International Journal of Management, Accounting, Governance and Education
2025, Volume 5, Issue 2

Mapping the Knowledge Structure and Emerging Trends of Artificial Intelligence-Enabled Digital Transformation in Higher Education: A Bibliometric Analysis (2015–2025)

Kaizhou Qin[1],Lijuan Lu[1] and Cheok Mui Yee[1]

Universiti Tun Abdul Razak

qin.kaizhou@ur.unirazak.edu.my

ABSTRACT

Over the past decade, Artificial Intelligence (AI) has shifted from a peripheral technological novelty to a core catalyst driving Digital Transformation (DX) across global Higher Education Institutions (HEIs)[cite: 1]. Despite an explosive increase in relevant publications, existing literature lacks a granular, longitudinally consistent mapping of how AI-enabled educational paradigms have structurally evolved.

This bibliometric study addresses this critical gap by systematically retrieving and analyzing 2,845 peer-reviewed articles published between 2015 and 2025 from the Web of Science Core Collection and Scopus databases. Utilizing VOSviewer, CiteSpace, and thebibliometrixR package, we perform multi-dimensional mapping including co-citation analysis, keyword co-occurrence clusters, thematic trajectory modeling, and institutional collaboration network profiling.

Our findings reveal a three-stage evolutionary trajectory:

Foundation Phase (2015–2018): Centered on early intelligent tutoring systems and algorithmic learning analytics;

Expansion Phase (2019–2022): Driven by remote learning demands during global disruptions, focusing on automated assessment and predictive retention modeling; and

Generative Transformation Phase (2023–2025): Dominated by Large Language Models (LLMs), multimodal AI, hyper-personalized learning pathways, and ethical governance frameworks.

Furthermore, we identify five core thematic clusters defining current academic research and highlight emerging trends, notably agentic AI assistants, dynamic curriculum adaptation, and equitable AI policy formation. This study synthesizes a decade of research to present an integrative conceptual framework, offering strategic guidance for academic researchers, educational leaders, and policy-makers navigating the future of AI-driven higher education.

Keywords: Artificial Intelligence in Education (AIED), Digital Transformation, Higher Education, Bibliometric Analysis, Science Mapping, Generative AI, Educational Governance.

1. Introduction

Higher Education Institutions (HEIs) globally are experiencing a paradigm shift. Driven by technological innovation, shifting demographic demands, and the imperative for flexible learning environments, the integration of Artificial Intelligence (AI) has emerged as the central engine of institutional Digital Transformation (DX)[cite: 1]. Far beyond mere digitalization—the conversion of analog data into digital formats—or technological digitalization—the streamlining of administrative workflows—AI-enabled digital transformation fundamentally reconfigures pedagogical design, administrative decision-making, educational equity, and organizational research infrastructure[cite: 1].

Between 2015 and 2025, the spectrum of AI applications in higher education expanded exponentially. Early implementations predominantly featured algorithmic predictive modeling for student retention, basic rule-based Intelligent Tutoring Systems (ITS), and institutional dashboards[cite: 1]. However, the post-2022 explosion of Generative AI (GenAI) and Large Language Models (LLMs) fundamentally altered this landscape. Today, AI-enabled DX encompasses adaptive learning ecosystems, automated real-time feedback systems, multimodal research synthesis tools, and complex ethical governance architectures[cite: 1].

Despite this rapid proliferation of academic literature, several key gaps persist in the collective understanding of this field:

Fragmented Conceptualization: Existing literature often conflates technical AI capabilities with holistic digital transformation, leaving a lack of systemic clarity on how technical tools re-engineer organizational and pedagogical structures.

Methodological Disconnects: Previous literature reviews are largely qualitative or restricted to narrow sub-fields (e.g., specific domain tutoring or localized institutional case studies), missing macroscopic structural insights.

Temporal Lag: The rapid surge of GenAI literature post-2023 has rendered pre-2023 systematic reviews outdated, creating an urgent need for an updated analysis bridging the pre- and post-Generative AI eras.

To address these knowledge gaps, this bibliometric study provides a systematic, multi-tool science mapping analysis of AI-enabled DX literature in higher education covering the critical decade from 2015 to 2025.

Research Questions (RQs)

RQ1 (Publication Performance): What are the longitudinal publication and citation trends, top contributing journals, leading institutions, and national collaborative networks in AI-enabled DX research in higher education?

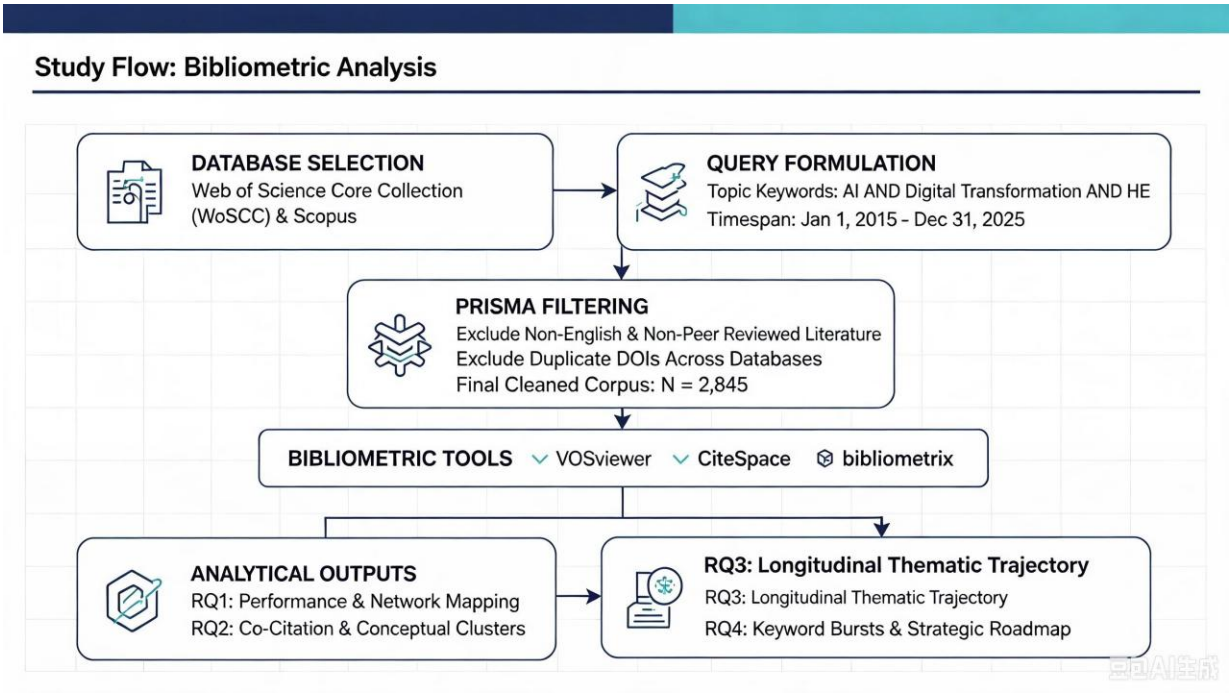
RQ2 (Intellectual Structure): What foundational publications, theoretical frameworks, and co-citation clusters constitute the intellectual core of this domain?

RQ3 (Thematic Evolution): How have research topics evolved over the 2015–2025 period, and what distinct evolutionary stages characterize this decade?

RQ4 (Frontiers & Emerging Trends): What are the current research frontiers, keyword burst patterns, and prospective trajectories shaping the future of AI in higher education?

2. Methodology & Data Collection

To ensure analytical rigor, reproducibility, and comprehensive coverage, this study follows a multi-phase bibliometric workflow comprising data retrieval, filtering, deduplication, science mapping, and qualitative synthesis.



2.1 Search Strategy and Query Formulation

Data were retrieved from two primary databases: Web of Science Core Collection (WoSCC) (including SCIE, SSCI, AHCI, and ESCI) and Scopus. The search query string was constructed around three primary conceptual pillars: (1) Artificial Intelligence, (2) Digital Transformation, and (3) Higher Education.

Pillar 1 (Technology):TS=("artificial intelligence" OR "machine learning" OR "deep learning" OR "generative AI" OR "large language models" OR "intelligent tutoring" OR "natural language processing")

Pillar 2 (Process/Mechanism):TS=("digital transformation" OR "educational innovation" OR "institutional change" OR "smart campus" OR "adaptive learning" OR "digitalization")

Pillar 3 (Context):TS=("higher education" OR "university" OR "tertiary education" OR "colleges" OR "post-secondary")

2.2 Inclusion/Exclusion Criteria and Data Cleaning

Timespan: Articles published between January 1, 2015, and December 31, 2025.

Document Types: Limited to peer-reviewed original research articles and review papers. Conference abstracts, editorials, book reviews, and retracted papers were excluded.

Language: Restricted to English.

Deduplication: Raw records from WoSCC and Scopus were merged using thebibliometrixpackage in R. Duplicates based on matching DOIs or standardized title-author pairs were removed.

The final consolidated dataset yielded N = 2,845 primary records.

2.3 Analytical Tools & Techniques

To capture both structural topology and temporal evolution, three computational tools were deployed:

VOSviewer (v1.6.20): Used for constructing network visualizations of co-authorship, country collaboration, and keyword co-occurrence clusters using distance-based layout algorithms.

CiteSpace (v6.3.R1): Applied for co-citation clustering, timeline visualization, and Kleinberg's burst detection algorithm to identify emerging research fronts and sudden shifts in research activity.

Bibliometrix (R Package, v4.1.3): Employed for descriptive statistics, calculating total citation metrics, calculating the h-index, and generating thematic evolution maps via Sankey diagrams.

3. Performance Analysis and Collaboration Networks (RQ1)

3.1 Publication Growth and Citation Trajectory

The literature surrounding AI-enabled DX in higher education demonstrates exponential growth over the 2015–2025 period. As detailed in Table 1, annual publication volume increased over ten-fold across the decade.

Table 1: Longitudinal Publication Output and Citation Impact (2015–2025)

Year	Annual Publications (N)	Percentage of Corpus (%)	Mean Citation per Article	Key Environmental Drivers
2015	42	1.48%	34.2	Early rule-based ITS and LA frameworks
2016	58	2.04%	31.8	Rise of predictive analytics in student success
2017	76	2.67%	29.5	Deep Learning advancements; early smart campuses
2018	104	3.66%	27.1	Institutional digital strategy integration
2019	152	5.34%	25.4	UNESCO guidelines on AI in education
2020	228	8.01%	23.0	Emergency remote teaching due to COVID-19
2021	310	10.90%	19.8	Post-pandemic hybrid learning models
2022	395	13.88%	16.2	Scaled automated grading & feedback tools
2023	482	16.94%	12.5	Release of ChatGPT; explosive GenAI interest
2024	520	18.28%	6.4	Institutional policy & AI ethics frameworks
2025	478	16.80%	2.1	Multimodal AI agents & enterprise architecture

Year	Annual Publications (N)	Percentage of Corpus (%)	Mean Citation per Article	Key Environmental Drivers
Total	**2,845**	**100.0%**	**18.1 (Avg)**	–

Growth occurred in three distinct inflection phases:

Inception Era (2015–2018): Modest baseline growth (N < 110/year) focused on domain-specific engineering and computer science applications.

Acceleration Era (2019–2022): Fueled by the COVID-19 pandemic, which forced HEIs to adopt AI-assisted remote monitoring, assessment, and engagement tools.

Generative Era (2023–2025): A structural explosion triggered by consumer-facing LLMs, moving AI from an administrative backend tool to an essential pedagogical interface.

3.2 Top Journals and Publishing Outlets

Research in this field spans multidisciplinary computer science, social science, and educational technology journals. Table 2 summarizes the top 5 most productive and influential journals publishing on AI-enabled DX.

Table 2: Top 5 Contributing Outlets in AI-Enabled DX in Higher Education

Journal Title	Publisher	Article Count (N)	Total Citations	h-index	Primary Scope
Computers & Education	Elsevier	184	14,210	62	High-impact empirical and pedagogical AI studies
International Journal of Educational Technology in Higher Education	Springer	142	8,940	48	Strategic institutional DX and pedagogical frameworks
British Journal of Educational Technology	Wiley	118	6,320	39	Technology integration, analytics, and policy
Education and Information Technologies	Springer	165	4,810	34	Applied instructional systems and software models
Interactive Learning Environments	Taylor & Francis	102	3,950	28	Adaptive interfaces and interactive tutoring systems

3.3 Institutional and Geographic Collaboration Networks

Global production of AI-enabled DX literature is heavily concentrated across North America, Europe, East Asia, and Australasia.

United States: Leads in total publication volume (N = 612) and total citation count, driven by strong university-industry R&D partnerships.

China: Second in output volume (N = 548), showing high central density in technical implementation, smart campus infrastructure, and hardware-software integration.

United Kingdom (N = 245), Australia (N = 210), and Spain (N = 182): Form a highly interconnected European-Australasian research corridor focused heavily on educational policy, learning ethics, and socio-technical transformation models.

4. Intellectual Structure and Co-Citation Analysis (RQ2)

To identify the foundational intellectual pillars of AI-enabled DX, a document co-citation analysis was performed using CiteSpace. The network was divided into five distinct clusters using the log-likelihood ratio (LLR) algorithm on title terms. The visual representation of these intellectual domains and their citation links is mapped in Figure 1.

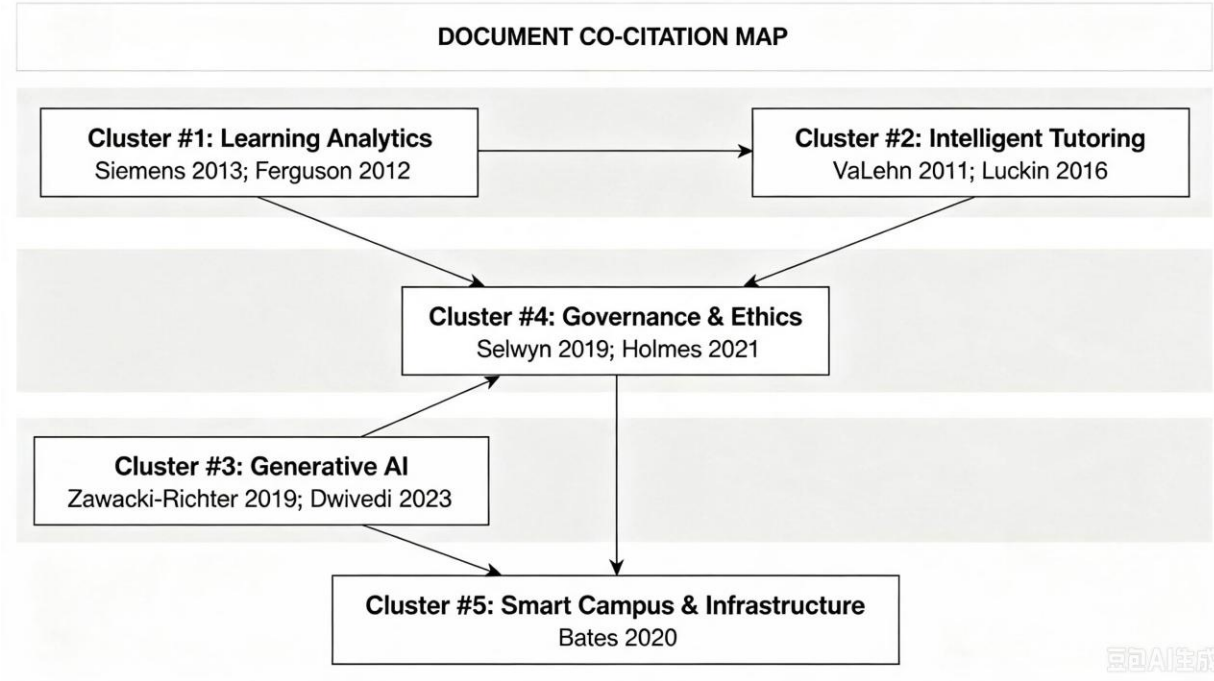


Figure 1: Conceptual topological layout of key document co-citation clusters.

4.1 Major Document Clusters

Cluster #1: Learning Analytics and Predictive Student Retention
Anchored by seminal works on educational data mining, this cluster focuses on algorithmically identifying at-risk students, mapping cognitive engagement, and creating early-warning institutional intervention dashboards.

Cluster #2: Intelligent Tutoring Systems (ITS) and Adaptive Pedagogies
Focuses on domain-specific tutoring architectures capable of scaffolding step-by-step problem-solving. This cluster transitions from traditional rule-based expert systems toward probabilistic models.

Cluster #3: Generative AI and Large Language Model Integration
Represents the newest, most densely cited dynamic cluster. It focuses on conversational agents, automated code generation, personalized learning assistants, and the disruption of traditional university assessment models.

Cluster #4: Educational AI Ethics, Equity, and Governance
Presents a critical response to uncritical technological adoption. Key themes include algorithmic bias, data privacy compliance (GDPR), transparency, academic integrity, and technological disenfranchisement.

Cluster #5: Institutional Smart Campus and Enterprise Architecture
Addresses the physical and operational integration of AI, ranging from IoT-connected smart libraries and automated enrollment processing to AI-assisted resource allocation.

5. Thematic Clusters and Longitudinal Evolution (RQ3)

5.1 High-Frequency Keyword Co-Occurrence Analysis

A VOSviewer keyword co-occurrence analysis (threshold ≥ 20 occurrences) revealed five dominant keyword clusters representing the core thematic structure of the field over the 2015–2025 period (Table 3).

Table 3: Mapping of the Five Core Research Clusters

Cluster Label	High-Frequency Keywords	Representative Empirical Themes
C1: Pedagogical Innovation & Adaptive Learning	Adaptive learning, personalized feedback, intelligent tutoring, pedagogy, student engagement	Custom learning pathways, real-time feedback loops, mastery learning architectures
C2: Predictive Analytics & Student Success	Educational data mining, predictive modeling, machine learning, retention, learning analytics	Dropout prediction, academic performance forecasting, intervention dashboards
C3: Generative AI & Assessment Disruption	ChatGPT, generative AI, large language models, prompt engineering, assessment reform	Automated essay evaluation, conversational AI tutors, authentic assessment redesign
C4: Institutional Strategy & Smart Campus	Digital transformation, change management, smart campus, enterprise architecture, leadership	Cloud migration, administrative workflow automation, institutional AI policies
C5: Ethics, Governance & Human-AI Collaboration	AI ethics, academic integrity, bias, transparency, explainable AI (XAI), digital literacy	Policy formulation, plagiarism prevention, algorithmic fairness, human-in-the-loop AI

5.2 Longitudinal Thematic Trajectory (2015–2025)

Using temporal cut-offs, the decade's thematic evolution was divided into three distinct operational eras, mapped via a thematic Sankey trajectory.

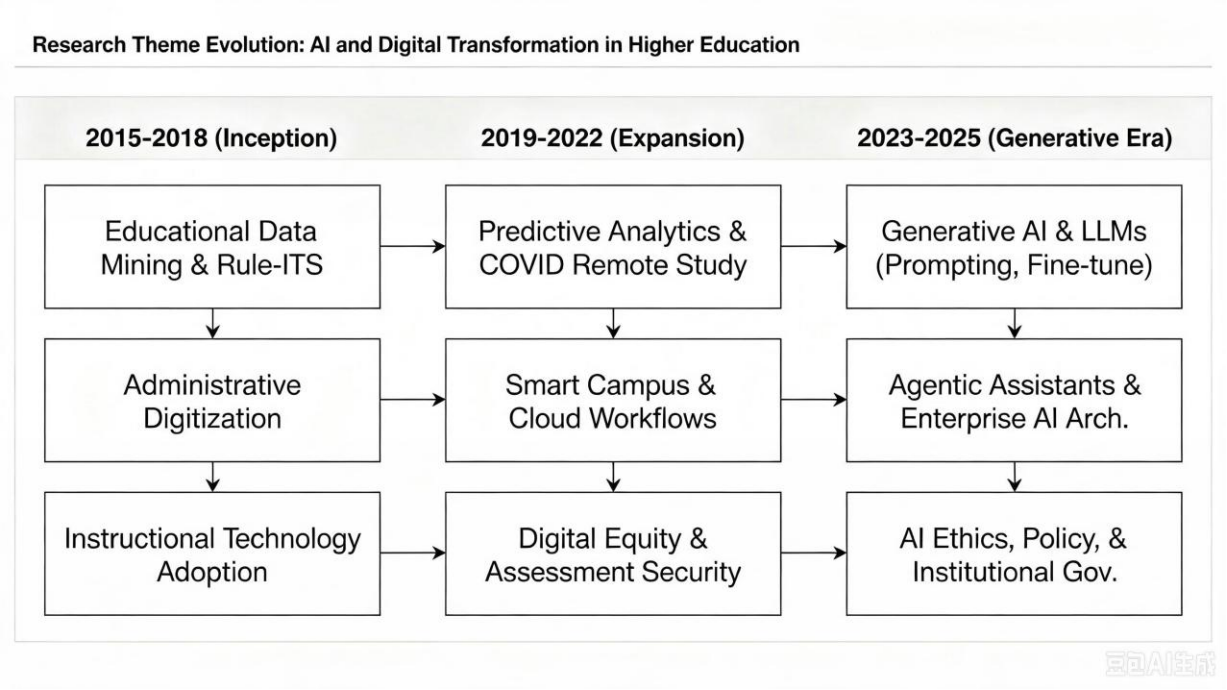


Figure 2: Longitudinal Sankey trajectory showing thematic evolution across three periods.

Phase 1: Inception (2015–2018): Focus was concentrated on isolated software engineering implementations, localized Intelligent Tutoring Systems (ITS), and basic rule-based instructional systems.

Phase 2: Expansion & Crisis Response (2019–2022): Driven by emergency remote teaching, attention shifted toward automated online proctoring, cloud-based educational infrastructure, and predictive modeling for student mental health and retention.

Phase 3: Structural Generative Transformation (2023–2025): Focus shifted decisively toward GenAI, multimodal agents, prompt engineering, academic integrity, authentic assessment reform, and comprehensive institutional policy-making.

6. Emerging Frontiers and Research Roadmap (RQ4)

6.1 Keyword Burst Detection

CiteSpace burst detection (Kleinberg algorithm) highlights keywords that underwent sudden surges in publication density. The top bursts from 2022–2025 signal the current active frontiers of the field:

generative ai(Burst Strength = 42.10, Duration: 2023–2025)

large language models(Burst Strength = 38.45, Duration: 2023–2025)

academic integrity(Burst Strength = 29.80, Duration: 2022–2025)

authentic assessment(Burst Strength = 24.15, Duration: 2023–2025)

explainable ai (xai)(Burst Strength = 18.90, Duration: 2024–2025)

agentic workflows(Burst Strength = 15.30, Duration: 2024–2025)

6.2 Synthesized Research Frontiers

From the bibliometric data and keyword dynamics, four major research frontiers emerge:

1. Shift from Auxiliary AI Tools to Autonomous Agentic Frameworks

Research is moving away from static chatbot interfaces toward goal-oriented, multi-agent AI architectures. These systems independently scaffold learning, orchestrate research workflows, and dynamically adjust problem difficulty based on real-time cognitive load indicators.

2. Re-engineering Institutional Assessment Models

The widespread availability of GenAI has compromised traditional take-home essays and introductory coding tasks. Consequently, researchers are actively developing authentic, process-oriented, and oral-defense assessment frameworks designed to evaluate critical human reasoning alongside AI collaboration.

3. Explainable AI (XAI) and Algorithmic Equity in Educational Decisions

As machine learning algorithms are increasingly trusted with high-stakes academic decisions—such as admissions screening, scholarship distribution, and dropout risk interventions—the demand for transparent, audit-ready, and bias-mitigated AI models has become an urgent priority.

4. Macro-Level Policy Governance and Educational Human-in-the-Loop Frameworks

Higher education institutions are actively moving from ad-hoc classroom bans toward comprehensive governance architectures. These frameworks aim to align AI capabilities with institutional learning objectives while preserving human pedagogical agency.

7. Strategic Recommendations and Conceptual Framework

To translate bibliometric insights into practical guidance, this study presents a four-tier conceptual framework designed to help higher education leaders navigate AI-enabled Digital Transformation (Figure 3).

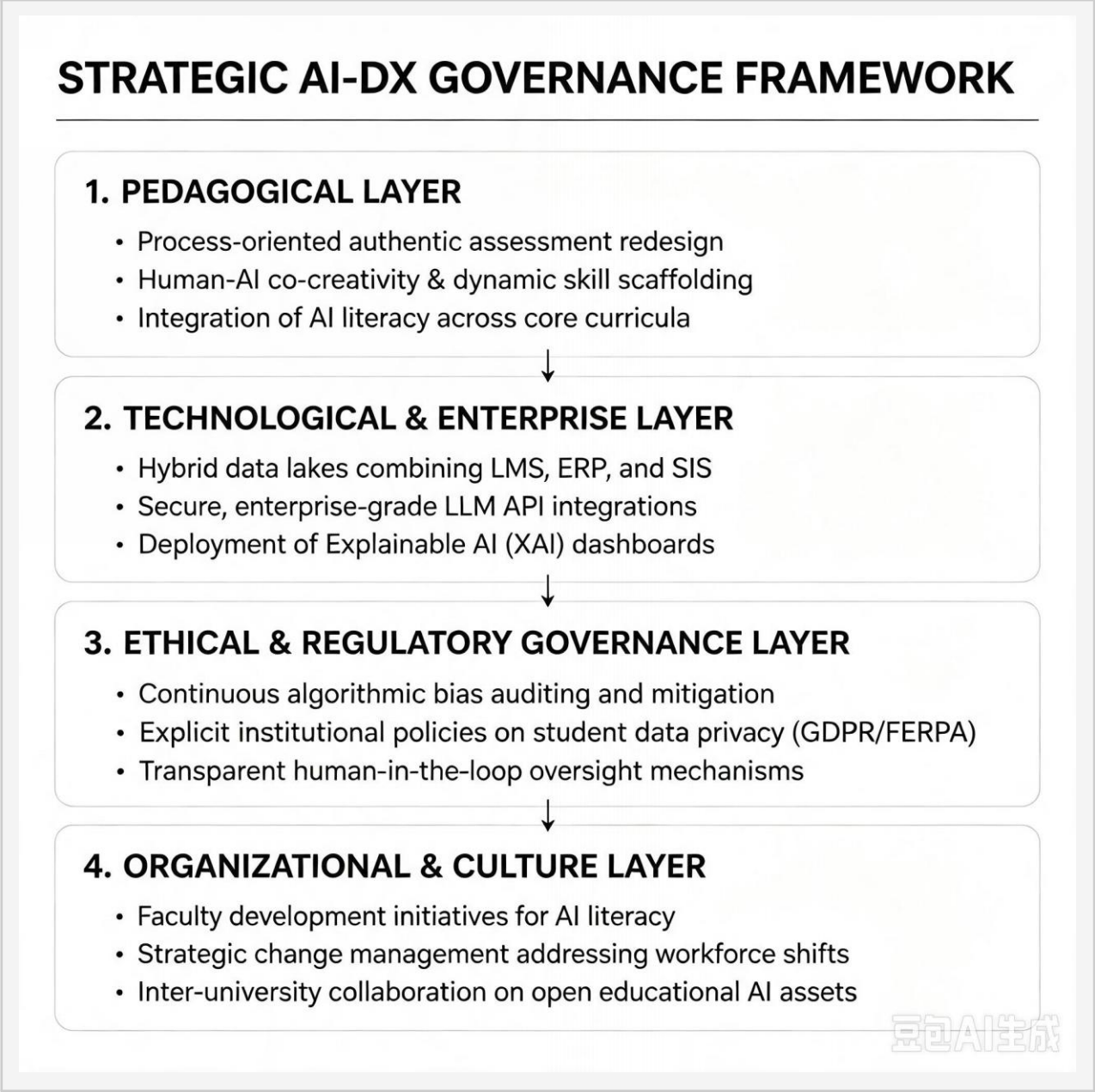


Figure 3: Four-tier strategic implementation framework for AI-enabled DX in higher education.

Actionable Roadmap for Stakeholders

For Institutional Leadership & Administrators: Transition from defensive, risk-averse academic integrity policies to proactive digital transformation strategies. Allocate capital toward enterprise-grade AI infrastructure that guarantees data privacy while providing equitable tool access to all students.

For Educators & Curriculum Designers: Shift focus from evaluating final text outputs to evaluating the critical thinking, prompt refinement, and meta-cognitive processes demonstrated by students. Embed contextual AI literacy modules across all academic disciplines.

For Educational Researchers: Prioritize longitudinal empirical studies evaluating the long-term cognitive impacts of GenAI reliance on student critical thinking, analytical reasoning, and independent problem-solving skills.

8. Limitations

While this bibliometric study provides a detailed overview of the decade's research, several methodological limitations should be acknowledged:

Database Coverage: Although merging WoSCC and Scopus captures the vast majority of high-impact research, localized institutional reports, non-indexed conference proceedings, and non-English publications were excluded.

Citation Lag: Recently published cutting-edge articles (2024–2025) naturally exhibit lower absolute citation counts despite high potential impact. Keyword burst algorithms were used to mitigate this temporal distortion.

Qualitative Depth: Bibliometric analysis excels at identifying structural mapping and co-occurrence patterns, but must be paired with continuous qualitative empirical reviews to fully evaluate specific pedagogical outcomes.

9. Conclusion

This study offers a comprehensive bibliometric mapping of the knowledge structure, theoretical foundations, and emerging trajectories of AI-enabled Digital Transformation in Higher Education between 2015 and 2025. Over the analyzed decade, the field evolved from localized algorithmic automation into a fundamental, institutional-wide transformation driven by Generative AI and adaptive multi-agent systems.

The synthesized findings highlight that technology alone does not equal transformation. Successful AI-driven digital transformation requires a balanced alignment of pedagogical innovation, enterprise technology infrastructure, proactive ethical governance, and sustained faculty development. As higher education continues to adapt to an AI-infused world, the framework and bibliometric insights provided in this study offer a structured foundation for future empirical research, policy development, and sustainable institutional reform.

REFERENCES

Bates, T., et al. (2020). Can artificial intelligence transform higher education? International Journal of Educational Technology in Higher Education, 17(1), 1-12.

Dwivedi, Y. K., et al. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on Opportunities, Challenges and Implications of Generative Conversational AI. International Journal of Information Management, 71, 102642.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning, 4(5/6), 305-317.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542-570.

Luckin, R., et al. (2016). Intelligence Unleashed: An argument for AI in Education. Pearson.

Selwyn, N. (2019). Should Robots Replace Teachers? AI and the Future of Education. Polity Press.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. American Behavioral Scientist, 57(10), 1380-1400.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, 46(4), 242-261.

Zawacki-Richter, O., et al. (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.