Generative Artificial Intelligence in Education From 2021 to 2025: A Scientometric Review

Junmin Guo[1], Enio Kang Mohd. Sufian Kang[1,*] & Norliza Ghazali[1]

[1]Faculty of Educational Studies, Universiti Putra Malaysia, Selangor, Malaysia

*Correspondence: Faculty of Educational Studies, Universiti Putra Malaysia, Selangor, Malaysia. Tel: 60-3-9769-8199. E-mail: enio@upm.edu.my

Received: November 4, 2025 Accepted: November 26, 2025 Online Published: December 31, 2025

doi:10.5430/jct.v15n1p102 URL: https://doi.org/10.5430/jct.v15n1p102

Abstract

This study presents a bibliometric analysis of 965 peer-reviewed articles on generative artificial intelligence (GenAI) in education published from 2021 to 2025 in the Web of Science Core Collection. Through keyword co-occurrence, co-citation, and collaboration network analyses, it identifies core research themes, intellectual structures, and developmental trends. Findings reveal an exponential rise in GenAI-related publications, with dominant themes centred on technological applications of GenAI in teaching and assessment—especially ChatGPT—alongside technology acceptance mechanisms and learner outcomes such as motivation and self-efficacy. Three major thematic clusters emerge: GenAI educational applications, user adoption theories, and learning impacts. Co-citation patterns show strong reliance on traditional acceptance models like TAM, indicating limited development of GenAI-specific theoretical frameworks. Collaboration analyses reveal fragmented author networks and uneven global participation, concentrated mainly in North America and East Asia. The study highlights research gaps, including ethical governance, creativity development, interdisciplinary applications, and insufficient qualitative or mixed-method studies. It recommends developing theoretical models tailored to GenAI's interactive and multimodal characteristics, strengthening ethical and cross-cultural frameworks, expanding interdisciplinary innovation, and enhancing global research cooperation to support the sustainable and responsible integration of GenAI in education.

Keywords: bibliometric analysis, generative artificial intelligence, education, Web of Science database, bibliometrix(R language)

1. Introduction

1.1 Introduce the Problem

The global education landscape is undergoing a profound transformation with the rapid development of digital technologies. Artificial Intelligence is playing a key role in this transformation, particularly Generative Artificial Intelligence (GenAI), a type of AI that uses machine learning to create new content such as text, images, music and video. It learns from existing data and then uses that knowledge to generate new outputs. GenAI is rapidly gaining popularity in education, impacting many aspects of teaching practice, learning resource development, and educational assessment. The significance of GenAI in education is now widely validated in academia. Gruenhagen et al. (2024) stated that GenAI can potentially serve as a powerful educational tool, offering a personalised, adaptive learning experience that complements traditional teaching methods. Khalid Alshahrani and Qureshi (2024) suggested that ChatGPT plays a multifaceted role in redefining educational landscapes, ranging from enhancing programming proficiency and fostering creativity in writing to augmenting student engagement. In addition, some scholars have mentioned that it can trigger behaviours, such as academic misconduct. Su and Yang (2023) also found that challenges such as the untested effectiveness of the technology, limitations in data quality, and ethical and safety concerns must also be considered. Against this backdrop, there is an urgent need for a systematic understanding of GenAI to gain an objective and comprehensive grasp of the current state of research, the developmental landscape, and future trends of GenAI in the educational sphere.

1.2 Explore the Importance of the Problem

The explosive growth of GenAI has sparked an urgent need to systematically examine its impacts on education.

While GenAI tools have become widely used to support teaching, automate assessments, and create educational content, the field still lacks a cohesive and well-organised knowledge framework. It is therefore critical to map out the intellectual landscape of GenAI research in education. To make things more complex, research on GenAI in education is expanding at an unprecedented pace. This makes it all the more important to identify the core hotspots, thematic clusters, and emerging frontiers that are shaping scholarly discussions today. Furthermore, as educational institutions and governments around the world work to develop guidelines for responsible and sustainable AI integration, there is a growing demand for insights based on evidence to guide decision-making. A comprehensive mapping of this field is thus essential: it can provide useful references for future research and policy development in educational technology.

1.3 Objectives of the Study

This study presents an overview of research on GenAI in education by reviewing 965 papers from 2021 to 2025. It aims to clarify the knowledge structure of the field, understand major trends in publications, and identify key research horizons and notable research fronts. Through keyword co-occurrence, bibliometric co-citation analysis, and collaboration network mapping, the study investigates the main thematic clusters and intellectual frameworks that constitute contemporary GenAI research. In addition, it analyses author, institutional, and country collaborations to provide an organisation of research communities. Overall, this study provides insights that can strengthen future research and support policymakers in educational technology.

2. Literature Review

2.1 Overview of GenAI

GenAI refers to technologies that create new content by interpreting patterns in raw materials such as text, images, audio, and video. These technologies utilise advanced machine learning techniques, including deep learning and neural networks, to generate outputs that reflect human imagination and intelligence (Hughes, Zhu, & Bednarz, 2021; Uzun, 2023; Yu & Guo, 2023; Alier et al., 2024; Andrade-Girón et al., 2024). Within this category, large language models (LLMs), such as GPT and BERT, are a major subset of GenAI. They play an especially significant role because of their capacity to process and generate human language. Not only that, LLMs trained with large datasets can perform tasks like text completion, summarisation, translation and Q&A with remarkable accuracy. They are thereby absolutely essential tools in educational contexts, where effective communication is central.

GenAI can be grouped into unimodal and multimodal models. Unimodal models generate content in a single modality, such as text-only or image-only outputs. Multimodal models combine multiple perceptual modalities, such as text and images, to build interactive and context-specific learning materials (Cao et al., 2025). The models enhance comprehension and interaction by fulfilling diverse learning styles and demands (Gimpel et al., 2023).

2.2 Current Status of GenAI in Education

2.2.1 Applications for Teaching and Learning

GenAI is becoming more and more integrated into students' everyday learning activities, including producing scholarly publications, translating articles, responding to challenging queries, and offering grammar and composition correction (Chen et al., 2024; Maphoto et al., 2024; Wang & Ren, 2024; Tan et al., 2024). Through individualised explanations and example-based reasoning, GenAI assists students in text-intensive topics in comprehending abstract concepts and overcoming learning obstacles (Jiang & Jiang, 2024; Sirnoorkar et al., 2025; Wan & Chen, 2024). In order to enhance inclusive learning for multilingual learners, its cross-linguistic capabilities have also been included in language and cross-cultural instruction (Creely, 2024).

In order to streamline assessment and feedback, educators utilise generative AI tools for lesson planning, instructional design, resource building, and the generation of customised learning trajectories (Kim & Koo, 2024). These uses show that GenAI is developing into a dual-purpose educational facilitator that helps teachers make decisions about their lessons while also promoting students' acquisition of knowledge.

2.2.2 Career Advancement

Teachers' professional practices are being impacted by GenAI. Teachers report increased productivity and less time pressure when they utilise it to create instructional materials, assessment tasks, and better idea development (Liu et al., 2025; Karpouzis et al., 2024). GenAI has shown promise in specialised fields like medicine, nursing, engineering, mathematics, and language instruction in addition to general education (Chen et al., 2024; Maphoto et al., 2024; Wang & Ren, 2024; Tan et al., 2024). For instance, GenAI facilitates the creation of case simulations, the automatic creation of exam items, and communication practice in nursing and clinical education, all of which enhance students' professional competency (Hale et al., 2024).

Notwithstanding these advantages, concerns remain about dependability, openness, and incorporation into customary procedures. Teachers need new skills and training to use AI responsibly as GenAI becomes more integrated into educational workflows. According to studies, professional development and AI literacy are essential for enhancing instructors' readiness and minimising potential abuse (Tan, 2024).

2.2.3 Academic Integrity and Trust

Academic integrity has become a major concern in academia due to the quick adoption of GenAI. When completing assignments and academic writing, students might rely more on AI technologies, which would reduce originality and raise the possibility of plagiarism (Creely, 2024). Furthermore, there are notable disparities in how teachers view and embrace GenAI, and this "technological trust gap" has emerged as a key obstacle to its widespread adoption (Zhang et al., 2025).

2.3 Review of Relevant Econometric Analysis Studies

In the current research boom around the integration of GenAI and education, literature review studies have principally adopted a series of mainstream analytical tools and methodologies. A review of 33 literature review papers in the Web of Science Core Collection revealed that methodologically, systematic reviews are the most prevalent method, followed by bibliometric analysis and meta-analysis. In addition, qualitative analysis methods such as thematic analysis, content analysis, or typology frameworks (e.g., the UTAUT model) are also included in these studies. Of these, Jensen et al. (2025) employed systematic searching coupled with manual coding to aggregate 10 typical discourses of ChatGPT in higher education in its initial launch phase. Whereas Deng et al. (2025) extended further in employing meta-analysis methods to quantify the ChatGPT intervention impact on learners' learning outcomes, employing tools such as the JBI quality assessment and effect size calculation.

At the tool level, existing studies have suggested visualisation and bibliometric tools, with VOSviewer, CiteSpace and Bibliometrix being the most common tools to construct and visualise literature co-occurrence networks, including keyword co-occurrence analysis, research hotspot detection and co-authorship networks (Chang et al., 2024). Qualitative analysis tools such as NVivo were not explicitly stated but inferred to be used in coding and thematic categorisation procedures. In addition, open-source programming languages such as R and Python are being utilised to manage advanced data analysis and visualisation (for example, the pymeta, bibliometrix, and Gensim packages).

In terms of metrics, commonly used indicators are categorised into six kinds. First are literature features. These are referred to as research trends and international contributions, i.e., year distribution, research fields, journal sources, and co-authorship networks. Secondly are keywords. These indicate hotspots and evolutionary patterns, i.e., keyword co-occurrence and burst detection analysis. Thirdly is influence. These indicate the influence of literature, authors, or institutions. This includes citation frequency, oft-cited literature, H-index, etc. Fourth, structure metrics. These build research structures and future direction knowledge graphs. This includes clustering (Cluster), path analysis (Main path), etc. Fifth, methodology metrics. These indicate method distribution and reliability tests. This includes research method proportions (qualitative, quantitative, mixed), sampling approaches, etc. Sixth, thematic metrics. These indicate content variation and future trends. Particular categories include topic development, research paradigms, and classification of fields of educational use. Furthermore, some studies apply the SWOT analysis approach to analysis and evaluation, for example, ChatGPT's potential application in K-12 education (Zhang & Tur, 2024).

Briefly, recent review studies of the pedagogical use of GenAI indicate tendencies of methodological variety, tool datafication, and indicator systematisation. Not only are these instruments and methods helpful in revealing gaps and trends in research, but they also provide a well-grounded theoretical and technical foundation for subsequent empirical research.

2.4 Research Gaps

Even while research on GenAI in education is developing quickly, there is still a dearth of pertinent literature. The majority of studies focus on short-term applications in teaching and learning, such as writing assistance and automated feedback, while the long-term implications for future professional development and practice receive scant attention. Furthermore, current reviews usually rely on smaller datasets or shorter time periods, even though they use a variety of approaches and tools to depict early trends. Given the rapid evolution of GenAI, larger-scale, more comprehensive analyses are required to understand research themes, author networks, and citation patterns.

3. Method

3.1 Research Design

Bibliometric analysis is a research analytical technique used with the aim to quantify and map the structure, evolution, and trends of a given scientific discipline based on the analysis of many scientific publications. This study design was utilised for this study with the aim to map influential authors, institutions, nations, and emerging topics in order to acquire an overall understanding of the scholarly productivity, collaboration network, and influence of generative AI in education. It also facilitates systematic determination of knowledge gaps, keeps track of topic development under scrutiny, and informs subsequent studies. This study adopts an observational and descriptive design. No experimental manipulation was performed. Scientometric indicators are used to interpret structural patterns rather than infer causality.

3.2 Data Source and Search Strategy

A literature search was conducted using the Web of Science Core Collection to discover papers on GenAI in educational contexts. The search was conducted on May 30, 2025. And the search approach was as follows: Set the search terms to 'generative ai', 'chatgpt', 'aigc', 'education', and more; define the publication time span from 2021 to 2025; and utilise topic fields like titles, author keywords, and abstracts. The terms employed by researchers were combined using Boolean operators (AND & OR) to form generic terms encompassing the research subject. The search keywords were further refined by identifying additional relevant vocabulary and retrieving synonyms from other papers and lexical databases. To ensure the scientific merit of the dataset while guaranteeing all data underwent peer review, this dataset exclusively comprises journal articles. Other publication sources (such as books, conference papers, editorials, white papers, etc.) were excluded. Subsequently, the search results were exported as a CSV file for subsequent processing. See Table 1 for details.

Table 1. Data Source and Search Strategy

Category	Items	Specific Condition
Topic	Ts1	"GenAI" OR "generative AI" OR "generative artificial intelligence" OR "ChatGPT" OR "AIGC" OR "large language model"
	Ts2	"education" OR "teaching" OR "learning"
Inclusion criteria	Publication type	Peer-reviewed journal articles
	Time span	2021–2025
	Language	English
	Research area	Education and educational research
Exclusion criteria	Type	Conference proceedings, book chapters, editorial materials, letters
	Content	Papers unrelated to educational application or conceptualization of GenAI

3.3 Data Cleaning and Preprocessing

To obtain correct and accurate information for the dataset, this study undertook data cleansing prior to analysis using the Bibliometrix (R language) tool. Duplicates were identified via Web of Science access codes and DOIs and subsequently removed manually. Author names and journal titles underwent standardisation to eliminate inconsistencies; all incomplete references were excluded, while duplicate citations and works with multiple versions were merged. Keyword fields were harmonised by merging synonyms (e.g., 'GenAI', 'Generative AI', 'Generative Artificial Intelligence'), standardising ChatGPT-related keywords, and unifying punctuation and capitalisation. To minimise morphological variation, all noun forms and multi-word phrases were standardised according to structured lexicon rules. To enhance clustering stability and optimise co-occurrence analysis visualisation, a minimum occurrence threshold of five instances was established for co-occurring terms. Collectively, these measures ensure the dataset's consistency, completeness, and methodological reliability for subsequent bibliometric analysis.

3.4 Sampling Procedures

The sample in this study consists of all relevant articles indexed in WoS that meet the inclusion criteria. Since no human subjects were involved, demographic characteristics, sampling frames, and ethical approvals were not applicable.

3.5 Measures and Covariates

As scientometric studies rely on bibliographic metadata rather than human participants, the primary measures in this study include keyword frequencies, citation counts, co-citation and bibliographic coupling, co-authorship relationships, source impact (journals), and cluster modularity and density. No covariates or psychometric measures were used.

3.6 Validity, Reliability, and Limitations of Method

This study took several steps to enhance methodological reliability. To improve reliability, this study included only peer-reviewed journal articles, performed rigorous data cleaning and normalisation, used multiple analytic tools for triangulation, and reported clustering parameters transparently. This study has several limitations, including reliance on a single database (WoS), potential bias toward English publications, and the inability of scientometric analyses to assess the quality of argumentation within individual papers.

4. Results and Discussion

4.1. Main Information Analysis

The data for this study was extracted on May 30, 2025. The bibliometric analysis of generative AI research in education covers the period from 2021 to 2025 (Table 2). After keyword integration and restricting the source type to journal articles and the subject area to education, a total of 965 articles were identified. The average age of the documents is 0.557 years. The average number of citations per document is 11.75, and the cumulative number of references is 38,240. Regarding publication types, 642 articles are regular academic papers, accounting for 66.6%, and 288 articles are early access papers, accounting for 30.9%. Additionally, there are 26 review articles and 7 early access reviews, which together represent 3.4% of the total.

The data indicate that the field has experienced explosive growth, with an annual growth rate of 371.45%, suggesting that generative AI applications in education have emerged as a cutting-edge topic of high academic interest. The very low average document age reflects rapid updating of research findings and a highly dynamic developmental trajectory of the field. The relatively high average citation count and cumulative number of references demonstrate the strong academic influence and knowledge accumulation in this area. The high proportion of early access publications suggests an urgent need for rapid dissemination of research results and also reflects the timeliness associated with the iterative development of generative AI technologies and their educational applications. The comparatively small share of review articles indicates that current research remains dominated by original practical exploration and that systematic review work has yet to be strengthened, which provides direction for future integrative research.

Table 2. Main Information of the Dataset (2021 – 2025)

Metric	Value
Timespan	2021–2025
Sources	127
Documents	964
Annual Growth Rate (%)	3.7145
Authors	2569
Authors of Single-Authored Documents	135
International Co-Authorship (%)	0.221
Co-Authors per Document	3.3
Author's Keywords (DE)	2468
References	38240
Document Average Age	0.557
Average Citations per Document	11.75

4.2 Keyword Co-Occurrence Analysis

4.2.1 Network Structure and Cluster Distribution

The GenAI keyword co-occurrence network in education contains 50 nodes (only the top 10 nodes are displayed in Table 3 to improve clarity and readability). Using the Louvain community detection algorithm, three major thematic clusters were identified (Table 4). Network centrality measures show that ChatGPT has the highest betweenness centrality (419.592), followed by “artificial intelligence” (125.512) and “generative AI” (73.098). Average closeness centrality in the network is 0.014, and the average PageRank score is 0.016.

The central positioning of ChatGPT indicates its role as a bridge between technical terminology (e.g., “artificial intelligence”) and educational contexts (e.g., “higher education”). The relatively low closeness and PageRank averages reflect a core–periphery structure, suggesting that a limited number of high-centrality keywords dominate knowledge exchange and thematic formation.

Table 3. Top 10 Keywords by Betweenness Centrality in the Co-Occurrence Network

Rank	Keyword	Betweenness	Closeness	PageRank	Cluster
1	ChatGPT	419.592	0.020	0.140	1
2	Artificial intelligence	125.512	0.020	0.090	1
3	Generative AI	73.098	0.019	0.062	1
4	Higher education	28.773	0.017	0.047	1
5	Education	20.324	0.015	0.040	1
6	Students	21.280	0.016	0.038	3
7	Technology	16.053	0.015	0.030	3
8	Motivation	2.049	0.013	0.016	3
9	Acceptance	3.757	0.013	0.022	2
10	User acceptance	2.248	0.013	0.020	2

Table 4. Thematic Clusters in the Keyword Co-Occurrence Network

Cluster	Node Count	Representative Themes
1	33	GenAI technology and educational applications
2	7	Technology acceptance model and user adoption mechanisms
3	10	Educational impact subjects and learning outcomes

4.2.2 Core Cluster and Thematic Analysis

Cluster 1: GenAI technology and educational applications

This cluster links the keywords “artificial intelligence,” “generative AI,” and “higher education,” with ChatGPT as the central connector (betweenness = 419.592; PageRank = 0.140). Co-occurrences emerge around academic writing, assessment, and educational integration of large language models.

Cluster 1 represents the core of the research landscape, emphasising how GenAI tools—especially ChatGPT—are embedded in instructional practices. The appearance of peripheral keywords such as “academic integrity” indicates limited exploration of plagiarism detection, ethical norms, and governance mechanisms in current research.

Cluster 2: Technology acceptance model and adoption mechanisms

Cluster 2 is centred around “acceptance” (3.757), which links to “information technology” and “user acceptance.” Within this cluster, “technology acceptance model” appears with low centrality.

The cluster reflects a theoretical chain connecting technology characteristics, user cognition, and adoption behaviour. However, weak connections with teacher-related terms show insufficient validation of acceptance frameworks for educators and limited adaptation of traditional theories to generative, dialogic AI technologies.

Cluster 3: Educational impact and learning outcomes

The core node of Cluster 3 is “students” (21.280), linking to psychology-related keywords such as “motivation” (2.049) and “self-efficacy” (2.188). Keywords related to higher-order cognition—“creativity” and “critical thinking"—show very low centrality.

This cluster represents the learner-centred dimension of GenAI research. The marginal presence of creativity and critical thinking indicates a clear evidence gap regarding how GenAI contributes to higher-order cognitive development, highlighting a misalignment between technology adoption and educational learning goals.

4.2.3 Cross-cluster Relationships and Research Gaps

Strong connections exist between ChatGPT in Cluster 1 and "students" in Cluster 3. Meanwhile, no direct linkage is observed between the "technology acceptance model" in Cluster 2 and "higher education" in Cluster 1. Keyword visualisation (Figure 1) places "ChatGPT," "artificial intelligence," and "higher education" at the centre, while "ethics" and "creativity" remain peripheral.

The network structure shows that GenAI applications are predominantly directed toward immediate student-orientated use, rather than institutional or systemic contexts. The absence of connections between acceptance theory keywords and real educational environments implies a disconnect between theoretical modelling and authentic teaching practice. Peripheral positioning of ethics and creativity reflects a lack of systematic research on technological risk, academic integrity, and pedagogical strategies for creativity enhancement, suggesting that future work must tackle ethical governance and align AI use with long-term educational objectives.

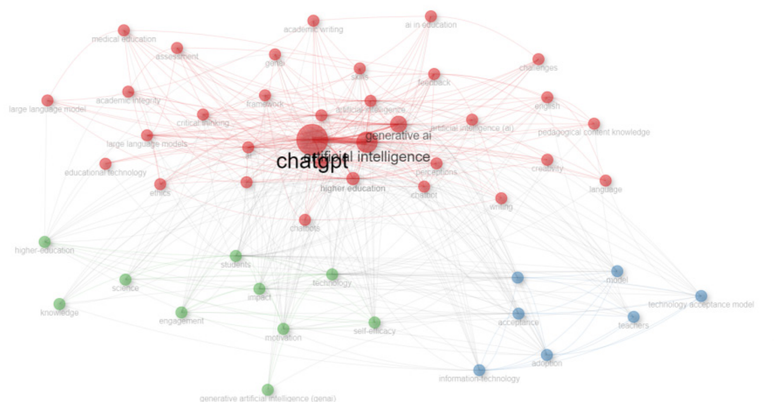

Figure 1. Co-occurrence Network

4.2.4 Implications for Research Trends

The keyword network reveals three dominant research paradigms: Technical applications of GenAI in education (Cluster 1). User acceptance and adoption theories (Cluster 2). Learning effects and psychological outcomes (Cluster 3).

Future research should: Develop ethical frameworks for GenAI use to ensure transparency, privacy, and academic integrity. Investigate how GenAI can support critical thinking and creativity, reducing the gap between technological practices and learning goals. Explore cross-disciplinary innovations, such as in medical and scientific education, where GenAI shows high potential but low current centrality.

4.3 Literature Co-Citation Analysis

4.3.1 Network Structure and Cluster Distribution

The GenAI education co-citation network consists of 50 author–year nodes (only the top 15 nodes are displayed in Table 5 to improve clarity and readability). Using the Louvain algorithm, three major clusters are identified (Table 6). The centrality metrics show the following characteristics: Strzelecki A (2024) has the highest betweenness (102.734). Davis FD (1989) ranks second in betweenness (75.313). Kohnke L (2023) also has comparatively high betweenness (68.076). Low betweenness values (12.35) and PageRank scores (0.018) are observed for most nodes.

The co-citation structure displays a hierarchical citation pattern, where early foundational work (e.g., Davis 1989) continues to hold central influence. Recent studies with high betweenness (e.g., Strzelecki 2024) indicate emerging research directions that extend existing frameworks, producing new lines of inquiry.

Table 5. Top 15 Most Frequently References in the Co-Citation Network

Rank	Keyword	Betweenness	Closeness	PageRank	Cluster
1	Strzelecki A 2024	102.734	0.017	0.017	2
2	Davis FD 1989	75.313	0.017	0.019	2
3	Kohnke L 2023	68.076	0.016	0.023	3
4	Yan D 2023	50.336	0.016	0.02	3
5	Kasneci E 2023	37.479	0.013	0.054	1
6	Venkatesh V 2003	32.733	0.016	0.019	2
7	Braun V 2021	32.519	0.014	0.014	3
8	Jeon J 2023	27.547	0.015	0.016	3
9	Barrot JS 2023	23.338	0.015	0.017	3
10	Hair JF. 2010	19.349	0.016	0.015	2
11	Fornell C 1981	18.382	0.015	0.015	2
12	Su YF 2023	17.869	0.015	0.017	3
13	Cotton DRE 2024	17.448	0.013	0.044	1
14	Huang WJ 2022	15.578	0.015	0.015	3
15	Dwivedi YK 2023	15.546	0.013	0.038	1

Table 6. Thematic Clusters in the Keyword Co-Citation Network

Cluster	Node Count	Core Thematic Focus	Representative Authors (Betweenness)
1	30	Empirical applications of GenAI in education	Kasneci E (2023, 37.479), Cotton DRE (2024, 17.448), Tlili A (2023, 11.076)
2	7	Theoretical frameworks for technology acceptance	Davis FD (1989, 75.313), Venkatesh V (2003, 32.733), Fornell C (1981, 18.382)
3	13	Learning theories and GenAI integration	Kohnke L (2023, 68.076), Braun V (2021, 32.519), Vygotsky LS (1978, 4.261)

4.3.2 Core Cluster and Thematic Analysis

Cluster 1: Empirical Applications of GenAI in Educational Contexts

Cluster 1 contains 30 nodes, constituting the majority of the network.

The central node is Kasneci E (2023) (betweenness = 37.479).

Other nodes with high PageRank include Cotton DRE (2024) (0.044) and Rudolph J (2023) (0.041). Nodes such as Tlili A (2023) and Chan CKY (2023) focus on applications in higher and medical education. Methodological approaches appear in works like Cooper G (2023) and Lo CK (2023). Nodes such as Kung TH (2023) and Bockting CL (2023) address ethical topics but have low centrality.

Cluster 1 reflects practical investigations of how GenAI tools are implemented in educational settings. Domain-specific studies show differentiated adoption patterns across disciplines. The presence of peripheral ethical literature indicates that research on integrity and governance remains less integrated within empirical work.

Cluster 2: Theoretical Roots of Technology Acceptance

Cluster 2 is anchored by Davis FD (1989) (betweenness = 75.313). Supporting nodes include Venkatesh V (2003) (32.733) and Fornell C (1981) (18.382). Model-validation works such as Hair JF (2010) (19.349) and Henseler J (2015) (7.608) appear in this cluster.

Cluster 2 consolidates foundational theories of technology acceptance and their analytical methods. Centrality patterns show sustained reliance on TAM-related frameworks. The absence of recent GenAI-specific theoretical extensions indicates limited adaptation of existing models to generative technologies.

Cluster 3: Learning Theories and Cognitive Effects

Cluster 3 is centred on Kohnke L (2023) (betweenness = 68.076). Methodological foundations appear in Braun V

(2021) (32.519). Sociocultural theory is represented by Vygotsky LS (1978) (betweenness = 4.261). Large language model literature includes Brown TB (2020) (9.461) and Bender EM (2021) (1.106). Studies focusing on cognition and affect include Yan D (2023) (50.336) and Jeon J (2023) (27.547).

Cluster 3 links Cluster 3 connects the use of GenAI with educational psychology and cognitive mechanisms. Conceptual and empirical work is organised around how GenAI may influence student motivation, self-efficacy, and learning processes. The low centrality of classic sociocultural sources suggests that theoretical frameworks are not fully incorporated into current research.

4.3.3 Cross-cluster Relationships and Research Gaps

Nodes such as Strzelecki A (2024) show high betweenness, linking Cluster 1 and Cluster 2. Yan D (2023) connects Cluster 1 and Cluster 3. Differences in centrality show disproportionate influence of TAM-related literature (e.g., Davis 1989) compared with lower-centrality theories (e.g., Vygotsky 1978). Peripheral ethical studies (e.g., Kung TH 2023) remain low in centrality metrics.

Cross-cluster bridging indicates integration of empirical research with acceptance models and learner-centred constructs. Imbalances in theoretical representation show that research predominantly references TAM, while sociocultural or ethical frameworks remain underdeveloped. Methodological disparities between quantitative model-driven studies and qualitative approaches result in inconsistent coverage of cognitive factors.

4.3.4 Implications for Research Trends

The co-citation network shows a three-layered structure, including theoretical acceptance foundations (Cluster 2), empirical applications (Cluster 1), and learning theories and cognitive perspectives (Cluster 3).

Future research directions include extending technology acceptance frameworks to generative interaction, integrating sociocultural perspectives, and incorporating ethical discussions into core research agendas.

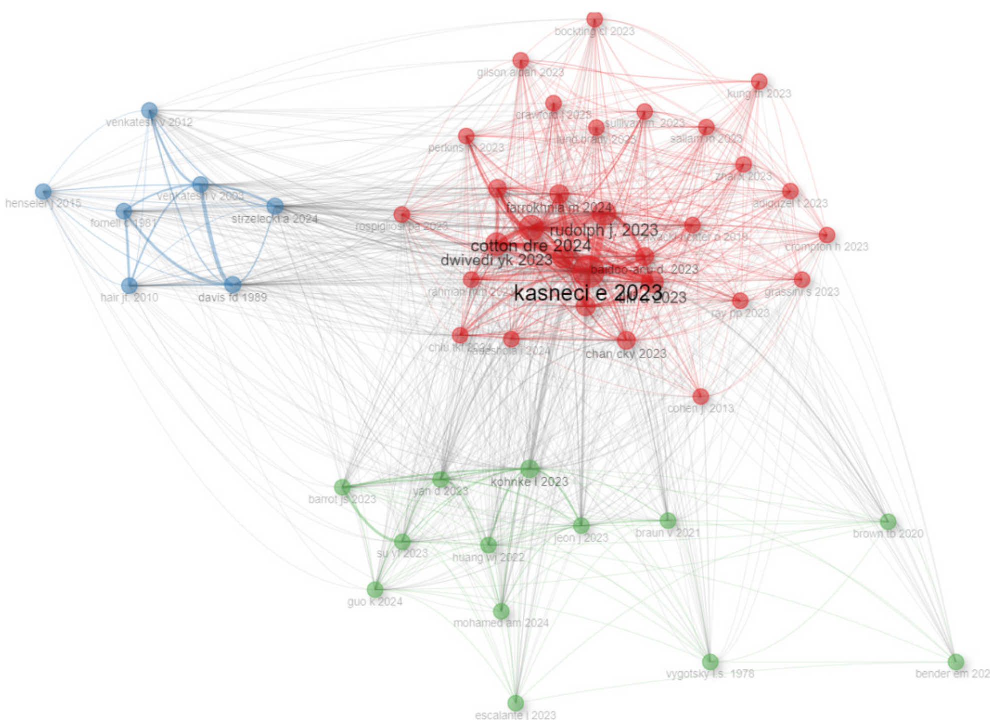

Figure 2. Co-citation Network

4.4 Authors' and Countries' Collaboration Analysis

4.4.1 Author Collaboration Network Structure

The author collaboration network consists of 44 nodes and 14 clusters (only representative author collaborations are displayed in Table 7). Clusters 1–6 and 14 exhibit zero betweenness centrality, reflecting independent collaboration groups. Nodes such as Hwang GJ (betweenness = 3.000) and Kim J (2.000) serve as bridges within their clusters. The network demonstrates low mean betweenness (0.36) and high modularity.

These structural properties indicate fragmented research communities. Bridge authors connect local subgroups, yet their limited centrality suggests insufficient integration across clusters, resulting in weak knowledge circulation.

Table 7. Representative Author Collaboration Clusters

Author	Cluster	Betweenness	Closeness	PageRank	Role type
Shahzad MF	1	0	1	0.026	Isolated group (zero-betweenness)
Kim J	7	2	0.333	0.03	Bridging author
Hwang GJ	9	3	0.333	0.049	Key connector
Guo K	8	1	0.5	0.037	High-impact
Li Y	10	1	0.5	0.037	High-impact
Kim M	7	0	0.2	0.009	Peripheral
Author	Cluster	Betweenness	Closeness	PageRank	Role type

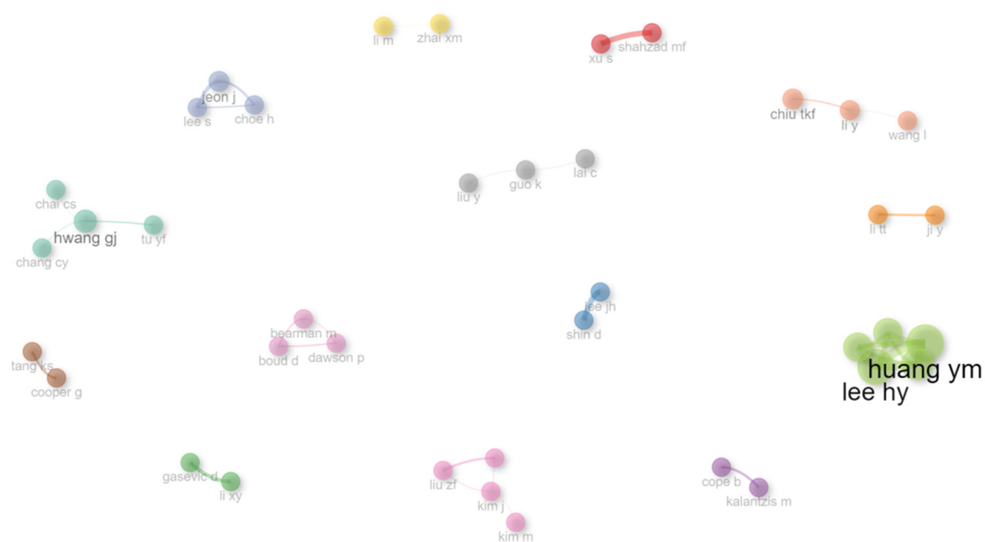

Figure 3. Author Collaboration Network

4.4.2 Author Collaboration Patterns

Clusters 1–6 and 14 display perfect closeness (1.000) and zero betweenness. Bridge authors such as Hwang GJ and Kim J connect subclusters internally. High-impact authors include Guo K (PageRank = 0.037) and Li Y (0.037). Peripheral authors such as Kim M (PageRank = 0.009) show minimal influence.

Isolated clusters reflect limited cross-group collaboration, impeding interdisciplinary integration. Bridge authors support knowledge exchange but are not sufficient to unify the network. Differences in PageRank indicate unequal research output and influence across author groups.

4.4.3 Country Collaboration Network Analysis

The country network contains 63 countries. Collaboration frequency ranges from 1 to 28 (only the top 5 pairs are displayed in Table 8). High-frequency collaborations include China–USA (28), China–Australia (19), and USA–Korea (15). Other collaborations include China–Malaysia (12) and Australia–UK (10). Regions such as Africa and South America show low frequency (e.g., Nigeria ≤ 2).

The network displays regional concentration, with stronger ties within North America and East Asia. Peripheral regions demonstrate lower connectivity, reflecting disparities in participation and research capacity.

Table 8. Leading Country Collaboration Pairs

Rank	Country Pair	Frequency	Research Focus
1	China-USA	28	Large language models in education
2	China-Australia	19	Academic integrity and AI assessment
3	USA-Korea	15	Student motivation in AI classrooms
4	China-Malaysia	12	Cross-cultural AI education comparisons
5	Australia-UK	10	Technology acceptance model validation

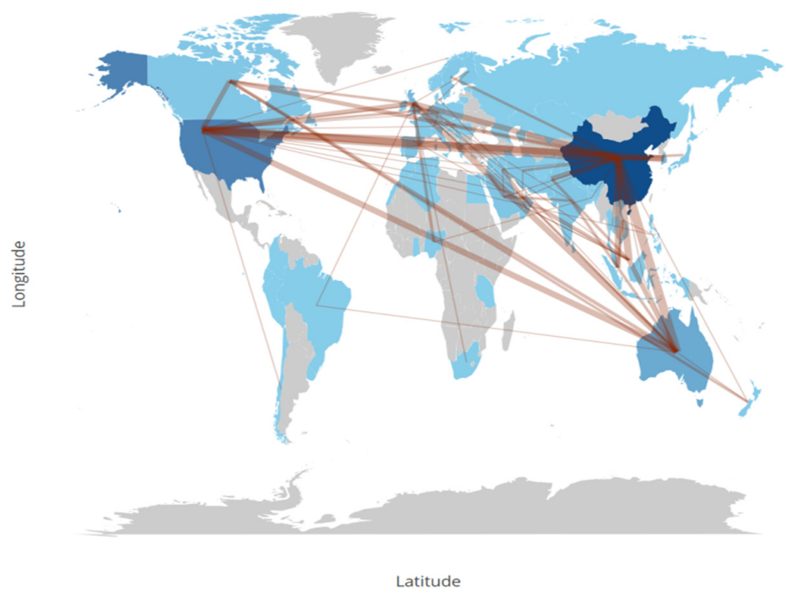

Figure 4. Country Collaboration Network

4.4.4 Collaboration Gaps and Implications

In total, 84% of author nodes exhibit zero betweenness centrality. Regional partnerships are concentrated in North America–East Asia pairs. Collaborations in ethical and cultural studies appear less frequent.

Fragmented author networks limit interdisciplinary knowledge transfer. Global collaboration patterns are uneven, with central regions dominating research output. Underrepresentation of ethics-orientated partnerships suggests unbalanced thematic development.

5. Summary and Limitation

5.1 Key Findings

This study conducted a bibliometric and network analysis of generative AI (GenAI) research in education from 2021 to 2025, covering 965 journal articles. The results indicate rapid growth in the field, with high citation impact and a predominance of original research, while review studies remain limited. Keyword and thematic analyses reveal that research focuses on technical applications of GenAI in educational contexts, with ChatGPT as a central connector, as well as on technology acceptance and adoption mechanisms guided by established frameworks such as the Technology Acceptance Model. Learner-centred outcomes, including motivation and self-efficacy, are emphasised, although higher-order cognitive skills, such as creativity and critical thinking, are less explored. Co-citation analysis confirms the influence of foundational theories and empirical studies, while author and country collaboration networks show fragmented research communities and uneven global participation.

5.2 Limitation

Several limitations arise from both the dataset and the scientometric approach. First, database coverage is restricted, as only articles indexed in the Web of Science Core Collection were included, which may omit relevant studies from other databases such as Scopus or ERIC. Second, language bias is present, with English-language publications dominating the sample, thus limiting insights derived from non-English research contexts. Third, citation-based scientometric methods have inherent constraints, as they cannot assess the conceptual rigour or empirical validity of individual studies. Fourth, incomplete institutional metadata reduces analytical precision, particularly for mapping inter-institutional collaboration. Finally, the field remains in an early developmental stage, and the rapid evolution of generative AI technologies may render current findings time-sensitive. These limitations do not diminish the value of the findings but instead signal areas that require caution when interpreting the results.

5.3 Implication

The findings highlight several implications for research and practice. First, there is a need to integrate ethical considerations, academic integrity, and governance frameworks into GenAI applications in education. Second, educational interventions should not only focus on getting people to use technology, but also on helping them develop higher-order cognitive skills like creativity and critical thinking. Third, collaboration across disciplines and regions should be encouraged to enhance knowledge exchange and reduce global disparities. Finally, extending theoretical models, including technology acceptance and learning theories, to the context of generative AI can guide both future empirical studies and practical instructional design, promoting more effective and responsible integration of GenAI tools in educational settings.

6. Conclusion

This study adopts scientometric methods to overview research on GenAI in education from 2021 to 2025, drawing on 965 peer-reviewed journal articles included in the Web of Science Core Collection. It aims to identify the primary research topics, knowledge structures, collaboration patterns, and emerging trends in this domain. The findings reveal an explosive growth of educational research in this field. Research topics primarily centre on technological applications (e.g., ChatGPT as a teaching aid), user acceptance mechanisms (such as TAM), and students' learning outcomes. The topic distribution demonstrates GenAI's potential to enhance educational processes while also reflecting the current limitations of theoretical innovation. Co-citation analysis results show that GenAI research still relies heavily on traditional theoretical models like TAM. To address GenAI's characteristics, such as interactivity, generativity, and multimodal affordances, new frameworks need to be established. Collaboration analysis indicates a high degree of fragmentation in the author network and an uneven geographical distribution of research contributions.

These findings collectively suggest that the field is in a transitional stage toward the maturity of theory and practice. Future research can focus on several directions, including developing theoretical models for GenAI's interactive features, deepening ethical frameworks or AI literacy systems, promoting interdisciplinary integration, facilitating global collaboration, and expanding mixed research methods. To ensure GenAI drives meaningful, equitable, and ethical educational transformation, sustained academic attention is necessary, and this scientometric review can serve as a foundational reference.

References

Alier, M., Francisco-José García-Peñalvo, & Camba, J. D. (2024). Generative Artificial Intelligence in Education: From Deceptive to Disruptive. *International Journal of Interactive Multimedia and Artificial Intelligence*, *8*(5), 5-5. https://doi.org/10.9781/ijimai.2024.02.011

Alsharefeen, R., & Al Sayari, N. (2025). Examining academic integrity policy and practice in the era of AI: a case study of faculty perspectives. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1621743

Andrade-Girón, D., Marín-Rodriguez, W., Sandivar-Rosas, J., Carreño-Cisneros, E., Susanibar-Ramirez, E., Zuñiga-Rojas, M., Angeles-Morales, J., & Villarreal-Torres, H. (2024). Generative artificial intelligence in higher education learning: A review based on academic databases. *Iberoamerican Journal of Science Measurement and Communication, 4*(1), 1-16. https://doi.org/10.47909/ijsmc.101

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT. *ArXiv (Cornell University)*, *37*(4). https://doi.org/10.48550/arxiv.2303.04226

Chang, C., Chen, I.-h., & Tang, K. (2024). A bibliometric analysis and systematic review. *Educational Technology & Society, 27*(4), 471-486. https://doi.org/10.2307/48791567

Chen, K., Tallant, A. C., & Selig, I. (2024). Exploring generative AI literacy in higher education: student adoption, interaction, evaluation and ethical perceptions. *Information and Learning Sciences*. https://doi.org/10.1108/ils-10-2023-0160

Creely, E. (2024). Exploring the role of generative AI in enhancing language learning: Opportunities and challenges. *International Journal of Changes in Education, 1*(3), 158-167. https://doi.org/10.47852/bonviewIJCE42022495

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2024). Does chatgpt enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education, 227*, 105224. https://doi.org/10.1016/j.compedu.2024.105224

Gimpel, H., Hall, K., Decker, S., Eymann, T., Lämmermann, L., Mädche, A., Röglinger, M., Ruiner, C., Schoch, M., Schoop, M., Urbach, N., & Vandrik, S. (2023). *Unlocking the power of generative AI models and systems such as GPT-4 and ChatGPT for higher education: A guide for students and lecturers*. Universität Hohenheim. https://doi.org/10.13140/RG.2.2.20710.09287/2

Hale, J., Alexander, S., Wright, S. T., & Gilliland, K. (2024). Generative AI in undergraduate medical education: A rapid review. *Journal of Medical Education and Curricular Development, 11*. https://doi.org/10.17615/xvfq-6m17

Hughes, R. T., Zhu, L., & Bednarz, T. (2021). Generative adversarial networks-enabled human-artificial intelligence collaborative applications for creative and design industries: A systematic review of current approaches and trends. *Frontiers in Artificial Intelligence, 4*. https://doi.org/10.3389/frai.2021.604234

Jensen, L. X., Buhl, A., Sharma, A., & Bearman, M. (2024). Generative AI and higher education: A review of claims from the first months of ChatGPT. *Higher Education, 89*(4), 1145-1161. https://doi.org/10.1007/s10734-024-01265-3

Jiang, Z., & Jiang, M. (2024). *Beyond Answers: Large Language Model-Powered Tutoring System in Physics Education for Deep Learning and Precise Understanding*. ArXiv.org. https://arxiv.org/abs/2406.10934?utm_source=chatgpt.com

Karpouzis, K., Pantazatos, D., Taouki, J., & Meli, K. (2024, February 12). *Tailoring Education with GenAI: A New Horizon in Lesson Planning*. ArXiv.org. https://doi.org/10.48550/arXiv.2403.12071

Kinskofer, F., & Tulis, M. (2025). Motivational and appraisal factors shaping generative AI use and intention in Austrian higher education students and teachers. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1677827

Liu, A., Esbenshade, L., Sarkar, S., Tian, V., Zhang, Z., He, K., & Sun, M. (2025). *Decoding Instructional Dialogue: Human-AI Collaborative Analysis of Teacher Use of AI Tool at Scale*. ArXiv.org. https://arxiv.org/abs/2507.17985

Maphoto, K. B., Sevnarayan, K., Mohale, N. E., Suliman, Z., Ntsopi, T. J., & Mokoena, D. (2024). Advancing students' academic excellence in distance education: Exploring the potential of generative AI integration to improve academic writing skills. *Open Praxis, 16*(2), 142-159. https://doi.org/10.55982/openpraxis.16.2.649

Shrivastava, P. (2025). Understanding acceptance and resistance toward generative AI technologies: a multi-theoretical framework integrating functional, risk, and sociolegal factors. *Frontiers in Artificial Intelligence*, *8*. https://doi.org/10.3389/frai.2025.1565927

Sirnoorkar, A., Allen, W., Hashmi, & Sanjay, R. N. (2025). *Examining Student and AI Generated Personalized Analogies in Introductory Physics*. ArXiv.org. https://arxiv.org/abs/2511.04290?utm_source=chatgpt.com

Tan, X., Cheng, G., & Ling, M. H. (2024). Artificial Intelligence in Teaching and Teacher Professional Development: A Systematic Review. *Computers and Education Artificial Intelligence*, *8*, 100355-100355. https://doi.org/10.1016/j.caeai.2024.100355

Uzun, L. (2023). ChatGPT and Academic Integrity Concerns: Detecting Artificial Intelligence Generated Content. *Language Education & Technology, 3*(1), 45-54. Retrieved from https://www.researchgate.net/publication/370299956_ChatGPT_and_Academic_Integrity_Concerns_Detecting_Artificial_Intelligence_Generated_Content

Wan, T., & Chen, Z. (2024). Exploring generative AI assisted feedback writing for students' written responses to a physics conceptual question with prompt engineering and few-shot learning. *Physical Review. Physics Education Research*, *20*(1). https://doi.org/10.1103/physrevphyseducres.20.010152

Wang, L., & Ren, B. (2024). Enhancing academic writing in a linguistics course with generative AI: An empirical study in a higher education institution in Hong Kong. *Education Sciences, 14*(12), 1329. https://doi.org/10.3390/educsci14121329

Yu, H., & Guo, Y. (2023). Generative artificial intelligence empowers educational reform: current status, issues, Su, J., & Yang, W. (2023). Unlocking the power of chatgpt: A framework for applying generative AI in education. *ECNU Review of Education*, *6*(3), 355-366. https://doi.org/10.1177/20965311231168423

Zhang, J., Song, W., & Liu, Y. (2025). Cognitive bias in generative AI influences religious education. *Scientific Reports*, *15*(1), 15720. https://doi.org/10.1038/s41598-025-99121-6

Zhang, P., & Tur, G. (2023). A systematic review of ChatGPT use in K-12 education. *European Journal of Education, 59*(2), e12599. https://doi.org/10.1111/ejed.12599

Acknowledgments

The successful completion of this study was made possible by the support and collaboration of many people, for which I would like to express my sincere gratitude. First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Enio Kang, and co-supervisor, Dr. Norliza, for their invaluable guidance, continuous support and constructive feedback throughout the study. Secondly, thanks to Universiti Putra Malaysia for providing the necessary resources and facilities for this study. Meanwhile, the study also benefited from the cutting-edge explorations of GenAI educational applications in recent years, and the related literature provided important references for this scientometric analysis. Finally, I would like to thank my family and fellow students for their understanding and support during the research period, whose tolerance and encouragement enabled the researcher to stay motivated in the interdisciplinary exploration. We would like to invite colleagues to correct the shortcomings and limitations of this study to contribute more insights to the academic development and practical innovation of GenAI educational applications.

Authors contributions

Junmin Guo was responsible for the conceptualization of the study, literature review, data analysis, and drafting of the manuscript. Dr. Enio Kang contributed to the research design, methodological guidance, and critical revision of the manuscript. Dr. Norliza provided academic supervision, reviewed and refined the manuscript, and offered substantive intellectual input. All authors read and approved the final manuscript.

Funding

Not applicable.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal’s policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.