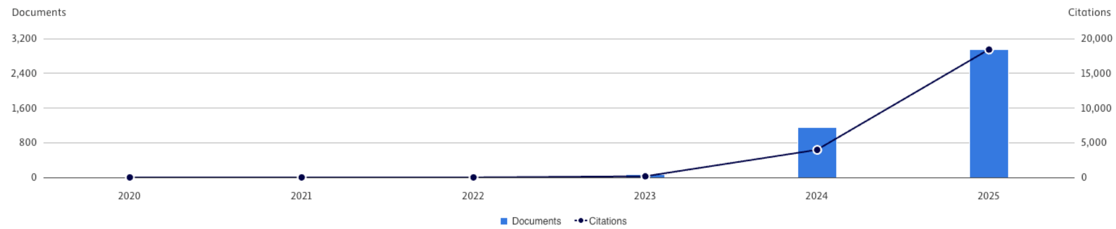

Fig. 11. Citation overview

  - **Web of sciences (WoS)**
    The tree map (Citation Topics Meso) shows an overwhelming dominance of the Communication discipline covering more than half of the total. The second-largest area, Education and Educational Research, is significantly smaller. The rest of the areas contribute little, indicating a clear bias of the sample towards communication and education topics.

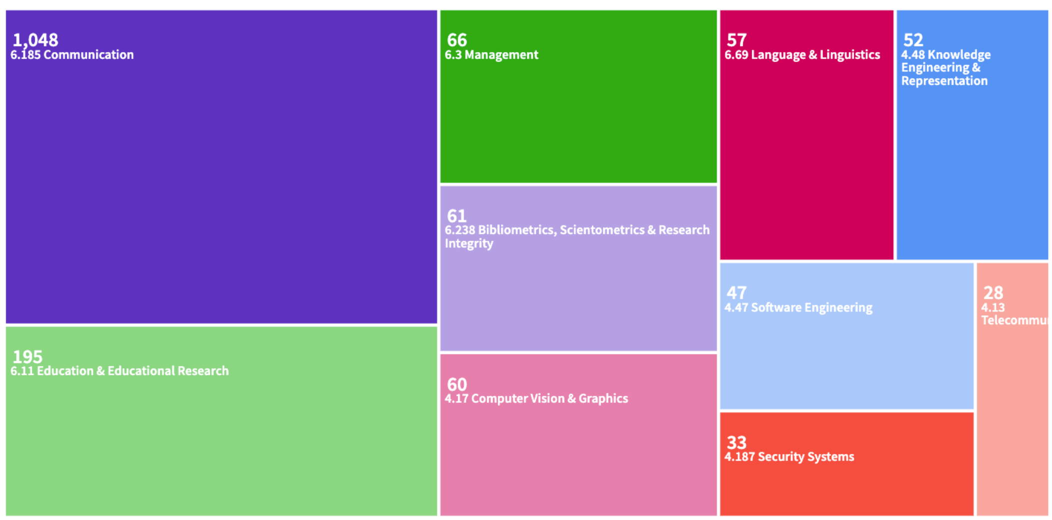

Fig. 10. Citation topics

- Bibliometric network map
  - **Scopus**
    The map reveals a highly centralized global collaboration network. The United States has the largest circle, indicating the highest document production, and is at the center of a strong collaboration cluster with Germany and the United Kingdom. The network also shows a second large cluster of Anglophone (Australia and Canada) and Asian (China, Hong Kong, India) countries, suggesting that research in this area is highly interconnected between English-speaking countries and emerging Asian economies, with few Latin American or Central African countries showing strong links or centrality.

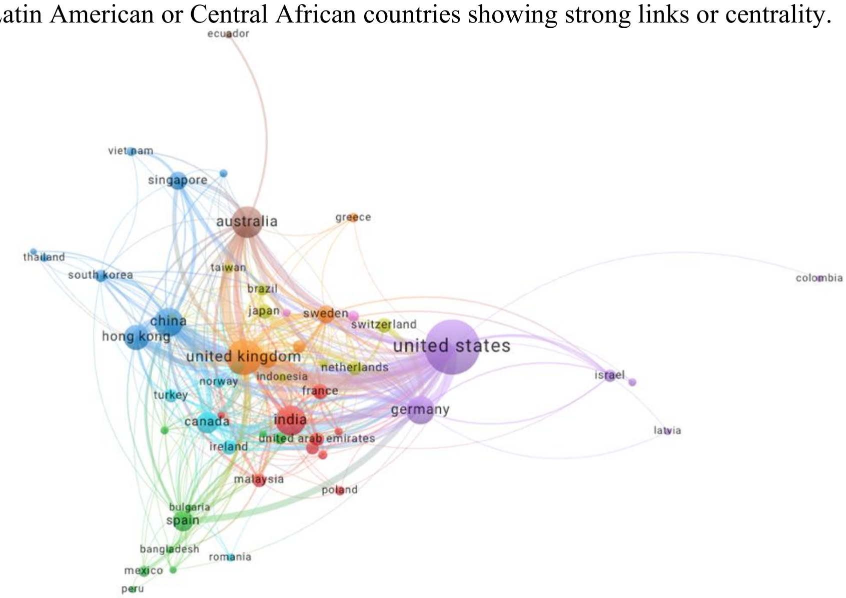

Fig. 11. Co-authorship map

  - **Web of sciences (WoS)**
    The collaboration network is clearly dominated by an Asia-Pacific/North America axis, with the USA and People's Republic of China as the two main centers of production and collaboration. Two large clusters are observed, Western/European and another Asian. The map highlights the high global interconnection of research, with particularly strong collaboration between leading countries (USA and China) despite belonging to distinct geographic clusters.

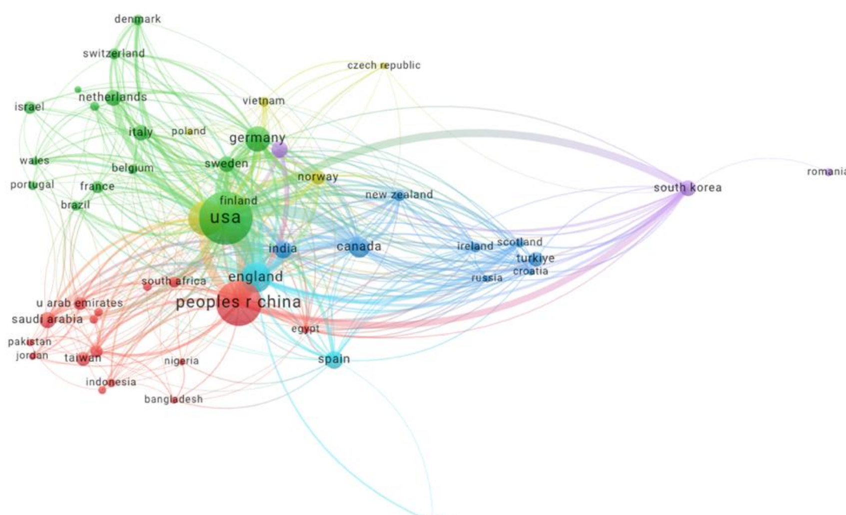

Fig. 12. Co-authorship map

5 Conclusions

Generative Artificial Intelligence (GenAI) has undergone a transformation from niche to dominant global phenomenon in an extremely short period, evidenced by exponential growth.

More than 95% of the total document production is concentrated in the 2020-2025 period, with the peak activity and impact projected for 2025, a year in which document production and citations double or triple compared to the previous year, confirming that the area is in its phase of maximum impact.

This massive growth is thematically biased towards Communication and Technology. It is observed that most of the production is concentrated in areas such as Computer Science and Social Sciences (Scopus) or Educational Research (WoS). The most influential disciplines are Communication and Educational Research, reflecting a marked interest in how technology is redefining communication and learning globally.

Geographically, the North America–Asia-Pacific axis leads the field: the United States tops documentary production, while institutions in Hong Kong and Australia are among the most productive. Collaboration maps confirm this asymmetry, with the US and China as central nodes of a highly interconnected global network, acting as the main drivers of cooperation and scientific generation in this emerging area.

References

Ding, L., Lawson, C., & Shapira, P. (2025). *Rise of generative artificial intelligence in science*. *Scientometrics, 130*(9), 5093–5114. https://doi.org/10.1007/s11192-025-05413-z

Doshi, R., & Kaleel, A. (2025). *Bibliometric analysis of generative AI and large language models in the Scopus database: Trends, insights, and research landscape*. *Applied Data Science and Analysis, 2025*, 7–18. https://doi.org/10.58496/ADSA/2025/003

Dwivedi, R., & Elluri, L. (2024). Exploring Generative Artificial Intelligence Research: A Bibliometric Analysis Approach. *IEEE Access, 12,*119884-119902. https://doi.org/10.1109/ACCESS.2024.3450629

Eke, D. O. (2023). *ChatGPT and the rise of generative AI: Threat to academic integrity? Journal of Responsible Technology, 13*, 100060. https://doi.org/10.1016/j.jrt.2023.100060

Ganjavi, C., Eppler, M. B., Pekcan, A., Biedermann, B., Abreu, A., Collins, G. S., … & Cacciamani, G. E. (2024). *Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: Bibliometric analysis*. *BMJ, 384*, e077192. http://dx.doi.org/10.1136/bmj-2023-077192

Khan, N., Khan, Z., Koubaa, A., Khan, M. K., & Salleh, R. B. (2024). *Global insights and the impact of generative AI-ChatGPT on multidisciplinary: A systematic review and bibliometric analysis*. *Connection Science, 36*(1), 2353630. https://doi.org/10.1080/09540091.2024.2353630

Orchard, T., & Tasiemski, L. (2023). The rise of Generative AI and possible effects on the economy. *Economics and Business Review*, *9*(2), 9–26. https://doi.org/10.18559/ebr.2023.2.732

Wallach, H., Desai, M., Cooper, A. F., Wang, A., Atalla, C., Barocas, S., Blodgett, S. L., Chouldechova, A., Corvi, E., Dow, P. A., Garcia-Gathright, J., Olteanu, A., Pangakis, N. J., Reed, S., Sheng, E., Vann, D., Vaughan, J. W., Vogel, M., Washington, H., & Jacobs, A. Z. (2025). Position: Evaluating Generative AI Systems Is a Social Science Measurement Challenge. En A. Singh, M. Fazel, D. Hsu, S. Lacoste-Julien, F. Berkenkamp, T. Maharaj, K. Wagstaff & J. Zhu (Eds.), Proceedings of the 42nd International Conference on Machine Learning (pp. 82232–82251). Proceedings of Machine Learning Research. https://proceedings.mlr.press/v267/wallach25a.html

Wang, N., Li, S., Wang, C., & Zhao, L. (2024). Current status and emerging trends of generative artificial intelligence technology: A bibliometric analysis. Journal of Internet Technology, 25(3), 477–485.