IJEMD-SS, 5 (3) (2026) https://doi.org/10.54938/ijemdss.2026.05.3.750

International Journal of Emerging Multidisciplinaries: Social Science

Research Paper
Journal Homepage: *www.ojs.ijemd.com*
ISSN (Online): *2958-0277*

Artificial Intelligence Tools and English Writing Development in Higher Education: A Scopus-Based Bibliometric Analysis Using Biblioshiny and VOSviewer

1. Dr. Tasneem Akhter, 2. Dr. Safia Begum, 3. Iqra Mehmood, 4. Aqsa Ramzan

1. Assistant Professor of Economics, Faculty of Management Sciences, University of Central Punjab, Lahore, Pakistan
tasneem.akhter@ucp.edu.pk **(Corresponding author)**

2. Associate Professor of Economics &Principal Government Girls Degree College at Higher Education Department
safiakakakhail@gmail.com

3. MS Economics, Department of Economics, International Islamic University Islamabad, Pakistan
Iqramehmood07025@gmail.com

4. Department of Economics, Government College University Faisalabad
aqsaamn241@gmail.com

Abstract

The use of generative AI has transformed the planning, drafting, revising, and assessment of English writing in higher education. Although research on the area has expended rapidly, there is remains a lack of a coherent understanding of the fields intellectual structure, publications patterns, key researchers, collaboration networks, and the thematic development. This study examines the evolution of research on AI tools and English writing in higher education and draws inferences for BS English programmes in Pakistani universities. A Scopus-based dataset was filtered based on year, subject-area, document-type, language, publication-stage, and source-type. From the 2,458 records initially identified, 1,034 English-language journal articles published in the Social Sciences at the final publication stage were included for the period 2020-2026. Performance analysis and science mapping were performed using Biblioshiny/bibliometrix and VOSviewer, respectively. The analysis examined annual scientific production and citations; productive sources and authors; locally cited documents; Bradford's and Lotka's laws; h-index indicators; country collaboration; keyword co-occurrence; thematic mapping; trend topics; Reference Publication Year Spectroscopy (RPYS); and a logistic model of the publication life cycle. Scientific production increased from 1 article in 2020 to 431 in 2025 and 354 in 2026; however, the 2026 figure should be interpreted with caution because the year may be unfinished. The conceptual structure was dominated by ChatGPT (257 occurrences), artificial intelligence (223 occurrences), generative AI (114 occurrences), and EFL writing (59 occurrences). Arab World English Journal was the most productive source, while Guo K and Wang Y were the most prolific authors. Bradford's Law analysis displayed that 18 journals accounted for 34.3% of the literature, while Lotka's Law showed a highly attentive authorship pattern. Thematic and temporal studies revealed a shift from general research on AI tools and chatbots toward generative AI, large language models, and EFL writing. Overall, the field is rising swiftly but remains uneven across methodological and geographical dimensions. The outcomes highlight the importance of curriculum-integrated AI literacy, process-oriented writing assessment, transparent AI-use disclosure, faculty development, and locally grounded empirical research in Pakistani universities.

Keywords: Artificial intelligence, ChatGPT, Generative AI, English writing, EFL, Higher education, Bibliometric analysis, Biblioshiny, VOSviewer, Scopus, Pakistan

1. Introduction

Artificial Intelligence (AI) is now an integral part of modern academic writing. While some writing-support systems focused on correcting spelling, grammar, punctuation, and stylistic issues at surface level, generative AI systems generate outlines, paragraphs, explanations, revisions, and feedback based on natural-language prompts. The wide-seeing of ChatGPT in late 2022 increased the interest in education regarding the use of LLMs and prompted discussions on how they can be utilized in education, learning and assessment, and research (Kasneci et al., 2023; Miao & Holmes, 2023). In the field of English language learning, they are especially significant as the learning of writing involves repetition, formative feedback, lexical and grammatical assistance, understanding of genre, and revision.

AI tools for writing have been reported to offer a number of educational affordances. ChatGPT may help learners to brainstorm, structure ideas, to provide language revision, and to reflect on written arguments (Su et al., 2023). Mixed methods and intervention studies have also correlated structured use of ChatGPT with positive academic writing outcomes, motivation, and feedback uptake by EFL/ESL students (Mahapatra, 2024; Song & Song, 2023). Meanwhile, the advantages are not guaranteed and are contingent. When AI-generated content is merely substituting the student's own cognitive and linguistic efforts, it can hinder learning by providing generic or inaccurate feedback, presenting fabricated information, increasing reliance, reducing authorial agency, posing privacy issues, and causing academic integrity concerns (Imran & Almusharraf, 2023; Kasneci et al., 2023; Miao & Holmes, 2023).

This is particularly relevant where English is an 'additional language' and where students' starting point for university is unevenly academic literate. In BS English, students should be able to acquire skill and competence in grammar, composition, literary analysis, research writing, argumentation and scholarly communication in the Pakistan context. While AI tools can help overcome language barriers and offer immediate support, their pedagogical significance relies on the institutional policy, teacher mediation, students' AI literacy, and the structure of the assessment. Bibi and Atta (2024) observed a generally positive perception of ChatGPT as an English writing assistant amongst the students, while also noting the necessity for studying its longer-term impacts on students' autonomous writing and ethical utilization. The national debate is also evolving from a focus purely on banning generative AI to responsible governance, as the Higher Education Commission of Pakistan (HEC) has turned its attention to laying down a framework of ethical norms for generative AI in higher education (HEC, 2025).

While there is a growing number of individual empirical studies and systematic reviews, the literature has been expanding so fast that it cannot be clearly discerned by a simple narrative review of its publication trends, focus, networking, conceptual underpinnings, or themes. The bibliometric analysis solves this problem by using quantitative methods to analyse bibliographic data and the relationships found in the citations. It can determine who the key authors and journals are, identify co-occurring concepts, explore collaboration, and track the emergence of new research fronts (Donthu et al., 2021; Zupic & Čater, 2015). The analysis and visualization of large bibliographic datasets are possible by using science-mapping tools like bibliometrix/Biblioshiny (Aria & Cuccurullo, 2017) and VOSviewer (van Eck & Waltman, 2010).

Thus, the present study investigates the global Scopus-indexed knowledge base on AI tools and English writing in higher education and makes a careful delimitation of implications for the BS English programmes of Pakistani universities. This is significant because it is an international dataset, and not just from Pakistani institutions. The bibliometric results are therefore representative of the global research landscape while the Pakistan-centric sections are aimed at unpacking the potential implications for Pakistan's policy, pedagogy and research agendas. This separation makes it possible to understand what is being said about students globally without assuming that the same applies to Pakistani students.

This study has four contributions. First, it combines the elements of performance indicators with conceptual and social and intellectual mapping. Second, it explores classic bibliometric regularities as described by Bradford's law and Lotka's law, as well as the use of h-index indicators and reference-spectroscopy indicators. Third, it details how the general AI tools and chatbots have progressed to ChatGPT, generative AI, large language models, and EFL writing over time. Fourth, it makes the worldwide evidence structure into practical priorities for Pakistani BS English programmes rather than exaggerating the conclusions that the bibliometric evidence can draw about local learner outcomes.

1.1.Research Questions

1. **RQ1**. How did scientific production and citation impact evolve between 2020 and 2026?
2. **RQ2**. Which journals, authors, and documents were the most productive or influential within the dataset?
3. **RQ3**. How concentrated was publication output across journals and authors according to Bradford's and Lotka's distributions?
4. **RQ4**. What geographic collaboration patterns and conceptual structures characterized the field?
5. **RQ5**. Which topics were established, emerging, declining, or transitioning over time?
6. **RQ6**. What implications do the global bibliometric patterns have for AI-supported English writing development in Pakistani BS English programmes?

2. Literature Review

2.1.AI Tools and the English Writing Process

AI-assisted writing is not just one application, but a group of applications. Corrective tools can detect grammatical and lexical errors, punctuation errors, and style errors; paraphrasing tools paraphrase sentences; translation systems can help with cross-linguistic transfer of ideas; plagiarism and similarity systems can aid source-use monitoring; and generative AI systems can engage throughout the writing process. This diversity is important because the effectiveness of a tool's educational impact will vary according to its purpose. A grammar suggestion might lead to noticing and self-correction, while the automatic use of a generated paragraph may skip planning, building arguments and language production. AI can best be used as a scaffold in an iterative process to foster writing development. Students have the opportunity to compare alternatives, ask for clarification of mistakes, test the coherence of the paragraph, produce counterarguments, and assess revisions with a tool. This practice keeps the learner accountable for their objectives, evidence, evaluation, and expression. ChatGPT could help with outlining, revising content, proofreading, and reflection after writing, as described by Su et al. (2023); however, teachers need to explicitly mention the limitations. Likewise, Song and Song (2023) concluded that guided instructional design would facilitate AI-supported language learning in improving writing ability and motivation. The issue of product substitution and process support is key. Product substitution is when a student gives the main intellectual/linguistic task to the system. Process support is when the student maintains the authorship and uses the system to explain, give feedback, compare, or revise. The second model is more in line with the developmental objectives of English programs, which can make thinking and revising more visible. It also allows students to challenge AI output, examine weak evidence, and explain why a suggestion was accepted or rejected.

2.2.Generative AI in EFL and ESL Writing

The amount of research on ChatGPT in EFL/ESL education has grown at an unprecedented pace. In a recent literature review, Lo et al. (2024) identified 70 empirical studies published in the first year and a half following the release of ChatGPT, where they determined that writing was one of the most prevalent application areas. It is only natural that this is where they focus in this early stage of concentration, as written output is the most common medium used by LLMs, and tasks that involve writing enable the direct comparison of student drafts and AI feedback with revised texts. Empirical results are generally quite favorable and mixed. In the context of undergraduate ESL writing, Mahapatra (2024) investigated the potential of ChatGPT as a formative feedback system and found it to be beneficial for academic writing. Consistent with the results of Song and Song (2023), the results revealed that AI-supported instruction resulted in improvement in writing performance and motivation. In classrooms with a high ratio of students to teachers, this study suggests that in-person, immediate feedback can be useful for students who are not able to receive personalized feedback on each of their many drafts. However, the success of AI feedback hinges on how good the prompt is, how experienced the learner is, what type of task they are doing, and how well the student can assess the suggestions. Systematic reviews warn against positive attitudes serving as an indicator of enduring learning. While various benefits have been reported in the literature, including idea generation, language support, efficiency, and confidence, there are also concerns about plagiarism, overreliance, misinformation, and the lack of critical engagement (Imran and Almusharraf, 2023). The key research question is then whether or not students can become better writers with the help of ChatGPT. Answering that question will require a longitudinal design, process data, a delayed post-test, and a comparison of guided and unguided use.

2.3.Academic Integrity, Human Agency, and AI Literacy

The use of Generative AI in academic writing poses challenges to traditional notions of authorship and academic honesty. A student could use AI to correct some of the language acceptably, questionably paraphrase some of the text, or generate the entire text, which means that this does not require 100% AI usage. Ethical judgment involves considering the task instructions, disclosure, attribution, learning outcomes, and the extent of human control. UNESCO's guidance emphasizes a human-centred approach, which values agency, inclusion, privacy, transparency, and accountability (Miao & Holmes, 2023). AI literacy in writing should go beyond prompt construction. Students should be aware that fluent output might contain factual inaccuracies, poor reasoning, false references, bias and/or the wrong register. They ought to be able to identify sources, keep track of previous drafts, reveal whether they have received any allowed help, and state the reasons for their changes. Assessment designs for teachers should also focus on process evidence, oral defence, source verification, reflection, and discipline-specific judgement, rather than just a finished product. This would be better than adoption without restrictions or a ban for Pakistani Universities. It may be hard to enforce a blanket ban, and it may not allow students to learn responsible usage, and unchecked adoption can exacerbate dependency and inequity. Permissible use by task, disclosure (where appropriate), protection of student data, and not using AI to replace the learning outcome assessed are to be considered in institutional guidance (HEC, 2025).

2.4.Bibliometric Analysis and the Research Gap

Bibliometric analysis is a mix of two: performance analysis and science mapping. While publication and citation patterns are considered performance analysis, document and concept relationships among authors, sources, countries, documents, and concepts are described as science mapping. Research influence may be identified through citation analysis, collaboration through co-authorship analysis, conceptual structures through co-word analysis, research front and intellectual foundations through bibliographic coupling and co-citation analysis, respectively (Cobo et al., 2011; Zupic & Čater, 2015). Bibliometrix is an R environment for comprehensive science mapping, while Biblioshiny offers a graphical interface for numerous of its functions (Aria & Cuccurullo, 2017). VOSviewer has been developed to build and display bibliometric networks such as co-authorship, citation, bibliographic coupling, co-citation networks, and term co-occurrence maps (van Eck & Waltman, 2010). Together, they can enhance analysis: Biblioshiny can help perform large-scale descriptive and thematic analysis, whereas VOSviewer can offer detailed network and overlay visualizations. The current body of literature on AI and English writing consists of empirical studies and systematic reviews. Still, the provided dataset allows for a more holistic understanding of publication trends, the concentration of publications in a particular source, author productivity, collaboration, conceptual development, and citation patterns. The present study aims to fill this gap and relates the findings to the needs of BS English education in Pakistan.

3. Methodology

3.1.Research Design

The study was designed as a descriptive and science-mapping bibliometric study. Given the size of the research field and the fact that it had grown to a point that manual reading would not be sufficient, bibliometric analysis was chosen. The process used was in accordance with the guidelines of transparent data selection, performance analysis, network construction and theoretically meaningful interpretation (Donthu et al., 2021; Lim et al., 2024; Mukherjee et al., 2022). The workflow combined Scopus metadata outputs with outputs from Biblioshiny/bibliometrix and VOSviewer.

3.2.Database, Search Fields, and Eligibility Criteria

Scopus was used as the bibliographic database. A search was made on the title, abstract and keywords. The retrieved records were gradually narrowed down by year, topic, document type, language, and publication status and source type. The final inclusion criteria were: publication period from 2020 to 2026; classification in Social Sciences; journal article; English language; final publication stage; and the type of journal source. Reporting requirement: The exact Boolean search string, search date, format of the Scopus export and software version numbers were not provided in the analytical file supplied. The details must be provided in the final submitted manuscript as they are needed to reproduce the experiment. In this draft, there is no single query invented.

Table 1: Bibliometric screening and dataset refinement

Stage	Criterion	Records retained	Records excluded
Initial search	Titles, abstracts, and keywords	2,458	
Publication-year filter	2018-2027 retained initially	2,430	28
Final study-period filter	2020-2026	2,335	95
Subject-area filter	Social Sciences	1,517	818
Document-type filter	Articles only	1,175	342
Language filter	English only	1,159	16
Publication-stage filter	Final publications only	1,035	124
Source-type filter	Journals only	1,034	1

The final dataset contained 1,034 articles, representing 42.07% of the 2,458 initially identified records.

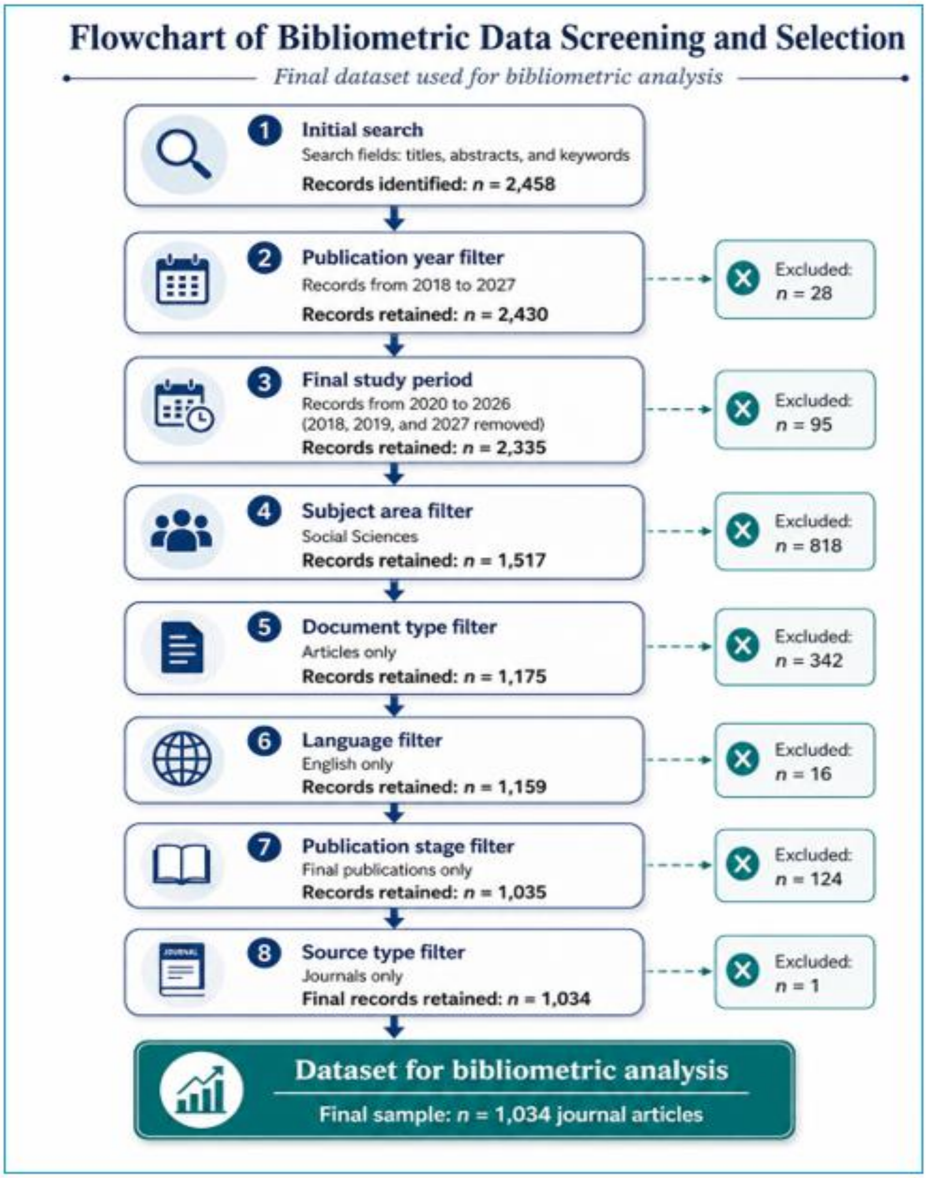

***Figure 1**: Flowchart of bibliometric data screening and selection.*

3.3.Screening Outcome

One hundred and forty-four records were removed from the screening process. The most significant drop was at the subject-area level. When the data set was limited to the Social Sciences, there were 818 fewer records, representing about 35.03% of the 2,335 records at that level. A total of 1,034 English-language, final-stage journal articles were used in the final dataset, which were published between 2020 and 2026.

3.4.Analytical Techniques

Performance was analyzed in terms of annual scientific article production, average number of citations per year, productive sources, productive authors, local citations, and local h-index indicators. The h-index is a measure that reflects the maximum number of papers h published by an author or source that has been cited at least h times (Hirsch, 2005). Recent publications were treated with caution, as citation counts build over time. Articles' concentration across sources was assessed using Bradford's Law. The Law predicts that a small number of highly productive journals will be surrounded by progressively larger numbers of journals generating similar numbers of articles (Bradford, 1934). Lotka's Law was applied to measure the phenomenon of disproportionate authorship in which several authors publish just one article while a few publish more than once (Lotka, 1926). The geographic distribution of research was explored through country collaboration mapping, as was the cross-national distribution in terms of links. Conceptual relationships were captured by keyword co-occurrence analysis, and a temporal aspect was added with overlay visualization. The thematic map was used to categorize clusters as motor, niche, basic and emerging or declining themes, based on their centrality and density. The Trend-topic analysis determined the time frames of the highest activity of big terms. The historical years with abnormally large numbers of cited references were identified using Reference Publication Year Spectroscopy (RPYS). Annual and cumulative publication output was also modeled using the logistic life-cycle model. The logistic results were considered to be conditional descriptive projections only, as only seven annual observations were available.

3.5.Keyword Network Parameters

Table 2: Parameters used for keyword co-occurrence analysis

Parameter	Value	Meaning
Field	ID	Keywords Plus were analysed
Number of terms	250	Maximum number of terms considered
Minimum frequency	5	A term had to occur at least five times
N-grams	1	Only individual words were analysed
Stemming	False	Related word forms were not merged
Size	0.3	Compact visual size of network elements
Number of labels	3	Three prominent labels were displayed
Community repulsion	0.5	Moderate separation between clusters
Repel	False	Automatic label repulsion was disabled
Clustering method	Louvain	Modularity-based community detection

The single-word setting and absence of stemming preserve indexed forms but may separate closely related expressions into different nodes.

3.6.Reliability, Transparency and Interpretation

This analysis examined multiple indicators, not simply a ranking. Each of these dimensions of a research field (productivity, citation impact, network position, thematic centrality, and temporal emergence) is not interchangeable. Database indexing and disambiguation might also apply to source and author names. No causal statements about student learning were made from bibliometric relationships, and figures and tables were interpreted within the analytical framework outlined above.

4. Results

4.1.Annual Scientific Production

Results 4.1 Annual Scientific Production The volume of publications increased significantly from the beginning to the end of the study period. The dataset contained one article in 2020, four in 2021, 13 in 2022, 41 in 2023, 190 in 2024, 431 in 2025, and 354 in 2026. The number of publications jumped from 59 at the end of 2023 to 680 at the end of 2025, indicating that the field accelerated its growth after the public release of generative AI. However, an apparent decrease in 2026 should not be viewed as an actual decrease, as the data may have been collected before the close of the calendar year.

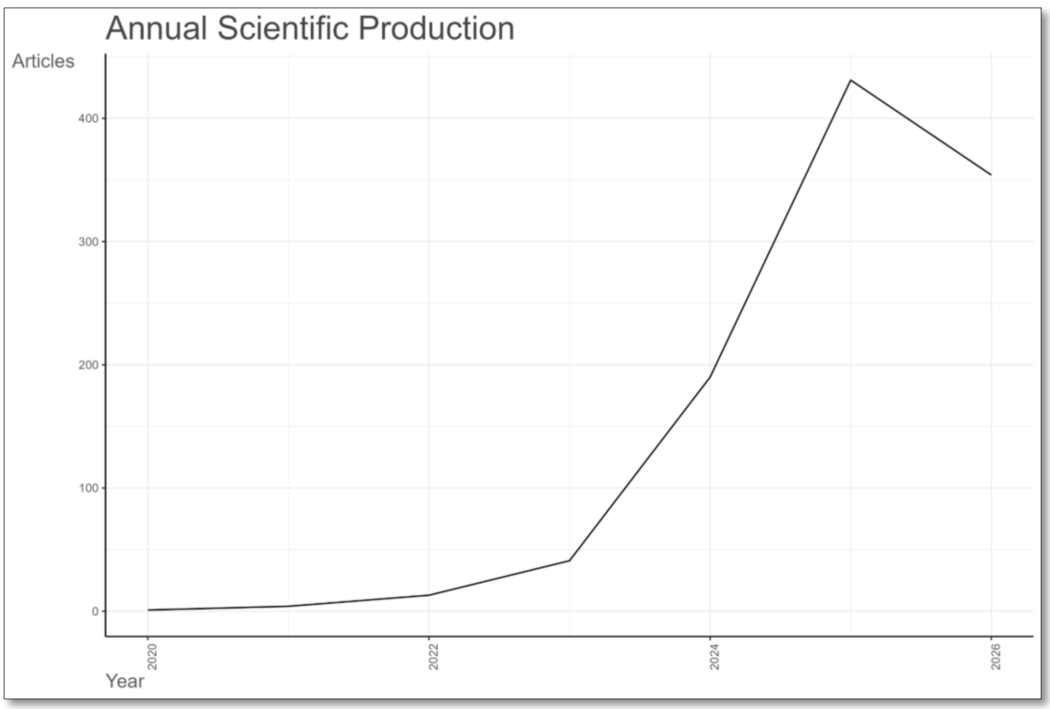

Figure 2*: Annual scientific production, 2020–2026.*

4.2.Average Citations by Publication Year

There was a large range in average citations across years of publication. The curve peaked in 2023 publications and decreased for the following year. The pattern is partially attributed to citation-window effects: the papers published in 2025 and 2026 are recent papers, while the older ones have more time to receive citations. It also implies that early publications on ChatGPT and generative AI were reference points for the subsequent development of the field in 2023.

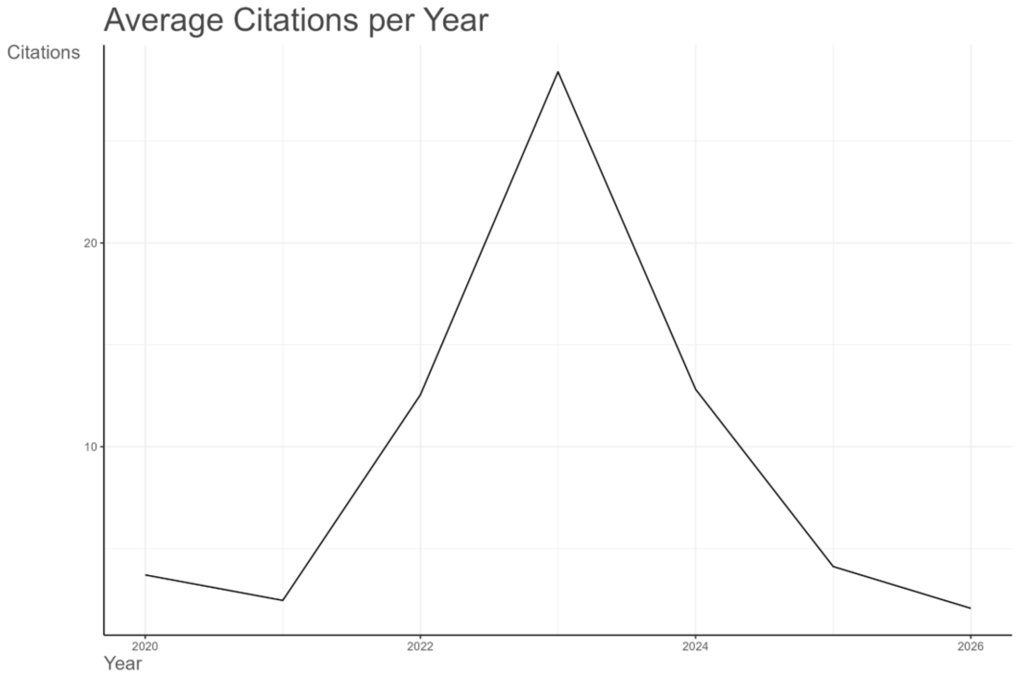

Figure 3: *Average citations per publication year.*

4.3.Most Relevant Sources

Table 3: Ten most productive publication sources

Source	Documents
Arab World English Journal	43
System	32
World Journal of English Language	29
International Journal of Learning, Teaching and Educational Research	25
Education and Information Technologies	24
Computers and Education: Artificial Intelligence	23
Assessing Writing	20
Journal of English for Academic Purposes	19
Forum for Linguistic Studies	15
Frontiers in Education	15

The source distribution indicates that the field is applied linguistics, language education, educational technology, writing assessment, and general education research. The document with the highest number of leads was Arab World English Journal, with 43 documents, followed by System, with 32, and World Journal of English Language, with 29. Language-focused and technology-focused journals indicate the interdisciplinary nature of writing research with the support of AI.

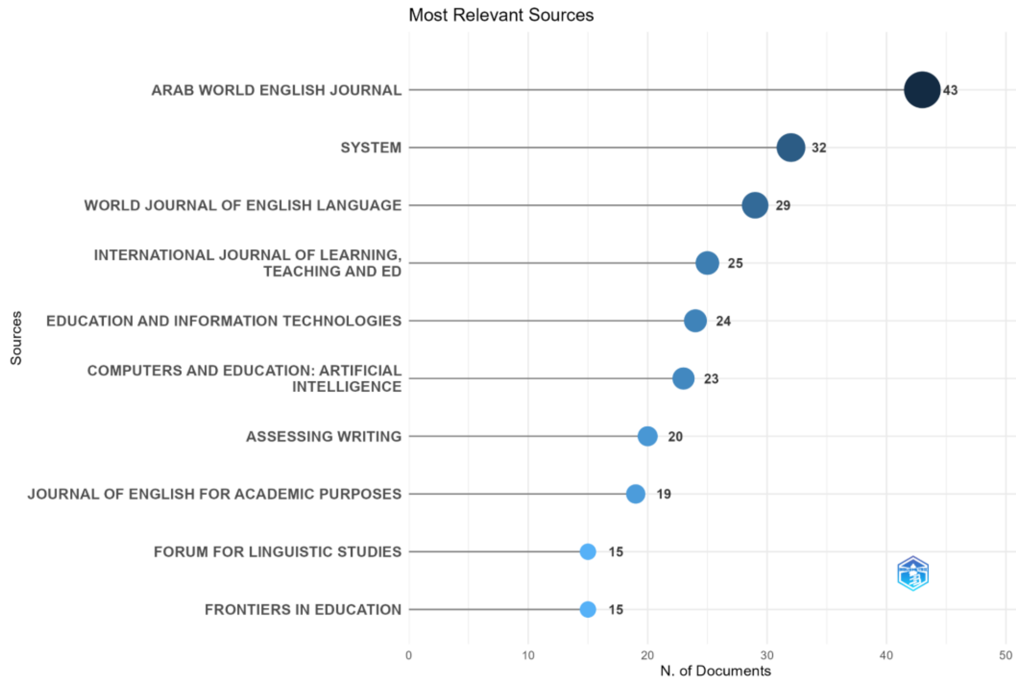

Figure 4: *Most relevant sources by number of documents.*

4.4.Most Relevant Authors

Table 4: Ten most productive authors

Author	Documents
Guo K	17
Wang Y	16
Teng MF	11
Mizumoto A	10
Huang J	8
Wang J	8
Zhang Y	8
Zou D	8
Basthomi Y	7
Li Y	7

Wang Y and Guo K were the two most prolific authors with 16 and 17 documents, respectively. Teng MF and Mizumoto A contributed 11 and 10 documents, respectively. The other top authors created 7-8 documents. The distribution shows a small number of regular contributors among a larger number of occasional contributors.

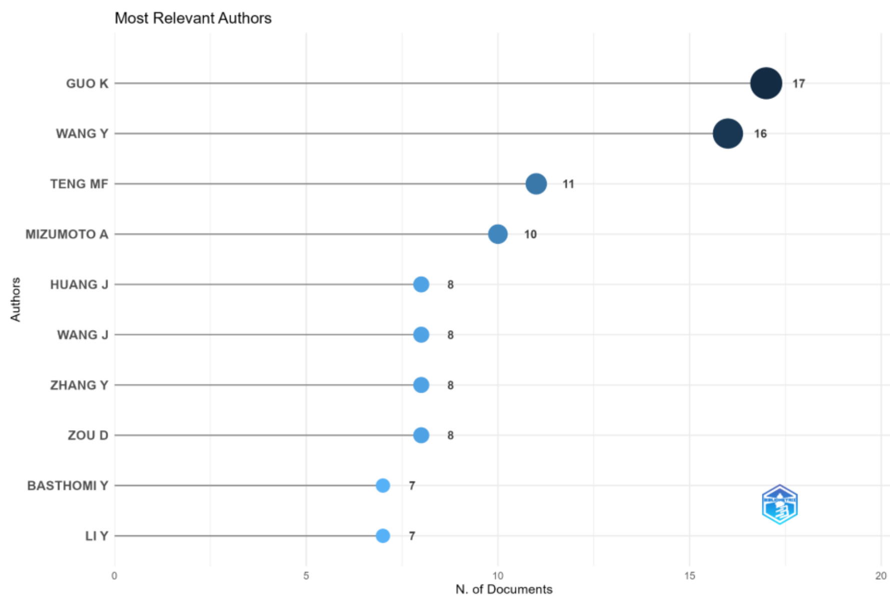

Figure 5: *Most relevant authors by number of documents.*

4.5.Most Locally Cited Documents

Table 5: Most locally cited documents in the dataset

Document as reported in the analysis	Local citations
Carless D (2026), International Journal of TESOL Studies	147
Marzuki (2023), Cogent Education	108
Mahapatra S (2024), Smart Learning Environments	78
Escalante J (2023), International Journal of Educational Technology in Higher Education	73
Guo K (2024), Education and Information Technologies	72
Gayed JM (2022), Computers & Education	70
Steiss J (2024), Learning and Instruction	68
Teng MF (2024), Computers & Education	62
Mizumoto A (2023), Research Methods in Applied Linguistics	57
Teng MF (2024), International Journal of TESOL Studies	42

Local citations refer to citations received from other documents within the analysed dataset. Full bibliographic details should be checked against the Scopus export before submission.

The locally cited-document ranking identifies works that function as internal reference points for the field. Carless D (2026) led with 147 local citations, followed by Marzuki (2023) with 108. The prominence of Mahapatra (2024), Guo (2024), Teng (2024), and other recent works illustrates how quickly a new research area can establish a concentrated internal canon.

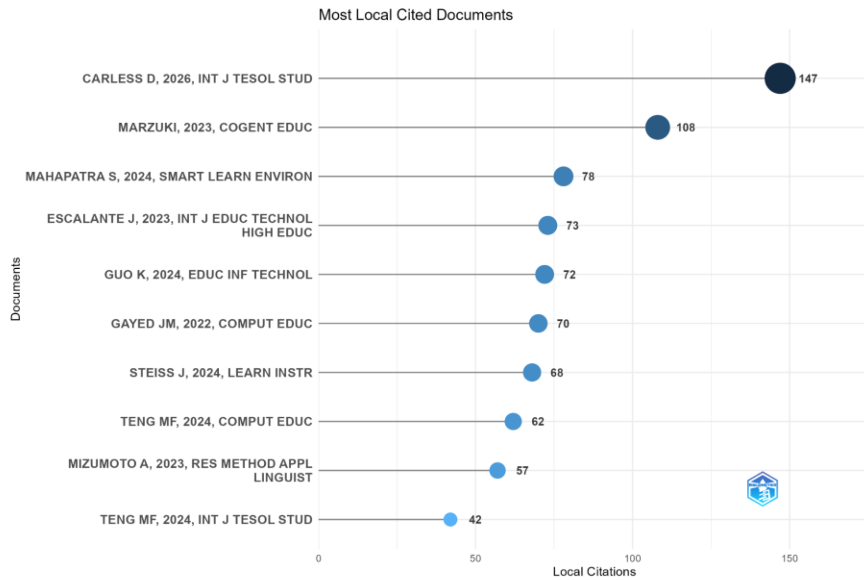

***Figure 6**: Most locally cited documents.*

4.6.Bradford's Law and Source Concentration

Table 6: Bradford's Law distribution of sources and articles

Zone	Sources	Sources (%)	Articles	Articles (%)
Zone 1	18	5.9	355	34.3
Zone 2	55	18.1	337	32.6
Zone 3	231	76.0	342	33.1
Total	304	100.0	1,034	100.0

The 304 publication sources were split into three zones, with roughly one-third of the articles in each. Zone 1 included just 18 sources (5.9% of all sources), but accounted for 355 articles (34.3%). Zone 2 contained 55 sources and 337 articles (32.6%). There were 231 sources (76.0% of all sources) in zone 3 that produced 342 articles (33.1% of all articles). This distribution broadly reflects Bradford's Law: a few journals concentrate the major proportion of the field's literature, with the rest of the journals, articles and publications spread out in a long tail.

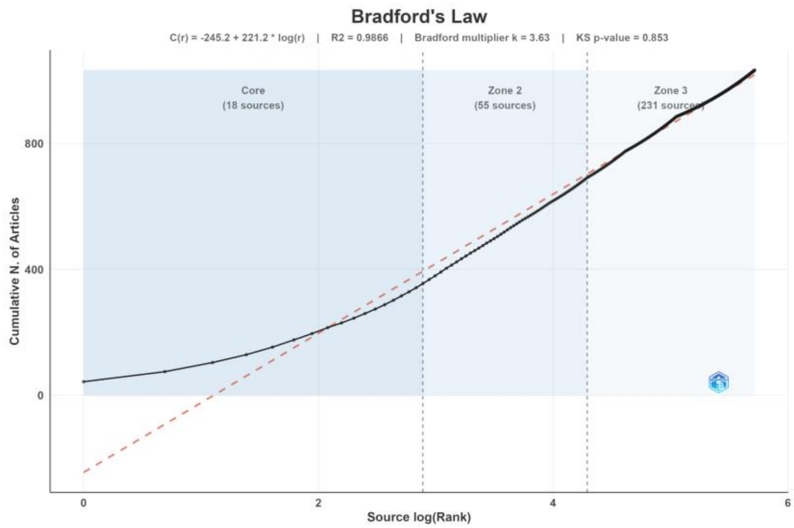

Figure 7: Bradford's Law and source concentration

4.7.Country Collaboration

The country-collaboration map shows an unevenly distributed network, with pockets of collaboration in various parts of the world. The major centres seem to be China and the United States, with some visible connections to Canada, Europe, the Middle East, South and Southeast Asia and Australia. The map indicates that the research activity is concentrated in a few countries that are very productive, while many other countries are involved in the research activity in smaller quantities in terms of publications or in collaboration with other countries. The network is not visually dominant in Pakistan, indicating potential for more local production and collaboration in English writing research through the lens of AI.

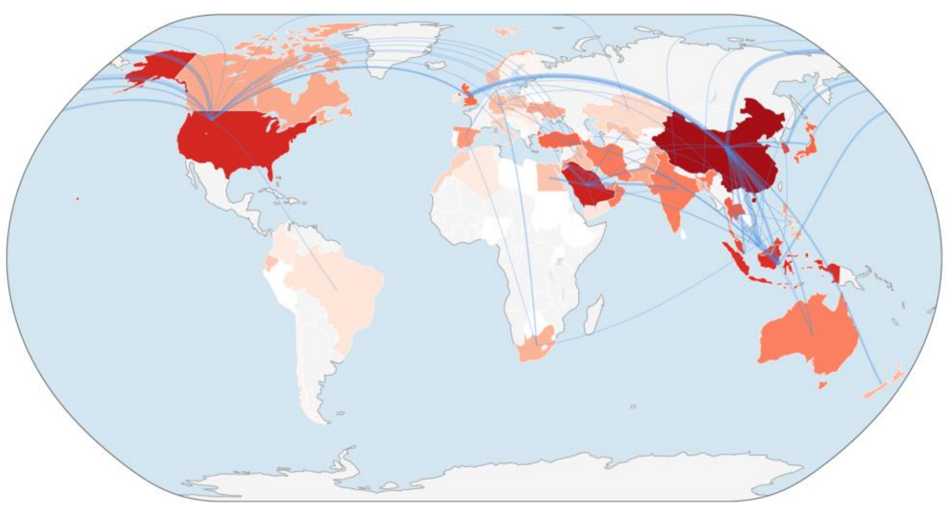

Figure 8: *Country collaboration map.*

4.8.Keyword Co-occurrence and Temporal Evolution

Table 7: Temporal distribution of major research terms

Term	Frequency	Year (Q1)	Median year	Year (Q3)
ChatGPT	257	2024	2025	2025
Artificial intelligence	223	2024	2025	2026
Generative AI	114	2025	2025	2026
Generative artificial intelligence	61	2025	2026	2026
EFL writing	59	2025	2026	2026
AI tools	21	2024	2024	2025
Chatbots	18	2024	2024	2025
Large language models	17	2025	2026	2026
English-as-a-foreign-language	12	2024	2024	2025

The most common term was ChatGPT with 257 mentions, followed by artificial intelligence with 223 mentions. The number of times Generative AI was mentioned was 114, and EFL writing was mentioned 59 times. The use of AI tools, chatbots, and English-as-a-foreign-language were focused earlier, primarily in 2024–2025. Generative AI, EFL writing, and large language models were focused on in 2025–2026, on the other hand. The temporal pattern shows a shift from general terms for AI applications to more specific studies on generative systems, LLMs, and second-language writing. The overlay network places ChatGPT and AI at the center of a vast network of nodes. Thematic regions are connected areas of academic writing, generative AI, higher education, automated feedback, language learning, writing assessment, and large language models. The colour gradient shows that more recent work tends to be more focused on generative AI and LLM-specific terms.

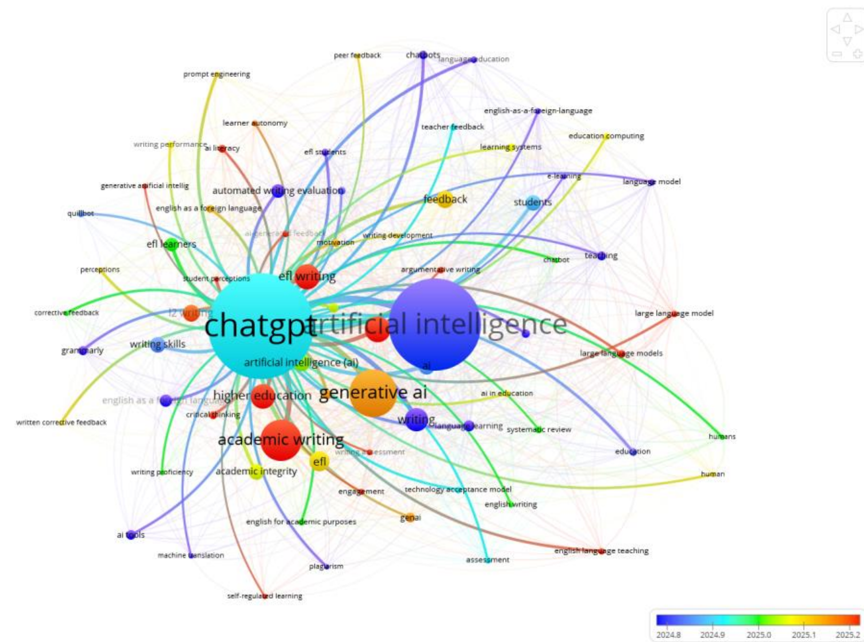

Figure 9*: Keyword co-occurrence and temporal overlay network.*

4.9.Thematic Structure

The thematic map differentiates clusters according to their centrality and density. The basic-theme quadrant includes the areas of artificial intelligence, EFL, and writing, which are highly relevant to the field but show relatively low internal development. The connectedness and development of ChatGPT, artificial intelligence, and L2 writing are close to the motor-theme region. Academic writing and higher education are in the middle, pointing to themes that are growing in importance as a result of generative AI. Writing proficiency, syntactic complexity, and activity theory are niche themes, meaning that they are internally developed but not as well integrated into the network as the other themes. Motivation, AI-generated feedback and the technology acceptance model are in the emerging or declining quadrant, where their future is unclear.

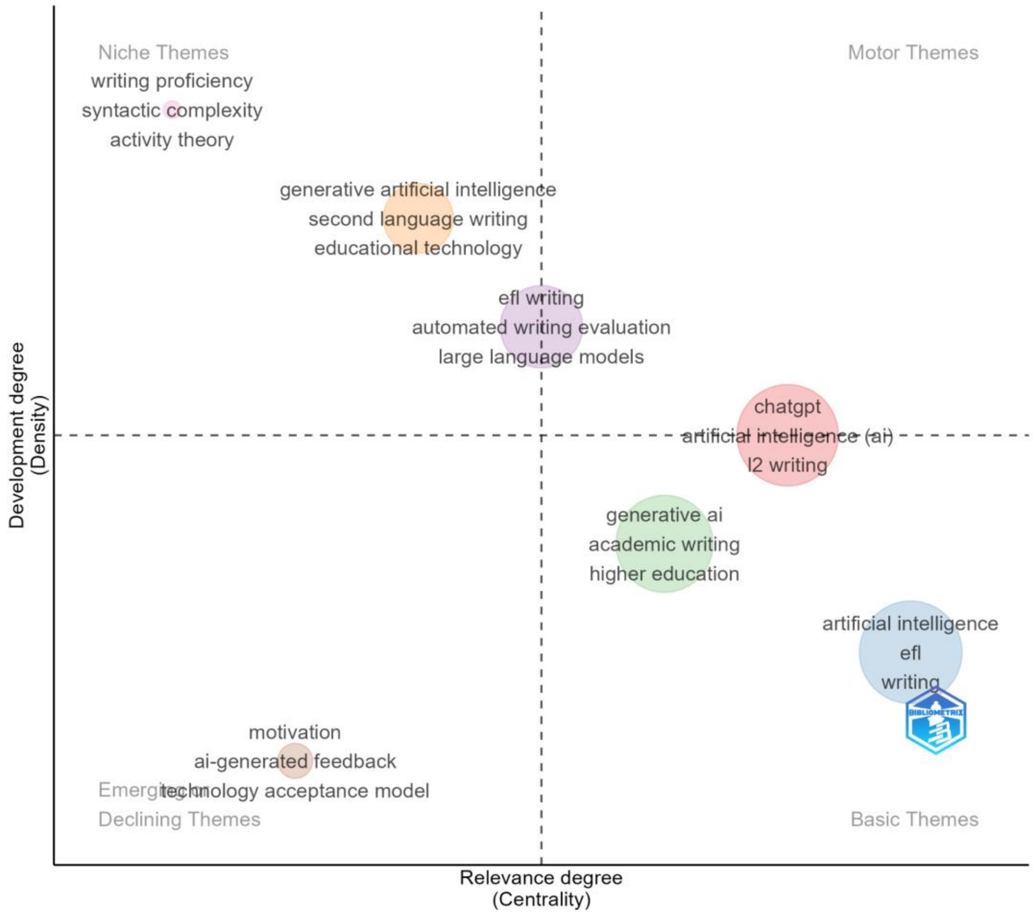

Figure 10: *Thematic map of the research field.*

4.10. Trend Topics

The trend-topic analysis confirms the temporal change that is seen in the keyword table. The previous stage of the field was linked with AI tools, chatbots, and English-as-a-foreign-language. Artificial Intelligence and ChatGPT and generative AI emerged in the spotlight around 2024-25. Generative artificial intelligence, EFL writing, and large language models are the most advanced of the current research fronts that will continue to expand in 2026. The stages show a gradual shift from generic technology integration to pedagogical design, feedback, writing performance and the integration of human and AI working together.

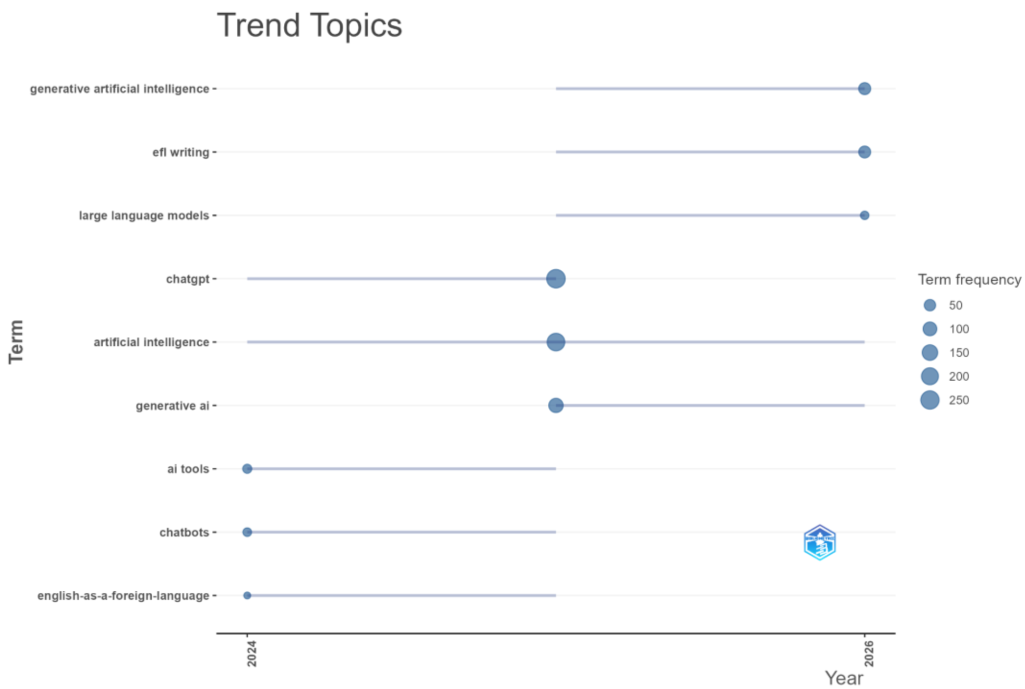

***Figure 11**: Trend topics by period and frequency*

4.11. The sources and author impact

Artificial Intelligence and Education and Information Technologies were the top sources, with local h-indexes of 15 each. System followed with 12 and Arab World English Journal with 10. The h-index of Language Testing in Asia was 9, and that of Language Learning & Technology was 8. The remaining displayed sources recorded local h-indices of 7. This ranking is not based on the number of publications, but it takes into account the impact of the publications.

4.12. Lotka's Law of Author Productivity

The Lotka curve shows (Supplementary Figure S4) a steep decline in the proportion of authors as the number of documents written increases. The majority of the contributors had only one paper published, and a very small number had more than one. The fitted exponent was around 2.77, with a high R^2 value and a non-significant goodness-of-fit test. The outcome reinforces a highly polarized productivity pattern, but must be viewed as an empirical approximation and not as a general rule.

4.13. Reference Publication Year Spectroscopy

The RPYS curve (supplementary Figure S5) shows that there are peaks in cited references for each publication year, and a recent increase in cited references. The earlier peaks represent intellectual antecedents that were cited repeatedly by the contemporary literature, and the buildup of the series near the end is due to the recent expansion of the field and the extensive use of the most recent publications. The black curve shows the number of citations that are counted, while the deviation curve emphasizes years that are above or below the local 5-year median.

4.14. Logistic Life-Cycle Model

Supplemental Tables S1 and S2 indicate that the total capacity of the documents could be around 1,189.66 documents, and the inflection point was estimated to occur around 2025.32. The average number of publications produced in a year was estimated at 459.50. The field grew from 10% to 90% of its estimated capacity over 2.84 years, which is a period of rapid growth. The fitted annual values closely matched the observed output, especially in 2024-2026, and the biggest difference was in 2023, with excess production of approximately seven publications. The following estimates should be viewed with great care. Only seven annual observations were used for the model, and the 2026 cumulative value was below the observed cumulative total, and 2026 may be missing. The apparent drop after 2025 may, therefore, be because only part of the year was covered, and not to any real saturation. The estimate of 2028.30 should be treated as a conditional model projection, rather than as an end-point of research.

5. Discussion

5.1.Rapid Growth and the Emergence of a Distinct Research Field

Firstly, the remarkable acceleration is research following 2022. Annual production increased from 13 articles in 2022 to 41 in 2023, 190 in 2024, and 431 in 2025. This path is in line with the educational disruption brought by easily accessible G.A.I. With ChatGPT, the technical barrier to interacting with a large language model has been lowered, creating a view of writing support for students and teachers from any discipline. The ensuing literature progressed rapidly from speculation and comment to empirical research, systematic review, instructional design and policy analysis. The citation pattern supports this growth pattern. Studies from 2023 saw the highest average citations, indicating that the early studies and conceptual papers provided pillars for later research. The lower averages in more recent years are mostly due to the shorter citation windows. So, direct comparison of recent productivity with older citation impact is not appropriate without normalization. The logistic model implies that the field is nearing a stage of transition from explosive growth to a mature stage. But the bibliometric data is not sufficient to make a strong case for saturation. New releases of models, institutional policies, assessment practices and changes in regulations can trigger additional publication waves. It is better to think of the model as a short description of the curve seen in the field in 2020-2026 than as a good prediction of the actual size of the field.

5.2.Interdisciplinary Source Structure and Concentrated Scholarly Influence

AI-powered writing research is not limited to applied linguistics, as shown by leading sources. It is published in journals in educational technology, language learning, writing assessment, academic purposes and general education. This interdisciplinary relation is fruitful because technical knowledge of AI, pedagogical knowledge of writing instruction, and linguistic knowledge of second-language development are needed in the field. Figure A shows that just 18 journals supplied over a third of the articles in the distribution of Bradford journals. This can be a useful point for researchers to begin their literature monitoring. It also signals to editors and reviewers that a relatively limited number of outlets influence methodological expectations and concept vocabulary. At the same time, it is clear from the widespread nature of the broad third zone that the topic has spread widely and could be understood differently across disciplinary communities. The productivity and h-index of authors also show some centralization. While a handful of authors have emerged as highly visible, the trend of Lotka suggests that the majority of authors have a single publication. This is typical of a very dynamic and interdisciplinary field that draws in other scholars from adjacent fields. There is a need for ongoing research programs rather than cross-sectional research or one-shot classroom interventions in the next phase of development.

5.3.Conceptual Transition: From AI Tools to Generative AI and LLM-Mediated Writing

A clear conceptual transition is identified from the keyword analysis. In previous research, general terms like AI tools, chatbots, and English-as-a-foreign-language were employed. Later work started to increasingly employ Chat GPT, generative AI, generative artificial intelligence, large language models, and EFL writing. This is more than a name change. It marks a move away from technologies that offer limited corrections or scripted interaction towards extended discourse technology that can produce and transform extended discourse. The thematic map lists the broad themes of artificial intelligence, EFL and writing as basic themes. Motor-theme is becoming a reality for ChatGPT and L2 writing, and generative AI, academic writing, and higher education are becoming more intertwined. Specialized directions in analysis are suggested by niche themes, including syntactic complexity and activity theory. New themes in the use of AI for feedback, motivation, and technology acceptance indicate that researchers are growing more interested in what students do with AI rather than just whether or not they do use it. A conceptual shift has been observed, which is in line with empirical literature. The research journey from using ChatGPT for brainstorming, revision and motivation has begun, and the research trajectory is turning to the challenges of the quality of feedback, critical evaluation, authorial voice and human–AI collaboration (Mahapatra, 2024; Song & Song, 2023; Su et al., 2023). The questions are more pedagogically useful, related to the learning processes, and not to novelty.

5.4.Geographic Concentration and the need for Contextual Diversity

The collaboration map shows high activity in China and the USA and significant connections throughout Europe, the Middle East, Asia and Australia. This distribution is based both on the strength of English language research and on the importance of the English language in the world. However, the visibility of the book is still not even. The contexts of countries that have a high EFL population can be very weak, if research is published in journals that are not indexed in Scopus, in local language, or in other publications that are not available in the journal format. The low visibility of Pakistan on the map does not mean that there is no use of AI by the students in Pakistan. Instead, it is a publication and evidence gap that is being highlighted. Multi-institutional studies are needed at the local level to investigate real writing development, variations between public and private institutions, between urban and regional access, teacher preparedness, between disciplines, and students' socioeconomic factors. Attitudes toward longitudinal evidence of independent writing should be addressed in Pakistan-based research.

5.5.Implications for BS English Students in Pakistani Universities:

Global evidence structure has direct but well-defined implications for BS English programmes. AI must be embedded in writing pedagogy in the context of clear learning outcomes. Pupils should be taught how to use AI to generate ideas, diagnose issues, explain concepts and revise texts, but be reminded that they are still responsible for evidence, reasoning and expression. Assignments may involve submitting prompts, AI-generated responses, revision notes, and reflective justification. Second, there should be a focus on process and oral accountability. Learning becomes more visible, and draft portfolios, in-class writing, source annotation, viva-style explanation, peer feedback, and reflective commentaries limit the value of undisclosed text generation. These approaches also enable teachers to differentiate between productive assistance and substitution. Thirdly, the integration of AI literacy must be a component of the BS English program and not just a workshop. Language courses can discuss grammar feedback and vocabulary choice; composition courses can discuss outlining and revision; literature courses can comment on AI interpretation and cultural bias; research-writing classes can comment on source verification, citation integrity, disclosure, and falsified sources. Fourth, faculty development is necessary. AI practical experience and knowledge of design, prompt assessment, privacy and ethics are required. Institutional policies must be clear and allow flexibility so they can differentiate between courses and assessments. While the HEC framework can offer a national direction, universities need to apply general principles to the course level and make rules (HEC, 2025). Last but not least, equity of access should be taken into account. Equal benefits can result from paid versions, internet quality, device availability, and proficiency in English. AI-based assessment models should not be designed assuming that students are all equally capable of accessing advanced AI systems. To integrate responsibly, institutional tools, data protection measures, and other means of accessing the systems must be available and accessible, as well as alternative paths for students who do not or cannot use commercial systems.

6. Contributions and Implications

The study views the use of AI tools in English writing as a dynamic discipline with three interrelated aspects: technological functionality, writing pedagogy, and learner agency. The shift from general AI tools to generative AI and LLMs in the bibliometric field demonstrates the swift pace of technology progress. The importance of the feedback theme, EFL writing, motivation and academic writing as themes indicates that the pedagogical and human dimensions are increasingly gaining significance. Theory needs to be further developed to better account for how learners regulate, evaluate, and transform AI contributions during the writing process in the future. Methodologically, the study uses descriptive performance indicators together with Bradford and Lotka distribution, mapping of collaboration, mapping of keywords and themes, trend analysis and

RPYS, and logistic modelling. This multi-method design helps to minimize the possibility of associating the number of publications with the significance of the work. It also shows how Biblioshiny and VOSviewer can be used in combination, which is also recommended for a synergistic bibliometric workflow (Lim et al., 2024). A global bibliometric map is converted into a context-sensitive agenda for BS English programmes in Pakistan. The principle of guided augmentation is that AI should support feedback, reflection, and revision but not supplant the student's authorship or the learning outcome. This principle underpins transparency of disclosure, process-based assessment, faculty training, data protection, and ongoing review of learning outcomes.

7. Conclusions, Recommendation, and Limitations

This study investigated the development of research on AI tools and English writing in higher education, drawing on 1,034 Scopus-indexed journal articles published between 2020 and 2026. The results indicate the rapid expansion of this research field, with annual scientific output increasing from 1 article in 2020 to 431 in 2025. The apparent drop to 354 publications in 2026 should be interpreted with caution, as the 2026 publication record may be incomplete. The field is interdisciplinary, with research appearing across applied linguistics, language education, writing assessment, and higher education. ChatGPT, AI, generative AI, and EFL writing conquered the conceptual structure of the literature. Temporal and thematic analyses further specified a shift from general research on AI tools and chatbots toward more specialized research on generative AI, large language models, automated feedback, academic writing, and EFL writing. The source and authorship analyses also displayed a concentrated research structure, while the country collaboration analysis specified uneven geographical participation. Bradford's and Lotka's studies further recognized concentration in a relatively small number of sources and authors. The findings have several implications for higher education, particularly for BS English programmes in Pakistan. AI literacy should be integrated into writing education through clearly defined learning outcomes and responsible-use guidelines. Students can be encouraged to use AI tools for activities such as brainstorming, language feedback, revision, and reflection while retaining responsibility for evidence, reasoning, authorship, and final decisions. Assessment should place greater focus on the writing process through drafts, reflective explanations, in-class writing, and other process-based evaluation methods. Faculty development is also essential for helping instructors understand AI capabilities and limitations, assessment design, privacy, and ethical use. At the institutional level, universities should establish transparent, periodically updated guidelines for responsible AI use while accounting for disparities in access to devices, internet connectivity, and paid AI services.

However, several limitations should be recognized. First, the study relies solely on the Scopus database; therefore, relevant publications indexed in other databases may be underrepresented. Second, the findings depend on the search strategy and inclusion criteria employed to construct the dataset. Third, author and source names may be inconsistent due to differences in database indexing and name disambiguation. Fourth, the 2026 publication data may be incomplete because the year had not ended at the time of data collection. Fifth, the logistic life-cycle model is based on only seven annual observations and should therefore be analyzed as a descriptive projection rather than evidence of research-field saturation. Finally, although the study draws implications for Pakistani BS English programmes, the underlying bibliometric dataset is global in scope. Therefore, these findings should not be interpreted as direct proof of AI use or writing outcomes among Pakistani students. Future research should conduct locally grounded empirical and longitudinal studies to examine how AI-supported writing practices affect students' writing development, learning processes, and academic integrity in Pakistani higher education.

References

[1] 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

[2] Bibi, Z., & Atta, A. (2024). The role of ChatGPT as AI English writing assistant: A study of students' perceptions, experiences, and satisfaction. *Annals of Human and Social Sciences*, *5*(1), 433–443. https://doi.org/10.35484/ahss.2024(5-I)39

[3] Bradford, S. C. (1976). Sources of information on specific subjects.

[4] Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, *62*(7), 1382–1402. https://doi.org/10.1002/asi.21525

[5] 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

[6] Higher Education Commission of Pakistan. (2025). *Draft framework on the ethical use of generative AI tools in higher education institutions.* HEC Pakistan.

[7] Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, *102*(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102

[8] Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, *15*(4), Article ep464. https://doi.org/10.30935/cedtech/13605

[9] Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

[10] Lim, W. M., Kumar, S., Donthu, N., Mukherjee, D., & Goh, K. W. (2024). How to combine and clean bibliometric data and use bibliometric tools synergistically: Guidelines using metaverse research. *Journal of Business Research*, *182*, 114760. https://doi.org/10.1016/j.jbusres.2024.114760

[11] Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, *11*, Article 50. https://doi.org/10.1186/s40561-024-00342-5

[12] Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, *16*(12), 317–323.

[13] Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, *11*, Article 9. https://doi.org/10.1186/s40561-024-00295-9

[14] Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research.* UNESCO. https://doi.org/10.54675/EWZM9535

[15] Mukherjee, D., Lim, W. M., Kumar, S., & Donthu, N. (2022). Guidelines for advancing theory and practice through bibliometric research. *Journal of Business Research*, *148*, 101–115. https://doi.org/10.1016/j.jbusres.2022.04.042

[16] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

[17] Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

[18] Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, *57*, 100752. https://doi.org/10.1016/j.asw.2023.100752

[19] van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

[20] Waltman, L., van Eck, N. J., & Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, *4*(4), 629–635. https://doi.org/10.1016/j.joi.2010.07.002

[21] Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, *18*(3), 429–472. https://doi.org/10.1177/1094428114562629

SUPPLEMENTARY TABLES AND FIGURES

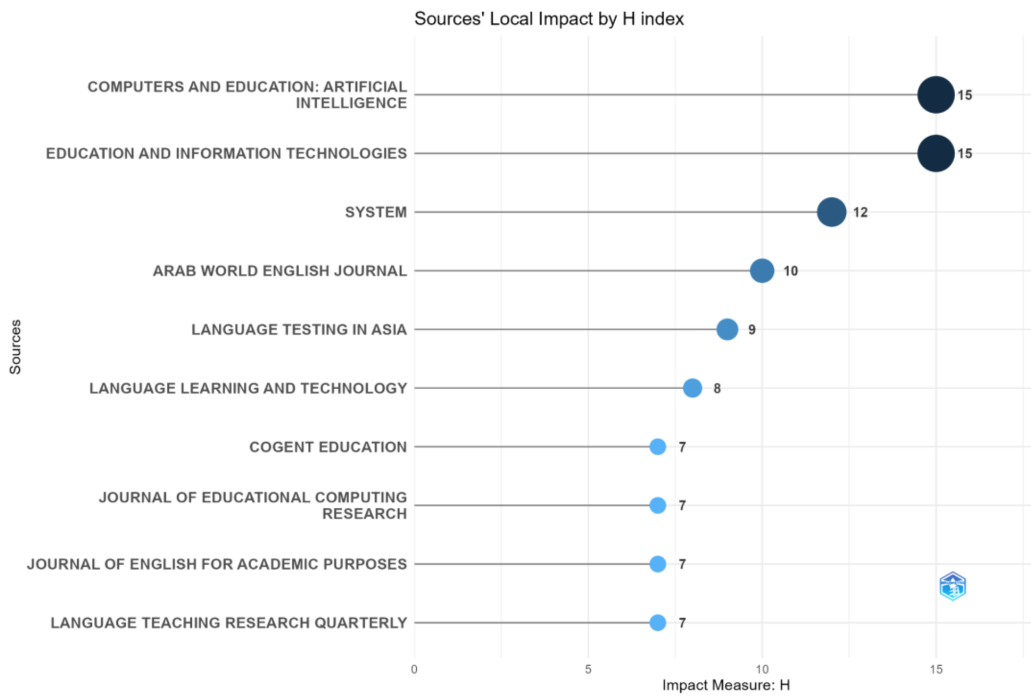

Figure S1: *Sources' local impact by h-index.*

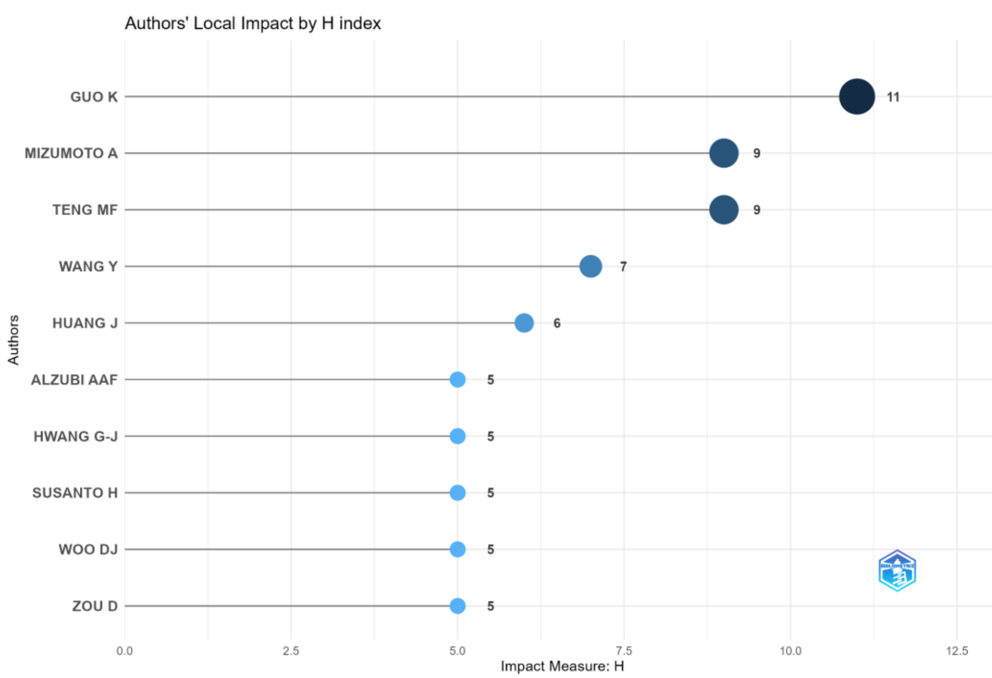

Figure S2: *Authors' local impact by h-index.*

At author level, Guo K had the highest local h-index (11), followed by Mizumoto A and Teng MF (9 each), Wang Y (7), and Huang J (6). Several authors recorded an h-index of 5. Author production over time shows that the most visible contributors became increasingly active during the rapid expansion from 2023 onward, with publication intensity concentrated in 2024–2026.

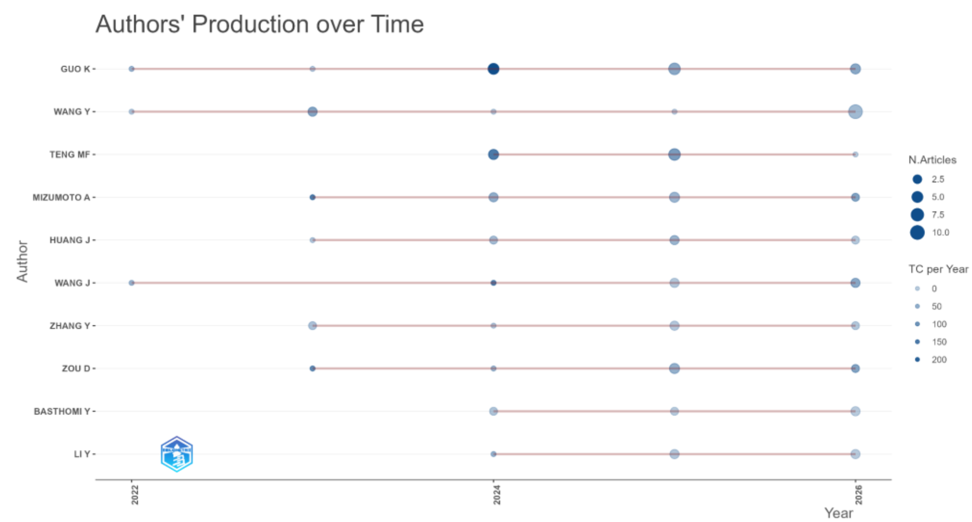

Figure S3: *Authors' production over time.*

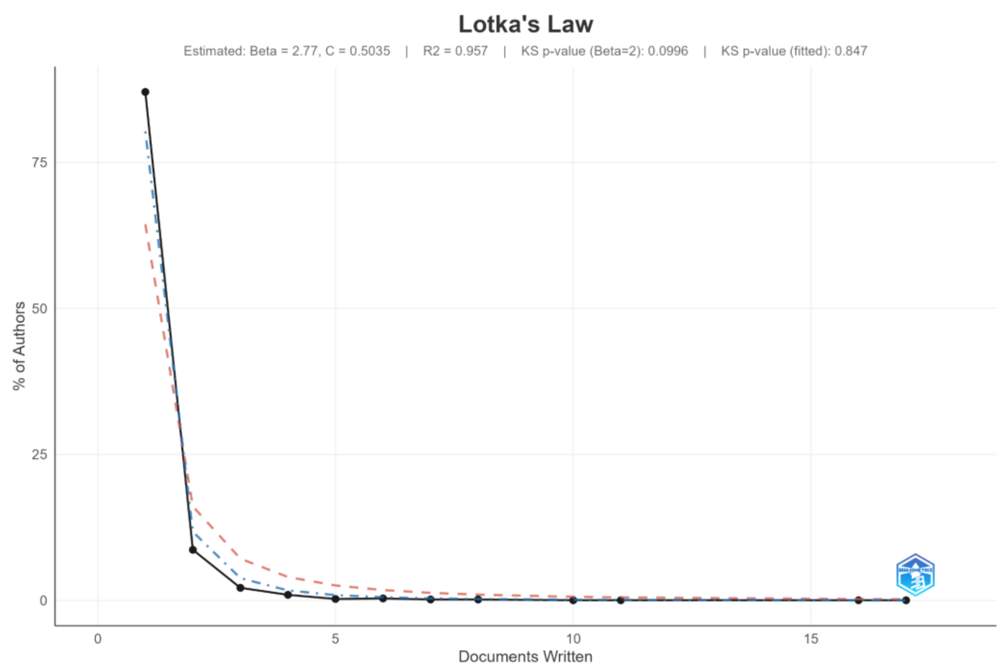

Figure S4: *Lotka's Law of author productivity.*

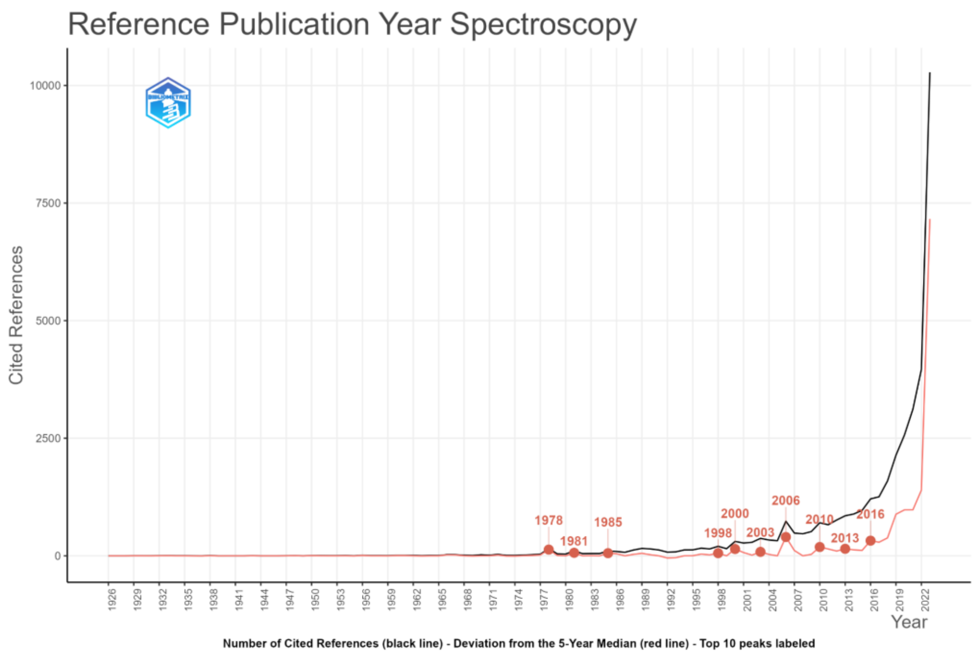

Figure S5: *Reference Publication Year Spectroscopy.*

Table S1: Logistic life-cycle model parameters and milestones

Metric	Estimated value
Estimated cumulative capacity (K)	1,189.66469 documents
Midpoint/inflection (tm)	6.3227328 (approximately year 2025.32)
Growth-period width (Δt)	2.8443573 years
Estimated peak annual production	459.499322 documents
10% level	2023.90055
90% level	2026.74491
99% level	2028.29698
Model R^2	0.99963
RMSE	3.19546 publications
Base year	2020

Table S2: Observed and fitted annual and cumulative scientific production

Year	Observed annual	Index	Fitted annual	Fitted cumulative	Observed cumulative	Residual
2020	1	1	0.49284	0.31908	1	0.50716
2021	4	2	2.30579	1.49432	5	1.69421
2022	13	3	10.70970	6.97284	18	2.29030
2023	41	4	48.10331	31.99594	59	−7.10331
2024	190	5	186.64073	136.45731	249	3.35928
2025	431	6	432.08277	449.53470	680	−1.08277

Year	Observed annual	Index	Fitted annual	Fitted cumulative	Observed cumulative	Residual
2026	354	7	353.56455	880.44102	1,034	0.43545