J Data Sci. Info. Citation Studies., 2026; 5(3):431-446.
https://www.jcitation.org

Research Article

Mapping the Intellectual Landscape of Generative AI in Academic Libraries: A Scientometric Analysis and Science Mapping of Scopus Literature (2021-2026)

Vishnudas A V[1,*], Arathy G S[2]

[1]Library and Knowledge Resource Center (LKRC), Presidency University, Bengaluru, Karnataka, INDIA.
[2]School Library, The KPR Academy, Arasur Campus, Coimbatore, Tamil Nadu, INDIA.

ABSTRACT

This study maps the intellectual landscape of Generative Artificial Intelligence (GenAI) research in academic libraries through a scientometric analysis of Scopus-indexed literature published between 2021 and 2026. The study aims to examine publication growth, influential sources, authorship patterns, institutional and geographical contributions, citation impact, and keyword trends. Bibliographic data were retrieved from Scopus on 26 September 2026, yielding 2,753 records for analysis. Quantitative bibliometric indicators and science-mapping techniques were applied using R version 4.5.1 and the Biblioshiny tools. The findings reveal a substantial increase in publication output, particularly during 2024-2026, which accounted for 94.55% of the retrieved records. China and the United States emerged as leading contributors in corresponding-author publications, while the collaboration analysis revealed differences in international co-authorship patterns across countries. The keyword analysis highlighted the prominence of language-model terminology, artificial intelligence, digital libraries, ChatGPT, and human-related concepts. The results demonstrate the rapid expansion and multidisciplinary character of the retrieved literature, while also indicating the importance of interpreting publication and citation patterns in light of database coverage, keyword variations, and the incomplete coverage of 2026. This study provides a structured overview of the publication landscape and offers a basis for identifying emerging topics and research opportunities at the intersection of generative AI and academic libraries.

Keywords: Academic libraries, Bibliometric analysis, Generative artificial intelligence, Scientometrics, Science mapping.

Correspondence:
Mr. Vishnudas A V
Assistant Librarian, Library and Knowledge Resource Center (LKRC), Presidency University, Bengaluru-560119, Karnataka, INDIA.
Email: vishnudasav2000@gmail.com

Received: 28-10-2026;
Revised: 03-11-2026;
Accepted: 19-11-2026.

INTRODUCTION

Generative Artificial Intelligence (GenAI), particularly Large Language Models (LLMs), has emerged as an important area of technological development, with growing implications for higher education, academic research, and information services. Academic libraries have a particular interest in these developments because they support information discovery, research assistance, scholarly communication, information literacy, and access to scholarly resources. GenAI offers potential applications in several of these areas, including conversational reference services, metadata generation, search assistance, and research guidance. At the same time, its use raises important questions concerning the accuracy of generated information, algorithmic bias, transparency, data privacy, copyright, academic integrity, and the continuing role of professional judgement in library services.

The development of GenAI has created new opportunities for examining changes in scholarly communication and research activity. Understanding this emerging field requires more than measuring the number of publications produced each year. It is also necessary to examine the journals in which research is published, the contributions of authors and institutions, patterns of international collaboration, relationships among research themes, and the intellectual connections that shape the field. Scientometric analysis provides a systematic approach to examining these dimensions through publication and citation indicators, while science-mapping techniques help in understanding relationships among authors, keywords, institutions, and countries (Aria and Cuccurullo, 2017; Donthu *et al.*, 2021).

Previous investigations of GenAI have established a useful foundation for understanding publication growth, research themes, influential contributors, and patterns of scholarly communication. However, studies covering earlier periods of

DOI: 10.5530/jcitation.20260016

Copyright Information :
Copyright Author (s) 2026 Distributed under Creative Commons CC-BY 4.0

Publishing Partner : Manuscript Technomedia. [www.mstechnomedia.com]

digital-library development cannot, by themselves, explain the more recent emergence of GenAI-related research in academic library settings. Furthermore, research on artificial intelligence in libraries encompasses a broader range of technologies than generative AI alone. A focused examination of GenAI-related literature is therefore necessary to distinguish the publication patterns, thematic relationships, and collaboration structures represented in this developing area of scholarship. Against this background, the present study examines Scopus-indexed literature retrieved using search terms related to generative AI and academic libraries for the period 2021-2026. It combines publication and citation analysis with science-mapping techniques to investigate annual research output, source dispersion, author productivity, institutional and country contributions, international collaboration, keyword co-occurrence, thematic development, and co-authorship networks. The study also examines selected bibliometric patterns using Bradford's Law of Scattering and Lotka's Law, where the dataset and analytical results permit appropriate assessment.

By bringing these dimensions together, the study seeks to provide a structured account of the development and intellectual organisation of GenAI-related scholarship in academic library research. The findings are intended to contribute to a clearer understanding of the publication channels, research communities, geographical contributions, and conceptual themes associated with this emerging field. They may also help identify areas requiring further scholarly attention. Publication and network indicators describe patterns in the indexed literature; they do not independently establish the extent to which GenAI has been implemented successfully in academic libraries.

LITERATURE REVIEW

Foundational scholarship conceptualized digital libraries as evolving sociotechnical systems supporting information access and scholarly communication (Borgman, 1999), while subsequent studies have examined the development of digital- library research through bibliometric and scientometric approaches (Ahmad e*t al*., 2018; Borgohain *et al*., 2021; Mahadevagouda and Pavithrabai, 2022; Mondal and Surendran, 2024). Investigated publication productivity, citation patterns, influential contributors, and research networks, while (Wattanasiri *et al.,* 2025) analysed research on digital-library services Lestari and Dewanti (2025) explored digital-library research related to information literacy, identifying themes such as digital literacy, information services, academic libraries, and education. With the growing development of artificial intelligence, research has also examined its potential applications in academic libraries Kulkanjanapiban *et al.* (2025) investigated AI-driven innovation and services in academic libraries. These developments are relevant to understanding the possible applications of Generative Artificial Intelligence (GenAI), including information discovery, reference services, research support, and information-literacy assistance. However, AI research encompasses technologies beyond GenAI, and its broader findings cannot automatically be interpreted as evidence of GenAI adoption. Issues such as inaccurate information, algorithmic bias, privacy, copyright, and academic integrity also remain important considerations.

The integration of emerging technologies into academic libraries is closely associated with institutional research-support responsibilities, professional competencies, infrastructure, and organisational preparedness. Perkins and Slowik (2013) demonstrated the institutional value of research in academic libraries, Padhan and Naidu (2022) examined research-support services in academic libraries, and Adeniran and Oyovwevotu (2019) discussed research- support activities in an institutional context. Although these studies do not necessarily provide direct evidence of GenAI implementation, they offer a broader perspective on the role of academic libraries in supporting research and scholarly communication. Understanding these developments requires attention to staff training, digital infrastructure, institutional policies, structural service readiness, and responsible technology use (Ncube and Ngulube, 2025). Methodologically, bibliometric indicators and science-mapping techniques provide complementary means of examining research productivity, citation patterns, influential sources, institutional and geographical contributions, keyword relationships, and collaboration networks (Aria and Cuccurullo, 2017; Donthu *et al.*, 2021). Nevertheless, the interpretation of such findings depends on data quality, keyword normalisation, counting procedures, and analytical settings, and bibliometric evidence alone cannot establish institutional readiness or the practical effectiveness of library services.

Although existing studies have examined digital libraries, information literacy, and AI-related developments in academic libraries, the literature reviewed here does not establish a comprehensive account of the intellectual structure and collaboration patterns of GenAI-related research in academic libraries during 2021-2026. This gap provides the rationale for the present study, which combines descriptive bibliometric indicators with selected science-mapping techniques to examine publication growth, influential sources, author and institutional contributions, geographical distribution, keyword relationships, and research collaboration in the literature retrieved through the study's Scopus search strategy. The analysis aims to provide a structured overview of the development and thematic organisation of this research area without treating publication patterns as direct evidence of technology adoption, institutional readiness, or practical impact.

Research Gap

Previous studies have examined digital-library research, AI-driven library services, publication trends, and scholarly

contributions; however, differences in their time coverage and technological focus, particularly the limited attention to GenAI following its recent expansion, highlight the need for a focused scientometric analysis of GenAI-related research in academic libraries. Publication counts and keyword frequencies alone cannot adequately explain the intellectual and social structure of this emerging field, making an integrated examination of publication and citation trends, author productivity, leading sources, institutional and geographical contributions, international collaboration, co-authorship networks, keyword co-occurrence, source dispersion, and thematic development necessary to understand its scholarly relationships and evolution. To address this gap, the present study analyses Scopus-indexed literature on GenAI and academic libraries published during 2021-2026, combining descriptive bibliometric indicators with selected science-mapping techniques to provide a structured overview of the field's publication patterns, collaboration networks, conceptual relationships, and intellectual landscape.

Research Objectives

The primary objective of this study is to map the development, intellectual structure, and collaboration patterns of Scopus-indexed literature on generative artificial intelligence in academic libraries during 2021-2026. The study is guided by the following specific objectives:

RO1. To examine publication and citation trends in GenAI-related academic library research during 2021-2026, including annual research output and citation patterns.

RO2. To identify the major publication sources and examine source dispersion, including the applicability of Bradford's Law of Scattering to the distribution of publications across journals.

RO3. To analyse author productivity and identify leading contributors, including an assessment of the applicability of Lotka's Law to the observed distribution of author productivity.

RO4. To map institutional and geographical contributions and examine international collaboration patterns using country-level publication output, single country publication, multiple country publication and relevant citation indicators.

RO5. To investigate the conceptual structure and thematic development of the field through keyword co-occurrence analysis, thematic relationships, and the temporal trends of selected research topics.

RO6. To examine the social structure of research collaboration through co-authorship network analysis, including the identification of connected research groups and the use of relevant network centrality measures where supported by the analytical procedures.

METHODOLOGY

Research Design and Analytical Framework

The study adopts a quantitative scientometric and science-mapping approach to examine the development and intellectual structure of research on Generative Artificial Intelligence (GenAI) in academic libraries. The analytical framework combines descriptive bibliometric indicators with relational analyses to investigate publication output, citation patterns, source productivity, author contributions, geographical distribution, keyword relationships, and research collaboration. Scientometric analysis provides a systematic means of examining the characteristics and development of scholarly literature, while science-mapping techniques help identify relationships among research themes and contributors (Aria and Cuccurullo, 2017; Donthu *et al.*, 2021). The present study build as an expansion on the authors' preliminary bibliographic exploration undertaken in connection with a book chapter entitled "Evolution of Publication Trends in Generative AI and Academic Library Research: A Scopus-Based Analysis (2021-2026)," which is a basic analysis study of scopus result analysis and is currently under consideration for publication. The present article extends this preliminary work through a broader scientometric analysis and science-mapping approach.

Data Source and Search Strategy

Bibliographic records were retrieved from Scopus on 26 September 2026. Scopus was used as the sole bibliographic data source because it provides multidisciplinary coverage and relevant bibliographic and citation metadata, including author information, institutional affiliations, abstracts, keywords, and cited references. The database was considered suitable for the study's scientometric and science-mapping objectives. The search aimed to retrieve literature related to generative artificial intelligence and libraries published or indexed between 2021 and 2026. Using a single database also provided a consistent source for data extraction and analysis. However, the findings are limited to records indexed in Scopus and do not represent all relevant publications available in other bibliographic databases. The search expression used in Scopus was: TITLE-ABS-KEY ("generative AI" OR "chatgpt" OR "large language models" AND "library" OR "libraries" OR "Academic Library" OR "Academic Libraries") AND PUBYEAR > 2020 AND PUBYEAR < 2027.

Inclusion and Exclusion Criteria

The analysis included English-language journal articles, conference papers, review articles, and book chapters published or indexed between 2021 and 2026. The reason to limit to this year range is because, when comparing the years, an increase in publication is observed from 2021 onwards. Editorial materials, conference front matter, errata, letters, and non-scholarly notes

were excluded. These criteria defined the document types, language, and publication period considered in the analysis.

Data Export, Cleaning, and Software

Following retrieval, screening, and cleaning, a total of 2,753 bibliographic records were taken for analysis. The records were exported from Scopus in BibTeX format. The exported data included citation counts, author affiliations, abstracts, author keywords, index keywords, and cited references. The analysis was conducted using Bibliometrix and its graphical interface, Biblioshiny, in R version 4.5.1.

Data Analysis

The analysis examined annual publication output and citation patterns, publication-source productivity and dispersion, author productivity, institutional and geographical contributions, international collaboration, keyword co-occurrence, topic trends, and co-authorship networks. Bradford's Law of Scattering and Lotka's Law were considered in relation to source dispersion and author productivity, respectively. Single-Country Publications (SCP) and Multiple-Country Publications (MCP) were used to examine international collaboration patterns. Keyword co-occurrence and topic-trend analyses were used to explore relationships among research concepts and changes in the prominence of selected topics. Co-authorship network analysis was used to examine the structure of collaboration among contributing researchers. As the records were retrieved on 26 September 2026, publication data for 2026 represent an incomplete calendar year. Consequently, the findings for this year are interpreted and are not treated as directly equivalent to the results for complete calendar years.

RESULTS

Annual Scientific Production and Growth Trajectory (2021-2026)

The annual distribution of publications on Generative Artificial Intelligence (GenAI) in academic libraries shows a substantial increase in research output between 2021 and 2026. As shown in Table 1, the dataset comprises 2,753 records, with relatively few publications in the initial years. In 2021, only 2 publications were recorded (0.07%), followed by 4 publications in 2022 (0.15%).

Publication output increased markedly in 2023, reaching 144 records (5.23%). This growth coincided with the wider adoption of generative AI tools, particularly ChatGPT and other large language models, which brought increased attention to their potential applications and implications for academic libraries. The upward trend continued in subsequent years. The dataset recorded 563 publications in 2024 (20.45%) and 993 in 2025 (36.07%), while 2026 accounted for the largest number, with 1,047 publications (38.03%). Together, the years 2024-2026 contributed approximately 94.55% of the total records, indicating that research in this area is concentrated in the most recent years covered by the study. This pattern suggests growing scholarly interest in the applications of GenAI in academic libraries, including its potential role in information services, research support, digital resource discovery, and institutional practices.

Citation Impact and Annual Citation Dynamics (2021-2026)

The distribution of citations across publication years indicates a clear relationship between publication age and citation impact. As shown in Table 2, publications from the earlier years generally received higher average citations per document than those published more recently. Publications from 2021 recorded a mean of 97.50 citations per document, equivalent to an annualised rate of 16.25 citations per year. The mean citation count for 2022 was 48.00 per document, with an annualised rate of 9.60 citations.

Publications from 2023 also demonstrated substantial citation impact, averaging 44.85 citations per document and 11.21 citations per year. These figures indicate that publications from the earlier stages of the study period have attracted considerable scholarly attention and may have contributed to the initial development of research discussions on generative AI in academic libraries. In contrast, the average citation count declined for publications from the more recent years. The mean was 16.00 citations per document for 2024, equivalent to 5.33 citations per year, followed by 5.17 citations per document for 2025, or 2.58 citations per year. Publications from 2026 recorded a mean of 0.70 citations

Table 1: Annual Scientific Production.

Publication Year	Number of Articles (NP)	Annual Percentage (%)	Cumulative Articles	Cumulative Percentage (%)
2021	2	0.07%	2	0.07%
2022	4	0.15%	6	0.22%
2023	144	5.23%	150	5.45%
2024	563	20.45%	713	25.90%
2025	993	36.07%	1,706	61.97%
2026	1,047	38.03%	2,753	100.00%
Total	2,753	100.00%	-	-

Table 2: **Annual citation indicators.**

Year	Mean TC per Article	Number of Articles (N)	Mean TC per Year	Citable Years
2021	97.50	2	16.25	6
2022	48.00	4	9.60	5
2023	44.85	144	11.21	4
2024	16.00	563	5.33	3
2025	5.17	993	2.58	2
2026	0.70	1,047	0.70	1

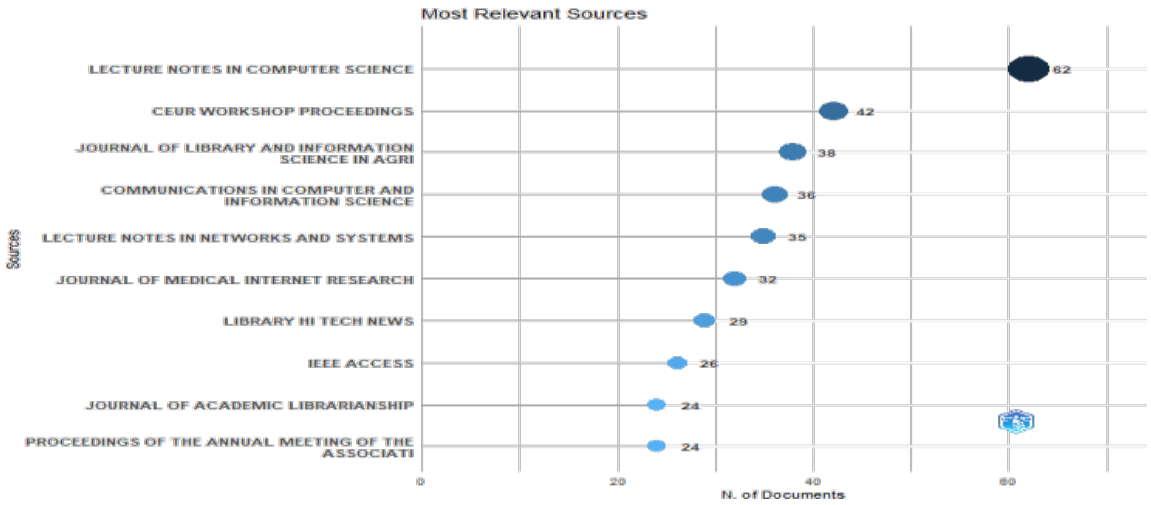

Figure 1: Most relevant publication sources.

per document. The findings highlight the distinction between publication growth and citation impact. Although the volume of research on generative AI in academic libraries has increased considerably in recent years, the citation influence of these newer publications cannot yet be assessed on the same basis as that of earlier studies. Further analysis over a longer period would help establish whether the recent growth in publication output translates into sustained scholarly influence.

Most relevant source

The distribution of the most productive publication sources indicates that research on Generative Artificial Intelligence (GenAI) in academic libraries is disseminated across a range of disciplines, including computer science, Library and Information Science (LIS), and specialised applied fields. As shown in Figure 1, *Lecture Notes in Computer Science* is the most productive source, with 62 publications, followed by *CEUR Workshop Proceedings* with 42, *Communications in Computer and Information Science* with 36, and *Lecture Notes in Networks and Systems* with 35. The prominence of conference proceedings and computer science publication series highlights the contribution of technology-oriented research outlets to this emerging field. These sources provide platforms for presenting technical developments, emerging applications, and research on the use of generative AI across different information environments.

Library and information science journals also feature among the leading sources. *Library Hi Tech News* contributed 29 publications, while *The Journal of Academic Librarianship* accounted for 24. Their presence reflects the relevance of professional practice, library services, information literacy, and the organisational implications of AI adoption within academic libraries. Other prominent sources include *IEEE Access*, with 26 publications, and the *Journal of Medical Internet Research*, with 32. The inclusion of these multidisciplinary and specialised journals suggests that research related to generative AI and academic libraries intersects with broader discussions of digital technologies, information services, and applied AI. Overall, the leading publication sources demonstrate the interdisciplinary nature of this research area. Contributions from computing and technology-oriented outlets coexist with publications in library-focused and applied research journals. This distribution suggests that the development of generative AI in academic libraries is being examined from both technical and service-oriented perspectives. The findings also indicate the importance of considering publication sources across disciplinary boundaries when assessing the development and dissemination of research in this field.

Source dispersion and Bradford's Law of Scattering

Bradford's Law of Scattering was used to examine the distribution and concentration of publications across journals, conference

proceedings, and other scholarly sources. Bradford's Law analysis in Figure 2 explains how publications within a specific research field are distributed across sources, with a relatively small core group producing a substantial share of the literature and the remaining publications scattered across a larger number of sources. The distribution is commonly represented through three zones, with the number of sources in successive zones following an approximate geometric progression of 1: k: k^2, where *k* is the Bradford multiplier.

The dataset comprised 2,753 publications distributed across 1,379 unique publication sources. To examine source dispersion, the sources were ranked according to their publication output and classified into three zones based on the Bradford distribution. This classification helps identify the core sources contributing to research on generative artificial intelligence in academic libraries and distinguish them from sources that publish on the topic less frequently. The resulting distribution provides a basis for understanding the concentration of scholarly output and the extent to which research in this field is dispersed across different publication outlets.

The Bradford's Law analysis shows that the 2,753 articles in the dataset are distributed across 1,379 sources, indicating a substantial dispersion of publications across journals and other scholarly sources. As shown in Table 3, Zone 1 comprises 65 core sources (4.71% of all sources), accounting for approximately 918 articles. Zone 2 contains 397 moderately productive sources (28.80%), contributing approximately 1,835 cumulative articles, while Zone 3 consists of 917 peripheral sources (66.49%), accounting for the remaining articles up to the total of 2,753. This distribution indicates that a relatively small proportion of sources forms the core of the literature, whereas a much larger number of sources contributes to the broader research landscape of generative AI and academic libraries.

The mean Bradford multiplier (k) of 4.21 suggests an uneven distribution of publications across the three zones, with the number of sources increasing considerably from the core to the peripheral zone. This pattern is consistent with the general principle of Bradford's Law, which describes the concentration of relevant literature in a limited number of core sources alongside its dispersion across a wider range of publications. This result highlights the multidisciplinary and dispersed nature of research on generative AI and academic libraries, suggesting that researchers should consider sources beyond the core group when conducting comprehensive literature searches.

Source production over time

The analysis of source production over the period 2021-2026 reveals changes in the publication activity of the leading sources contributing to research on Generative Artificial Intelligence (GenAI) in academic libraries. As shown in Figure 3, publication activity among the selected sources was limited during the early years of the study period. No publications were recorded for these sources in 2021 or 2022. In 2023, the *Journal of Library and Information Science* contributed 9 publications, while *CEUR*

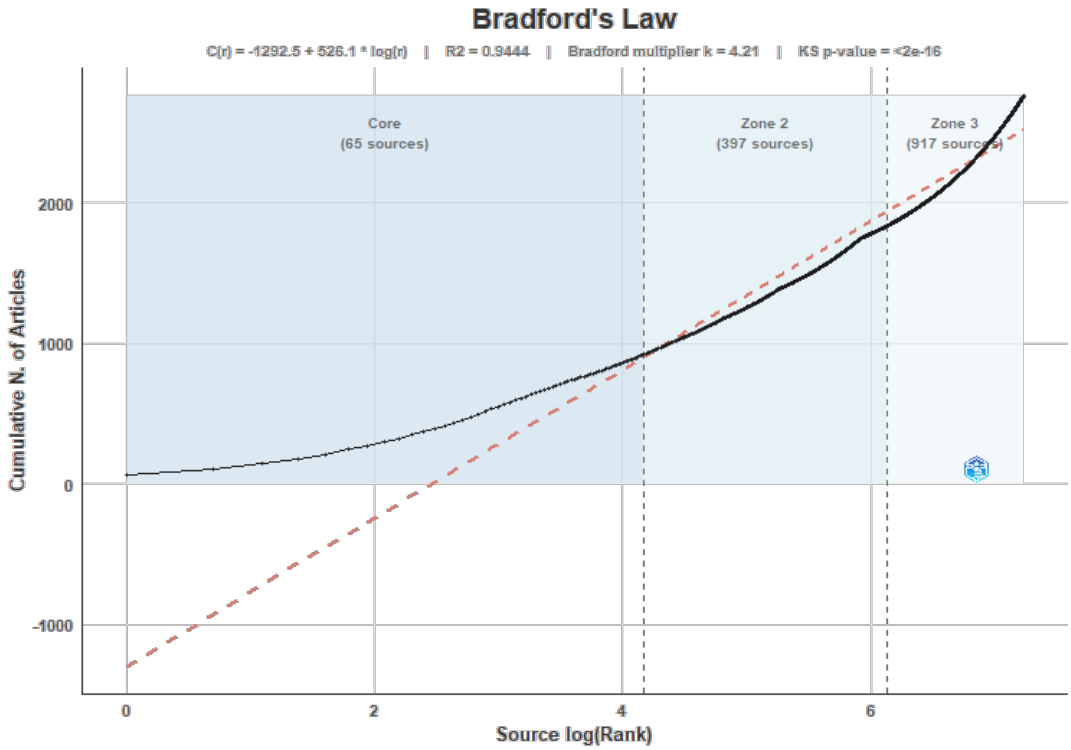

Figure 2: Source dispersion and Bradford's Law of Scattering.

Table 3: Source dispersion under Bradford's Law.

Zone	Classification	Number of Sources (r)	Share of Sources (%)	Cumulative Articles	Bradford Ratio
Zone 1	Nuclear / Core Sources	65	4.71%	~918	1
Zone 2	Moderately Productive Sources	397	28.80%	~1,835	K
Zone 3	Dispersed / Peripheral Sources	917	66.49%	2,753	$_K2$
Total	-	1,379	100.00%	2,753	Mean k = 4.21

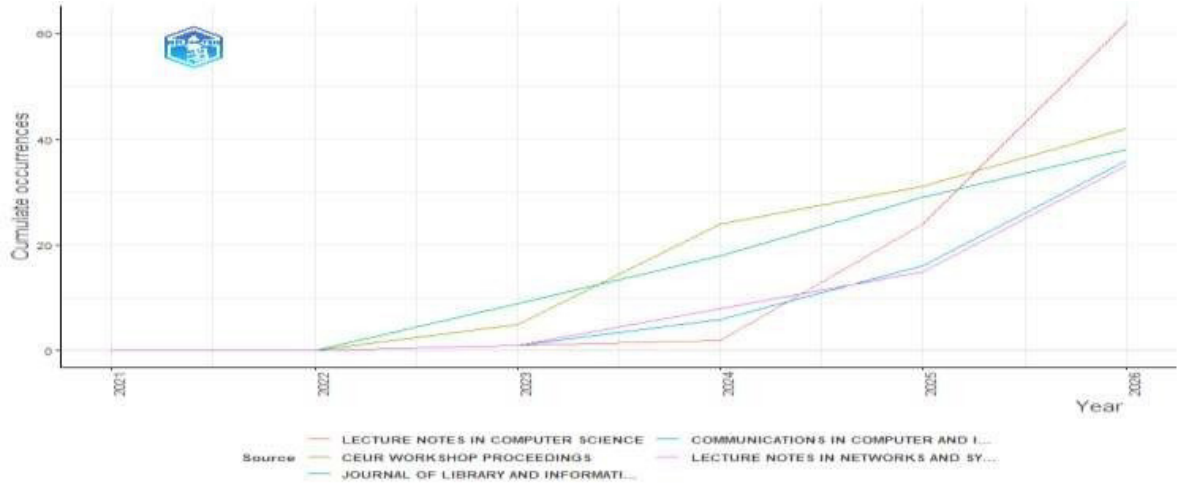

Figure 3: Productivity of Sources over time.

Workshop Proceedings recorded 5. This early activity indicates that the topic began receiving attention through both library and information science publications and computer science workshop proceedings.

Publication output increased in 2024. *CEUR Workshop Proceedings* reached 24 cumulative publications, while the *Journal of Library and Information Science* recorded 18. During the following years, the Springer series *Lecture Notes in Computer Science* showed substantial growth, increasing from 2 publications in 2024 to 24 in 2025 and 62 by 2026, making it the leading source in the reported analysis. *Communications in Computer and Information Science* and *Lecture Notes in Networks and Systems* also recorded notable contributions, reaching 36 and 35 publications, respectively, by the end of the study period. This result indicates that publication activity among the leading sources expanded considerably as research on GenAI in academic libraries developed. Library and information science journals and workshop proceedings contributed to the earlier publication activity, while major computer science proceedings series became prominent in the later years. This pattern highlights the involvement of both library-focused and technology-oriented publication outlets in disseminating research on the topic. However, the source-level trends alone do not establish that particular outlets focused specifically on conceptual research or technical implementation; such distinctions would require examination of the content of the individual publications.

Authors' Productivity over time and Lotka's Law

The analysis of author productivity between 2021 and 2026 indicates that research on Generative Artificial Intelligence (GenAI) in academic libraries has expanded from a relatively small group of early contributors to broader author participation. As illustrated in Figures 4 and 5, publication activity among the selected authors emerged in 2023 and increased in subsequent years. Researchers such as Lund Brady D, Gupta Varun, Michalak Russell, Auer Sören, and Oladokun.

Bolaji David contributed during this developing period. Differences in node size and colour indicate variations in the measures represented by the visualisation; however, citation differences should be interpreted cautiously because earlier publications have had more time to accumulate citations. The overall productivity distribution reinforces this pattern. Of the 10,982 authors who contributed to the 2,753 publications, 10,272 (93.53%) published only one document, while 560 (5.10%) published two. Only a small proportion contributed three or more publications, indicating that the literature is dominated by occasional contributors rather than a large, consistently productive author group. The fitted productivity model yielded

an exponent of (beta = 2.20), a constant of (C = 0.0844), and (R2 = 0.7129). The reported Kolmogorov-Smirnov test produced a *p*-value of 0.124 for both the classical Lotka model and the fitted model. Since this exceeds the

0.05 significance level, the null hypothesis cannot be rejected. Overall, the findings suggest a productivity distribution broadly consistent with a Lotka-type pattern, although they do not establish definitive conformity with Lotka's Law.

Most relevant affiliation and affiliation production over time

The institutional analysis reveals that publications are concentrated among a group of research-intensive universities. As shown in Figure 6, Peking University ranked first with 57 publications, followed by Tsinghua University (44), Zhejiang University and the University of California system (43 each), and the University of Chinese Academy of Sciences (42). Other notable contributors included the University of Hong Kong (33), Sun Yat-sen University (32), Nanjing University (30), and Nanyang Technological University (26). The growth in publication output among leading institutions from 2023 onwards, with Peking University recording a substantial increase by 2026. The strong representation of Chinese institutions highlights their significant contribution to GenAI-related research, alongside contributions from the United States and Singapore. However, these publication counts reflect output within the retrieved dataset and do not independently establish the influence of institutional policies, funding, or investment in AI infrastructure and library services.

Country-Wise Publication Output and International Collaboration

The geographical distribution of research on Generative Artificial Intelligence (GenAI) in academic libraries reveals substantial contributions from China and the United States, alongside participation from countries across Asia, Europe, and other regions. As shown in Table 4, China had the highest number of corresponding-author articles, with 567 publications (20.60%), followed by the United States with 385 (13.98%) and India with 104 (3.78%). Germany (95), Italy (63), the United Kingdom (61), and South Korea (60) were also notable contributors. Together, China, the United States, and India accounted for 38.36% of the corresponding-author publications, highlighting their prominent positions in the literature.

The country-affiliation distribution provides a complementary perspective on geographical contributions. China recorded 1,665 affiliation occurrences (28.52%), followed by the United States with 1,467 (25.13%) and India with 318 (5.45%). Germany and the United Kingdom recorded 288 and 244 occurrences, respectively. These figures represent author-affiliation occurrences rather than unique publications and therefore should not be directly compared with corresponding-author article counts. Taken together, the results indicate that China and the United States account for substantial shares of the recorded research affiliations, while contributions extend across a wider range of countries.

The analysis of Single-Country Publications (SCP) and Multiple-Country Publications (MCP) further highlights differences in international collaboration. China recorded 483 SCP and 84 MCP, giving an MCP ratio of 14.81%, while

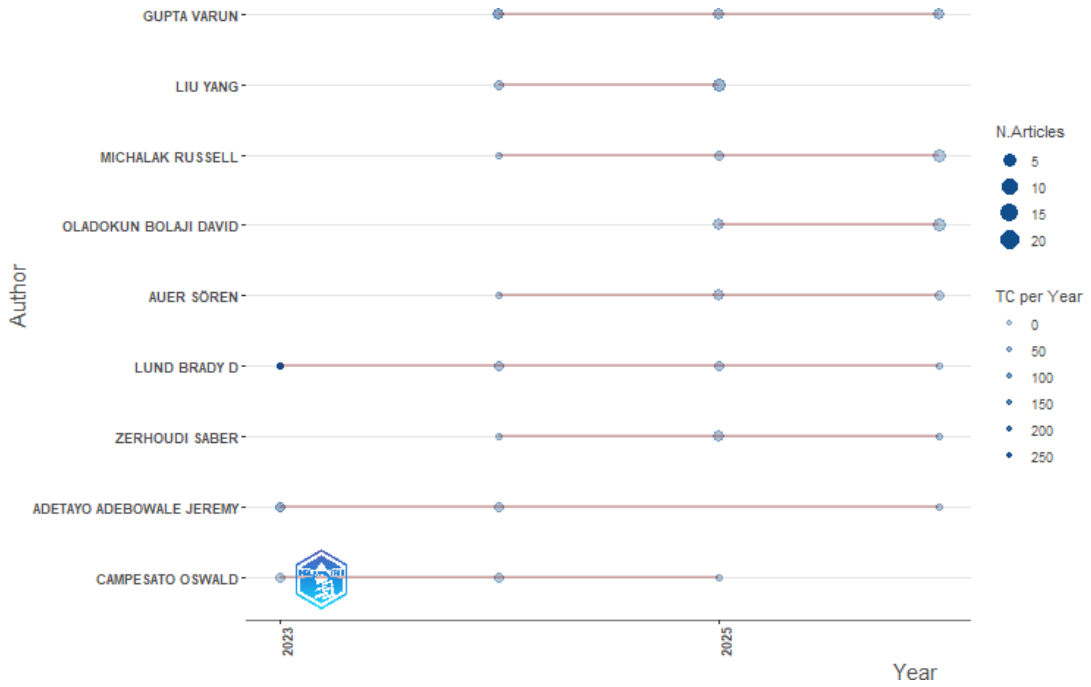

Figure 4: Author production over time.

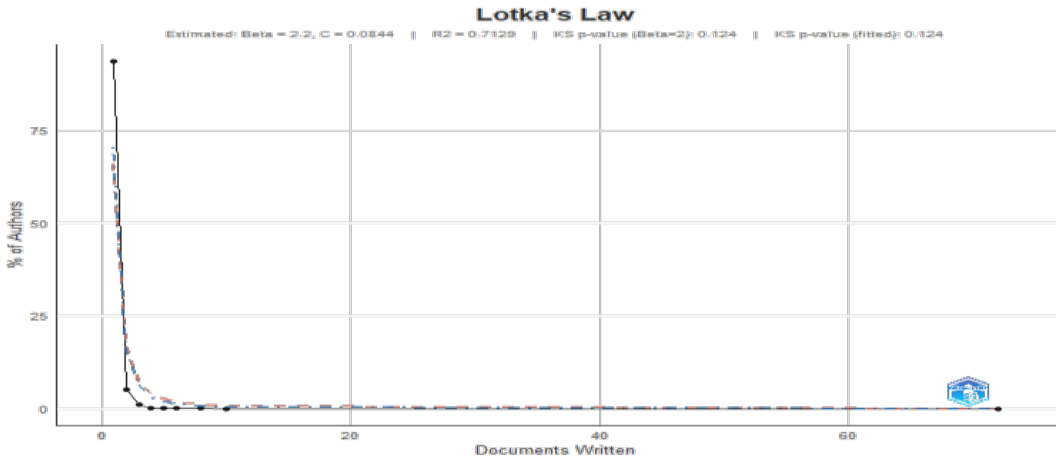

Figure 5: Authors' Productivity Distribution and Lotka's Law.

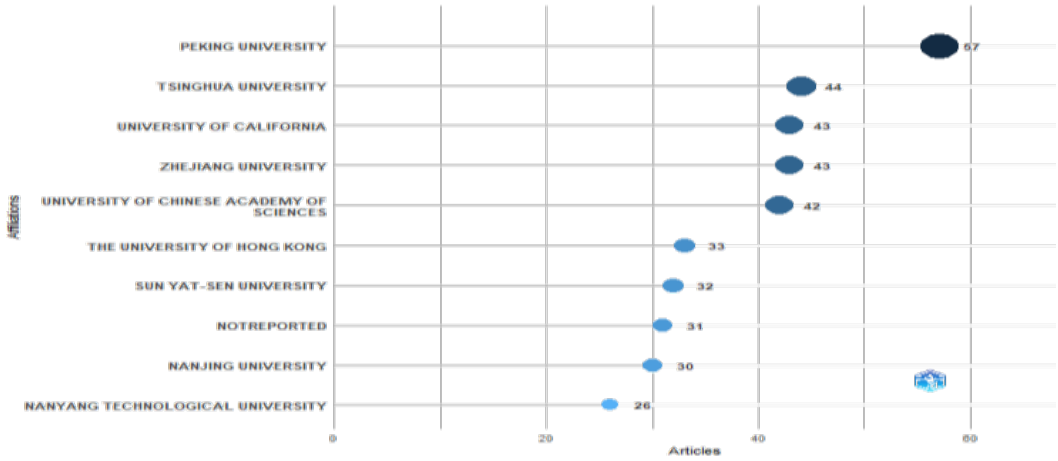

Figure 6: Most relevant institutional affiliations.

the corresponding ratios for the United States and India were 15.84% and 10.58%, respectively. In contrast, higher MCP ratios were observed for Pakistan (64.71%), Hong Kong (63.33%), Singapore (52.63%), the Netherlands (47.06%), and the United Kingdom (40.98%). These patterns suggest that internationally co-authored research constitutes a larger proportion of the corresponding-author output in some countries and territories than in others.

Overall, the findings reveal both a concentration of research output among leading contributors and variation in international collaboration patterns. However, publication and affiliation counts alone cannot explain the factors underlying these differences or establish the extent of institutional cooperation, research investment, or technology transfer. Furthermore, the annual country-production figures require verification against the original Scopus export before changes over time can be interpreted reliably.

Most-cited countries

The country-level citation analysis from the above Table 5 provides a complementary perspective on publication output by highlighting differences in citation impact among leading contributors. The United States recorded the highest total number of citations, with 4,260, accounting for 39.28% of citations among the top 10 countries. Its Average Article Citation (AAC) was 11.06. China ranked second, with 2,607 citations and an AAC of 4.60. The United Kingdom ranked third in total citations, with 1,555, despite having 61 corresponding-author articles. Its AAC of 25.49 was higher than that reported for the United States and China, indicating a comparatively high citation rate per article. This result highlights the distinction between the volume of a country's research output and its average citation impact. However, citation counts alone do not establish the specific themes or characteristics responsible for this impact.

Among the other countries listed, Germany received 673 citations, with an AAC of 7.08, while South Korea received 485 citations, with an AAC of 8.08. Sweden also demonstrated a comparatively high AAC of 11.07, based on 321 citations. India recorded 379 citations and an AAC of 3.64, while Nigeria received 261 citations, with an AAC of 4.42. These results show variation in citation impact across the countries represented in the dataset. The overall findings indicate that the countries producing the largest volume of research do not necessarily achieve the highest average citation rates.

Keyword Distribution, Frequency, and Emerging Trends

The keyword frequency and temporal analysis provide an overview of the principal concepts represented in the retrieved literature. As shown in Figure 7, language model was the most frequent keyword (1,162 occurrences), followed by large language model (956), artificial intelligence (731), large language models (560), human (386), digital libraries (351), ChatGPT (345), computational linguistics (314), generative AI (299), and natural language processing systems.

(264). The trending-keyword analysis indicates greater prominence and diversity of terminology from 2023 onwards. In 2024, terms such as problem oriented languages (69 occurrences), adversarial machine learning (38), and program debugging (31) featured prominently, while language-model and AI-related terms remained important in later years. Human-related terms, including human (386) and humans (246), had median years of 2026, reflecting their presence in the more recent portion of the dataset.

The findings show that the retrieved literature encompasses language-model technologies, broader AI concepts, digital-library contexts, technical computing topics, and human-related terminology. However, keyword frequencies indicate the vocabulary represented in the corpus rather than conclusively demonstrating thematic shifts or the intellectual evolution of the field. The separate occurrence of singular and plural variants also affects frequency comparisons; notably, large language model appears with frequencies of 956 in the keyword analysis and 954 in the trend output. These values should be verified, and related terms should be normalised or explicitly treated as separate indexed keywords. Moreover, because the search strategy may retrieve publications that do not explicitly address both GenAI and academic libraries, the findings should be interpreted in the context of the retrieved dataset. Further examination of keyword co-occurrence patterns and the underlying publications would be needed to establish more specific thematic relationships. The word-cloud representation of the most frequent words is presented separately in Figure 8.

Table 4: Country-Wise Publications Output and International Collaboration.

Rank	Country / Territory	Corresponding Articles (NP)	Articles Share (%)	Single Country Publications (SCP)	Multiple Country Publications (MCP)	MCP Ratio (%)	Affiliation Occurrences	Affiliation Share (%)
b1	China	567	20.60%	483	84	14.81%	1,665	28.52%
2	USA	385	13.98%	324	61	15.84%	1,467	25.13%
3	India	104	3.78%	93	11	10.58%	318	5.45%
4	Germany	95	3.45%	65	30	31.58%	288	4.93%
5	Italy	63	2.29%	53	10	15.87%	189	3.24%
6	United Kingdom	61	2.22%	36	25	40.98%	244	4.18%
7	Korea	60	2.18%	43	17	28.33%	144	2.47%
8	Canada	50	1.82%	38	12	24.00%	164	2.81%
9	Australia	32	1.16%	20	12	37.50%	128	2.19%
10	Spain	32	1.16%	22	10	31.25%	109	1.87%
11	Brazil	30	1.09%	21	9	30.00%	91	1.56%
12	Hong Kong	30	1.09%	11	19	63.33%	-	-
13	Switzerland	23	0.84%	16	7	30.43%	71	1.22%
14	Turkey	22	0.80%	19	3	13.64%	64	1.10%
15	Poland	19	0.69%	15	4	21.05%	40	0.69%
16	Singapore	19	0.69%	9	10	52.63%	88	1.51%
17	Netherlands	17	0.62%	9	8	47.06%	47	0.81%
18	Pakistan	17	0.62%	6	11	64.71%	62	1.06%
19	Thailand	17	0.62%	12	5	29.41%	-	-
20	Greece	16	0.58%	12	4	25.00%	-	-

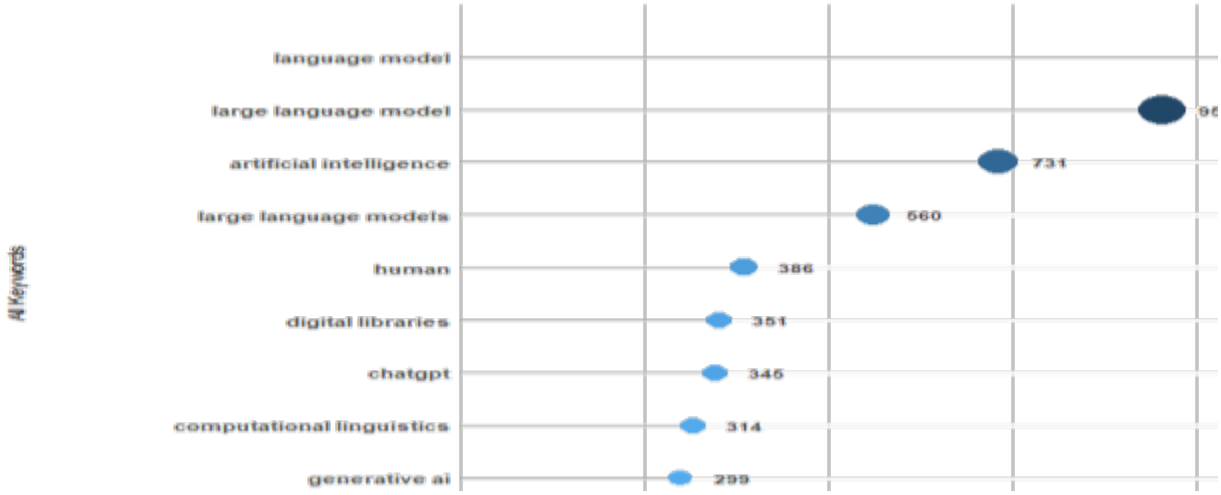

Figure 7: Most frequent words.

Keywords co-occurrence network

The keyword co-occurrence network from Figure 9 reveals three major clusters that represent related conceptual areas within the literature on generative artificial intelligence and academic libraries. The clusters bring together terminology associated with language-model architectures and digital library infrastructure, AI applications and human-related considerations, and natural language processing. The network therefore provides a view of the relationships among technical, application-oriented, and language-processing concepts in the retrieved corpus.

Cluster 1: Model Architecture and Digital Library Infrastructure

Cluster 1 represents the technical and infrastructural dimension of the network. The term "large language model" has the highest reported betweenness centrality (40.7218), while "language model" has a betweenness centrality of 25.9050. Their PageRank values are 0.0819 and 0.0840, respectively. "Large language models" also appears as a prominent node, with a betweenness centrality of 11.5251 and a PageRank of 0.0477. These measures indicate that language-model terminology occupies an important position in the network. In particular, high betweenness centrality suggests that certain terms connect otherwise less directly connected parts of the keyword network. Other terms in this cluster include "digital libraries" (betweenness centrality = 3.1885; PageRank = 0.0320), "libraries" (betweenness centrality = 1.7448), and "information retrieval" (betweenness centrality = 1.1262). Their co-occurrence with terms such as "retrieval-augmented generation," "open source software," "human computer interaction," "semantics," and "codegeneration" indicates connections between language-model research, digital information environments, and computational methods. These relationships suggest that the literature considers generative AI alongside information retrieval and digital library technologies. However, the network alone does not establish the extent to which specific systems, such as retrieval-augmented generation, have been implemented in academic libraries.

Cluster 2: AI Applications and Human-Related Research

Cluster 2 captures a set of terms associated with artificial intelligence, conversational tools, human-related concepts, and research processes. "Artificial intelligence" has a reported betweenness centrality of 19.5330 and a PageRank of 0.0533, indicating its prominent position in the network. The terms "human" (betweenness centrality = 3.5705; PageRank = 0.0390) and "humans" (betweenness centrality = 1.1452; PageRank = 0.0278) also appear in this cluster. Other connected terms include "ChatGPT," "chatbots," "students," "decision making," and "procedures." The cluster also includes "article" (betweenness centrality = 0.8358) and "systematic review" (betweenness centrality = 0.7922), linking AI-related terminology with scholarly publication and research activity. These co-occurrence patterns indicate that the corpus includes human-related and research-oriented terminology alongside AI applications. Nevertheless, the presence of these keywords does not independently demonstrate that human oversight, AI literacy instruction, algorithmic bias mitigation, or librarian-led evaluation are dominant themes. Such interpretations require support from the underlying publications or additional thematic analysis.

Cluster 3: Linguistic Foundations and Computational Text Processing

Cluster 3 is centred on language-processing terminology. "Natural language processing" has a betweenness centrality of 1.3396 and a PageRank of 0.0215, while "natural languages" and "language processing" have betweenness centrality values of 0.4504 and 0.3226, respectively. The grouping of these terms indicates a connected area of the network concerned with language and computational text processing. The cluster provides evidence of the representation of natural language processing concepts within the retrieved literature. Its relationship to bibliographic text mining, semantic parsing, classification, or other specific

library applications should be established through the associated publications rather than inferred from the cluster labels alone.

The overall keyword co-occurrence network identifies three connected areas: language-model architectures and digital library infrastructure, AI applications and human-related research, and natural language processing. The centrality measures help identify terms that occupy prominent or connecting positions in the network, while the clusters show how related concepts are grouped. Together, these findings provide a structural overview of the concepts represented in the corpus without, on their own, establishing the implementation status or maturity of generative AI in academic libraries.

Authors collaboration network

The author collaboration network presented in Figure 10 illustrates the structure of co-authorship among researchers contributing to the literature on generative AI in academic libraries. The network contains nine identified communities, with the reported patterns indicating a fragmented structure characterised by small collaborative groups and limited connections between some components. Rather than forming one extensively interconnected network, the authors appear in several localised collaboration groups.

Clusters 3-9 illustrate this pattern through several small co-authorship groups. The reported dyads include Auer Sören and D'Souza Jennifer (Cluster 3), Zerhoudi Saber and Granitzer Michael (Cluster 4), Wang Bo and Wang Wei (Cluster 5), Wang Xin

Table 5: **Most-cited corresponding-authors' countries.**

Rank	Country	Total Citations (TC)	Articles (Corresponding)	Average Article Citations (AAC)	Share of Top 10 Citations (%)
1	USA	4,260	385	11.06	39.28%
2	China	2,607	567	4.60	24.03%
3	United Kingdom	1,555	61	25.49	14.34%
4	Germany	673	95	7.08	6.21%
5	Korea	485	60	8.08	4.47%
6	India	379	104	3.64	3.49%
7	Sweden	321	29	11.07	2.96%
8	Italy	290	63	4.60	2.67%
9	Nigeria	261	59	4.42	2.41%
10	Hong Kong	216	30	7.20	1.99%



Figure 8: Word Cloud of Most Frequent Words.

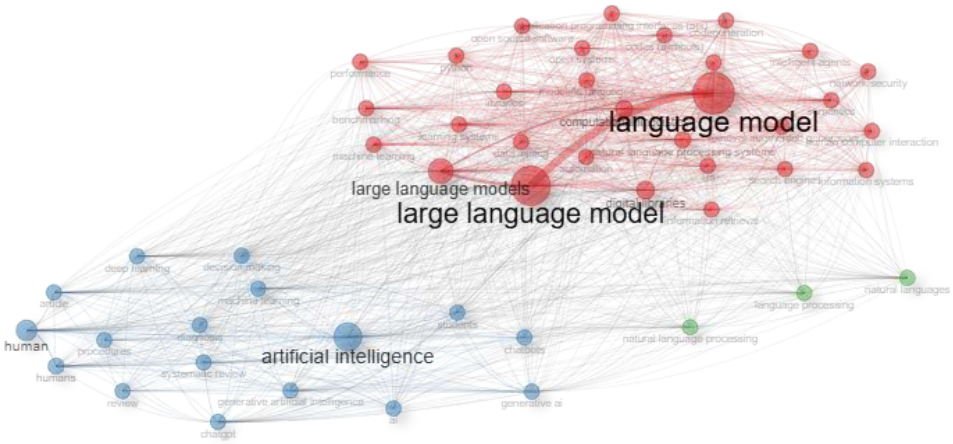

Figure 9: Keywords co-occurences network.

and Chen Hong (Cluster 6), Zhang Lingming and Xia Chunqiu Steven (Cluster 7), and Ingram William A and Banerjee Bipasha (Cluster 9). Cluster 8 comprises Godoy William F, Teranishi Keita, and Vetter Jeffrey S. The reported centrality results indicate limited brokerage by these groups within the wider network. However, the extent of their isolation should be confirmed by inspecting the complete network and its component structure.

The first two clusters contain more interconnected collaboration patterns. In Cluster 1, Liu Yang is identified as a prominent node, with a betweenness centrality of 2.0000 and a PageRank of 0.0667. The reported network structure connects Liu Yang with Liu Wei, Wang Bin, and Chen Kai. In Cluster 2, Oladokun Bolaji David has a betweenness centrality of 1.0000, a closeness value of 0.5000, and a PageRank of 0.0663, linking the group's collaboration with Ajani Yusuf Ayodeji and Ayinde Lateef. These results suggest that certain authors occupy more central positions within their respective collaboration groups. Nevertheless, their centrality within a local group should not automatically be interpreted as evidence of brokerage across the entire research field. The overall co-authorship analysis indicates that collaboration in the retrieved corpus is organised into several relatively small research groups, with the reported network offering limited evidence of extensive connections between communities. This pattern points to opportunities for strengthening collaboration across institutions and disciplines. Future research could encourage partnerships between computer science researchers and academic librarians, as well as multi-institutional studies examining the applications, limitations, and evaluation of generative AI in academic library settings. Broader collaboration may also help researchers compare findings across institutional contexts and develop more consistent approaches to evaluating AI-enabled library services.

DISCUSSION

The findings address the six research objectives by identifying a sharp increase in publication output after 2022, a diverse publication-source landscape, a highly skewed author-productivity distribution, concentrated institutional and country participation, keyword clusters connecting language models with library-related concepts, and a fragmented co-authorship structure. These results provide an overview of the records retrieved through the search strategy adopted in this study. However, because the Boolean search expression may not consistently require both Generative Artificial Intelligence (GenAI)-related and library-related concepts, the findings should be interpreted as describing the retrieved corpus rather than the entire field of GenAI research in academic libraries.

First, the annual publication analysis indicates substantial growth from 2023 onward, with particularly high record counts in 2025 and 2026. The 2026 figure is provisional because the Scopus search was conducted on 26 September 2026 and therefore does not represent a complete calendar year. The citation analysis also highlights the importance of citation windows when comparing publication years. Older publications have had more time to accumulate citations, so differences in mean citation counts should not be interpreted directly as differences in research quality or influence without accounting for publication age and other relevant factors.

Second, the source analysis shows that the retrieved records are distributed across a large number of publication outlets, while a smaller group of sources accounts for a notable share of the output. The Bradford-zone analysis suggests a degree of source concentration. However, the reported goodness-of-fit test is statistically significant and indicating a departure from the fitted distribution under the test as reported. Consequently, the result should not be presented as definitive confirmation of Bradford's Law. The source counts, zone boundaries, fitted curve, and

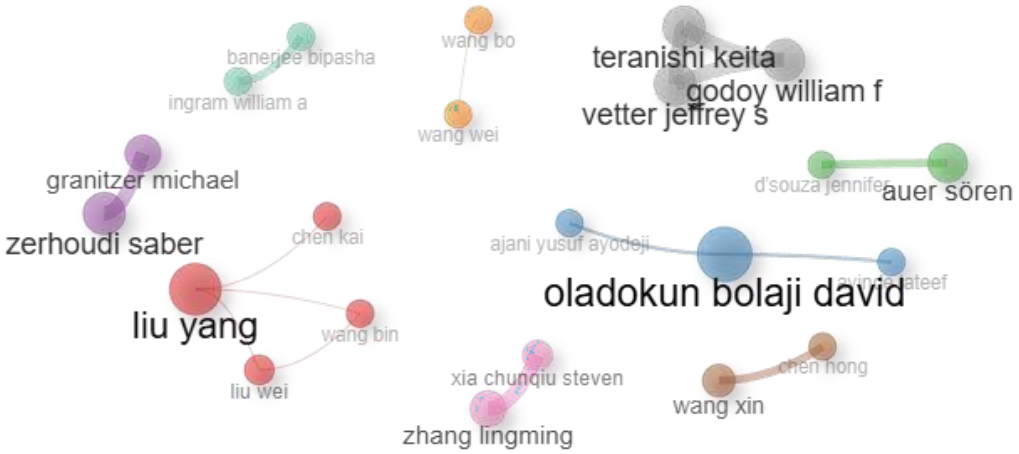

Figure 10: Authors collaboration network.

statistical procedure require verification against the underlying data before a firm conclusion can be drawn.

Third, the author-productivity results reveal a strongly skewed distribution. Most authors contributed one publication to the retrieved corpus, whereas a comparatively small proportion contributed multiple publications. The reported Kolmogorov-Smirnov test result ((p=0.124)) does not provide sufficient evidence to reject the specified Lotka model at the 0.05 significance level. Nevertheless, a non-significant result does not prove that the observed distribution conforms exactly to Lotka's Law. The model specification, treatment of authors with four or more publications, and test procedure should be documented and checked before the result is interpreted as evidence of conformity.

Fourth, the institutional and country analyses show substantial publication representation from China and the United States. However, publication volume and citation impact do not produce identical rankings. The analyses also use different counting bases: country-production figures represent author-affiliation occurrences across the retrieved records, whereas the corresponding-author analysis assigns articles according to the country of the corresponding author. These indicators must therefore be labelled and interpreted separately rather than compared as though they shared the same denominator. Citation differences may also be influenced by publication age, disciplinary composition, and database coverage. The present results do not establish the reasons for the differences in publication volume or citation impact among countries and institutions.

Fifth, the keyword-frequency and co-occurrence analyses identify language-model terminology, artificial intelligence, digital libraries, and natural-language-processing concepts as prominent elements of the retrieved vocabulary. Their co-occurrence suggests a connection within the corpus between technical AI-related concepts and library or information-discovery topics. However, the appearance of terms such as "human," "article," "program debugging," and "adversarial machine learning" may also reflect the broad search expression and the multidisciplinary coverage of the retrieved records. These terms should not automatically be treated as evidence of specific trends in academic libraries without examining the relevant publications and confirming their subject relevance.

Finally, the co-authorship analysis indicates a fragmented collaboration structure, with several small groups and limited reported betweenness centrality among the mapped authors. This suggests that the observed network does not form a single densely interconnected collaboration structure. Nevertheless, network measures depend on decisions concerning author-name disambiguation, link thresholds, counting methods, and the treatment of disconnected components. Reporting these settings is essential for reproducibility and for determining how far the network structure reflects the underlying publication data.

Overall, the findings provide a descriptive account of publication patterns, sources, authorship, institutional and country participation, keyword relationships, and collaboration within the retrieved corpus. Their interpretation as a map specifically of GenAI research in academic libraries depends on improving the search strategy and verifying the relevant data-processing and analytical procedures.

IMPLICATIONS AND RECOMMENDATIONS

The findings suggest four priorities for future research and scholarly reporting: first, researchers should adopt a fully parenthesised search strategy that explicitly combines GenAI-related and academic-library-related concepts and report the complete search expression, database, search date, filters, and screening

procedures to ensure reproducibility. Second, academic-library-specific studies should be distinguished from broader research on digital libraries, education, information science, and computer science through transparent screening criteria and, where appropriate, a validated search strategy. Third, institutional and country comparisons should clearly specify the counting methods and denominators used, reporting country-affiliation occurrences, corresponding-author article counts, total citations, and average citations per article as separate indicators; where suitable, citation windows and field-normalised indicators should also be considered to improve comparability. Finally, future science-mapping and network studies should report software and package versions, counting and normalisation methods, network thresholds, and community-detection settings, while sharing search strategies, data-cleaning procedures, and analysis code or other reproducibility materials wherever licensing conditions permit, thereby enabling verification and further development of the findings.

LIMITATIONS

This study has few limitations. First, it relies exclusively on Scopus, excluding publications not indexed in the database. Second, the data were retrieved on 26 September 2026; therefore, the publication count for 2026 is provisional. Citation comparisons across publication years are also affected by differences in citation accumulation time. Furthermore, variations in keyword terminology and bibliographic metadata, along with the search strategy, may influence the results and limit their representativeness of research specifically focused on generative AI in academic libraries.

CONCLUSION

This study mapped the intellectual landscape of Generative Artificial Intelligence (GenAI) research related to academic libraries through a scientometric analysis of 2,753 Scopus-indexed records published between 2021 and 2026. The findings reveal substantial growth in publication output, particularly during 2024-2026, indicating increasing scholarly attention to this emerging research area. The analysis also identified leading publication sources, productive authors and institutions, geographical contributions, and patterns of scholarly collaboration. Keyword analysis highlighted the prominence of language-model technologies, artificial intelligence, digital libraries, ChatGPT, and related computational concepts, reflecting the diverse scope of the retrieved literature. The findings further demonstrate the value of combining descriptive bibliometric indicators with science-mapping techniques to examine research productivity, conceptual relationships, and collaboration patterns within an emerging field. Despite these contributions, the results are interpreted in light of the study's reliance on Scopus, the partial coverage of 2026. Moreover, publication and citation patterns do not independently establish the actual adoption or effectiveness of GenAI in academic library services. Future research should employ transparent and reproducible search strategies, refine the identification of academic-library-specific studies, and investigate the practical experiences of librarians and users. Overall, this study provides a structured overview of the available literature and a foundation for further research into the development and responsible application of GenAI in academic libraries.

ABBREVIATIONS

AAC: Average Article Citations; **AI:** Artificial Intelligence; **GenAI:** Generative Artificial Intelligence; **LIS:** Library and Information Science; **LLMs:** Large Language Models; **MCP:** Multiple Country Publications; **NP:** Number of Articles; **R^2:** Coefficient of Determination; **RO:** Research Objective; **SCP:** Single Country Publications; **TC:** Total Citations; **β:** Beta Exponent Parameter in Lotka's Productivity Model.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORS CONTRIBUTION

Vishnudas A V: Data curation (data retrieval), formal analysis, interpretation of results, and writing - original draft. **Arathy G S:** Supervision, development and structuring of the article content, content review, and writing - review and editing. Both authors are responsible for the final manuscript.

REFERENCES

Adeniran, P. O., & Oyovwevotu, L. (2019). Academic library research support services: A review of Redeemer's University and the Nigeria Natural Medicine Development Agency's research activities. *Library Philosophy and Practice*, 2753. https://digitalcommons.unl.edu/libphilprac/2753

Ahmad, K., Ming, Z. J., & Rafi, M. (2018). Assessing the digital library research output: Bibliometric analysis from 2002 to 2016. *The Electronic Library*, *36*(4), 696–704. https://doi.org/10.1108/EL-02-2017-0036

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Borgman, C. L. (1999). What are digital libraries? Competing visions. *Information Processing & Management*, *35*(3), 227–243. https://doi.org/10.1016/S0306-4573(98)00059-4

Borgohain, D. J., Zakaria, S., & Verma, M. K. (2021). Cluster analysis and network visualization of global research on digital libraries during 2016–2020: A bibliometric mapping. *Science & Technology Libraries*, *41*(3), 266–287. https://doi.org/10.1080/0194262X.2021.1993422

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, *133*, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Kulkanjanapiban, P., Silwattananusarn, T., & Lambovska, M. (2025). Research on AI-driven innovations and services in academic libraries: A bibliometric and systematic literature review. *Journal of Data and Information Science*, *10*(4), 146–196. https://doi.org/10.2478/jdis-2025-0036

Lestari, P., & Dewanti, S. S. (2025). Bibliometric analysis of digital library trends using information literacy in Scopus database. *Jurnal Kajian Informasi & Perpustakaan*, *13*(1). https://doi.org/10.24198/jkip.v13i1

Mahadevagouda, R., & Pavithrabai, M. S. (2022). Recent trends in digital library publications: A scientometric analysis. *Library Philosophy and Practice*, 7058. https://digitalcommons.unl.edu/libphilprac/7058

Mondal, I., & Surendran, S. (2024). A scientometric analysis on global digital library research during 1991–2022. *Annals of Library and Information Studies*, *71*(3), 302–309. https://doi.org/10.56042/alis.v71i3.8966

Ncube, M. M., & Ngulube, P. (2025). Beyond service inventories: A three-dimensional framework for diagnosing structural barriers in academic library research dataset management. *Information*, *16*(12), Article 1046. https://doi.org/10.3390/info16121046

Padhan, D. K., & Naidu, G. H. (2022). Research support services in academic libraries. In [Eds. missing], *Envisioning digital transformation in libraries for NextGen academic landscape* (pp. 197–207). Academic Publishers.

Perkins, G. H., & Slowik, A. J. W. (2013). The value of research in academic libraries. *College & Research Libraries*, *74*(2), 143–158. https://doi.org/10.5860/crl-308

Wattanasiri, P., Nguyen, L. T., & Kwiecien, K. (2025). The changing landscape of digital library services research: A comparative study of two periods (1994–2023). *Journal of Multidisciplinary Academic Research and Development*, *7*(3), 46–70.

Cite this article: Vishnudas AV, Arathy GS. Mapping the Intellectual Landscape of Generative AI in Academic Libraries: A Scientometric Analysis and Science Mapping of Scopus Literature (2021-2026). Journal of Data Science, Informetrics, and Citation Studies. 2026;5(3):431-46.