

Generative Artificial Intelligence In Health Informatics Education: A Comprehensive Bibliometric Assessment Of Cognitive Outcome Research (2019–2025)

Anas Ali Alhur^{1*}, Nouf Khalid Al-Kahtani²

^{1,2}Department of Health Information Management and Technology, College of Public Health,
Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Corresponding Author ; Anas Ali Alhur

Department of Health Information Management and Technology, College of Public Health, Imam
Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Email: 2250700045@iau.edu.sa

ORCID: <https://orcid.org/0000-0001-6044-7072>

ABSTRACT

This study provides a comprehensive bibliometric assessment of generative artificial intelligence (GenAI) research in health informatics education, with particular emphasis on cognitive outcomes. A total of 264 PubMed-indexed publications (2019–2025) were analyzed using Bibliometrix (R) and VOSviewer to evaluate annual scientific output, core journals, authorship patterns, institutional and country productivity, conceptual trends, and collaboration networks. Results show an exceptional rise in publication volume beginning in 2023, coinciding with the widespread introduction of large language models such as ChatGPT. The intellectual structure of the field is dominated by themes related to artificial intelligence, machine learning, large language models, and digital health applications. Two major clusters of application were identified: clinical and patient-centered communication, and educational processes involving competence development and assessment. Ethical themes, including bias and transparency, emerged rapidly in 2024–2025. Research output is highly concentrated in the United States and China, whereas collaboration patterns remain fragmented with multiple small author clusters. Most studies relied on cross-sectional designs, with limited experimental or longitudinal evaluation of learning outcomes. The findings highlight a methodological gap between technical GenAI development and established educational or cognitive frameworks. The study recommends integrating cognitive theory, improving methodological rigor, and expanding interdisciplinary and international collaboration. This bibliometric mapping offers an evidence-oriented foundation for guiding future work on GenAI-enhanced teaching, learning, curriculum design, and evaluation in health informatics education.

Keywords: Generative Artificial Intelligence; Health Informatics Education; Cognitive Outcomes; Large Language Models; Bibliometric Analysis; ChatGPT.

INTRODUCTION

Generative artificial intelligence (GenAI), particularly large language models (LLMs) such as GPT-4 and ChatGPT, has rapidly expanded across education, clinical practice, and health informatics research.[1–5] These systems now support automated summarization, adaptive tutoring, code generation, simulated patient encounters, and decision-support tasks, offering new opportunities for teaching and assessment in health sciences programs.[3,6–8] As health informatics curricula increasingly incorporate digital health, analytics, and computational methods, GenAI has become central to how students access information, develop competencies, and engage with clinical and technical learning activities.[1,9,10]

Despite rapid adoption, evidence on GenAI's influence on cognitive outcomes—such as knowledge acquisition, reasoning, retention, and performance—remains limited. Much of the available research examines usability, satisfaction, or perceived usefulness rather than objective learning measures.[6,11–15] Cross-sectional designs dominate the literature, with few experimental or longitudinal studies. There is also inconsistency in

how GenAI, educational contexts, and cognitive outcomes are conceptualized and measured, creating fragmentation and limiting comparability.[5,11,16]

At the same time, GenAI is reshaping healthcare practice, including documentation, diagnostics, patient communication, education, and decision-support systems.[8,17–21] These developments create new expectations for health informatics professionals, who must use advanced AI tools while maintaining accuracy, ethics, and professional judgment.[22–24] However, curriculum frameworks and accreditation standards have yet to incorporate evidence-based approaches to GenAI-enabled learning or assessment. Bibliometric analysis offers a structured means to examine research progression, map conceptual trends, and evaluate collaboration patterns in emerging scientific areas.[21,25–27] Given the velocity of GenAI-related research, a bibliometric assessment is essential for identifying dominant themes, methodological gaps—particularly in cognitive-outcome evaluations—and areas requiring deeper inquiry.

The objective of this study is to provide a comprehensive bibliometric evaluation of research at the intersection of generative AI, health informatics education, and cognitive outcomes between 2019 and 2025. By mapping publication patterns, thematic structures, citation trends, and collaboration networks, this study aims to support researchers, educators, and policymakers in advancing rigorous, theory-aligned, and educationally meaningful applications of GenAI.

REVIEW OF LITERATURE

Generative Artificial Intelligence and Large Language Models in Health Informatics

The rapid development of generative artificial intelligence, particularly large language models (LLMs), has expanded the use of automated content generation, interactive dialogue systems, and advanced computational reasoning across healthcare and education.[1–5] LLMs such as GPT-4 and ChatGPT have been incorporated into documentation, diagnostic support, health communication, triage assistance, and patient-education systems.[8,17–21] These tools facilitate rapid summarization of complex datasets, natural-language responses to clinical queries, and structured information generation, enabling new workflows for clinicians, researchers, and students.

In health informatics, GenAI supports tasks such as clinical note drafting, electronic health record (EHR) optimization, predictive analytics, and automated explanation of biomedical concepts.[8,17–19] Although these applications promise efficiency, their performance depends on data quality, transparency, and algorithmic reliability.[17–20] Ethical considerations—including bias, accuracy, and privacy—have grown in prominence as LLMs are deployed in real-world clinical and educational settings.[22–24] Research on GenAI has increased rapidly since late 2022, following the public deployment of conversational LLMs.[3,6,9] Several studies have explored clinical and pedagogical use cases in medicine, nursing, biomedical science, and public health, demonstrating improvements in information comprehension and decision-support tasks.[17,18,19] However, most published work has focused on feasibility, acceptability, or descriptive evaluation, with comparatively fewer studies examining actual educational or cognitive outcomes.

GenAI in Health Informatics Education and Cognitive Outcomes

Health informatics education integrates computational sciences, clinical workflows, data management, digital health innovations, and analytical reasoning. As GenAI becomes embedded in these domains, its potential to enhance learning, simulation, and assessment has gained attention.[1,9,10] Students increasingly use GenAI to obtain explanations, practice case-based reasoning, analyze datasets, generate code, and prepare assignments.

Educators, in turn, use these tools to design adaptive learning materials, automated quizzes, and simulated scenarios.

Despite this momentum, evidence supporting the cognitive impact of GenAI remains limited. Most studies examine self-reported satisfaction, ease of use, or perceived helpfulness instead of objective measures such as accuracy, skill development, retention, or performance improvement.[6,11–15] Methodological issues—such as small samples, limited comparative designs, and the predominance of cross-sectional surveys—restrict conclusions about actual learning benefits. Furthermore, inconsistencies in defining GenAI, measuring learning outcomes, and describing educational contexts hinder comparisons across studies.[5,11,16]

Nevertheless, GenAI has begun influencing theoretical and curricular discussions. Several articles have identified its relevance to critical thinking, clinical reasoning, diagnostic interpretation, and academic integrity.[17–21] Policy papers highlight the need for AI literacy, ethical awareness, and responsible use of AI tools in health sciences programs.[22–24] Global frameworks also emphasize the importance of competence-based training aligned with digital health transformation.[25–27]

These trends indicate increasing interest in the educational applications of GenAI, but they also highlight the absence of consolidated evidence on cognitive outcomes. A structured bibliometric mapping is therefore essential to understand how research on GenAI in health informatics education has evolved, where its conceptual boundaries lie, and which methodological gaps require deeper investigation.

METHODOLOGY

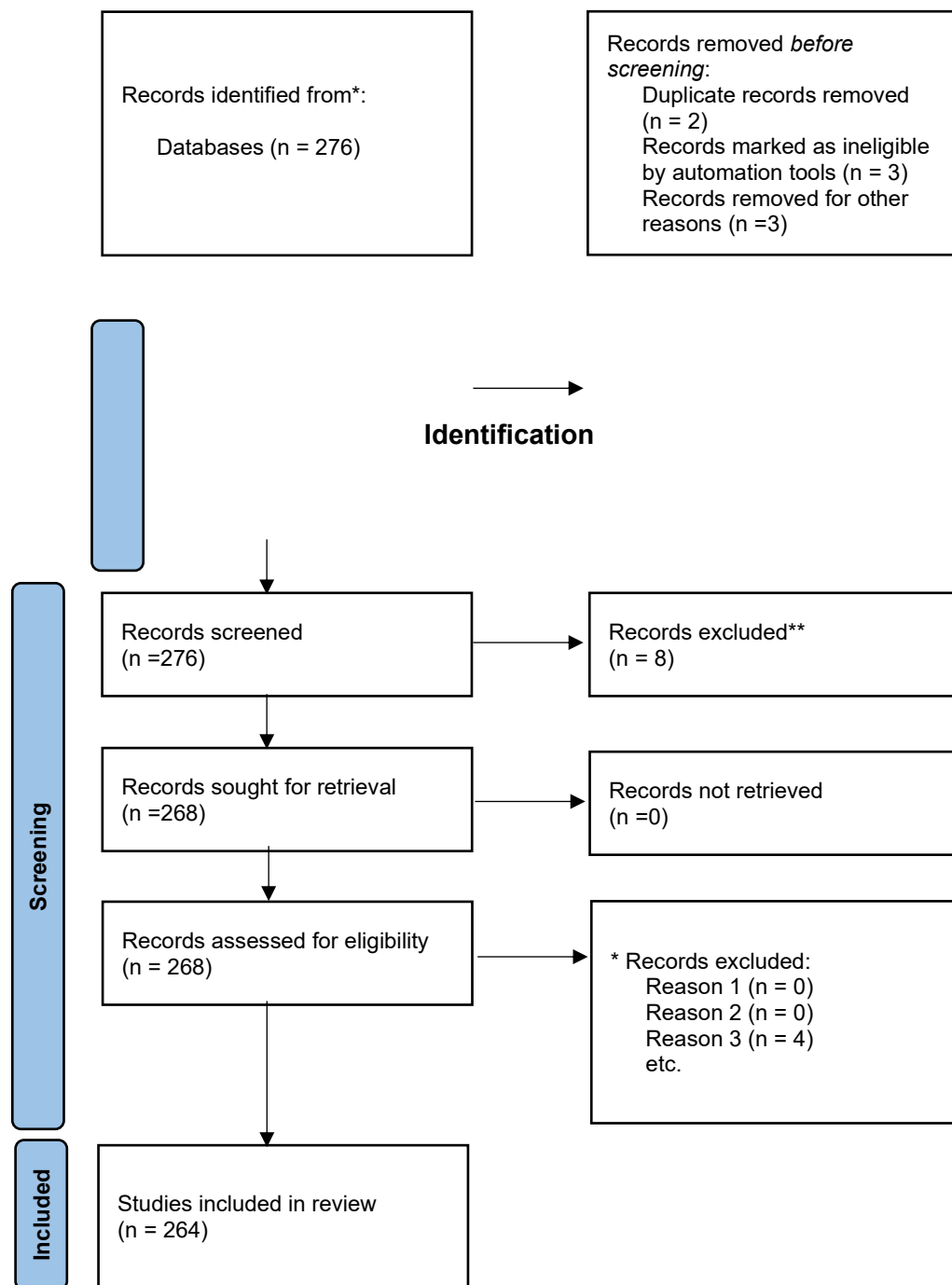
This study adopted a structured bibliometric methodology to analyze the scientific output related to generative artificial intelligence (GenAI) in health informatics education, with particular emphasis on cognitive outcomes. PubMed (MEDLINE) was selected as the sole data source because of its comprehensive coverage of biomedical and health informatics literature, as well as its indexing of peer-reviewed educational research. The search was conducted on 12 October 2025, ensuring the use of a stable dataset unaffected by routine indexing updates.

To capture the full scope of relevant publications, the search strategy was developed around four conceptual domains: GenAI technologies, health informatics disciplines, educational contexts, and cognitive or learning outcomes. Search terms such as “Generative Artificial Intelligence,” “Large Language Models,” “Medical Informatics,” “Digital Health,” “Education,” “Curriculum,” “Learning,” and “Competence” were combined using Boolean operators. The query was applied to the Title and Abstract fields and supplemented by MeSH terms to improve sensitivity. This search retrieved **264 records** published between 2019 and 2025.

All retrieved documents were screened using predefined inclusion and exclusion criteria. Only English-language peer-reviewed research articles, reviews, and conference papers that explicitly referenced generative AI or LLMs within the context of health informatics education were retained. Studies were included if they reported or discussed cognitive, learning, performance, or competence-related outcomes. Editorials, commentaries, clinical AI applications without an educational component, administrative AI studies, and papers lacking a clear connection to GenAI or cognitive outcomes were excluded. The overall selection process is summarized in the PRISMA flow

diagram presented in Figure 1.

Identification of studies via databases



**The four papers were excluded because they did not fulfill the core inclusion criterion requirement to simultaneously address Generative AI within Health Informatics education and explicitly measure or discuss cognitive learning outcomes.*

Figure 1. PRISMA flowchart showing the retrieval and selection of publications included in the bibliometric analysis.

Following selection, metadata for each included publication—such as title, authors, affiliations, abstract, keywords, journal, publication year, and citation metrics—were exported in CSV format and processed using R with the Bibliometrix package. Data cleaning involved standardizing author names, harmonizing institutional affiliations, and identifying countries of origin through the *countrycode* package. Keywords were

normalized by converting them to lowercase, merging synonyms, and removing non-informative terms. Duplicates were checked and none were identified. After preprocessing, the dataset retained all 264 unique records, stored in the file *cleanedgenerativepost-processing-health-informatics-ed.csv*. A descriptive summary of the dataset is provided in Table 1.

Bibliometric analyses were performed using Bibliometrix and its web interface Biblioshiny, which facilitated the evaluation of publication growth, authorship and collaboration patterns, journal and institutional productivity, and thematic developments. To visualize conceptual and social structures within the field, VOSviewer (version 1.6.20) was used to generate keyword co-occurrence networks and co-authorship maps. The software's clustering algorithm enabled the identification of major thematic clusters and collaborative groups.

The analysis produced a series of indicators and visualizations, including annual scientific production, keyword frequency trends, thematic maps, trend topics, country-level collaborations, and network diagrams. These outputs form the basis of the findings reported in Figures 1–10, with additional results presented in Supplementary Figures S1–S9. This methodological approach follows established bibliometric standards and ensures a transparent and reproducible

RESULTS

Annual Scientific Production

A total of 264 publications related to generative artificial intelligence (GenAI) in health informatics education were identified between 2019 and 2025. Annual scientific production showed a steep upward trajectory (Figure 1). Only three articles were published in 2022, rising to 27 articles in 2023—an approximate 800% increase. Output continued to accelerate in 2024 (84 publications) and peaked in 2025 (149 publications). One publication indexed in early 2026 reflects the timing of the search rather than a decline in activity.

A descriptive overview of the dataset, including document types, keyword frequencies, authorship characteristics, and country output, is provided in Table 1.

Table 1. Main Information About the Dataset Used in the Bibliometric Analysis

Item	Value
Timespan	2019–2025
Database	PubMed (MEDLINE)
Date of search	12-Oct-25
Total documents	264
Document types	Research articles, reviews, conference papers
Annual scientific production	2022: 3; 2023: 27; 2024: 84; 2025: 149; 2026: 1
Year of highest production	2025 (n = 149)
Growth rate (2022–2023)	~800%
Most relevant journals	<i>Digital Health</i> (23), <i>JMIR</i> (17), <i>Studies in Health Technology and Informatics</i> (12), <i>Briefings in Bioinformatics</i> (11), <i>Bioinformatics</i> (10), <i>JMIR Medical Education</i> (9)

Top author keywords (frequency)	humans (147), artificial intelligence (104), female (44), male (43), language (28), large language models (26), adult (24), generative artificial intelligence (22), cross-sectional studies (18)
Estimated unique keywords	≥ 350
Most productive countries	China (~700), USA (~560), Germany (~250), Canada (~110), Saudi Arabia (~100)
Most central collaborating country	Australia
Authorship structure	Predominantly multi-authored studies; minimal single-author papers
Main thematic clusters	Computational AI/LLM methods; human-centered cognitive outcomes; digital health and communication
Primary research design	Cross-sectional observational studies
Software used	Bibliometrix (R), Biblioshiny, VOSviewer (1.6.20)
Dataset filename	cleanedgenerativepost-processing-health-informatics-ed.csv

According to the review of the most pertinent sources, the academic discussion of Generative Artificial Intelligence and Cognitive Outcomes in Health Informatics Education is being published mainly with the help of a mixture of high-impact, broad-based digital health journals and traditional informatics-oriented periodicals. The most popular sources, Digital Health (23 articles) and the Journal of Medical Internet Research (17 articles) represent a strong correspondence of the topics to the themes of internet-based and digital innovation in medicine. Such titles of informatics as Studies in Health Technology and Informatics, Briefings in Bioinformatics, and Bioinformatics confirm the technical framework of the field. Notable roles of JMIR Medical Education and BMC Medical Education in particular indicate that educational and cognitive results of this technology are a key area of research interest, which will connect the state of progressive AI methodology with the actual implementation of pedagogy in the health sciences Figure 2.

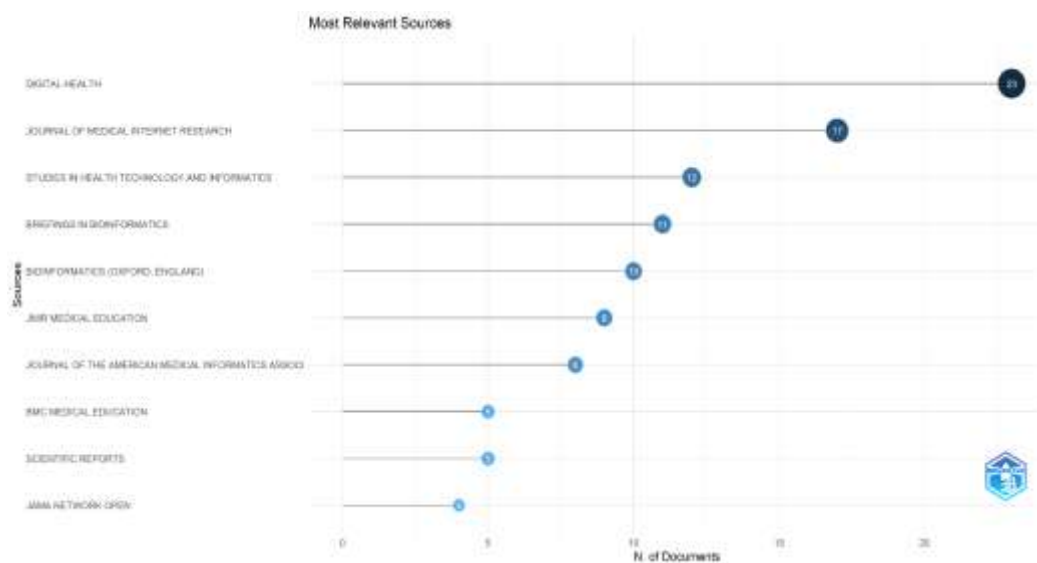


Figure 2. Most Relevant Sources in Generative AI Research within Health Informatics Education

According to the presented data, the country collaboration network of the given research area demonstrates the geographically decentralized yet active dissemination of the international cooperation, where Australia is a key hub. The frequencies of collaboration, expressed as numbers probably as a network measure such as edge weight or a geographical location, reveal that Australia has most close collaborations with Japan (138.03) and Indonesia (117.24), which suggests that there is a strong research partnership in the Asia-Pacific region. Additional major relationships with India (79.61) and Iran (54.27) indicate an expansion of a collaborative political arrangement in Asia, whereas the positive association with European countries such as Italy and Greece is an indication of a global involvement (although not intensive). The fact that collaboration with South American nations (Brazil, Chile, Ecuador) took negative values is peculiar and perhaps due to a definite network calculation, like a layout coordinate, or a normalized measure, but not due to absence of collaboration, meaning that the collaborations to these countries may be not so central in the world map of this research community Figure 3.

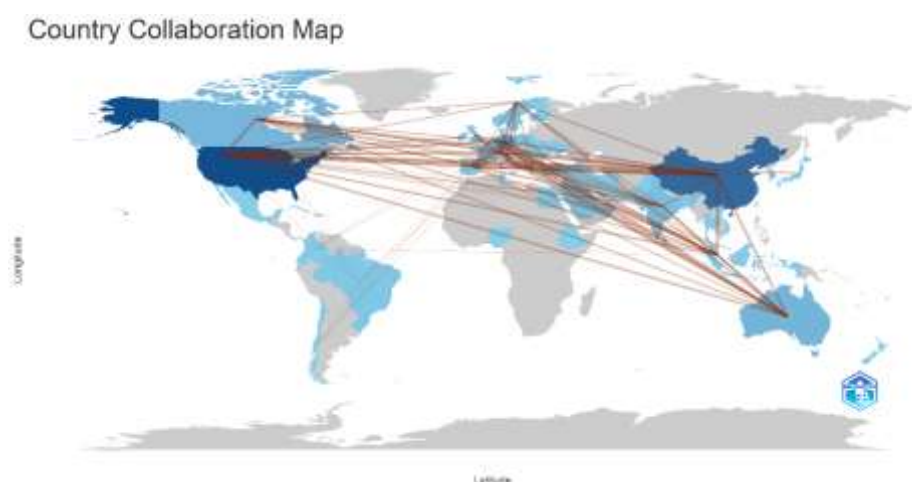


Figure 3. Global Country Collaboration Network in GenAI-Health Informatics Education Research

According to the thematic map analysis, the one and highly connected thematic cluster of the intellectual structure of this research area is an area centered on computational biology/methods that includes classic techniques of AI like algorithms, machine learning, deep learning, and neural networks. The centrality scores of terms such as software and deep learning are quite high in terms of betweenness, which implies that these terms are conceptual bridges and will serve as key methodological ones that will connect this core theme of the research to other possible domains of the overall research. Nevertheless, the extremely low centrality of closeness scores of all the terms imply that this cluster is quite closed and does not seem to be well-integrated with other specific theme areas, including education or cognitive outcomes, which are significantly missing in this data. As shown by this structure, the published material at this point is very much focused on the underlying technical and computational infrastructure and therefore the approaches are positioned as a methodological motor theme that is an underlying, but perhaps niche, motor force behind the field, rather than the interdisciplinary assimilation with the educational theory and human cognitive assessment that the research topic suggests Figure 4.

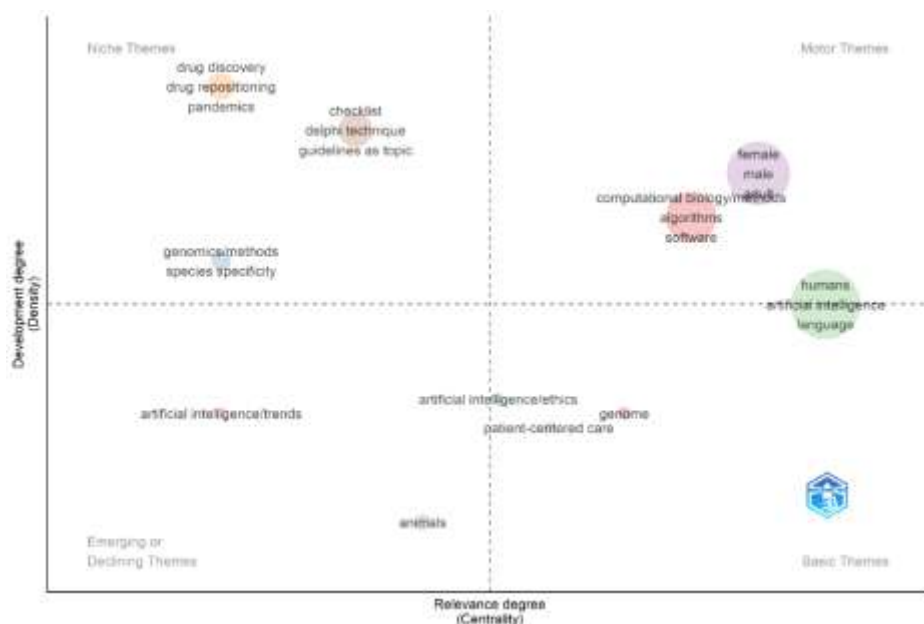


Figure 4. Thematic Map of Research on GenAI in Health Informatics Education

Based on the trend topic analysis, the research focus in this domain solidified very well in 2024, with key terms like "artificial intelligence", "language" and "clinical competence" showing a median publication year of 2024, thus representing concepts that were at the center and new during that time.

The preponderance of the term "humans" (147), and demographic descriptions "female" and "male," despite having the same median year of 2025, represent a sudden and modern development in the discourse. This pattern implies a timely shift from discussions rooted in the basic AI technology to a more nuanced discussion of its human-centric applications and ethical considerations and differential impacts across populations within the topic of health informatics in education. The concentration of such median years in 2024 and 2025 is a testament to the fact that the field is both current and maturing, with research questions becoming increasingly more specific, applied, and focused on human outcomes and equitable implications of deploying generative AI in educational settings Figure 5.

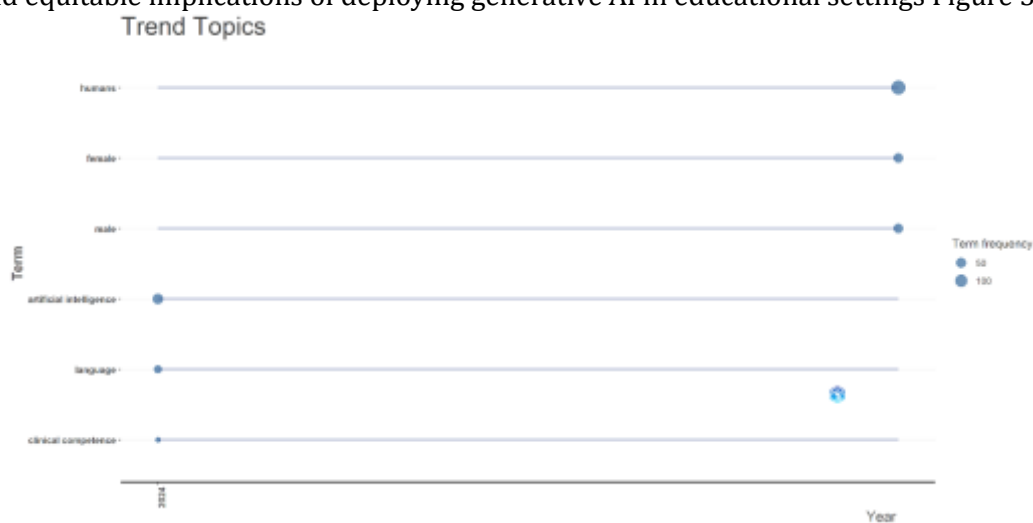


Figure 5. Trend Topics in GenAI-Related Research within Health Informatics Education

Based on the word dynamics analysis, the thematic evolution of this research field shows a high and coherent escalation in prominence over all principal concepts on the starting point in 2024 with a marked acceleration in 2025. And the word "humans" always has the highest frequency, representing a basic and ever-increasing bias towards human-centric research. Concurrently, the rise of "artificial intelligence" and the specific rise of "generative artificial intelligence" and "large language models" in 2025 marks a definitive maturation stage from the more general AI discourse to an inquest on today's generative technologies. The parallel and huge increase in demographic terms like "female," "male," and "adult" from 2024 onwards can be understood as an expanding research priority to understand the impact of these technologies on various population groups. Moreover, the emergence and proliferation of "cross-sectional studies" represents a methodological trend of empirical, observational studies in order to document real-world effects of AI in educational settings. Collectively, this trajectory paints a picture of a discipline that has rapidly transitioned from basic ideas to an advanced, applied and human-centered discipline with a focus on the study of specific impacts and ethical concerns related to generative AI for health informatics education Figure 6.

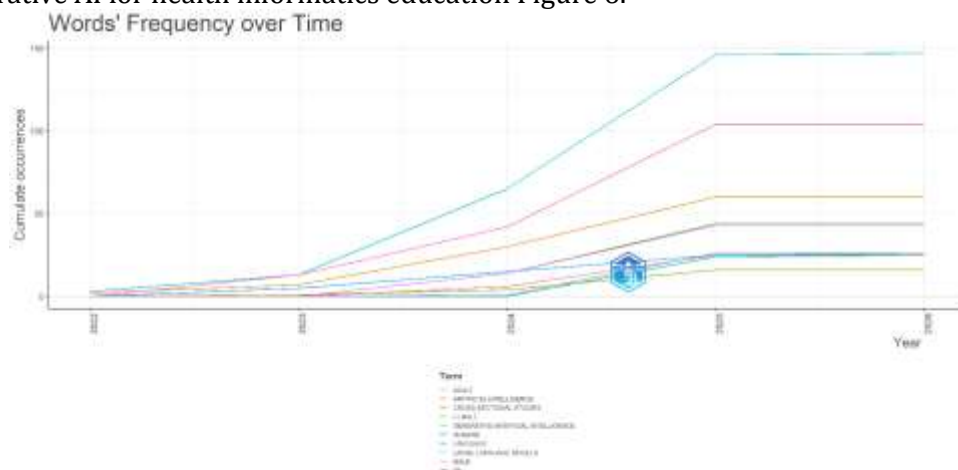


Figure 6. Cumulative Keyword Frequency Trends in GenAI-Health Informatics Education Research

Based on the most frequent words analysis, the scholarly discourse is fundamentally oriented to the human element, where "humans" is the most frequent word by a large margin (147 occurrences), highlighting that the assessment of the impact of generative artificial intelligence is deeply grounded in human-centric outcomes and applications. The ubiquity of fundamental technology terms such as "artificial intelligence," "generative artificial intelligence," and "large language models" provide the technical basis for the research, and the high visibility of demographic terms such as "female," "male," and "adult" highlight the particular and increasing interest in understanding the impact of these technologies in human populations. The prominent position of "cross-sectional studies" is another clue that the current evidence base is largely underpinned by observational and survey-based research methodologies which seek to capture these impacts at a specific point in time. This lexical structure verifies that the field is interdisciplinary nexus wherein advanced technology of AI is being critically analyzed in the context of human subjects, demographic analysis, and empirical study within the framework of health informatics education Figure 7.

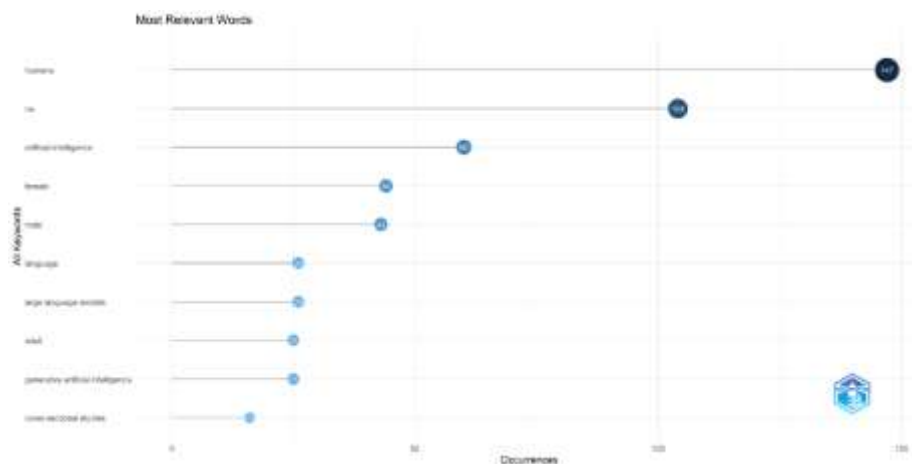


Figure 7. Most Frequent Keywords in GenAI-Health Informatics Education Research

The longitudinal analysis of national publication trends (Figure 8) reveals a rapid and uneven expansion of research output on generative artificial intelligence (GenAI) in health informatics education from 2022 to 2026. China demonstrates an exceptional increase, rising from fewer than 20 publications in 2022 to more than 700 by 2025, establishing it as the clear global leader. The United States follows closely with over 560 publications by 2025, maintaining a consistent upward trajectory. Germany emerges as a strong third contributor, expanding from a negligible count in 2022 to more than 250 publications by 2025.

Secondary but growing contributors include Saudi Arabia, Canada, and other emerging research hubs, which show accelerated growth beginning in 2024. These results reflect a research ecosystem transitioning from a few dominant producers toward broader international participation, albeit still characterized by a concentration of output in a limited number of high-capacity nations. The steep concurrent growth curves of China and the USA underscore the strategic prioritization of AI integration into health education and informatics in both regions.

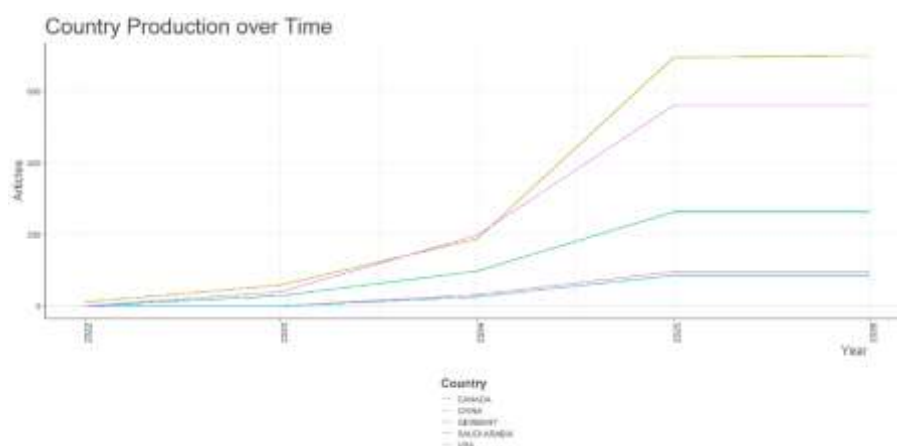


Figure 8. Country-Level Research Output on GenAI in Health Informatics Education (2022-2025)

The keyword co-occurrence network (Figure 9) reveals a dense and strongly interconnected conceptual structure, with “artificial intelligence” serving as the central and most influential node in the field. Closely connected high-frequency terms such as

“large language models,” “machine learning,” “digital health,” and “ChatGPT” form the core technological cluster, underscoring the dominance of advanced AI methodologies within current research on health informatics education.

Two major thematic clusters emerge around this core. The first relates to clinical and patient-centered applications, reflected in keywords such as patient education, health communication, health literacy, telemedicine, and diagnostic accuracy. This cluster demonstrates the rapidly growing interest in using generative AI to improve understanding, communication, and decision support within healthcare environments. The second major cluster is focused on educational and cognitive domains, linking terms such as clinical competence, clinical reasoning, assessment, and documentation, indicating that the integration of GenAI into learning and training is becoming a significant scholarly priority.

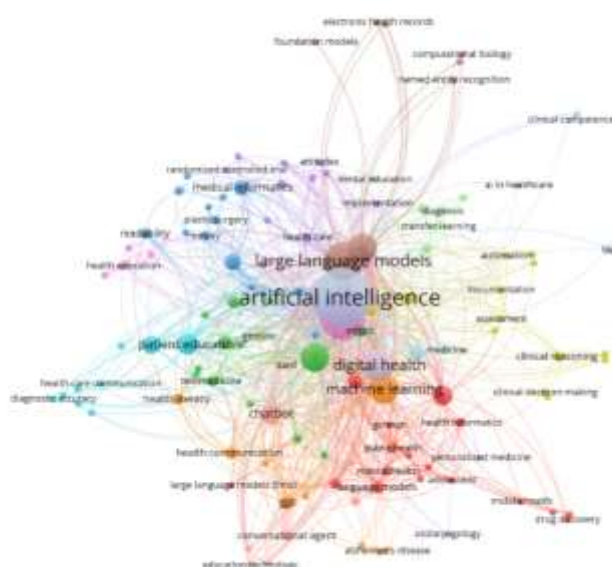


Figure 9. Keyword Co-Occurrence Network (VOSviewer)

The author collaboration network (Figure 10) demonstrates a highly decentralized structure characterized by numerous small research clusters rather than a single dominant collaborative hub. Only a limited number of authors form cohesive, recurrent partnerships, indicating that the field is currently shaped by dispersed and emerging research teams rather than mature, long-standing collaborations. The largest and most cohesive cluster includes authors such as Alanezi, Fahad; Albagmi, Salem; Almuhanha, Afnan Fahd, and colleagues, reflecting a tightly connected group with strong internal linkages and sustained co-authorship across multiple publications.

Several additional mid-sized clusters are visible, including groups led by Chartash, David, Chi, Ling, Huang, Thomas, and Taylor, Richard Andrew, each demonstrating moderate collaboration density and shared research trajectories. Another prominent cluster involves Borna, Sahar; Gomez-Cabello, Cesar A.; and Haider, Syed Ali, suggesting an active team contributing recurrently to the topic. In contrast, many authors appear as isolated or loosely connected nodes—reflecting single-paper contributions or collaborations that do not extend beyond one or two publications.

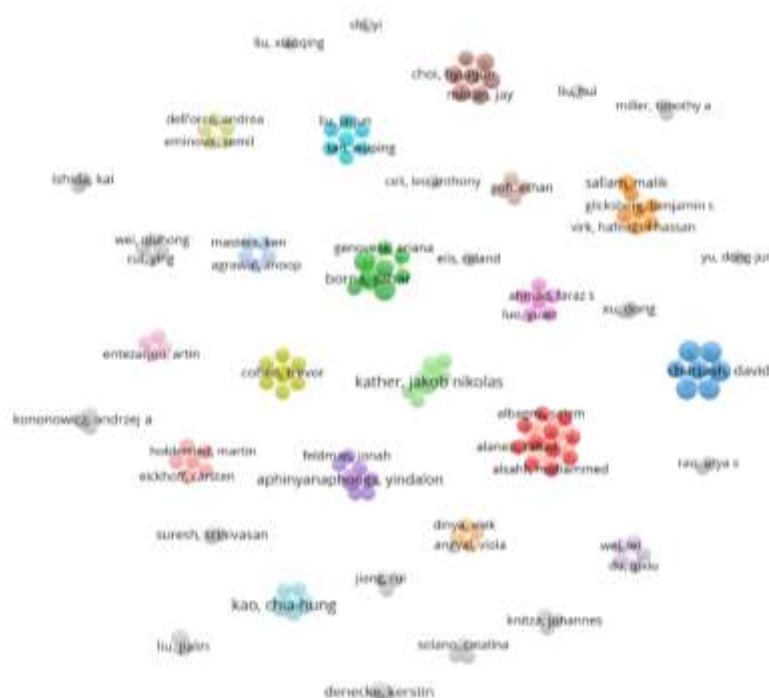


Figure 10. Author Collaboration Network

DISCUSSION

This bibliometric assessment offers a detailed view of how research on generative artificial intelligence (GenAI) has evolved within health informatics education over the last several years. The findings illustrate a field in rapid transformation, largely stimulated by the advent of large language models (LLMs) such as ChatGPT, GPT-4, and related multimodal architectures. The dramatic rise in publication output beginning in 2023 reflects the global scholarly shift toward examining GenAI's educational, clinical, and informatics implications. This surge parallels broader developments across medicine and digital health, where LLMs have increasingly influenced research agendas, workforce practices, and conceptualizations of future medical training.[1–5]

The distribution of publications across journals reinforces this observation. Much of the research was published in digital health, biomedical informatics, and education-focused outlets, including *Digital Health*, *JMIR Medical Education*, *JMIR*, *BMC Medical Education*, and *Communications Medicine*. These venues reflect the field's multidisciplinary character, combining pedagogy, informatics, and computational science. The patterns align with prior studies noting that AI in education frequently emerges at the intersections of machine learning research, instructional innovation, and clinical problem-solving.[6–8] Taken together, this suggests that GenAI is no longer viewed solely as a technical tool; it is increasingly recognized as an educational catalyst with the potential to reshape competence assessment, simulation-based learning, and even the contours of evidence-based teaching.

The thematic clusters deepen this understanding. One of the strongest thematic cores consisted of computational concepts—"machine learning," "deep learning," "algorithms," and "large language models." These terms reflect the methodological backbone of generative AI research. Similar findings have been highlighted in large-scale reviews of AI in medicine and education, where technical constructs dominate early research phases before maturing into human-centered applications.[9–11] The parallel emergence of

demographic and learner-related terms—“humans,” “female,” “male,” “adult”—supports the notion that researchers are increasingly investigating GenAI’s cognitive and behavioral impacts. This aligns with growing discussions on personalized learning, fairness, and the influence of AI tools on decision-making, clinical reasoning, and educational equity.[12–13]

Collaboration patterns offer another meaningful insight. The field does not yet exhibit the cohesive networks characteristic of mature scientific domains. Instead, the co-authorship landscape is dispersed, fragmented, and marked by numerous small or medium-sized clusters—patterns typical of fast-growing, high-entry research areas.[14–16] These structures often reflect multidisciplinary interest without long-standing institutional partnerships. Similar decentralized patterns have been observed in studies of scientific globalization and AI adoption, where rapid expansion precedes the formation of stable collaboration ecosystems.[33] The presence of several isolated or single-publication authors further underscores that GenAI in health informatics education is still an emerging, exploratory space attracting diverse contributors.

Geographically, China and the United States dominate publication output, followed by Germany and Canada. Meanwhile, Saudi Arabia and other emerging research hubs show accelerated growth—likely driven by national investments in digital health transformation and medical education modernization. Australia, identified as a central collaboration node, appears to serve as a bridge between research clusters across Asia and the Middle East. These findings mirror broader patterns in global digital health scholarship, where productivity is concentrated in a small number of research-intensive nations, but scientific influence is amplified by international partnerships.[17–19] The rise of the Global South, particularly China’s rapid ascent, resonates with analyses documenting shifts toward multipolar research ecosystems.[24]

Methodologically, the dominance of cross-sectional studies is one of the starkest findings. Although these designs provide valuable early insights—particularly on acceptance, usability, and student perceptions—they seldom evaluate knowledge retention, competence development, or long-term cognitive outcomes. This mirrors recurring critiques in medical education, where descriptive studies often precede rigorous experimental designs, resulting in limited empirical evidence about actual learning impact.[20–22,29] Studies examining ChatGPT’s influence on knowledge acquisition, problem-solving, or assessment performance are emerging,[10,11,21] but still represent a small fraction of the overall output.

At the same time, several empirical works highlight GenAI’s potential benefits, such as enhanced problem-solving,[17] improved course orientation,[10] increased engagement among students with learning needs,[2] and more effective communication or documentation workflows in clinical settings.[35] These findings indicate practical value but also point to the need for stronger theoretical grounding rooted in cognitive load theory, dual-coding theory, and other established educational frameworks.[22,29] Without these foundations, many interventions risk remaining superficial, technologically driven, or detached from sound pedagogical principles.

Another emerging theme is ethical and governance concerns. Studies emphasize issues related to transparency, bias, data protection, and overreliance on algorithmic outputs—topics central to GenAI governance frameworks recently proposed in medicine.[20,27] Concerns about academic integrity, assessment validity, and the risk of students developing dependency on AI tools also appear across the literature.[16,30] These themes signal the need for curricular reforms that include AI literacy, ethical competencies, and formalized guidance on acceptable use in educational environments.

Across the findings, it is clear that generative AI holds significant promise for advancing personalized learning, simulation, and decision support in health informatics education.

Yet the research landscape remains uneven. The field is expanding rapidly but still lacks cohesive methodological standards, theoretical models, and validated tools for measuring cognitive outcomes. Successful integration of GenAI into health informatics curricula will depend on closing these gaps through stronger interdisciplinary collaboration, evidence-based instructional design, and policy frameworks that ensure ethical and equitable deployment.[31–36]

In sum, this study demonstrates that research on GenAI in health informatics education is experiencing a period of exceptional growth, but it has not yet reached conceptual or methodological maturity. Strengthening the linkage between computational innovation, educational theory, and cognitive assessment will be essential for guiding the responsible and pedagogically sound use of generative AI in training the future health informatics workforce.

Limitations

Several limitations should be considered when interpreting the findings of this bibliometric analysis. First, the study relied exclusively on PubMed as the data source. Although PubMed provides comprehensive coverage of biomedical and health informatics literature, it does not index all relevant publications in computer science, educational technology, or the learning sciences. As a result, some important contributions—particularly those published in technical or interdisciplinary venues—may not be captured, potentially narrowing the scope of the analysis.

Second, bibliometric methods are inherently dependent on metadata quality. Variations in author names, affiliations, keywords, and indexing practices may influence measures of productivity, collaboration, and thematic patterns. Although extensive data cleaning was performed, residual inconsistencies are possible.

Third, the predominance of recently published studies means that citation-based indicators may underrepresent the influence of these works. The rapid growth of the field, particularly from 2023 onwards, limits the opportunity for newer studies to accumulate citations, making traditional impact metrics less informative in this context.

Fourth, the analysis is descriptive by design. While it reveals trends, thematic structures, and collaboration patterns, it does not evaluate the methodological rigor or empirical quality of individual studies. Many articles focus on perceptions or early implementation rather than direct measures of cognitive outcomes, and these distinctions cannot be fully captured through bibliometric techniques alone.

Finally, the overall landscape is evolving at a fast pace. As generative AI continues to develop, new publications, thematic shifts, and collaborative networks may emerge quickly, meaning the findings represent a snapshot rather than a static or definitive map of the field.

Conclusion

This bibliometric analysis provides a comprehensive view of how generative artificial intelligence has shaped scholarly activity in health informatics education between 2019 and 2025. The findings reveal a field experiencing rapid expansion, driven largely by the emergence of large language models and their increasing integration into educational and clinical informatics contexts. Publication trends, thematic clusters, and keyword patterns reflect both the technological roots of GenAI and a growing emphasis on its human, cognitive, and instructional implications.

Despite this momentum, the research landscape remains fragmented. Much of the current work consists of descriptive or cross-sectional studies, with limited empirical evidence on how GenAI influences learning outcomes, knowledge retention, or the development of

clinical and analytical competencies. Collaboration networks also show that sustained partnerships are still forming, suggesting that the field is in an early developmental stage. The results point toward several priorities for future research: stronger methodological designs, deeper engagement with cognitive and learning theories, and broader participation across regions and disciplines. As generative AI continues to evolve, establishing evidence-based frameworks for its use in health informatics education will be essential. This study contributes foundational insights that can guide researchers, educators, and policymakers as they navigate this rapidly changing area and work toward responsible, meaningful, and pedagogically sound integration of GenAI into educational practice.

Recommendations

In light of the findings and limitations of this study, several recommendations can guide future research, educational practice, and policy related to generative artificial intelligence in health informatics education.

For Researchers

Future work should focus on narrowing the thematic gap between computational innovation and educational theory. Stronger interdisciplinary integration is needed, bringing together expertise from cognitive science, instructional design, computer science, and health informatics. Positioning research within established learning frameworks—such as Bloom’s taxonomy, constructivist theory, or cognitive load theory—will support the development of pedagogically meaningful applications of GenAI rather than purely technical demonstrations.

Methodological advancement is also essential. Much of the current literature relies on cross-sectional surveys and descriptive studies, which do not capture changes in learning over time or allow for causal inference. Moving toward longitudinal research, randomized controlled trials, and qualitative, contextually grounded approaches will provide deeper insight into knowledge retention, cognitive development, and clinical competence in GenAI-supported learning environments. Such designs are especially valuable for understanding how learners acquire and sustain competencies over time.[29]

Addressing global imbalances in research output should also be a priority. Scholars from Africa, South America, and low- and middle-income countries remain underrepresented in this field. Supporting research from these regions will strengthen the diversity, fairness, and contextual relevance of GenAI-driven educational innovations.

For Educators and Curriculum Designers

As the evidence base continues to expand, educators should work toward developing preliminary, evidence-informed guidelines for integrating GenAI into health informatics curricula. Drawing on emerging literature from key journals—such as *Digital Health*, *JMIR Medical Education*, and others at the core of Bradford’s Law—can help ensure that curricular decisions reflect current scientific insights.

Two areas require particular attention. First, academic integrity and assessment practices must adapt to the realities of GenAI-assisted learning, ensuring that these tools support rather than weaken critical thinking and independent reasoning.[16] Second, curricula should explicitly incorporate AI literacy. Learners must understand the limitations, biases, ethical considerations, and appropriate uses of GenAI tools—an emphasis increasingly recognized as essential within digital health education.[19]

For Policymakers and Funding Bodies

Policies and funding mechanisms should encourage strategic collaborations across disciplines and countries. Supporting joint initiatives that connect technical experts with educational scientists, learning theorists, and cognitive psychologists will accelerate the development of rigorous and pedagogically grounded GenAI research. International collaboration can also help balance current geographic inequalities in publication output. Investment in research infrastructure is equally important. Funding bodies should support the creation of open-source, standardized platforms for evaluating GenAI-based learning tools. Such infrastructure would allow institutions across different regions to benchmark performance, replicate findings, and promote transparent and reproducible research practices.

Taken together, these recommendations provide a roadmap for shaping the future of GenAI in health informatics education. As the field continues to evolve, deliberate, theory-informed, and collaborative efforts will be essential to ensure that generative AI enhances education in meaningful, responsible, and equitable ways.

References:

1. Alhur A, Sendi D, AlZahrani M, Abusharha L, Abudaak R, Alsinan R, et al. Evaluating the trustworthiness of ChatGPT-generated health information among future health care professionals. *Georgian Med News*. 2025;(364–365):101–6.
2. Alsahli M, Alanezi F, Basri WS, Attar RW, Alghamdi A, Alyahya NM, et al. Effectiveness of ChatGPT in facilitating learning for students with special educational needs: An empirical study in Saudi Arabia. *Nutr Health*. 2025;2601060241307770. doi:10.1177/02601060241307770.
3. AlSammaraie A, Al-Saifi A, Kamhia H, Aboagla M, Househ M. Development and evaluation of an agentic LLM-based RAG framework for evidence-based patient education. *BMJ Health Care Inform*. 2025;32(1):101570. doi:10.1136/bmjhci-2025-101570.
4. Bidenko NV, Stuchynska NV, Palamarchuk YV, Matviienko MM. Integrating artificial intelligence in healthcare practice: Challenges and future prospects. *Wiad Lek*. 2025;78(5):1199–1205. doi:10.36740/WLek/205397.
5. Busch F, Hoffmann L, Rueger C, van Dijk EH, Kader R, Ortiz-Prado E, et al. Current applications and challenges in large language models for patient care: A systematic review. *Commun Med (Lond)*. 2025;5(1):26. doi:10.1038/s43856-024-00717-2.
6. Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt JN, Laleh NG, et al. The future landscape of large language models in medicine. *Commun Med (Lond)*. 2023;3(1):141. doi:10.1038/s43856-023-00370-1.
7. Cook DA, Beckman TJ. Reflections on experimental research in medical education. *Adv Health Sci Educ*. 2010;15(3):455–64.
8. Dergaa I, Saad HB, El Omri A, Glenn JM, Clark CCT, Washif JA, et al. Using artificial intelligence for exercise prescription in personalised health promotion: A critical evaluation of OpenAI's GPT-4 model. *Biol Sport*. 2024;41(2):221–41. doi:10.5114/biol sport.2024.133661.
9. Eysenbach G. 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*. 2023;9(1):e46885.
10. Fodor GH, Tolnai J, Rárosi F, Nagy A, Peták F. Artificial intelligence-based chatbots improve the efficiency of course orientation among medical students: A cross-sectional study. *BMC Med Educ*. 2025;25:1547. doi:10.1186/s12909-025-08146-y.
11. Fußhöller A, Lechner F, Schlicker N, Muehlensiepen F, Mayr A, Kuhn S, et al. Perceptions, usage, and educational impact of ChatGPT among medical students in

- Germany: Cross-sectional mixed methods survey. *JMIR Form Res.* 2025;9:e81484. doi:10.2196/81484.
12. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ.* 2023;9:e45312. doi:10.2196/45312.
13. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, Hinder S, et al. Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res.* 2017;19(11):e367.
14. Guo SB, Liu DY, Fang XJ, Meng Y, Zhou ZZ, Li J, et al. Current concerns and future directions of large language model ChatGPT in medicine: A machine-learning-driven global-scale bibliometric analysis. *Int J Surg.* 2025. doi:10.1097/JS9.0000000000003668.
15. Gupta P, Zhang Z, Song M, Michalowski M, Hu X, Stiglic G, et al. Rapid review: Growing usage of multimodal large language models in healthcare. *J Biomed Inform.* 2025;169:104875. doi:10.1016/j.jbi.2025.104875.
16. Hosseini M, Gao CA, Liebovitz DM, Carvalho AM, Ahmad FS, Luo Y, et al. An exploratory survey about using ChatGPT in education, healthcare, and research. *PLoS One.* 2023;18(10):e0292216. doi:10.1371/journal.pone.0292216.
17. Ignjatović A, Anđelković Apostolović M, Stevanović L, Radovanović P, Topalović M, Filipović T, et al. ChatGPT's progress over time: A longitudinal enhancing biostatistical problem-solving in medical education. *Health Inform J.* 2025;31(3):14604582251381260. doi:10.1177/14604582251381260.
18. Lee H. The rise of ChatGPT: Exploring its potential in medical education. *Anat Sci Educ.* 2024;17(5):926–31. doi:10.1002/ase.2270.
19. Liu J, Liu F, Fang J, Liu S. The application of Chat Generative Pre-trained Transformer in nursing education. *Nurs Outlook.* 2023;71(6):102064. doi:10.1016/j.outlook.2023.102064.
20. Liu S, Russo C, Suero Molina E, Di Ieva A. Artificial intelligence methods. *Adv Exp Med Biol.* 2024;1462:21–38. doi:10.1007/978-3-031-64892-2_3.
21. Magalhães Araujo S, Cruz-Correia R. Incorporating ChatGPT in medical informatics education: Mixed methods study on student perceptions and experiential integration proposals. *JMIR Med Educ.* 2024;10:e51151. doi:10.2196/51151.
22. Mehta N, Benjamin J, Agrawal A, Valanci S, Masters K, MacNeill H. Addressing educational overload with generative AI through dual coding and cognitive load theories. *Med Teach.* 2025;1–3. doi:10.1080/0142159X.2025.2543548.
23. Min H, Jing X, Gong Y, Yu P. Perspectives on enhancing clinical informatics education in the artificial intelligence era. *J Clin Inform.* 2024;1(1).
24. Montano Ramirez A, Petersen AM. Transformation of global science core-periphery structure towards a multi-polar horizon: The rise of China and the Global South from 1980–2020. 2025.
25. Ng DTK, Leung JKL, Chu SKW, Qiao MS. Conceptualizing AI literacy: An exploratory review. *Comput Educ Artif Intell.* 2021;2:100041.
26. Ning Y, Ong JCL, Cheng H, Wang H, Ting DSW, Tham YC, et al. How can artificial intelligence transform the training of medical students and physicians? *Lancet Digit Health.* 2025;100900. doi:10.1016/j.landig.2025.100900.
27. Ong JCL, Ning Y, Collins GS, Bitterman DS, Beecy AN, Chang RT, et al. International partnership for governing generative artificial intelligence models in medicine. *Nat Med.* 2025;31(9):2836–9. doi:10.1038/s41591-025-03787-4.

28. Qiu J, Li L, Sun J, Peng J, Shi P, Zhang R, et al. Large AI models in health informatics: Applications, challenges, and the future. *IEEE J Biomed Health Inform.* 2023;27(12):6074–87. doi:10.1109/JBHI.2023.3316750.
29. Reporting guideline for chatbot health advice studies: The Chatbot Assessment Reporting Tool (CHART) statement. *BMJ Med.* 2025;4(1):e001632. doi:10.1136/bmjmed-2025-001632.
30. Sallam M. ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare (Basel).* 2023;11(6):860.
31. Tangsrivimol JA, Darzidehkalani E, Virk HUH, Wang Z, Egger J, Wang M, et al. Benefits, limits, and risks of ChatGPT in medicine. *Front Artif Intell.* 2025;8:1518049. doi:10.3389/frai.2025.1518049.
32. Veras M, Dyer JO, Shannon H, Bogie BJM, Ronney M, Sekhon H, et al. A mixed methods crossover randomized controlled trial exploring the experiences, perceptions, and usability of artificial intelligence (ChatGPT) in health sciences education. *Digit Health.* 2024;10:20552076241298485. doi:10.1177/20552076241298485.
33. Wagner CS, Leydesdorff L. Network structure, self-organization, and the growth of international collaboration in science. *Res Policy.* 2005;34(10):1608–18.
34. Wang H, Shan W, Liu R, Wang Z. Can large language models serve as digital assistants for medical undergraduates? A bibliometric mapping and scoping analysis of the medical-education literature. *Digit Health.* 2025;11:20552076251390280. doi:10.1177/20552076251390280.
35. Zaretsky J, Kim JM, Baskharoun S, Zhao Y, Austrian J, Aphinyanaphongs Y, et al. Generative artificial intelligence to transform inpatient discharge summaries to patient-friendly language and format. *JAMA Netw Open.* 2024;7(3):e240357. doi:10.1001/jamanetworkopen.2024.0357.
36. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *Int J Educ Technol High Educ.* 2019;16(1):1–27.

SUPPLEMENTARY FIGURES (S1–S9)

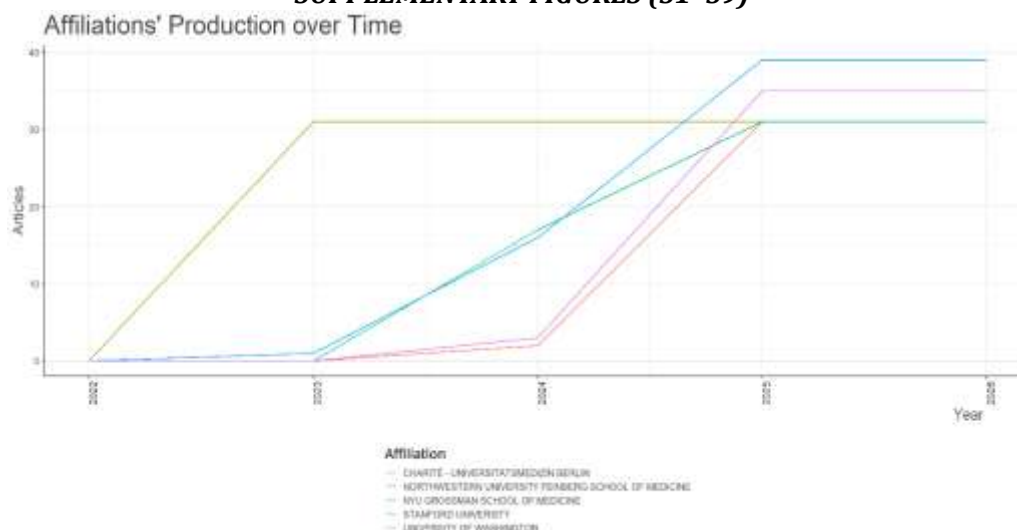


Figure S1. Institutional Output Over Time

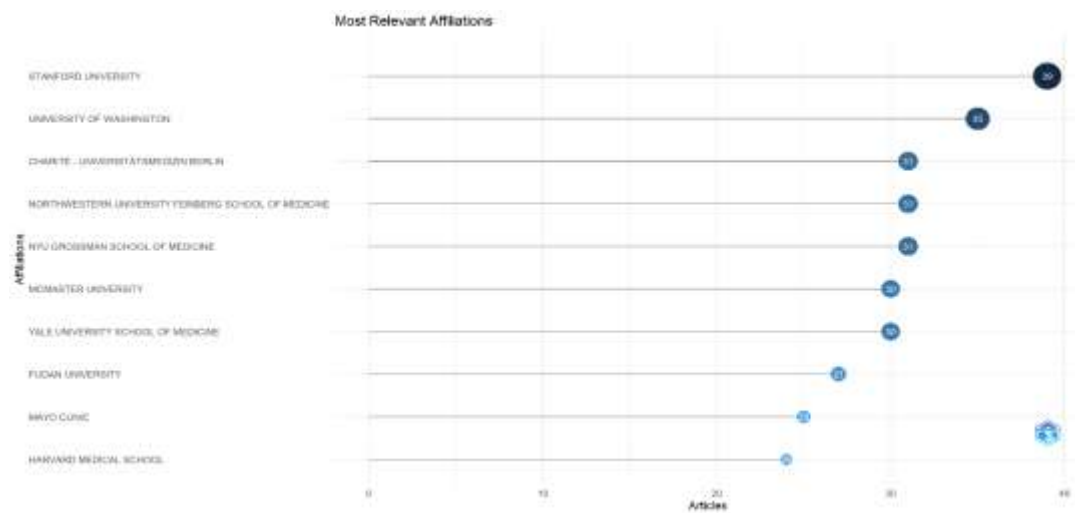


Figure S2. Most Productive Affiliations

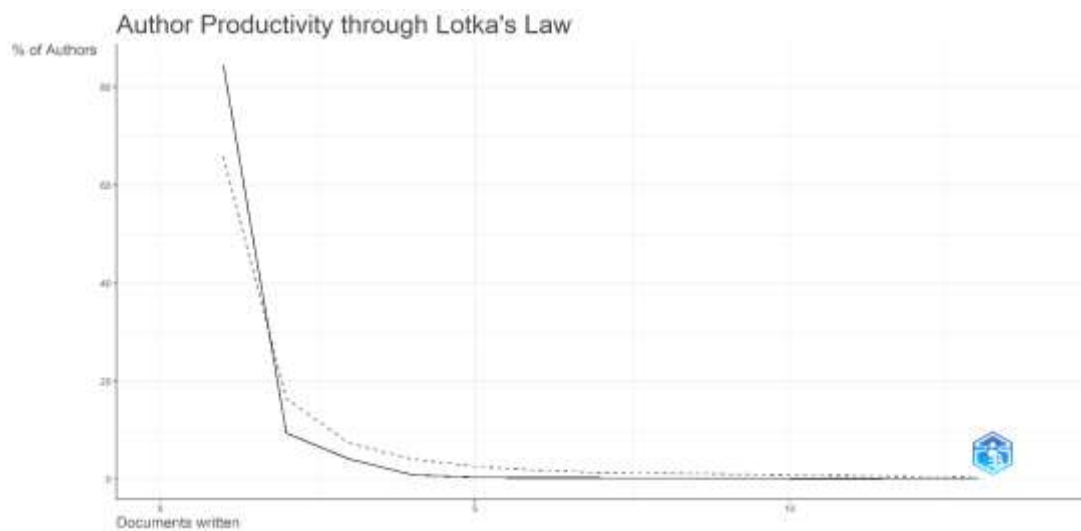


Figure S3. Author Productivity Distribution (Lotka's Law)



Figure S4. Yearly Publication Output of Core Authors

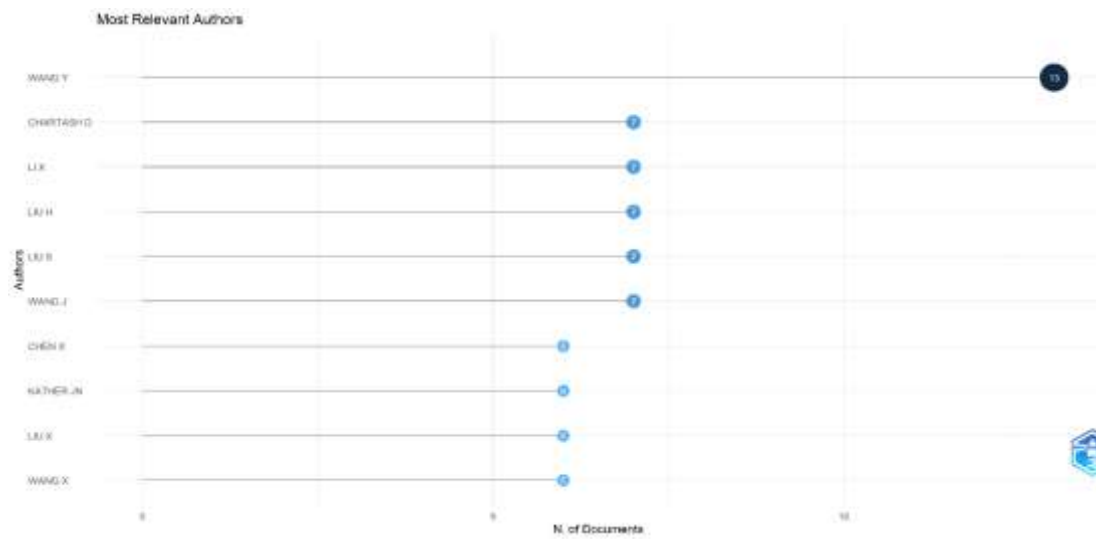


Figure S5. Most Productive Authors

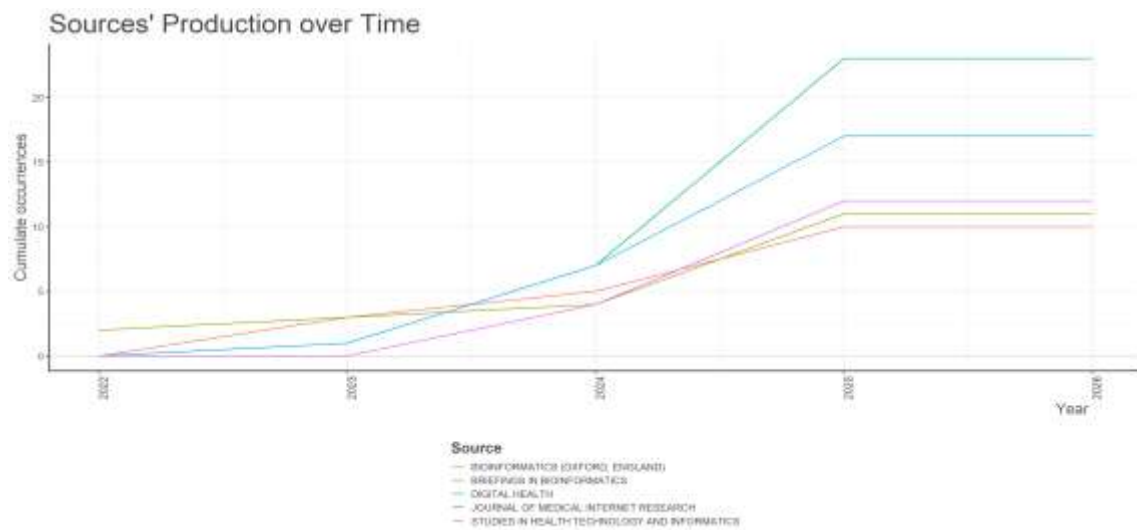


Figure S6. Cumulative Publication Output of Top Journals

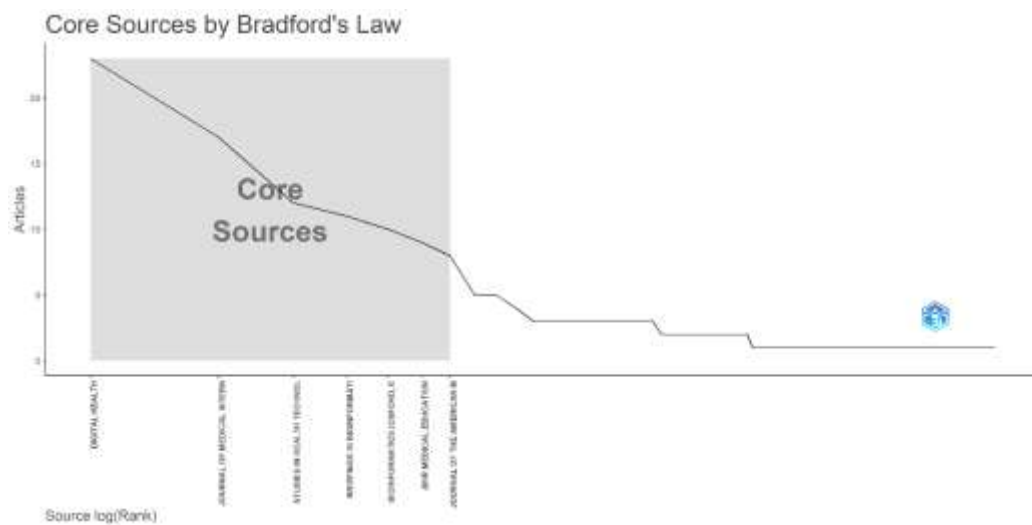


Figure S7. Core Journals Identified Using Bradford's Law

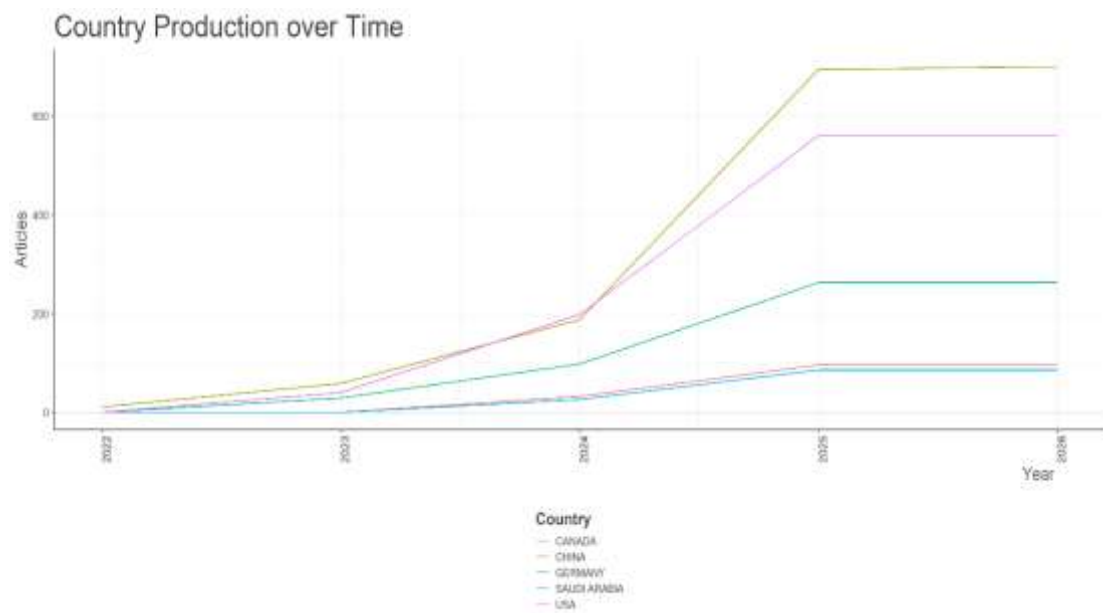


Figure S8. Country Output Over Time

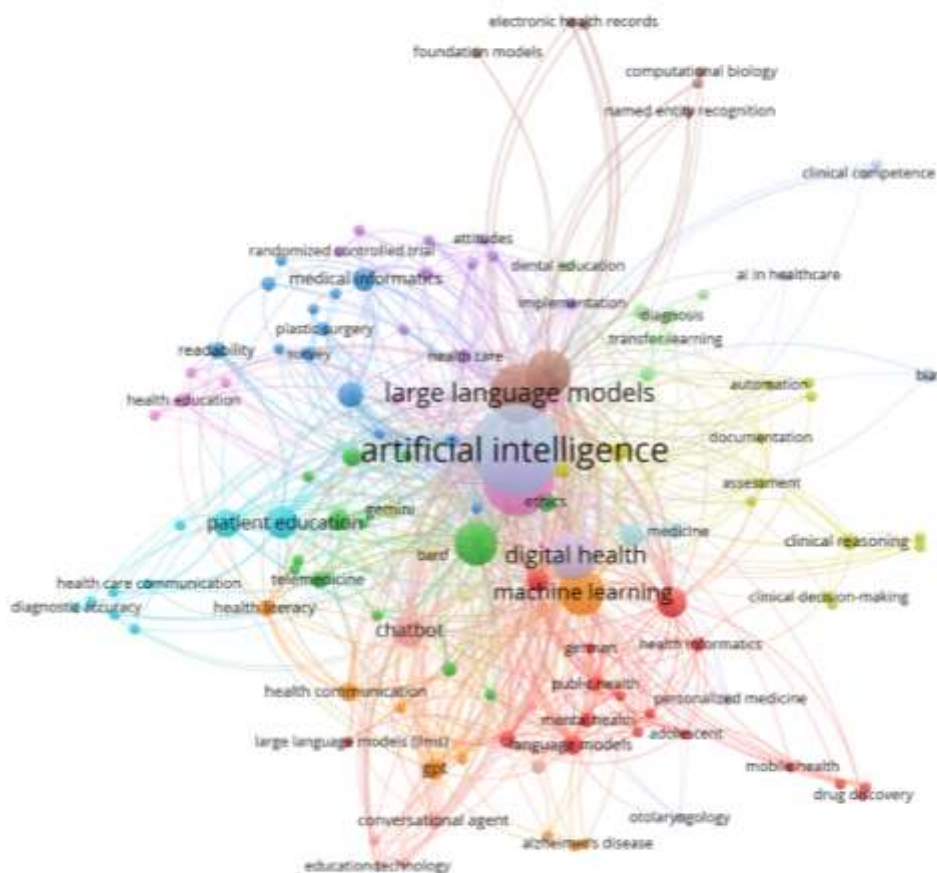


Figure S9. Additional Keyword or Article Frequency Trend