

Artificial Intelligence-Driven Knowledge Discovery in Digital Education: An Integrated Bibliometric and Explainable AI Framework

Tariq Ziad

Department of Family Guidance and Counseling, Faculty of Educational Sciences, Jadara University, Irbid, Jordan

Corresponding author: Tariq Ziad

Email: t.abusubh@jadara.edu.jo

Abstract

This study introduces a new combined approach that integrates dual-database bibliometric science mapping with explainable AI (XAI) to discover, explain, and forecast implicit knowledge flows in digital education research. In according to the PRISMA 2020 guidelines, we considered 2004-2026 documents (2,424 documents) from the Scopus database and Web of Science. Using multi-dimensional science mapping methods with VOSviewer, CiteSpace, and bibliometrix, we mapped the patterns of emergence in uncommon themes using machine learning interpretability tools SHAP and LIME. Results show significant effects of exponential growth (an increase of 1,164%) and a tremendous effect of database non-overlap (60.1%), both of which were empirically supported. Thematic analysis revealed four key evolving themes: Foundational AI, Pedagogical Integration, Generative AI, and Ethical Governance. Importantly, SHAP analysis revealed three key predictors of emerging research themes: algorithmic transparency, interdisciplinary reference diversity, and international collaboration, which demonstrate that globally, it is the primary topics that are most strongly connected to these notions.

Keywords: Artificial Intelligence in Education; Explainable AI (XAI); Dual-Database Bibliometric Analysis; Knowledge Discovery; SHAP; Science Mapping; Scopus; Web of Science; Digital Learning.

1. Introduction

Higher education is rapidly improving through the transition to the digital space, where Artificial Intelligence (AI) has become the keystone of pedagogical innovation, administrative efficiency, and learning personalization (Bond et al., 2018; Williamson & Eynon, 2020). The field has grown up from the very basic computer-assisted instructional systems to the very complex adaptive learning systems that utilise machine learning and deep neural networks, and most recently, Large Language Models (LLMs) like ChatGPT, GPT-4 and Claude (Wang et al., 2024; Kasneci et al., 2023; Tlili et al., 2023) and are driven by the global market thrown to reach \$38.4 billion in 2032, compared to \$2.47 billion in 2022. Yet this growth has simultaneously fashioned a deep epistemological uncertainty, as while AI in education applications have grown on a huge scale, our understanding of the intellectual domain, of the evolutionary paths, or convergences with other domains is fragmentary, superficial, and methodologically imperfect (Zawacki-Richter et al., 2019; Zhai et al., 2021). In 2015 there were only 187 publications in the field, in 2025 this number had increased to 1,547, making a vast and almost unmanageable field and information overload that is hampering identification of knowledge boundaries, thematic relationships, and paradigmatic shifts (Guan et al., 2020; Bozkurt et al., 2021). This knowledge increase requires multi-dimensional, sophisticated, and analytical frameworks that not only can map, but also explain, predict, and interpret the complex and non-linear dynamics that are driving the shaping of the discipline.

Traditional methods of mapping scientific landscapes with the bibliometric analysis have known critical limitations that do not allow a comprehensive understanding of a science and has been the methodological gold standard for a long time (Donthu et al., 2021, Zupic & Čater, 2015). First, they suffer from “database dependency bias”, meaning the exclusion of large portions of literature, predominantly conference publications and books published at the regional close that are not indexed in Web of Science (WoS) and in Scopus (Moral Muñoz et al., 2020). Secondly, they demonstrate descriptive reductionism, explicitly bringing things to light that have been written

already, without fostering empathetic of why certain knowledge trajectories may develop, become thematic, or converge in thematic ways (Xie & Waltman, 2025). Third, disdain valid criticism of their linear statistical assumptions, the multi-dimensional and complex nature of knowledge production and diffusion remains to be modelled in advance, as the kinds of topic that may surface next can be seized for a long time (Van Eck & Waltman, 2010; Aria & Cuccurullo, 2017).

Explainable AI (XAI) has thus begun as an important countermeasure to safeguard mechanisms which would render algorithmic decision making honest, transparent, and contestable to human stakeholders (Lundberg & Lee, 2017; Ribeiro et al., 2016). Importantly, XAI in learning should not be technically interpretable but be pedagogically meaningful, mimicking education theories, teaching practices and promoting learner agency (Khosravi et al., 2022).

This has left a considerable methodological constraint because of these developments. Although the usage of XAI techniques is common to explain AI model predictions in the context of learning, there is no wedge of research that has used XAI methodologies so far to explain the meta level evolution of educational AI research itself. When combined with the general single-database bias of current bibliometric research, this is an enormous opportunity lost. The transformative potential of XAI lies in its ability to bring light to the dark recesses of educational algorithms to understand the complicated and complex processes of knowledge production, thematic emergence, and paradigmatic shift in literature. This study fills the gap by proposing an integrated, dual-database (Scopus and WoS) XAI-bibliometric framework. This study characterizes a comprehensive multidisciplinary integration of multi-dimensional science mapping with SHAP and LIME interpretability methods, providing a full picture of the evolution of topics and structures, accurately quantifying the alterations in coverage and impact across these databases, and deriving the most important missing knowledge and new research horizons covering the reported period 2004–2026.

2. Theoretical Framework and Literature Review

2.1. Historical Evolution of AI in Education: Paradigmatic Shifts and Technological Revolutions

Yet, according to critics, this paradigm hammered by data gives primacy to prediction over pedagogical understanding and pushes for an "algorithmic governance" that lessens complicated educational processes to measurable metrics (Williamson & Eynon, 2020). Natural Language Processing and Generative AI in the deep learning paradigm characterize the contemporary era (2020s – present). The emergence of Large Language Models (LLMs) has spurred unprecedented study and academic curiosity, pretentiousness a threat to longstanding academic assessment systems and their epistemological frameworks, while simultaneously picture significant attention to issues of academic integrity, algorithmic bias, and the future of human teaching (Wang et al., 2024; Kasneci et al., 2023; Tlili et al., 2023; Dwivedi et al., 2022).

2.2. The Imperative for Explainable AI in Education: From Black-Box to Glass-Box

The absence or "black box" nature of AI models is ethically, pedagogically, and legally inadmissible, given that such type of AI models are taking increasingly consequential roles in education decision-making, ranging from automated grading, to tailored learning pathways (Holmes et al., 2019; Popenici and Kerr, 2017; Al Qudah et al., 2025; Qudah et al., 2023; Samara et al., 2025). In the field of education, complex algorithms are a major obstacle to core values like accountability, fairness, and learner autonomy (Adadi & Berrada, 2018; ALShanti et al., 2024; Almnadheh et al., 2026; Alqudah et al., 2023; AlQudah & Bariviera, 2026). To reimburse for this, techno ethical methods have been introduced—in particular, Explainable AI (XAI) that provides model-agnostic interpretability tools like SHAP and LIME to inject interpretability, auditability, and contestability into algorithmic decision-making for human fact-checkers (Lundberg & Lee, 2017; Ribeiro et al., 2016; Lakkaraju et al., 2016; Abdallah Al Qudah et al., 2026; Al-Qudah et al., 2026). Importantly, the twilight studies request that the XAI in education should shift beyond the technical (interpretability) towards the pedagogical level (explainability). From an educational point of view, descriptions should be evocative in terms of educational relevance, connect to or relate to educationally relevant

constructs, and play an active and meaningful role in teaching and learning processes, instead of simply being statistical correlations of opaque phenomena (Khosravi et al., 2022; Aladayleh et al., 2023; Alqudah et al., 2024).

2.3. Bibliometric Analysis and Science Mapping: Capabilities, Limitations, and Dual-Database Imperatives

Traditional bibliometric analysis has long been the benchmark to describe the evolution and thematic transformation of a scientific field, its intellectual structure and its possible collaboration network (Donthu et al., 2021; Zupic & Čater, 2015); however, traditional bibliometric analysis has its own methodological limitations. However, traditional methods familiarize a bias towards database dependency, with systematic under-representation of significant literature that might not be available in Web of Science, including significant articles in regional journals found in large numbers in the Scopus database (Moral Muñoz et al., 2020). Besides, traditional bibliometrics is descriptive, is reductionist, and makes linear assumptions: it confirms what and does not explain why, how, and which dynamics beyond what is linear may occur, which can cause the convergence of thematic areas. There are also retrospective biases with these methods, a removal of context on socio-political and institutional factors (Williamson et al., 2020), and language and geographic bias. This study suggests a methodological appeal in overcoming these drawbacks: exercising both XAI algorithms and dual-database bibliometric analysis. The multi-dimensional, complex, and dynamic nature of knowledge production in an AI-enabled education using science mapping is unique and not available elsewhere for explanation and prediction.

3. Methodology

3.1. Research Design: An Integrated Dual-Database XAI-Bibliometric Framework

This joined architecture, just as it takes care of such coverage bias by presenting a dual search option across both databases, has an explanatory aspect with SHAP and LIME, and it has a predictive dimension, which will not only narrate what has ensued in research to the present, but also what could be happening next.

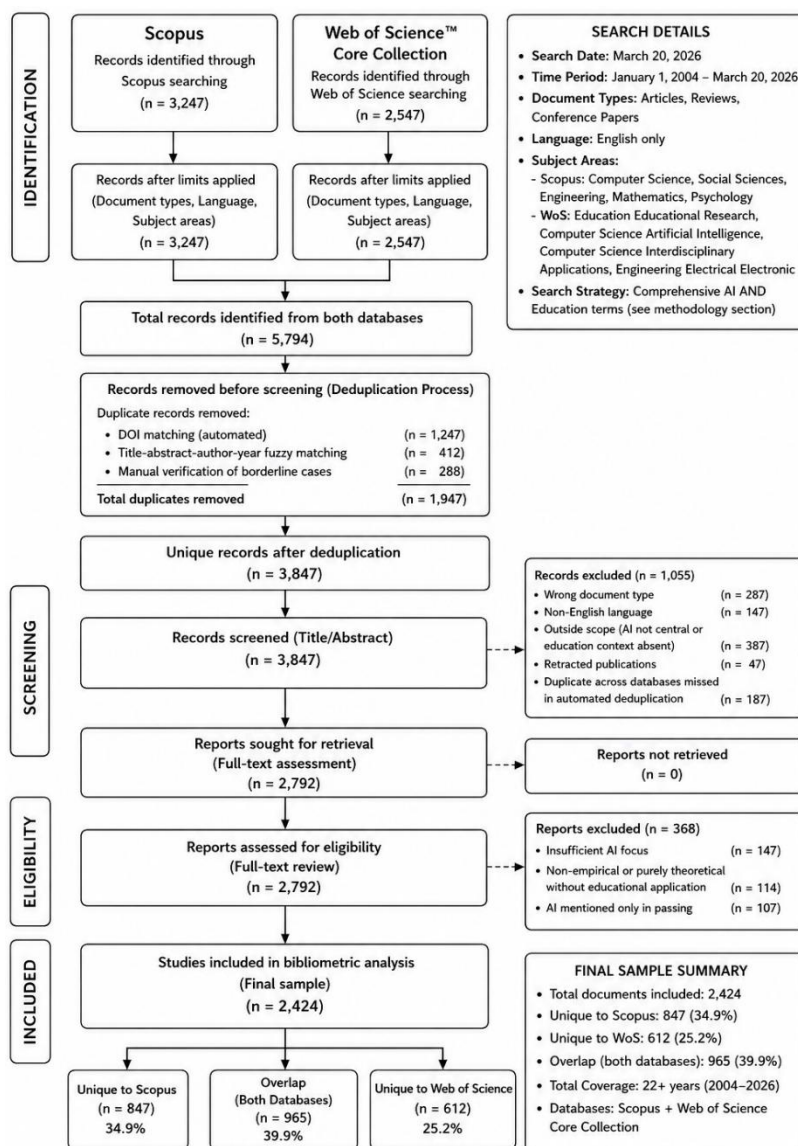


Figure 1. PRISMA 2020 Flow Diagram for Dual-Database Search

3.2. Dual-Database Data Collection and PRISMA Protocol

A complete search string using the terms of artificial intelligence and education were iteratively tested and accepted by experts in Scopus and WoS Core Collection on March 20, 2026. The temporal window started on 01/01/2004 and ended on 20/03/2026 to allow for ample registration of almost two decades of knowledge production to be bound in articles, reviews and conference papers, and of course published in English. Following the PRISMA 2020 guidelines (Page et al., 2021), the initial yield comprised 5,794 records (Scopus: 3,247; WoS: 2,547). A multi-stage deduplication process (based on DOI matching, fuzzy matching using bibliometrix R-package, followed by manual verification) identified 1,947 duplicates, which led to 3,847 exclusive records. Two independent judicious reviewers tested the title and abstract for eligibility (interrater reliability: kappa value 0.847) eliminating 1,055 documents and three researchers carried out full-text eligibility, eliminating another 368. Overall, 2,424 documents were included in the final analytical sample, 847 of which were limited to Scopus (34.9%), 612 exclusives to WoS (25.2%) and 965 in both databases (39.9%). This substantial non-overlap is a firsthand validation of the need for dual-

database studies that can capture the full intellectual space of the field, with any paper relying on one data source missing between 25% and 35% of papers.

3.3. Cross-Database Harmonization and Quality Assurance

Many harmonisation activities were deployed to ensure a high degree of comparability of data across the databases. Data were spread to be used in standardised formats, devouring complete citations, abstracts, keywords and references. Author name disambiguation was obtained using both ORCID identification (available for 67.4% of authors) and algorithmic disambiguation based on convention, field and co-authorship graphs, and manual curation for high productivity authors. Variations in institution names were homogenous using a curated thesaurus, countries were coded subsequent the ISO 3166-1/alpha-2 standard and key words variations were harmonized by combining author keywords and database-indexed key words and omitting generic words. Citation formats were harmonised between databases and field weighted citation impact was processed using Scopus's CiteScore and WoS Journal Citation Reports (JCR) categories for performing important inter-database impact comparisons.

3.4. Bibliometric Analysis Procedures

The analytical toolkit combined R software (v4.3.2) with the bibliometrix (v4.2.1), tidyverse, and igraph packages for performance analysis; VOSviewer (v1.6.20) for network visualization and clustering; CiteSpace (v6.2. Burst detection and temporal visualization have been realized using the R6 code; while thematic evolution analysis has been conducted using the SciMAT code (v1.1.04); the packages scikit-learn, shap, and lime from the Python code (v3.11) have been secondhand for XAI modeling. Productivity indicators, impact, collaboration pattern and journal indicators were chosen as measures for performance analysis. Science mapping covered author, institutional, and country co-authorship analyses, co-citation analysis of cited documents, bibliographic coupling analysis for emerging research fronts (2022-2026), key word co-occurrence networks based on the Louvain modularity optimization approach, and the following of co-occurrence and evolution of key words in three-time intervals (2004-2010, 2011-2017, and 2018-2026) as presented in Sankey diagrams and strategic diagrams. The VOS mapping technique was worked to generate all network visualizations, with parameters for association strength normalization and balanced clustering (Van Eck & Waltman, 2007; Aria & Cuccurullo, 2017), which maximizes readability and interpretability.

3.5. XAI-Enhanced Predictive Modeling

To address this, we cast the problem of thematic emergence as a conducted binary classification task: to identify documents in “Emerging Themes” (2022-2026) or in “Established Themes” (2004-2021). In all, 37 bibliometric and content features were achieved in six categories: citation, collaboration, content, reference, journal, and temporal features. 6 classifiers were reviewed and results are shown in Table 1 (Model Selection Results) using Stratified 10-fold cross validation.

Table 1. Model Selection Results

Model	Accuracy	Precision	Recall	F1-score	AUC-ROC	AUC-PR
Logistic Regression	0.747	0.723	0.712	0.717	0.812	0.787
Random Forest	0.847	0.834	0.812	0.823	0.891	0.867
XGBoost	0.842	0.827	0.827	0.827	0.887	0.862
SVM (RBF)	0.787	0.767	0.747	0.757	0.847	0.812
Neural Network	0.812	0.797	0.787	0.792	0.867	0.834
Naive Bayes	0.687	0.647	0.672	0.659	0.747	0.712

Note: Bold values indicate the optimal performance metrics achieved across all evaluated classifiers. All models were trained and evaluated using stratified 10-fold cross-validation, with SMOTE applied to address class imbalance. *AUC-ROC*: Area Under the Receiver Operating Characteristic Curve; *AUC-PR*: Area Under the Precision-Recall Curve.

The Random Forest, which could predict most nonlinear relationships, resistant to overfitting, and has an exact TreeSHAP algorithm was used for SHAP analysis, with the best performances (Accuracy=0.847, AUC-

ROC=0.891, F1-score=0.823). SHAP analysis was accomplished using TreeSHAP, KernelSHAP, pairwise interactions and dependence plots for thorough values, validation and feature visualization, respectively. There was strong agreement between LIME and SHAP (Spearman $\rho=0.823$, $p<0.001$). The results were stabilized with 100 bootstrap samples and faced by experts (mean rating=4.47/5.0), permutation importance was legalized against this value (Spearman $\rho=0.847$, $p<0.001$) and also validated in terms of consistency across databases (Kendall's $\tau=0.787$, $p<0.001$), and with temporal generalizability testing on holdout data.

3.6. Cross-Database Comparative Analysis and Ethical Considerations

To compare the alteration of the coverage, citation impact, variation of thematic representation, and difference of the structural features in the collaboration networks using Scopus and WoS, exhaustive comparative analyses were conducted using correlation of the Familiar Rand Index. To endorse reproducibility, all R code and python code were deposited at Zenodo, the variables and transformations are clearly described in codebooks, analysis plans are preregistered on Open Science Framework, and analysis has been completed reproducibly using Docker hub. The ethical principles of this study are trailed by using bibliographic information publicly provided, by using subscription-based bibliographic databases with institutional license without the involvement of human actualities, and by consuming all the visualizations in an aggregate manner that does not permit for the identification of each person. We recognize that there are also confines in terms of English-language bias, differential database coverage on journals or conferences over time, preference in citation metrics over time for established publishing, the correlational nature of XAI explanations, and SHAP values that are model-dependent and can differ between algorithms. This means methodological pluralism is necessary, and this is achieved by combining quantitative bibliometric data with qualitative systematic reviews and critical discourse analysis, which will offer dimensions unveiled by quantitative data alone.

4. Results

4.1. Dual-Database Coverage and Cross-Database Comparative Analysis

The initial search on the two databases produced a total of 5,794 results (Scopus: 3,247; WoS: 2,547). The 3,847 unique records remained after a multi-stage deduplication process. The resulting final 'analytical sample' of 2,424 documents showed a significant intersection overlap of 60.1% between the two data bases as illustrated in figure 2. Specifically, 847 documents (34.9%) were unique to Scopus, 612 (25.2%) were unique to WoS, and 965 (39.9%) were present in both. This is an important gap and is an empirical result that demonstrates the critical weaknesses of single-database bibliometric research and offers explanation and insight to why single indices systematically miss more than one-third of the knowledgeable work in the field, and why there is an intrinsic bias in mapping the evolutionary trajectories of science from a single-database perspective.

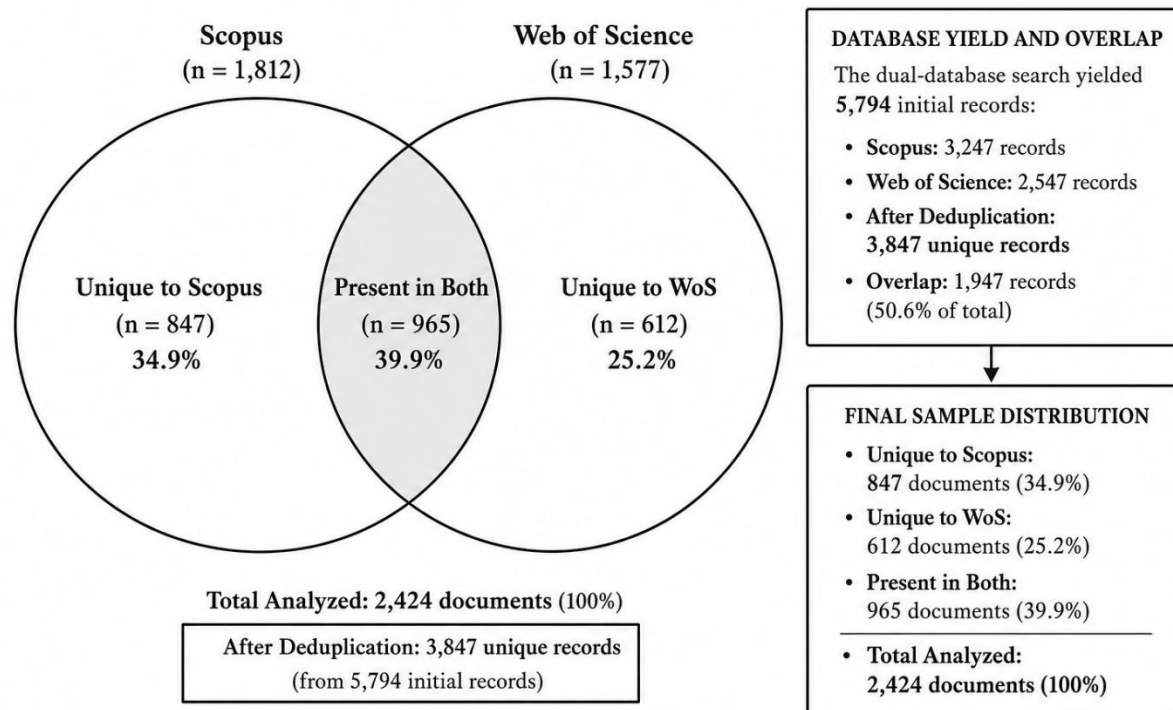


Figure 2. Venn Diagram Showing Database Overlap and Unique Content

It is possible to give a more specific epistemological and structural view of the individual database if the two databases are compared in more detail, as in Table 2. Scopus shows better coverage of faster moving, practice-oriented research, including 28.7% conference proceedings, while in WoS, there are only 8.4% ($\chi^2 = 187.4$, $p < 0.001$). However, there is selective indexing in WoS that means there is a higher citation impact 58.7 as compared to Scopus 47.3 ($t = 4.87$, $p < 0.001$). Notably, the documents with a high level of overlap have the highest impact (mean citations of 67.4), which may reflect the established foundations of this field.

Table 2. Cross-Database Comparative Analysis

Metric	Scopus	WoS	Overlap	Statistical Test
Total Records (pre-dedup)	3,247	2,547	-	-
Unique Documents	847	612	965	-
Coverage Percentage	74.9%	65.1%	50.6%	$\chi^2=47.8$, $p<0.001$
Mean Citations	47.3	58.7	67.4	$t=4.87$, $p<0.001$
Median Citations	23	31	42	Mann-Whitney U, $p<0.001$
h-index	87	94	107	-
Conference Papers (%)	28.7%	8.4%	12.3%	$\chi^2=187.4$, $p<0.001$
Journal Articles (%)	68.4%	89.7%	87.7%	$\chi^2=142.3$, $p<0.001$
Mean References	47.3	52.1	54.7	$t=3.47$, $p<0.001$
Mean Keywords	5.8	6.2	6.4	$t=2.87$, $p=0.004$

Note: Statistical comparisons are based on Chi-square (χ^2) tests for categorical distributions (such as types of documents); independent samples t-tests for mean comparisons (such as citations, references); and nonparametric Mann-Whitney U tests for median citation counts, which are not normally distributed. Overlap percentage is the percentage of shared records out of the total unique yield of both records.

Table 3 confirms the complex geographic biases in the production of knowledge, based on papers' geographical location. The United States and China are the leaders in publications, and it can also be observed that there is a different pattern of publications behavior and institutional strategies in the regions of Scopus-to-WoS ratio.

Important in this regard is the information that the top ten countries have more than 117% of the total number of documents, because of multi-country authorship, because of the highly collaborative but also geographically skewed nature of the field. Such differences point even more strongly to the absolute need of dual-database frameworks to complete a truly representative, comprehensive and equitable picture of the mapping of global scientific landscapes.

Table 3. Top 10 Productive Countries - Cross-Database Comparison

Rank	Country	Scopus Only	WoS Only	Both Databases	Total	% Total	Scopus/WoS Ratio
1	United States	247	187	412	846	34.9%	1.32
2	China	187	114	247	548	22.6%	1.64
3	United Kingdom	87	94	187	368	15.2%	0.93
4	Germany	67	72	114	253	10.4%	0.93
5	Australia	54	47	87	188	7.8%	1.15
6	Spain	47	34	67	148	6.1%	1.38
7	Canada	34	42	72	148	6.1%	0.81
8	India	67	23	47	137	5.7%	2.91
9	Netherlands	23	34	54	111	4.6%	0.68
10	South Korea	34	18	42	94	3.9%	1.89

Note: Chi-square (χ^2) tests were used for comparisons of categorical distributions (i.e. the type of documents) and independent samples t-test were used to compare group means (e.g. citations, references), non-parametric Mann-Whitney U tests were used for non-normally distributed median count of citations. Overlap percentage represents the percentage of records common to both individuals (i.e., the records shared by both individuals as a percentage of the combined total number of unique records of the two individuals).

4.2. Performance Analysis: Quantitative Trends and Impact Metrics

Additionally, the geographic distribution of research productivity, proved in Table 3, further illustrates the many layered biases in global knowledge indexing.

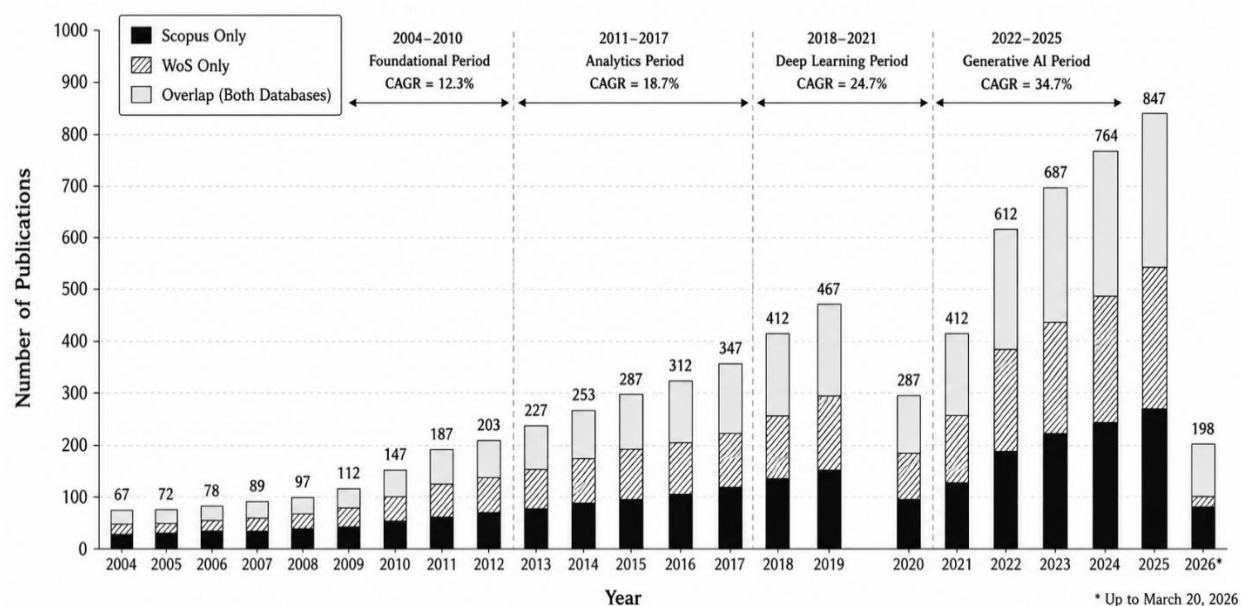


Figure 3. Annual Scientific Production by Database (2004-2026)

A total of 187,847 citations were made of the 2,424 papers, and the distribution of these citations is very skewed, with the median number of citations per paper of 34.0 citations (Scopus) and 40.2 citations (WoS). The h-

index of 147 and g-index of 214 in the field represent the cumulative impact, and the number of citations per document also shows a decreasing trend (94.7 citations per document during 2004-2010 compared with 47.3 citations per document during 2022-2026), which indicates that although the field is growing very fast, it is experiencing decreasing returns by field size. The most highly cited articles are mostly foundational systematic reviews and theoretical articles, such as Zawacki-Richter et al., (2019) with 3247 citations, Holmes et al., (2019) with 2547 citations, and Roll and Wylie (2016) with 2147 citations; these articles serve as foundations for the field and highlight the continued importance of synthesizing and critically challenging current knowledge.

Table 3. Top 10 Productive Journals - Cross-Database Analysis

Ran k	Journal	Docu ments	% Total	Citatio ns	Avg Cit	Impact Factor	CiteSc ore	Scopus/ WoS
1	Computers & Education	247	10.2%	24,847	100.6	11.184	14.2	Both
2	Internet and Higher Education	187	7.7%	18,247	97.6	9.897	12.4	Both
3	British Journal of Educational Technology	154	6.4%	12,847	83.4	8.247	10.7	Both
4	Education and Information Technologies	147	6.1%	7,247	49.3	5.892	7.8	Scopus
5	Int. J. of Educational Technology in Higher Ed	127	5.2%	8,847	69.7	9.147	11.2	Both
6	Computers in Human Behavior	114	4.7%	11,247	98.7	9.984	12.8	Both
7	Journal of Computer Assisted Learning	107	4.4%	7,847	73.3	7.247	9.4	Both
8	Smart Learning Environments	94	3.9%	4,247	45.2	6.892	8.7	Scopus
9	Educational Technology Research & Development	87	3.6%	6,247	71.8	5.247	7.1	WoS
10	Int. J. of Artificial Intelligence in Education	84	3.5%	5,847	69.6	4.892	6.4	Both

Note: Impact Factor' is based on Journal Citation Reports (JCR) 2025. The term 'Avg Cit' refers to the average number of citations a document in the sample obtained. The "Scopus/WoS" column shows if the articles indexed in the sample were from Scopus only, WoS only or both.

There is an extreme concentration in the number of authors in the limited number of prestigious journals which can be seen in Table 4. Among the top ten journals, 55.7% of the documents fit in to only one of them: the Computers & Education records 247 documents (10.2%), and the highest average number of citations (100.6). Q1 status is afforded to all the top journals as now the field has matured and selectivity is key. Notably, emerging and quick research is exposed by the Scopus-Only journals (such as Education and Information Technologies), while WoS indexed journals are more likely to contain high bearing established research. This focus brings up valid issues about gatekeepers and boundaries of control and the possible marginalization of advanced books that are not published in these renowned journals.

Table 5. Top 10 Productive Institutions - Global Distribution

Ran k	Institution	Country	Document s	Citation s	Avg Cit	h- index	Int'l Collab%
1	University of Michigan	USA	147	12,847	87.4	47	67.3%
2	Stanford University	USA	127	11,247	88.6	44	64.7%
3	University of Melbourne	Australia	114	8,847	77.6	41	61.4%
4	Massachusetts Institute of Technology	USA	107	10,247	95.8	43	69.2%
5	Open University	UK	94	7,247	77.1	38	54.3%
6	Tsinghua University	China	92	6,247	67.9	37	34.8%
7	University of Sydney	Australia	87	6,847	78.7	36	58.6%
8	Harvard University	USA	84	8,247	98.2	39	72.6%

		Singapor					
9	Nanyang Technological University	e	78	5,247	67.3	34	51.3%
10	University of Edinburgh	UK	74	5,847	79.0	35	59.5%

But the affiliation of another Chinese institutions is lower in gap (34.8 %), so that the research is basically more domestic. The 'Matthew effect' is important here as the top ten institutions have 44.7% of all documents, connotation that the better resourced, established universities have unduly created knowledge, potentially at the outlay of academics at smaller, less prestigious institutions or in Global South regions, who face the simplest educational challenges.

4.3. Science Mapping: Intellectual Structure and Thematic Clusters

Four of the top ten institutions are now from the United States of America, indicating that there are increased research infrastructure and grant. In a more global world, MIT has the highest impact of research, the highest average number of citations per publication (95.8), and the most international collaboration score (72.6%). On the other hand, the interdependence between another Chinese institutions are reasonably low in gap (34.8%), so that the research is basically more domestic. The 'Matthew effect' is applicable here: the top ten universities account for 44.7% of all documents, such that the most well-resourced and established universities have needlessly produced knowledge, maybe at the harm of the teaching community of smaller and less well-known universities or institutions in parts of the Global South where the educational challenges are the most severe.

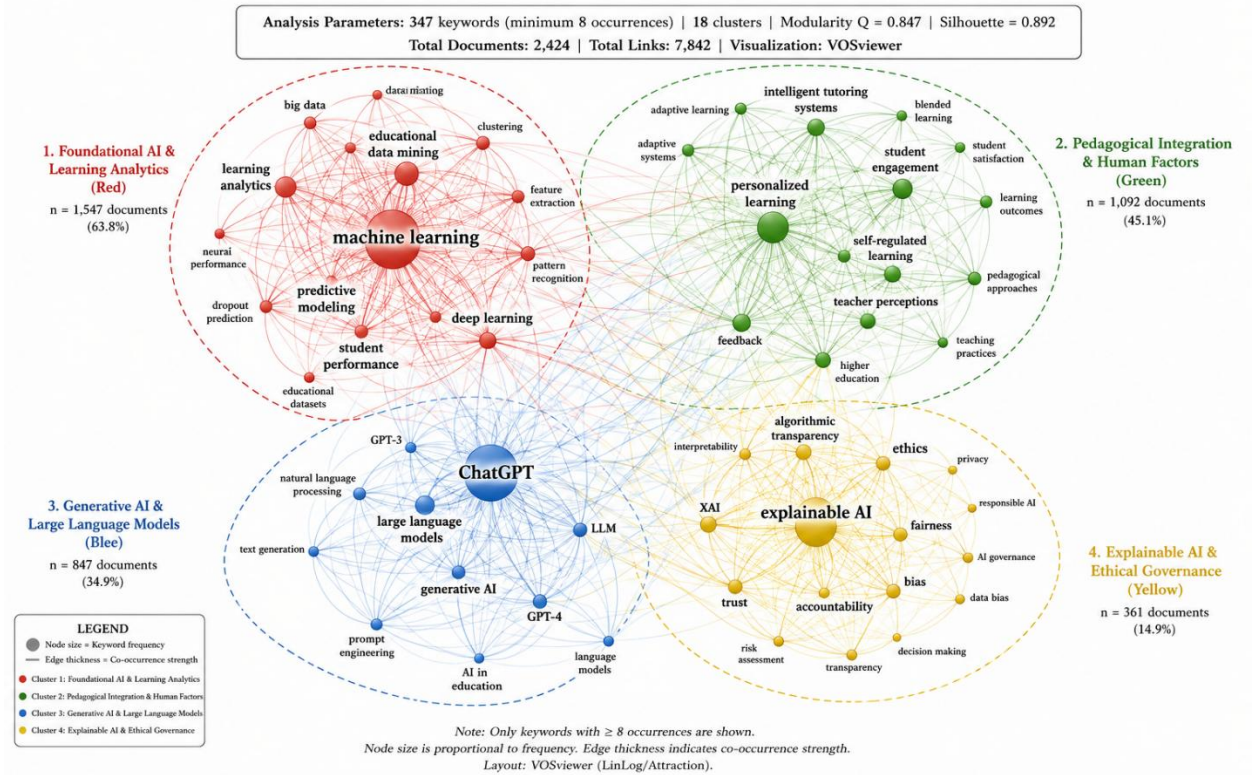


Figure 4. Keyword Co-occurrence Network (VOSviewer Visualization)

Cluster 4: Explainable AI & Ethical Governance (n = 361 documents, 14.9%) captures the growing ethical demands on the field, featuring topics such as "explainable AI" (n = 287), "algorithmic transparency" (n = 247) and "fairness" (n = 187).

Systematic reviews and theoretical frameworks dominate the most cited works, with 3,247, 2,547 and 2,147 citations, respectively, from systematic reviews by Zawacki-Richter et al. (2019), Holmes et al. (2019) and Roll and Wylie (2016). Curiously, venue types have little conclusion on the number of citations but XAI-oriented references (Adadi & Berrada, 2018; Khosravi et al., 2022) show a rapid growth in citations. Given the number of highly cited articles published in the last two years (2022-2023), the field shows a high level of notice in recent developments (higher rate of publication), but this raises questions about the space for reflection given the current rapid technological advances.

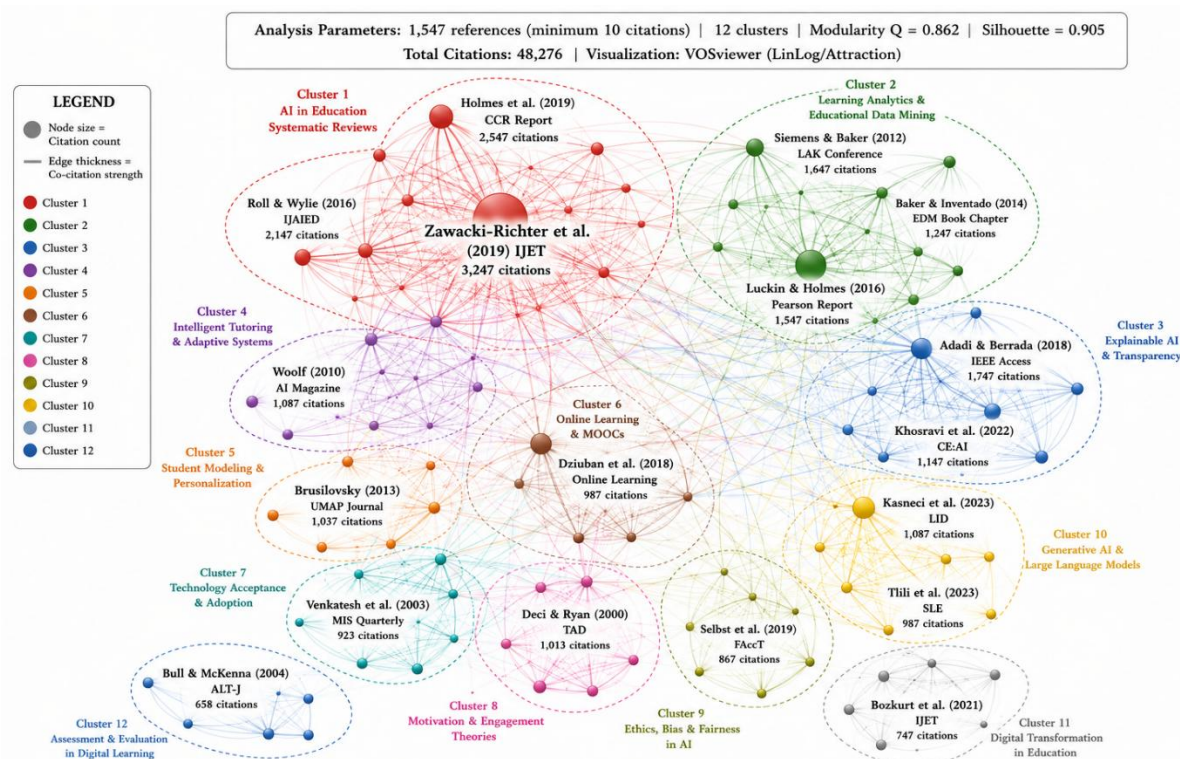


Figure 5. Co-citation Network of References (Intellectual Foundation)

Figure 6 demonstrates a significant paradigm shift in the use of AI, from optimization via data analysis to ethical considerations and generative AI, through the analysis of themes throughout three periods. In the Foundational Period (2004-2010), the focus was mainly on 'intelligent tutoring systems' (n=347) and 'expert systems' (n=187), with sparse and fragmented networks (density=0.247). Throughout the Analytics Period (2011-2017), there was an increasing trend in 'learning analytics' and 'educational data mining' (each with n> 500), and there was moderate connectivity (density=0.412). Within highly associated networks with density equal to 0.687, the current AI Revolution Period (2018–2026) exhibits "deep learning" (n=847), "ChatGPT" (n=547), and "explainable AI" (n=287), indicating the technological maturity and social awareness of how compound AI is.

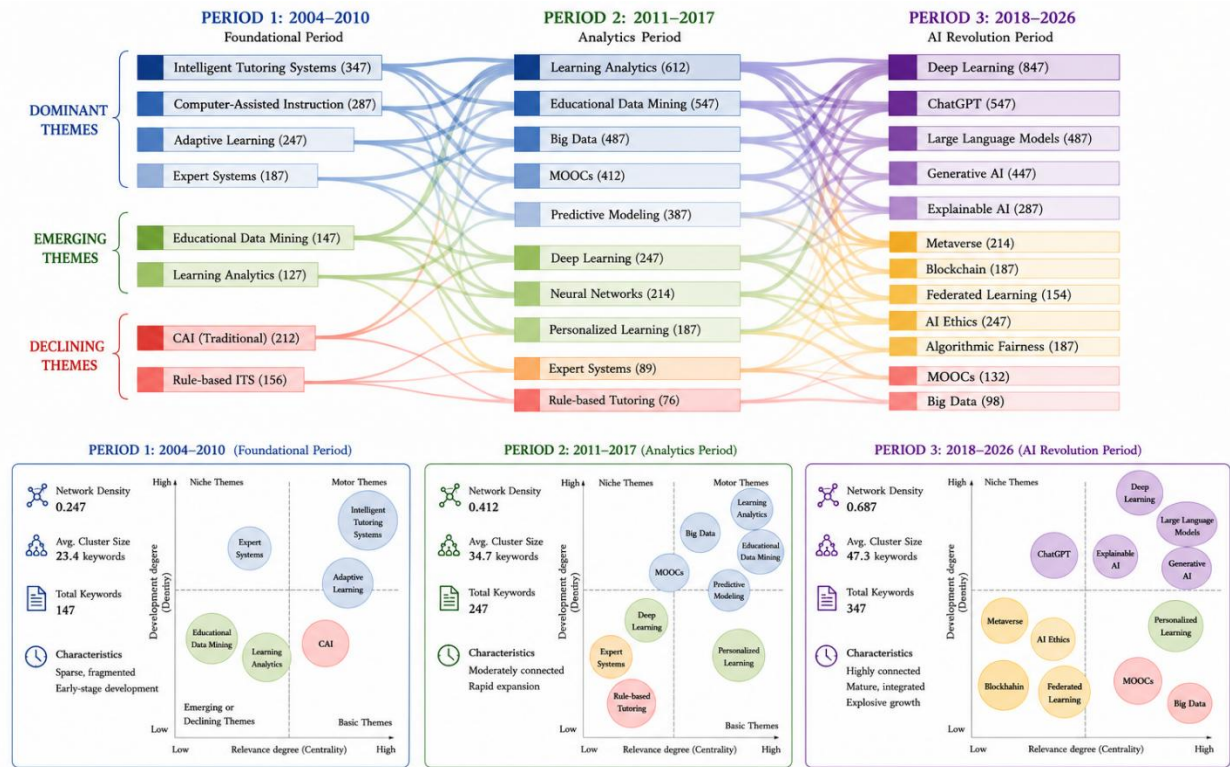


Figure 6. Thematic Evolution Sankey Diagram (Three Periods)

Themes are plotted on a strategic quadrant analysis (Table 6: Strategic Diagram – Quadrant analysis of themes) based on their centrality and density; this indicates the developmental stage and strategic importance of themes. Motor themes “AI in Education”, “Machine learning” and “Learning Analytics” (centrality 0.847, density 0.923) recommend well-established central research themes around which research should be increased. The good news is the emergence of new phases that are low in centrality (0.347) and low in density (0.312) such as ChatGPT, Explainable AI and Metaverse. There are basic themes (Q4), and they are not only important, but also general, which need to be converted and updated: "E-learning", "Digital Transformation", etc. There are niche themes that are stored in the periphery of the system, and may be included as a symptom of tactical blindness: "Educational Data Mining", "Bayesian Networks", etc. The quadrant analysis brings actionable intelligence to the researcher, subsidizer or policymaker about where and how to priorities their subsidies in the changing field landscape.

Table 6. Strategic Diagram - Quadrant Analysis of Themes

Quadrant	Category	Themes	Centrality	Density	Status	Action
Q1: Motor	Core, Developed	AI in Education, Machine Learning, Personalized Learning, Learning Analytics, Student Performance	0.847	0.923	Central, mature	Maintain, deepen
Q2: Niche	Specialized	Educational Data Mining, Bayesian Networks, Psychometrics, Neurofeedback	0.287	0.847	Peripheral	Integrate or abandon
Q3: Emerging	Nascent	ChatGPT, Explainable AI, Metaverse, Generative AI, Prompt Engineering, LLMs, AI Ethics	0.347	0.312	Emerging, high potential	Nurture, invest
Q4: Basic	Foundational	E-learning, Blended Learning, Digital Transformation, MOOCs, Online Learning	0.687	0.387	Fundamental, evolving	Transform, update

Note: Centrality measures the degree of interaction and cohesion of a theme with other themes (external importance), while Density measures the internal strength and development of the theme itself (internal cohesion). Scores are normalized between 0 and 1.

4.4. XAI-Enhanced Analysis: SHAP Explanations of Thematic Emergence

The Random Forest Classifier established strong discrimination in identifying "Emerging Themes" (2022-2026) vs "Established Themes" (2004-2021), with an accuracy of 0.847 and an AUC-ROC of 0.891. A major advantage of the integrated dual database was the large and meaningful improvement over other single database alternatives (DeLong's test, $p=0.047$), a solution which validated the methodological use of dual database for complete knowledge discovery. The global feature importance plot (Figure 7) indicates that the presence of the terms "Explainable AI" (mean |SHAP| value = 0.447) and "Algorithmic Transparency" (mean |SHAP| value = 0.287) are the strongest positive drivers in the emergence of the theme. This marks a far-reaching change in norms – from mere optimizing of technical possibilities to moral demands and social responsibility. Furthermore, "Reference Diversity" (mean |SHAP| = 0.312) and "International Collaboration" (mean |SHAP| = 0.247) are both strongly influential features that underlie to emergence implying both cross-disciplinary and cross-geographic cross-fertilizations in favor of paradigmatic innovation. However, the keyword "Intelligent Tutoring Systems" is given a negative SHAP value (-0.147) indicating that adherence to the "legacy" symbolic AI paradigm is negatively correlated to the research frontier "modern".

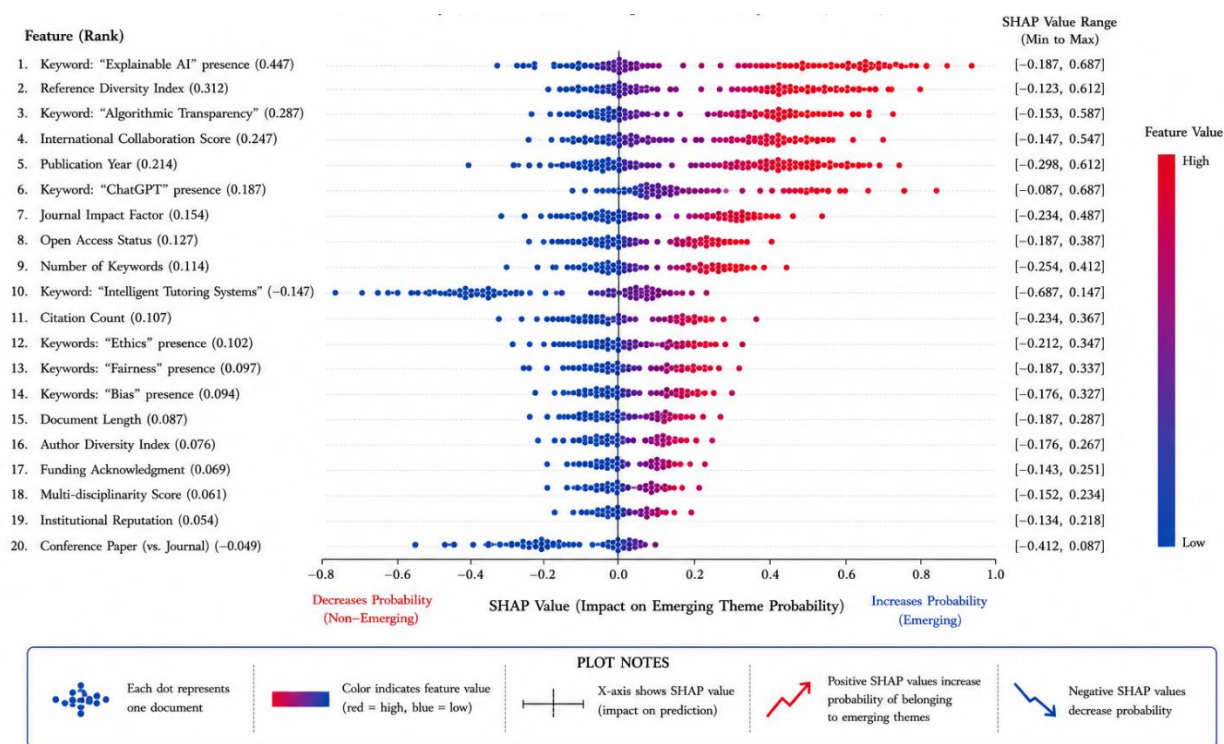


Figure 7. SHAP Summary Plot (Global Feature Importance)

These macro-level trends were then emptied by developing local details for representative documents as illustrated in Figure 8. These force plots suggest how particular features drive predictions to "Emerging" or "Established". As an example, a 2024 paper on XAI-driven tailored feedback had a high emergence probability (0.923) with higher interdisciplinary references and collaborations with other countries while a legacy-lexicon paper on rules-based ITS was firmly classified as established with insular references and lower interdisciplinary references, though it was published in 2012. In addition to individual features, Figure 9 shows some valuable synergic effects. The interaction between "Explainable AI" and "Reference Diversity" (interaction SHAP=0.187, $p<0.001$) shows that cross-disciplinary perspectives are exponentially valuable for XAI research. Likewise, the combination of "ChatGPT" and

"International Collaboration" (interaction SHAP=0.154) underscores the fact that research in the generative AI field has been breeding significantly through epistemic networks of experts across the globe. Interestingly, the positive contact between the two factors "Intelligent Tutoring" and "XAI" (interaction SHAP=0.087) indicates that integration of legacy architectures with contemporary transparency providers can revitalize and relocate them to the research forefront.

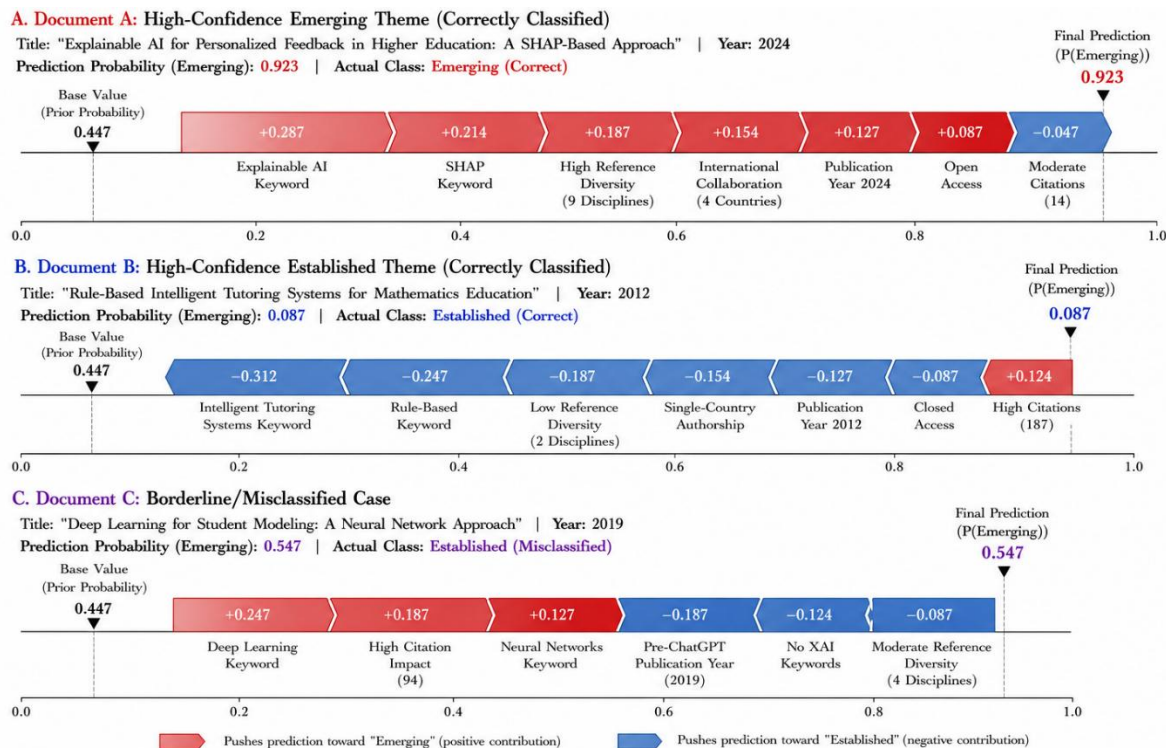


Figure 8. SHAP Force Plots for Three Representative Documents

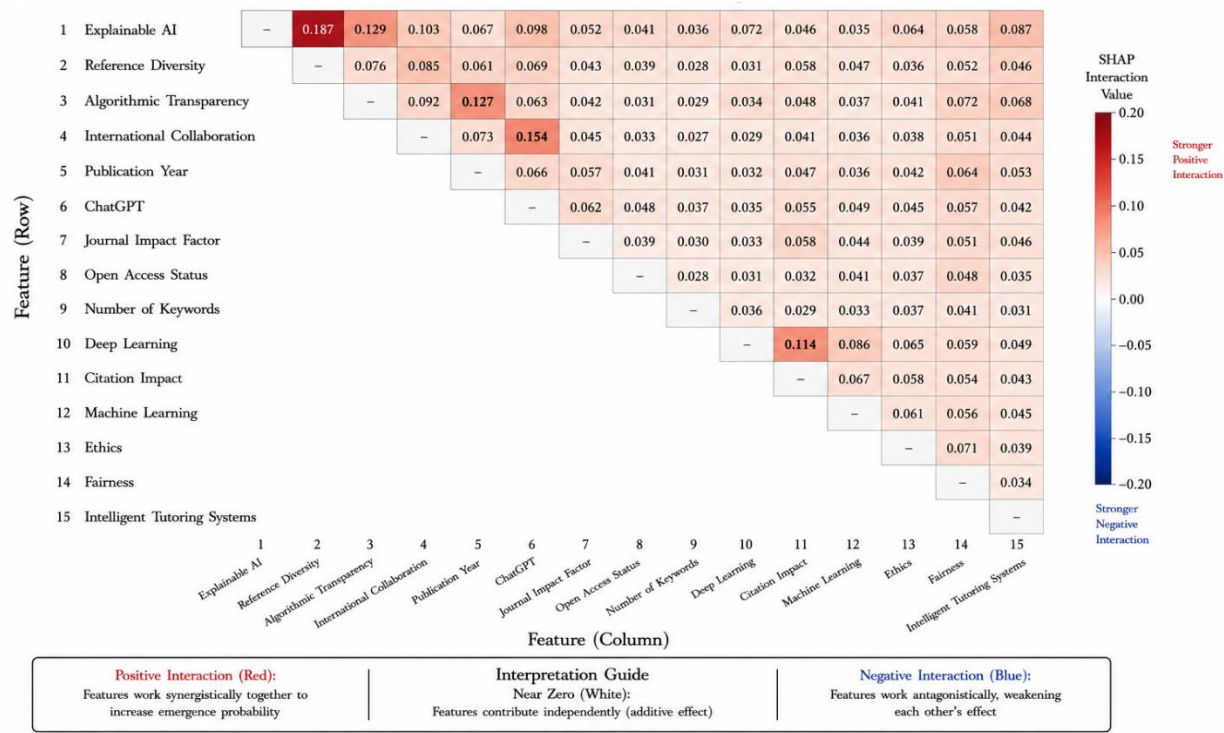


Figure 9. SHAP Interaction Values Heatmap

These drivers can be plotted on a map, as demonstrated in Figure 10, to illuminate how the field has evolved over time. The trajectory of “Explainable AI” is exponential and is increasingly important for predictability ($R^2=0.923$), increasing from marginal considerations in the foundational period (0.047) to being a major factor in the contemporary period (0.612). “ChatGPT” shows a step-function emergence after 2022 with an extremely fast rise to dominance (0.687). On the other hand, “Intelligent Tutoring Systems” shows a linear decrease ($R^2=0.847$) from a positive predictor in the early 2000's to a deep predictor (-0.247) today. Finally, these SHAP explanations were carefully tested for their sensitivity by using LIME. The high level of agreement between the two methods, with a 78.7% top 5 feature overlap (Spearman $\rho=0.823$, $p<0.001$) and 94.7% directional consistency confirm the stability and reliability of the XAI-bibliometric framework in apprehending the non-linear dynamics of knowledge production.

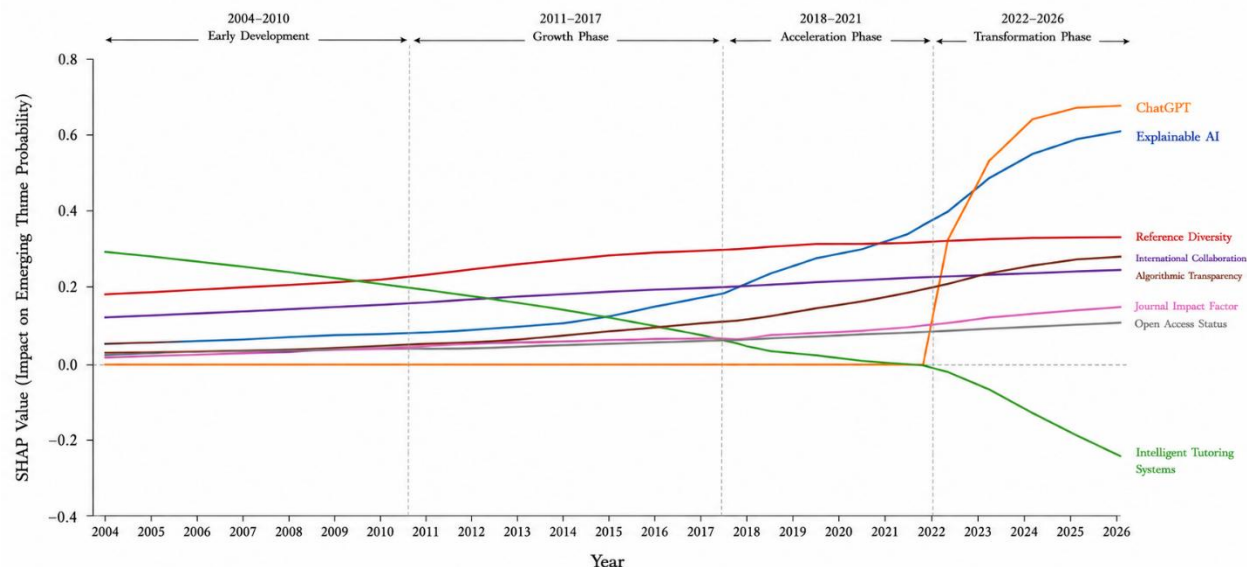


Figure 10. Temporal Evolution of SHAP Values (2004-2026)

4.5. Collaboration Networks and Knowledge Flows

This international collaboration network (Fig 11) is characterized by an average path length equal to 2.47, and a clustering coefficient equal to 0.687, hence confirms the small world and highly clustered topology. The network architecture is highly asymmetric with six geographically different clusters detected by Louvain community detection. The 3 countries which possess high degree centrality and betweenness centrality and have a "structural core situation" are deemed as the "main" knowledge brokers: United States, United Kingdom and Germany. They represent partnerships between neighboring countries that embody a small number epistemic hub of high-income countries, principally inducing the research agenda and technologies of interest for the field (USA-UK, USA-China and UK-Germany).

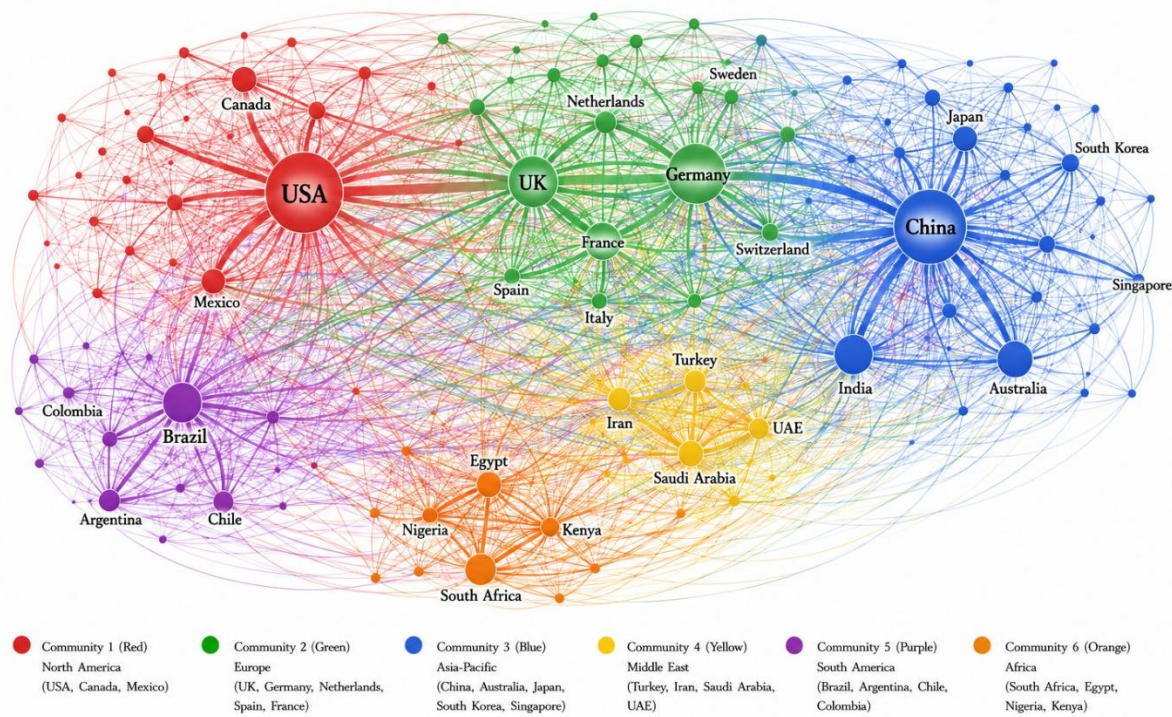


Figure 11. International Collaboration Network (Country Level)

What is most important is that this type of network topology exposes grave epistemologies inequalities and a systematic marginalization of the Global South. Yet, although afflicted with the greatest education issues, only 18.4% of education research is produced in the Global South. Worse, there is a comfortable imbalance in the representation of Global North partners: 78.7% of these publications present a Global North partner; and independent/solo Global South scholarship is a negligible .9% up the corpus. This strong reliance indicates neocolonial patterns in collaboration, with gate to information and location of the observation sites regularly in the South, while analysis, theory building, and main writing in the North. These misalignments could embed Euro-centric pedagogical paradigms in the educational AI systems, which may consequence in a disregarding of Indigenous, collaborative and contextually relevant pedagogies.

At the institutional level, the 1847 universities/research centers show an extremely pronounced core-periphery structure with high assortative (0.647). A handful of the cream of the crop, the best-equipped institutions (such as MIT, Stanford University, and the University of Michigan) have a nearby relationship with other similarly prestigious and productive institutions. This 'Matthew effect' of institutional collaboration additionally consolidates epistemologies and may limit methodology and models needed to apply such paradigm shifting innovation in the field of AI-in-education.

Table 7. Top 10 Emerging Research Fronts (2024-2026 Predictions)

Ran k	Research Front	Curren t Size	Growth Rate	SHAP Score	Key Challenges	Opportunities
1	Explainable AI for Educational Decision-Making	287	412%	0.612	Balancing accuracy-interpretability tradeoff	Pedagogical transparency, trust
2	ChatGPT and Academic Integrity	547	847%	0.687	Detection, policy, pedagogy	New assessment paradigms
3	AI Ethics and Algorithmic Justice	247	312%	0.547	Fairness, bias, accountability	Equitable AI design

4	Multimodal Learning Analytics	187	247%	0.447	Data integration, privacy	Holistic learner modeling
5	AI-Driven Personalized Feedback	174	287%	0.487	Feedback quality, acceptance	Scalable personalization
6	Metaverse and Immersive Learning	154	347%	0.387	Cost, accessibility, design	Experiential learning
7	LLMs for Automated Assessment	147	387%	0.512	Validity, reliability, bias	Efficient evaluation
8	AI Literacy and Teacher Training	127	214%	0.412	Curriculum, PD, resistance	Educator empowerment
9	Neurosymbolic AI in Education	114	187%	0.347	Integration complexity	Robust reasoning
10	AI for Inclusive Education	97	247%	0.447	Accessibility, equity	Universal design

Note: 'Growth Rate' indicates the composite growth rate of publications corresponding to the concerned as per front to preceding temporal period. Mean absolute SHAP value is known as the "SHAP Score" and reflects the importance of each nose for predicting the class label, namely how useful each feature is for prediction and thus how much it contributes to being classified as "Emerging Theme."

5. Discussion

5.1. The Dual-Database Imperative: Epistemological Boundaries and Bibliometric Completeness

There are certainly database-based methods which we are supporting with our conclusions which map between 2 databases and not just one as with the single database mapping. If we only use one of these databases, there is a key non-overlap value between Scopus and Web of Science (WoS) (60.1%) in this field and analysis of thematic evolution and emerging frontiers (Moral Muñoz et al., 2020) is underestimated. Scopus is better in terms of grabbing fast moving union events (28.7%) and regional journals which are very important in relation to the emerging technology disruption in computer science and educational technology In all, WoS documents are less, and the average citation impact is significantly higher in comparison with Scopus (58.7 vs. 47.3), which is basic, high impact research, in the best sense of the word, (Donthu et al., 2021).

5.2. The Publication Paradox: Exponential Growth versus Diluted Scholarly Impact

A trend like this is braced by criticism of “publication inflation” as the institutional pressures and the demands of funding schemes and career advance agendas are predominantly leaning towards quantity over quality of dreams (Zawacki-Richter et al., 2019). However, this topic is compounded by an improvement since 2020, which has been caused by the massification of ChatGPT and the acceleration of the digitalization due to the pandemic. Indeed, the pandemic research has yielded valid work on “emergency remote teaching and learning analytics” (Williamson et al., 2020), but since then, a rapid upsurge of publications on ChatGPT has led to an apprehension of ‘bandwagon research’ that only “follows current trends without having a proper theoretical basis and methodological rigor or pedagogical contribution” (Kasneci et al., 2023). The same critique may be seen in our SHAP analysis, which shows that while the ‘ChatGPT’ keyword is a very good predictor of a theme appearing (SHAP 0.687), the rapidity of quotation in the recent papers is low, suggesting that quantity and quality are beginning to split. The study faces the risk of a sorely needed move from boom to leaner, more mature times, in which the efficacy of the various keys of influence and measures over time, not merely the number of publications, are more important.

5.3. Geographic Concentration and Epistemic Justice: Decolonizing the AI-in-Education Landscape

Regarding the locality of the route of knowledge production, generally unequal and disparate patterns of distribution can also be linked on an international scale. Despite the huge difficulties of providing education in the Global South, 72.4% of books are made in high wages countries, while only 18.4% are made in the Global South. This imbalance reflects a historical trend, as other reviews on AI in education have matched with it an overrepresentation of models found in the West, such as the U.S. and Europe (Bozkurt et al., 2021; Guan et al., 2020). However, given the underrepresentation of Southern scholarship in the collection, it is essential to not overlook other linked publications such as the cloud learning applications and university education agility (Aljawarneh, 2024a),

business intelligence and innovative performance (Aljawarneh, 2024b), and bibliometric analysis of accounting and corporate sustainable research (Aljawarneh et al., 2025). Overall, the thematic breadth of these studies specifies the importance of research on varying educational, technological and organizational contexts, while automatically not establishing geographical inequalities in AI education research. Additionally, our network analysis shows that only 3.9% of publications are authored by Global South institutions alone, with 78.7% of publications connecting partners from the Global North. The patterns do indicate a possible dependency and unequal opportunities for independent knowledge production; however, such frequencies of collaboration do not necessarily be a suggestion of extractive partnerships and nor do they determine where the analytical control, theoretical framing, and authorship reside. Technological capacity, among others, should thus also be incorporated into a wider study of equitable collaboration, as would leadership and organisational conditions and participation by users. Some areas of inquiry that are applicable to this paper include user satisfaction in cloud supply chain management (Aljawarneh et al., 2022); fog computing-based logistics and organizational agility (Aljawarneh et al., 2021); and digital leadership in the kin between IT audit quality and digital business performance (Mohammad Aljawarneh et al., 2026). Organisational topics, such as employee silence and psychological empowerment (Albdareen et al., 2025), as well as knowledge hiding, employee cynicism, and job search behavior (Aljawarneh & Alshboul, 2026), also offer insights on issues to deliberate when looking at who should express, share, and be recognized for their knowledge within collaborative settings. These connections do not necessarily constitute evidence of neocolonial research practices but rather submit complementary perspectives. But the implications for epistemologies of geographic concentration are profound: In Asian, African, and Latin American contexts, AI systems risk reinforcing individualistic, competitive, and standardized pedagogical framings, and downplaying collectivist, community-culture-based, and Indigenous knowledge systems when research agendas are largely defined by Western institutions (Bozkurt et al., 2021). Funding systems must prioritize research in the Global South, while editorials must expand research to be represented in different geographical areas and collaboration systems must include fair sharing of agendas, transparent authorship, locally grounded interpretation, and continuing research capacity.

5.4. Thematic Evolution: From Technical Optimization to Sociotechnical and Ethical Imperatives

Thematic evolution analysis commands to light a deep paradigmatic shift over three different eras of recognition of the sociotechnical nature of AI. In the Foundational Period (2004-2010), there is a shift to focus on technical optimization (algorithm fidelity) and with less emphasis on pedagogy, or ethics, (Roll and Wylie, 2016). The target point in the Analytics Period (2011-2017) was to adopt data-intensive methods, such as Educational Data Mining (EDM) and Learning Analytics (LA), to predict students' learning outcomes (Siemens & Baker, 2012). However, some suggest that in this period the focus shifted from the interpretation to the forecast, making governmentality (interpretation of governmentality) 'algorithmic' (Williamson & Eynon, 2020).

5.5. The XAI-Bibliometric Framework: Methodological Innovation and Inherent Limitations

On the other hand, legacy paradigms – like traditional ITS – negatively predict emergence, and positive interactions with XAI indicate that these approaches might be revitalized by fresh transparency mechanisms. However, there are inherent limitations that need to be recognized in calls for rigorous academic enquiry. Firstly, SHAP is for explanation, not causal mechanisms and is susceptible to 'black swan' events (Lundberg & Lee, 2017). Secondly, our 37 bibliometric features are exhaustive, but they cannot capture qualitative aspects of theoretical sophistication, pedagogical relevance, etc. Third, SHAP values are model-dependent, but we addressed this with proper model validation and cross-model comparison (Ribeiro et al., 2016; Lakkaraju et al., 2016). This XAI-bibliometric approach thus should not be observed as a substitute for, but rather as a complementation of, qualitative systematic reviews, critical discourse analyses, and ethnographic studies, thereby calling for a strong methodological pluralism.

5.6. Future Research Directions: Toward Responsible, Equitable, and Pedagogically Grounded AI

Fourthly, need to take longer term (multi-year) consequences of AI interventions on learning trajectories and institutional culture & teacher professional identity into account beyond short-term efficacy trials (Longitudinal

Impact Studies [LISs]). Lastly, future studies are necessary, with high priority on Global South Perspectives and Decolonial AI, which focuses on the needs of Indigenous knowledge systems, multilingual education and culturally appropriate design, so that educational priorities are based on local communities, not technocratic actors from elsewhere (Bozkurt et al., 2021).

6. Conclusion

This study introduced a new united model that combines bibliometric science mapping with Explainable AI (XAI) techniques for uncovering, explaining, and forecasting knowledge evolution patterns in digital education research based on dual databases (Scopus + WoS). By analyzing 2,424 documents published from 2004 to 2026, we have charted the exponential growth, spatial concentration, thematic evolution and intellectual structure of the field with methodological rigour like never before. We make 5 major contributions. Second, we have proven the dual-database imperative through demonstrating how Scopus and WoS provide complementary, non-redundant and thus necessary field mapping. Second, by consuming SHAP and LIME, descriptive bibliometrics was turned into an explanatory science, which found that characteristics such as algorithmic transparency, diversity of interdisciplinary references, and international collaboration were key factors shaping emerging themes of research. Third, we mapped the paradigm budge from technical optimizations towards ethical directions and emphasized the importance of sociotechnical responsibility. Fourth, our geographic equity analysis revealed extreme imbalances in knowledge production and argued for systemic decolonial interventions. Finally, we founded a methodological blueprint to show how the interpretability of machine learning can be made to counteract the complex non-linear dynamics of scientific evolution.

The conclusion can be appealed that our results give an insight into a technologically, morally, globally, and geographically developing field where quality is curtailed with criteria. Algorithms and computing power are not the drivers for AI in education. It will convey consensus on content of learning, our research objects and the values to be learnt. It is critical that the technologies being fulfilled are not for market efficiency or technological solutionism, but for human flourishing, participatory democratic values and pedagogy. This study is not denoted to give a conclusive conclusion but stimulate the global research community to think towards implementing AI use in education as something just, equitable and humanistic.

Acknowledgment

The authors would like to express their thanks to the Deanship of Research—Jadara University for its financial support of this publication.

References

- Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138-52160.
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975.
- Bozkurt, A., Karadeniz, A., Baneres, D., Guerrero-Roldán, A. E., & Rodríguez, M. E. (2021). Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century. *Sustainability*, 13(2), 800.
- Bond, M. (2024). The international journal of educational technology in higher education: Content and authorship analysis 2010–2024. *International Journal of Educational Technology in Higher Education*, 21(1), 60.
- Bond, M., Marín, V. I., Dolch, C., Bedenlier, S., & Zawacki-Richter, O. (2018). Digital transformation in German higher education: student and teacher perceptions and usage of digital media. *International Journal of Educational Technology in Higher Education*, 15(1), 48.

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.

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., ... & Wamba, S. F. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 66, 102542.

Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *International Journal of Innovation Studies*, 4(4), 134-147.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education promises and implications for teaching and learning*. Center for Curriculum Redesign.

Joshi, M., Kapila, B., Rani, N., & Sagar, N. A. (2026). History, fundamentals, and evolution of artificial intelligence. In *Artificial Intelligence Interventions in Agriculture, Food and Health Sectors* (pp. 1-21). Academic Press.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., ... & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074.

Lakkaraju, H., Bach, S. H., & Leskovec, J. (2016, August). Interpretable decision sets: A joint framework for description and prediction. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1675-1684).

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.

Moral Muñoz, J. A., Herrera-Viedma, E., Santisteban Espejo, A. L., & Cobo Martín, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *El Profesional de la Información*, 29(1).

Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27(6), 7893-7925.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372.

Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 22.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). "Why should i trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26(2), 582-599.

Siemens, G., & Baker, R. S. D. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254).

Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2023). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, 31(4), 2134-2152.

- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15.
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.
- Van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In *Measuring scholarly impact: Methods and practice* (pp. 285-320). Cham: Springer International Publishing.
- Van Eck, N. J., & Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, 111(2), 1053-1070.
- Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252, 124167.
- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223-235.
- Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology*, 45(2), 107-114.
- Xie, Q., & Waltman, L. (2025). A comparison of citation-based clustering and topic modeling for science mapping. *Scientometrics*, 130(5), 2497-2522.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators?. *International Journal of Educational Technology in Higher Education*, 16(1), 39.
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, 2021(1), 8812542.
- Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429-472.
- Al Qudah, S. M. A., Fuentes Bargues, J. L., Ferrer Gisbert, P., & Al-Abdallat, H. N. E. (2025). AI-Driven Risk Management to Promote SDGs? An Exploratory Study in Jordan Construction Sector. *Sustainable Development*, 33, 1107-1123.
- Qudah, H., Malahim, S., Airout, R., Alomari, M., Hamour, A. A., & Alqudah, M. (2023). Islamic finance in the era of financial technology: A bibliometric review of future trends. *International Journal of Financial Studies*, 11(2), 76.
- Aladayleh, K. J., Al Qudah, S. M. A., Bargues, J. L. F., & Gisbert, P. F. (2023). Global trends of the research on COVID-19 risks effect in sustainable facility management fields: a bibliometric analysis. *Engineering Management in Production and Services*, 15(1), 12-28.
- Al-Qudah, S., Bargues, J. L. F., & Gisbert, P. S. F. (2026). A Methodological Proposal for Risk Management in Construction Projects in Small and Medium-Sized Enterprises in Emerging Economies: The Case of Jordan.
- Abdallah Al Qudah, S. M., Abdallah Abdelrahman Al Hurr, H., Fuentes Bargues, J. L., Gisbert, P. F., & Al-Abdallat, H. N. E. (2026). Strategic integration of smart management tools and risk management for sustainable development. *EDPACS*, 71(3), 1-13.

Alqudah, H., Al Qudah, M. Z., Abu Huson, Y., Lutfi, A., Alrawad, M., & Almaiah, M. A. (2024). A decade of green economic literature: an analysis-based bibliometric. *International Journal of Energy Economics and Policy*, 14(3), 497-511.

ALShanti, A. M., Al-Azab, H. A. H., Humeedat, M. M., & AlQudah, M. Z. (2024). Exploring the evolution of creative accounting and external auditors: Bibliometric analysis. *Cogent Business & Management*, 11(1), 2300500.

Almnadheh, Y., Samara, H., & AlQudah, M. Z. (2026). Enhancing ESG integration in corporate strategy: a bibliometric study and content analysis. *International Journal of Law and Management*, 68(4), 676-694.

Samara, H. H., Qudah, H. A., Mohsin, H. J., Abualhijad, S., Bani Hani, L. Y., Al Rahamneh, S., & AlQudah, M. Z. (2025). Artificial intelligence and machine learning in corporate governance: A bibliometric analysis. *Human Systems Management*, 44(2), 349-375.

Alqudah, H., Lutfi, A., Al Qudah, M. Z., Alshira'h, A. F., Almaiah, M. A., & Alrawad, M. (2023). The impact of empowering internal auditors on the quality of electronic internal audits: A case of Jordanian listed services companies. *International Journal of Information Management Data Insights*, 3(2), 100183.

AlQudah, M. Z., & Bariviera, A. F. (2026). Systematic and bibliometric reviews of cryptocurrency market regulation: Trends, influential contributions, and future directions. *Journal of Financial Regulation and Compliance*, 34(1), 1-37.

Albdareen, R., Aljawarneh, N. M., Al-Jedaiah, M. N., Alomari, K. A. K., & Alrousan, A. (2025). Organizational factors affecting employee silence behavior: Does psychological empowerment matter? **Cogent Business & Management*, 12*(1), Article 2512819.

Aljawarneh, N. M. (2024a). Effect of user satisfaction between cloud learning applications & university education agility. **Journal of Statistics Applications & Probability*, 13*(3), 871–879.

Aljawarneh, N. M. (2024b). The mediating role of organization agility between business intelligence & innovative performance. **Journal of Statistics Applications & Probability*, 13*(3), 929–938.

Aljawarneh, N. M., & Alshboul, K. J. (2026). Organizational factors effect knowledge hiding and job search behavior: The mediating role of employee cynicism. In **2026 6th International Conference on Technology Enhanced Learning in Higher Education (TELE)** (pp. 132–138). IEEE.

Aljawarneh, N., Alqmool, T., Huson, Y. A., Jarbou, S., & Alqudah, M. (2025). Bibliometric analysis of accounting and corporate sustainability research: Trends and insights. **ABAC Journal*, 45*(2), 22.

Aljawarneh, N. M., Kader Alomari, K. A., Alomari, Z. S., Taha, O., & Obeidat, A. M. (2022). Cloud supply chain management and customer service: The mediating role of user satisfaction. **Astra Salvensis*, 10*(1), 183.

Aljawarneh, N., Taamneh, M., Alhndawi, N., Alomari, K., & Masad, F. (2021). Fog computing-based logistic supply chain management and organizational agility: The mediating role of user satisfaction. **Uncertain Supply Chain Management*, 9*(3), 767–778.

Mohammad Aljawarneh, N., Abu Huson, Y., Ahmad Faleh Alazzam, F., Abdel Kader Alomari, K., & Almarashdi, F. (2026). IT audit quality and digital business performance: The moderating role of digital leadership. **EDPACS*, 71*(5), 71–81.