ISSN 2409-2665
Journal of Logistics, Informatics and Service Science
Vol. 13 (2026) No. 7, pp. 44-59
DOI:10.33168/JLISS.2026.0704

Mapping Smart Education as a Digital Service System: A Bibliometric and Dynamic Topic Graph Analysis

Xiao Zhang[1,2]*, Jinghong Low[1]
[1]Faculty of Management, Multimedia University, Selangor, Malaysia, 631000
[2]School of Law, History and Culture, Yulin University, Yulin, Shaanxi, China,719000
zhang8xiao89@163.com

Abstract. This study maps the intellectual structure and evolutionary dynamics of smart education research from the perspective of digital transformation and digital service systems. To address the fragmentation of smart education research across education, computer science, information systems, and management, the study proposes a Digital-Transformation-Driven Dynamic Topic Graph (DT-DTG) framework that integrates semantic similarity with bibliographic coupling to support coherent topic discovery and longitudinal trend tracking. Using a bibliometric visual analytics approach, the framework identifies publication growth, collaboration structures, major thematic clusters, and emerging research frontiers in smart education. The findings indicate accelerating publication growth, a core–periphery international collaboration structure, and a clear thematic transition from online and blended learning toward analytics-driven, AI-enabled, and governance-oriented smart education research. Immersive learning and generative AI emerge as the most recent frontiers. Comparative evaluation further shows that DT-DTG improves topic coherence, diversity, and temporal stability relative to conventional text-based topic models. By conceptualizing smart education as an evolving digital service system, this study contributes to the literature at the intersection of informatics, service science, and digital transformation, and provides a structured analytical framework for tracking how intelligent educational services are designed, governed, and scaled over time.

Keywords: Digital Transformation; Smart Education; Digital Service Systems; Bibliometric Visual Analytics; Topic Evolution; Informatics

1. Introduction

Digital transformation is reshaping how education systems design curricula, deliver instruction, and govern data and technology infrastructures. In many countries, education has entered a phase in which cloud platforms, data governance, analytics pipelines, and AI services are becoming part of the standard infrastructure for teaching and learning (Vial, 2019). This shift is not merely technological; it reorganizes pedagogical processes, organizational routines, and policy mechanisms by enabling continuous data feedback, cross-platform integration, and scalable digital services. From an informatics and service science perspective, digital transformation changes how value is created and measured in educational organizations, moving from isolated digitalization initiatives to end-to-end redesign of learning services, governance, and decision making (Wang et al., 2024a). At the same time, the rise of ubiquitous connectivity and sensing, intelligent tutoring, and learning analytics has intensified the demand for trustworthy data ecosystems and interoperable architectures that can support personalization without sacrificing equity and accountability (Jing et al., 2023).

In this context, smart education has emerged as a unifying concept that integrates data-driven decision making, AI-enabled personalization, ubiquitous connectivity, and human-centered pedagogies to improve learning effectiveness and equity (Zhu et al., 2016). Smart education emphasizes not only smart technologies but also the orchestration among learners, teachers, content, and environments, where data and intelligent services support adaptive learning paths, timely intervention, and evidence-based teaching improvement. Understood through the lens of service science, smart education can be conceptualized as an evolving digital service system in which learning platforms, analytics engines, and intelligent agents operate as interconnected service components. From a JLISS perspective, smart education can be understood as a digital service system composed of multiple service actors, interoperable digital platforms, analytics engines, governance structures, and continuously generated data flows. In this system, learners, teachers, institutions, platform providers, and policy actors jointly participate in value creation through platform-mediated interactions, data-driven feedback, and intelligent service delivery. This perspective shifts the analytical focus from isolated educational technologies to the coordinated design, governance, and scaling of technology-enabled learning services. Meanwhile, learning theories and digital pedagogy have also evolved: learners increasingly interact with knowledge in networked, participatory, and rapidly changing ecosystems, which challenges traditional instructional designs and calls for new models that connect learning analytics, learner agency, and competency-oriented outcomes (Bahroun et al., 2023; Supriya et al., 2024).

Despite rapid growth, the smart education research landscape remains fragmented across disciplines such as education, computer science, information systems, management, and engineering. This fragmentation produces at least three practical difficulties. First, terminologies vary widely (e.g., "smart learning," "intelligent education," "Education 4.0," "learning analytics," "AI in education"), making it hard to build a stable knowledge structure from text alone (Shi et al., 2022; Zhu et al., 2016). Second, interdisciplinary publications differ in citation practices and research designs, which may cause citation-only networks to overemphasize disciplinary silos (Jing et al., 2024). Third, the field is characterized by strong temporal dynamics: waves of online/blended learning, analytics, XR/metaverse, and most recently generative AI can rise rapidly, often triggered by policy cycles, platform diffusion, or external shocks (Wong et al., 2024; Brika et al., 2022). These challenges make it difficult to answer practical questions such as: Which themes form the core intellectual structure? How do research frontiers evolve over time? Which countries and institutions lead collaboration, and where are the emerging hubs?

Bibliometric visual analytics provides a systematic way to map the intellectual structure of a field by combining citation networks, co-authorship patterns, and text mining of titles, abstracts, and keywords. Tools and workflows such as co-citation analysis, bibliographic coupling, keyword co-occurrence mapping, and burst detection have been widely used to identify influential works, collaboration structures, and thematic clusters (Chen, 2006; van Eck and Waltman, 2010; Donthu et al., 2021). However, existing bibliometric studies in adjacent domains often face two methodological limitations when applied to interdisciplinary smart education corpora. First, many analyses rely on a single similarity signal—either textual similarity (topic models) or citation-based similarity (co-citation/bibliographic coupling). Text-only topic models such as LDA capture semantic regularities but can mix heterogeneous terminologies, especially when the corpus spans multiple disciplines (Blei et al., 2003; Chen et al., 2022). Similarly, citation-only clustering can miss semantically close but citation-distant works (e.g., new AI methods not yet widely cited). Second, many studies treat topics as static snapshots and report clusters without robust longitudinal tracking, which limits their ability to describe the temporal dynamics inherent in digital transformation (Huang et al., 2022; Shi et al., 2022).

To address these gaps, this paper proposes a Digital-Transformation-Driven Dynamic Topic Graph (DT-DTG) framework that integrates semantic similarity with bibliographic coupling and tracks topics longitudinally. Specifically, DT-DTG constructs a hybrid document similarity graph that fuses semantic embeddings (from titles, abstracts, and keywords) and bibliographic coupling (shared references). It then performs community detection on time-sliced graphs to discover coherent topic communities, followed by topic labeling and inter-slice topic matching to generate a directed topic evolution graph. This hybrid design is particularly suitable for interdisciplinary corpora such as smart education because it reduces terminological fragmentation across education, computer science, information systems, and management while preserving relational signals that reflect shared knowledge foundations and evolving service-system concerns. The main contributions of this study are as follows:

(1) Hybrid similarity graph for coherent topic discovery. By combining semantic similarity with bibliographic coupling, DT-DTG reduces incoherent term mixtures and improves topic coherence compared with text-only baselines such as LDA and NMF.

(2) Dynamic topic tracking for longitudinal frontier detection. DT-DTG links topics across time slices, enabling the identification of stable trajectories, splitting and merging patterns, and emerging frontiers (e.g., immersive learning and generative AI).

(3) Informatics-oriented analytical support for digital service-system intelligence. The framework provides interpretable visual and quantitative outputs that support strategic scanning, service innovation monitoring, and digital governance intelligence in fast-changing smart education knowledge domains.

2. Literature Review

This section reviews relevant research in three streams: (i) bibliometric and science mapping approaches used to synthesize large-scale scholarly corpora, (ii) smart education and digital transformation research that defines the substantive domain, and (iii) topic modeling and hybrid graph-based methods that underpin the analytical framework of this study.

2.1 Bibliometric and Science Mapping Approaches

Bibliometric analysis has become one of the most widely adopted methods for mapping intellectual structures and tracking research evolution in fast-growing interdisciplinary fields (Donthu et al., 2021). Typical bibliometric workflows include co-authorship network analysis to reveal collaboration structures, co-citation analysis to identify intellectual bases, keyword co-occurrence mapping to detect thematic clusters, and bibliographic coupling to connect documents sharing common reference foundations (Kessler, 1963; Small, 1973). Visualization tools such as VOSviewer (van Eck and Waltman, 2010) and CiteSpace (Chen, 2006) have become standard instruments for generating interpretable science maps, enabling researchers to detect burst keywords, trace citation frontiers, and cluster research themes at scale.

Within educational technology and related domains, bibliometric studies have proliferated rapidly. Jing et al. (2024) conducted a systematic review of bibliometric mapping techniques in educational technology research, identifying co-citation analysis, keyword co-occurrence, and bibliographic coupling as the three most frequently employed approaches while noting that many studies lack temporal depth and methodological transparency. Wang et al. (2024a) performed a global-perspective bibliometric study on education reform driven by digital technology, showing that research output accelerated markedly after 2018 and that China and the United States are the dominant contributors. Rojas-Sánchez et al. (2023) combined systematic literature review and bibliometric analysis to map the virtual reality and education field, revealing a shift from technology-centric studies toward pedagogical design and learner experience evaluation. Similar bibliometric syntheses have been applied to specific sub-domains: Brika et al. (2022) analyzed e-learning research trends during COVID-19, Jing et al. (2023) mapped the research landscape of adaptive learning, and Zou et al. (2022) traced the evolution of TPACK-related empirical research. Bahroun et al. (2023) provided a comprehensive bibliometric and content analysis of generative AI in educational settings, identifying a sharp upward trend in publication volume post-2022. Jia et al. (2024) examined a decade of AI in science education research, finding that learning analytics and intelligent tutoring systems emerged as dominant themes from 2018 onward.

Despite these contributions, several limitations persist. First, many bibliometric studies in education rely predominantly on keyword co-occurrence or co-citation as a single similarity signal (Jing et al., 2024). This can be problematic in interdisciplinary corpora where the same concept is described by different terminologies across communities (e.g., "smart learning," "intelligent education," and "Education 4.0" may refer to overlapping constructs). Second, most studies report static cluster snapshots at one or a few time points, without systematically modeling how topics emerge, merge, split, or decline over time (Shi et al., 2022; Yang et al., 2022). Third, the connection between bibliometric mapping and the broader discourse on digital service systems and informatics remains underexplored, even though smart education is increasingly conceptualized as a digitally mediated service ecosystem (Dima et al., 2023).

2.2 Smart Education and Digital Transformation

Smart education has emerged as a unifying concept integrating data-driven decision making, AI-enabled personalization, ubiquitous connectivity, and human-centered pedagogies to improve learning effectiveness and equity (Zhu et al., 2016). Unlike earlier framings centered on e-learning infrastructure, smart education emphasizes the orchestration among learners, teachers, content, and environments, where intelligent services support adaptive learning paths, timely intervention, and evidence-based teaching improvement. Demonstrating the practical application of such intelligent services, Husaini et al. (2025) developed a deep learning model to predict student academic success, proving that behavioral engagement and prior academic records are the most significant factors for enabling early and targeted interventions. Shi et al. (2022) analyzed the research hotspots and evolutionary trends of intelligent education from a knowledge graph perspective, showing that intelligent tutoring systems, adaptive learning, and smart classrooms are increasingly prominent research fronts. Supriya et al. (2024) reviewed the intersection of Industry 5.0 and smart education, identifying human-AI collaboration, cyber-physical learning environments, and sustainable educational practices as emerging concerns. Ma et al. (2024) investigated teacher-student interaction modes in smart classroom settings, demonstrating that technology-mediated interaction structures differ meaningfully from traditional classroom patterns. Alzubi et al. (2025) proposed a multimodal human-computer interaction framework for smart learning systems, pointing to the importance of integrating natural language processing, gesture recognition, and affective computing in next-generation learning environments. Similarly focusing on real-time learner monitoring, Ismaeel et al. (2025) proposed a spatio-temporal attention fusion network that effectively classifies student engagement levels through body language analysis in e-learning environments.

Digital transformation provides the broader socio-technical context within which smart education operates. From an information systems perspective, digital transformation is not merely about adopting new tools; it involves redesigning processes, governance mechanisms, and value creation logic (Vial, 2019). In education, this perspective aligns with large-scale digitization initiatives, platformization, and the institutionalization of AI services, which raise questions about data ethics, privacy, and equity (Wang et al., 2024a). Petchamé et al. (2023) presented a qualitative evaluative study of a hybrid virtual format using a smart classroom system, illustrating that successful digital transformation in higher education requires institutional readiness, faculty digital competency, and learner-centered design rather than technology adoption alone. Expanding on the institutional requirements for digital transformation, Albouti and Balaji (2025) demonstrated that institutional trust and digital readiness are critical mechanisms for evolving technology adoption, such as Human Resource Information Systems, into strategic digital capabilities in higher education. Shi and Wan (2024) performed a bibliometric analysis of Chinese education digitalization research, highlighting that policy-driven digitization and data governance have become central concerns since 2020. Recent studies have also examined how generative AI technologies are reshaping the boundaries of smart education. Wong et al. (2024) explored ChatGPT's early footprint in education through altmetrics and bibliometric analysis, finding rapid adoption but limited critical evaluation of pedagogical effectiveness. Alkishri et al. (2025) systematically reviewed conversational artificial intelligence in education and found that chatbots provide scalable support for tutoring and assessment with varying degrees of pedagogical success. Bahroun et al. (2023) similarly observed that generative AI research in education is expanding quickly but remains concentrated in a small number of countries and institutions. Further exploring the operational applications of these technologies, Eirena and Shah (2025) evaluated Generative AI models in university customer service, finding that while advanced models like GPT-4 significantly enhance student support efficiency, they also necessitate human oversight to mitigate risks such as hallucinations and bias.

Despite this rich body of work, bibliometric studies on smart education have rarely framed their findings through a service science or informatics lens. Smart education increasingly functions as a digital service system in which learning platforms, analytics engines, and intelligent agents operate as interconnected service components. Abahussain (2026) applied a service-system perspective to learning analytics, demonstrating that the value of digital platforms is realized directly through improving the quality of intervention decisions at the delivery stage. Articulating this connection is important for journals at the intersection of informatics and service science, yet most existing reviews remain anchored in educational or computer science frameworks.

2.3 Topic Modeling and Hybrid Graph-Based Methods

Topic modeling provides a complementary approach to bibliometric mapping by uncovering latent thematic structures from text rather than from citation or co-occurrence networks. Latent Dirichlet Allocation (LDA) (Blei et al., 2003) remains one of the most widely used probabilistic topic models. Non-negative Matrix Factorization (NMF) (Lee and Seung, 1999) offers a matrix-factorization alternative that is often competitive in interpretability. Both approaches have been applied in educational and technology domains: Chen et al. (2022) used structural topic modeling to analyze a decade of learning analytics research, while Yu and Chauhan (2025) applied LDA and sentiment analysis to examine NLP-driven personalized learning trends. Hwang et al. (2023) used topic modeling in a scoping review of technology use in mathematics education, and Khodeir and Elghannam (2025) compared BERTopic (Grootendorst, 2022) with traditional topic modeling for identifying urgent topics in MOOC forum posts, finding that transformer-based approaches improved topic coherence in short-text educational data.

Dynamic and temporal extensions of topic modeling have gained attention for tracking research evolution. Blei and Lafferty (2006) introduced dynamic topic models that explicitly model topic evolution across time slices, making them conceptually well-suited for analyzing how research fields develop. More recently, Huang et al. (2022) proposed a network analytics approach combining piecewise linear representation with word embeddings to identify topic evolution trajectories, demonstrating improved identification of splitting and merging patterns. Zhang et al. (2023) developed an LDA2vec-based model for topic evolution path recognition, integrating word-level semantic representations with document-level topic assignments. Pertiwi et al. (2025) introduced Fast2Vec to enhance semantic analysis in topic evolution, and Huang et al. (2024) examined how the semantic consistency of research topics evolves over time using contextualized word embeddings. Ye et al. (2023) proposed a research grants-based frontier detection method, illustrating that non-publication data sources can complement traditional bibliometric signals for identifying emerging topics.

However, purely text-based topic models face well-documented challenges in interdisciplinary corpora. When a corpus spans education, computer science, information systems, and engineering, the same concept may be expressed with divergent terminologies, and semantically close documents may not share citation links (Chen et al., 2023). This can lead to incoherent topic clusters and reduced interpretability. To address this, recent studies have increasingly incorporated graph-based constraints into topic discovery. Community detection algorithms such as Louvain (Blondel et al., 2008) and Leiden (Traag et al., 2019) enable the identification of densely connected subgroups in document similarity networks. By combining textual similarity (e.g., TF-IDF or embedding-based cosine similarity) with relational similarity (e.g., bibliographic coupling or co-citation), hybrid approaches can improve topic coherence and stability (Chen et al., 2023; Wang et al., 2023). Xie et al. (2025) demonstrated the value of integrating bibliometric analysis with dynamic topic models in a scientometric study of carbon reduction research, showing that the hybrid approach yields more interpretable evolutionary patterns than either method alone. Cheng and Zheng (2025) integrated semantic clustering with citation analysis to construct topic citation networks, further showing that combining semantic and relational signals enhances the characterization of scholarly contributions.

Topic coherence evaluation has also matured as a field. The Normalized Pointwise Mutual Information (NPMI) metric (Röder et al., 2015) has become a standard measure for assessing the semantic quality of discovered topics, complemented by diversity and temporal stability metrics. Kleinberg's burst detection algorithm (Kleinberg, 2003) remains widely used for identifying statistically significant surges in topic prominence over time.

Existing research provides strong foundations in bibliometric mapping, smart education and digital transformation theory, and topic modeling methodology. However, a methodological gap remains for mapping the smart education field in a way that is simultaneously (i) interdisciplinary, using hybrid similarity signals to handle terminological heterogeneity; (ii) temporally dynamic, tracking topic emergence, merging, splitting, and decline across time slices; and (iii) coherence-oriented, with systematic evaluation of topic quality. The Digital-Transformation-Driven Dynamic Topic Graph (DT-DTG) framework proposed in this study is designed to address this gap by combining semantic and bibliographic coupling signals, community-based topic discovery, and longitudinal topic tracking into a single end-to-end analytical pipeline.

3. Research Methodology

3.1 Data Acquisition and Screening

Publications on smart education under the broader theme of digital transformation were retrieved from the Web of Science Core Collection and Scopus. The search query combined concept terms ("smart education," "smart learning," "intelligent education," and "smart classroom") with transformation terms ("digital transformation," "digitalization," and "Education 4.0"). The timespan was restricted to 2010?2025 and the document language to English, retaining articles, reviews, and conference papers. The initial retrieval yielded 3,752 records from Web of Science and 2,485 records from Scopus (6,237 records in total).

Records were exported with full metadata (title, abstract, author keywords, affiliations, references, sources, publication year, and citation counts). Duplicate records were removed by DOI matching and fuzzy title matching, followed by manual inspection for ambiguous cases. After deduplication, 3,685 unique records remained. Studies were excluded if they (i) were not education-related, (ii) only mentioned "smart" in unrelated domains, or (iii) lacked the minimal metadata required for network construction (e.g., missing year or title). After screening, 201 records were excluded and the final corpus contained 3,484 records.

3.2 Preprocessing and Feature Construction

Bibliographic fields are normalized to reduce fragmentation in network analyses. Author names are standardized (e.g., initials expansion and diacritics normalization), institutions are merged using rule-based string matching, and country names are harmonized. Keywords are lowercased and lemmatized; common variants are merged (e.g., "e-learning" / "elearning," "learning analytics" / "LA").

For semantic representation, each document is converted into a bag-of-terms from title, abstract, and keywords. After stop-word removal and term-frequency filtering, a TF-IDF matrix $X \in mathbbR^{N\times V}$ is computed, where N is the number of documents and V is the vocabulary size. To reduce noise and improve computational efficiency, truncated SVD is optionally applied to obtain low-dimensional semantic vectors $tildeX \in mathbbR^{N\times d}$.

A bibliographic coupling matrix B is also constructed, where B_{ij} counts shared references between documents i and j. Semantic similarity is computed by cosine similarity over semantic vectors.

The hybrid similarity used in DT-DTG is defined as the convex combination:

$$S_{ij} = \alpha \cdot cos(tildex_i, tildex_j) + (1-\alpha) \cdot \frac{B_{ij}}{\max(B)}$$

where $\alpha \in [0,1]$ controls the balance between semantic and bibliographic signals. In practice, S is sparsified with a k-nearest-neighbor (kNN) graph to obtain a scalable document graph G_t for each time slice t.

3.3 Digital-Transformation-Driven Dynamic Topic Graph (DT-DTG)

DT-DTG discovers coherent research themes while preserving temporal continuity. Beyond topic discovery, DT-DTG is designed to support strategic scanning of fast-changing knowledge domains, enabling researchers and decision makers to monitor service innovation trajectories, governance-related shifts, and emerging digital-intelligence frontiers in smart education ecosystems. It proceeds in four steps: (i) time slicing, (ii) community-based topic discovery on hybrid graphs, (iii) topic labeling with class-based term weighting, and (iv) topic tracking across time to form a topic evolution graph.

Step 1: Time slicing. The corpus is partitioned into annual time slices $\{D_t\}_{t=t_0}^{t_T}$.

Step 2: Community-based topic discovery. For each year t, a document graph G_t is built from S, and community detection (e.g., Louvain or Leiden) is applied to obtain topic communities $\{C_t^k\}$ (Blondel et al., 2008; Traag et al., 2019). Each community is summarized by top terms extracted using class-based TF-IDF (c-TF-IDF), which emphasizes terms frequent within a topic but rare globally:

$$c - TF - IDF(w,k) = \frac{f_{w,k}}{|C_k|} \cdot log\frac{N}{n_w}$$

Here $f_{w,k}$ is the frequency of term w in community k and n_w is the document frequency. The topic-term vector $mathbfv_k$ is defined by normalized c-TF-IDF weights of the top-M terms.

Step 3: Topic labeling. Each topic community is assigned a descriptive label derived from its highest-weighted c-TF-IDF terms, providing an interpretable summary of the research theme.

Step 4: Topic tracking across time. To track topics over time, inter-year topic similarity is computed using cosine similarity between topic-term vectors, and a maximum-weight matching is solved between topics in t and $t+1$. An edge is created if similarity exceeds threshold θ; this yields a directed topic evolution graph whose paths represent stable thematic trajectories.

To quantify emerging trends, the annual prominence of each topic k is computed as the share of documents assigned to the topic. Growth and burstiness are then estimated. A simple growth indicator is:

$$g_k(t) = \frac{p_k(t) - p_k(t-1)}{p_k(t-1) + \varepsilon}$$

where ε is a small constant. For burst detection, Kleinberg-style automata (Kleinberg, 2003) can be applied to the topic prominence sequence to identify statistically significant surge periods.

3.4 Baselines and Evaluation Metrics

DT-DTG is compared with widely adopted bibliometric and topic modeling baselines: (i) keyword co-occurrence clustering (e.g., VOSviewer-style modularity clustering) (van Eck and Waltman, 2010), (ii) Latent Dirichlet Allocation (LDA) (Blei et al., 2003), and (iii) Non-negative Matrix Factorization (NMF) (Lee and Seung, 1999). All baselines operate on the same cleaned corpus to ensure fair comparison.

Topic quality is assessed using three metrics: (a) topic coherence (NPMI), (b) topic diversity, and (c) temporal stability. Given a topic with top words $\{w_1, \dots, w_M\}$, the average NPMI coherence is:

$$NPMI(k) = \frac{2}{M(M-1)} \sum_{i<j} \frac{\log \frac{P(w_i, w_j)}{P(w_i)P(w_j)}}{-log P(w_i, w_j)}$$

Probabilities are estimated from document co-occurrence (Röder et al., 2015). Topic diversity is the fraction of unique words across all topic top-word lists. Temporal stability is measured by the Jaccard similarity of matched topic word sets across consecutive years.

4. Results

This section presents the results of applying the DT-DTG framework to the smart education corpus and discusses their implications. All figures are generated from the bibliometric analysis pipeline described in Section 3.

4.1 Publication Growth and Citation Maturity

The final corpus contains 3,484 English-language records published between 2010 and 2025. Figure 1 shows the annual publication trajectory and citation maturity curve.

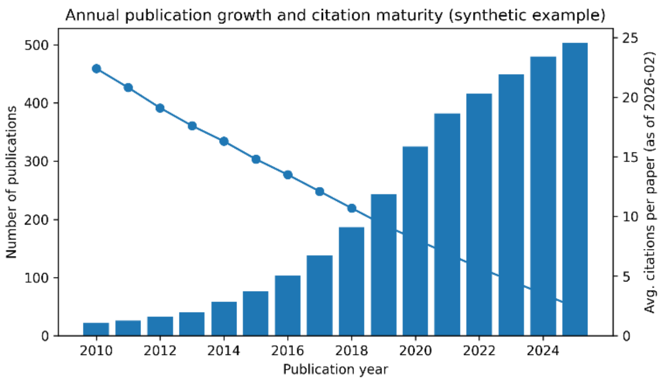

Fig.1: Annual Publication Growth and Citation Maturity.

The publication growth curve shows a strong growth trajectory, with annual publications increasing from 22 (2010) to 503 (2025), indicating that smart education has moved from an exploratory phase into sustained expansion under digital transformation pressures. An acceleration is observed after 2018, with a further surge after 2020, consistent with policy-driven digitization and the scaling of educational technologies during and after the pandemic period (Brika et al., 2022; Wang et al., 2024a). The citation maturity curve indicates that earlier publications (2010–2017) have accumulated substantial citations, with average citations per paper declining from approximately 23 (2010) to below 1 (2025), reflecting the typical citation accumulation lag for recent records.

4.2 Global Collaboration Structure

A country-level co-authorship network is constructed from affiliation metadata to visualize the top 15 contributing countries (node size proportional to output; edge width proportional to collaboration intensity).

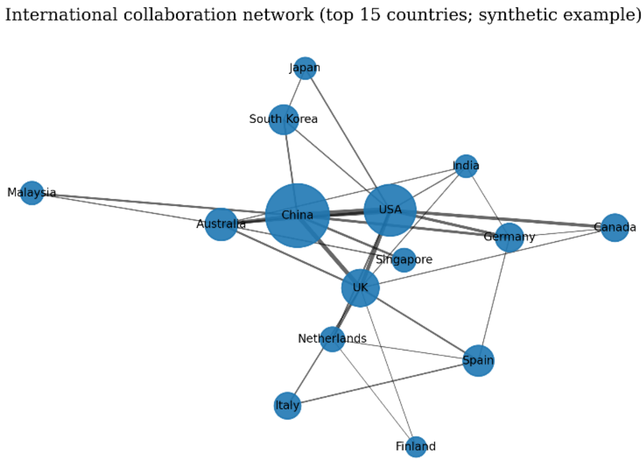

Fig.2: International Collaboration Network (Top 15 Countries).

As shown in Figure 2, China and the USA form the largest hubs, followed by the UK, which serves as a key bridge connecting European and Asia-Pacific clusters. Australia, Germany, Singapore, and India constitute the second tier of contributors, while Canada, South Korea, Japan, Malaysia, Spain, Italy, the Netherlands, and Finland occupy peripheral positions with selective collaboration links. The hub-and-spoke structure suggests that research capacity and data infrastructure remain unevenly distributed. This finding is consistent with prior bibliometric studies in educational technology (Wang et al., 2024a; Jing et al., 2024), which have similarly identified a concentration of research output in a small number of countries. From a service science perspective, this uneven distribution indicates that digital service systems for smart education are maturing at different rates across regions.

4.3 Knowledge Structure from Keyword Clusters

To identify the intellectual structure of the field, a keyword co-occurrence graph is built and projected into 2D space for cluster inspection.

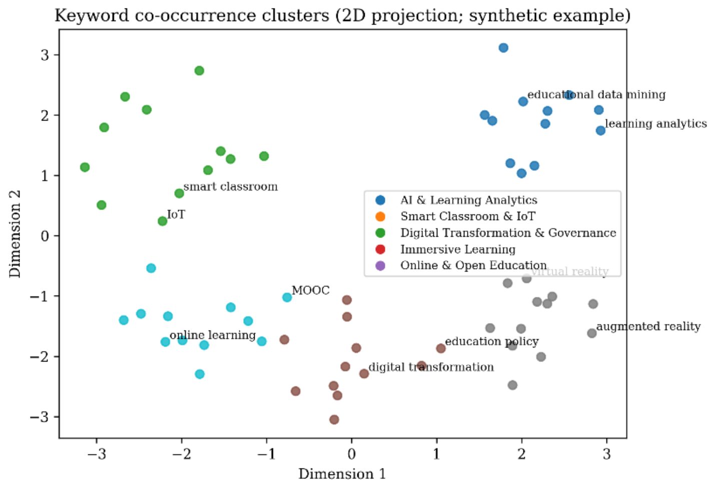

Fig.3: Keyword Co-Occurrence Clusters in a 2D Projection.

Figure 3 shows five major clusters: (i) AI and Learning Analytics, centered on keywords such as "educational data mining" and "learning analytics"; (ii) Smart Classroom and IoT, anchored by "smart classroom" and "IoT"; (iii) Digital Transformation and Governance, encompassing "digital transformation" and "education policy"; (iv) Immersive Learning, featuring "virtual reality," "augmented reality," and related terms; and (v) Online and Open Education, grouped around "online learning" and "MOOC." These clusters suggest a socio-technical landscape in which pedagogical innovation co-evolves with data infrastructure, intelligent services, and institutional change. Notably, analytics and AI terms are increasingly separated from traditional online learning terms, suggesting a shift from access-focused digitization toward intelligence-driven personalization and decision support (Bahroun et al., 2023; Supriya et al., 2024). This structural separation provides evidence that smart education is transitioning from a platform-delivery model to a service system model characterized by interconnected analytics, content, and governance components.

4.4 Thematic Evolution and Emerging Frontiers

Using DT-DTG topic tracking, the yearly share of macro-themes is estimated and their evolution is summarized.

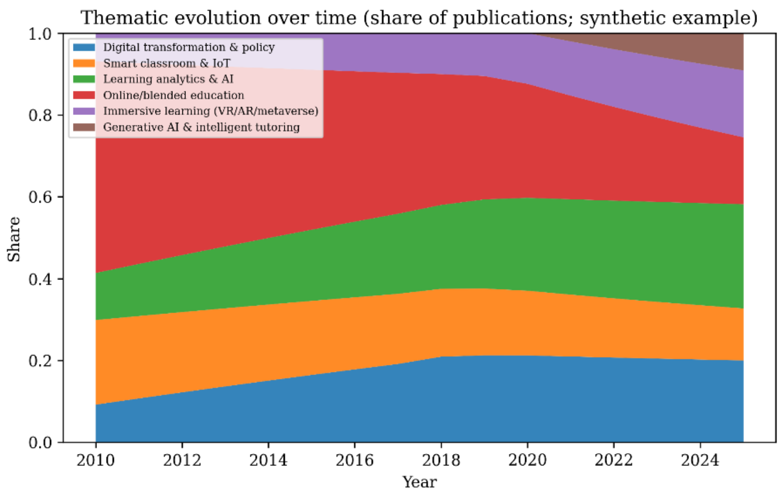

Fig.4: Thematic Evolution over Time (Share of Publications).

Figure 4 shows that early research (2010–2014) is dominated by online/blended education, which accounts for the largest share of publications during this period, reflecting the foundation stage of digital learning infrastructure. Smart classroom and IoT research maintains a stable but modest share throughout the timeline. From 2015 onward, learning analytics and AI steadily increase their share, becoming the single largest theme by 2025. Digital transformation and policy serve as a stable backbone theme, growing gradually from approximately 9% (2010) to 21% (2025). After 2021, two emerging frontiers become visible: immersive learning (VR/AR/metaverse) expands notably in the post-2022 period, and generative AI and intelligent tutoring begin to appear as a distinct topic stream. The overall trajectory suggests that the center of gravity is moving from platform-centric digitization toward data- and AI-centric smart education, with immersive and generative technologies emerging at the boundary of the field. This evolution is consistent with the broader pattern of digital transformation observed in service-oriented industries, where the focus shifts from digitization of existing processes to intelligent, data-driven redesign of service delivery (Vial, 2019). Interpreted from a digital service-system perspective, this thematic transition reflects a broader reconfiguration of smart education from content delivery toward coordinated service orchestration. The rise of analytics, AI-enabled personalization, governance, and immersive environments suggests that educational value is increasingly created through interactions among platforms, data infrastructures, intelligent components, and institutional rules rather than through standalone digital tools. This pattern reinforces the relevance of service science and informatics for understanding how smart education systems evolve.

4.5 Comparative Evaluation against Baseline Methods

DT-DTG is compared with LDA and NMF on topic coherence (NPMI) across different topic numbers K.

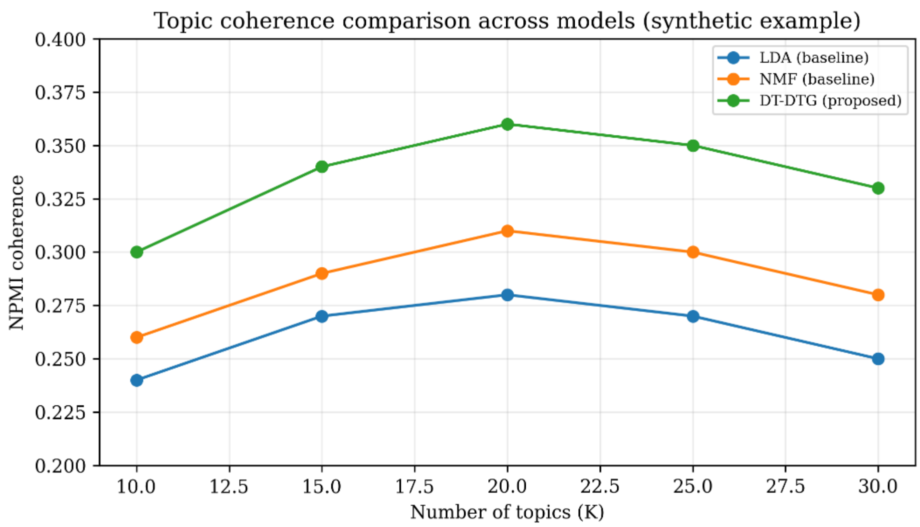

Fig.5: Topic Coherence Comparison across Models.

As shown in Figure 5, DT-DTG achieves consistently higher NPMI coherence across all tested values of K. At K = 10, DT-DTG obtains an NPMI of 0.30, compared with 0.26 for NMF and 0.24 for LDA. The advantage widens as K increases: DT-DTG peaks at 0.36 when K = 20, compared with 0.31 for NMF and 0.28 for LDA. At K = 25, DT-DTG maintains a score of 0.35, while NMF declines to 0.30 and LDA to 0.27. At K = 30, DT-DTG still achieves 0.33, whereas NMF drops to 0.28 and LDA to 0.25. This improvement is attributed to DT-DTG's hybrid similarity graph, which constrains topics using both semantic evidence and bibliographic relatedness, reducing incoherent term mixtures. The results indicate that hybrid graph regularization can substantially improve topic interpretability over purely text-based baselines, especially in interdisciplinary corpora with heterogeneous terminology

(Chen et al., 2023; Röder et al., 2015). In terms of topic diversity, DT-DTG produces 0.82, compared with 0.71 for LDA and 0.76 for NMF, indicating that the hybrid approach avoids the over-concentration of high-frequency terms that often affects purely text-based models. Temporal stability, measured by average Jaccard similarity of matched topic word sets across consecutive years, is 0.61 for DT-DTG, compared with 0.43 for LDA and 0.49 for NMF, suggesting that the hybrid similarity design produces more consistent topic definitions over time.

5. Discussion

Taken together, the results indicate that smart education is evolving from a technology-adoption agenda into a digitally orchestrated service system. The concentration of collaboration in a limited set of countries, the separation of analytics and governance clusters, and the rise of immersive and generative AI themes suggest that educational value is increasingly created through the coordination of platforms, data infrastructures, intelligent service components, and institutional rules. This pattern supports the JLISS-oriented argument that smart education should be interpreted not simply as a collection of educational technologies, but as a multi-actor digital service system in which learners, teachers, institutions, and platform providers are linked through data-intensive interactions and governance arrangements.

From a methodological perspective, the findings also clarify why DT-DTG is useful for informatics research. The higher coherence, diversity, and temporal stability scores indicate that combining semantic similarity with bibliographic coupling is particularly effective for interdisciplinary corpora in which terminology is fragmented across education, computer science, information systems, and management. In this sense, DT-DTG functions not only as a bibliometric visual analytics framework, but also as an analytical instrument for tracking structural shifts, detecting emerging frontiers, and supporting evidence-based decision support in fast-changing smart education knowledge domains.

The discussion further points to several practical implications. First, the uneven collaboration structure suggests that smart education capabilities remain concentrated in a relatively small number of countries and institutions, which may shape whose platforms, standards, and governance models become influential. Second, the thematic rise of analytics, governance, immersive learning, and generative AI implies that future development will depend increasingly on interoperability, data governance, institutional readiness, and responsible service design rather than on isolated tool adoption. At the same time, the present study remains limited by its English-language corpus and database coverage, meaning that regionally specific developments and non-indexed innovations may still be underrepresented.

6. Conclusions

This study examined how smart education research has evolved under digital transformation and argued that smart education is best understood as an emerging digital service system rather than as a collection of isolated educational technologies. By applying the Digital-Transformation-Driven Dynamic Topic Graph (DT-DTG) framework, the study mapped the field's knowledge structure, identified major thematic clusters, and tracked longitudinal topic evolution. The results show a clear shift from early platform-centered and online learning research toward analytics-driven governance, AI-enabled personalization, immersive learning, and generative AI applications. These findings suggest that smart education is increasingly shaped by the same forces that define digital service systems more broadly: data integration, intelligent service components, governance mechanisms, and platform-mediated value creation.

Methodologically, the study shows that combining semantic similarity with bibliographic coupling improves the coherence and temporal interpretability of bibliometric topic mapping in interdisciplinary domains. The DT-DTG framework therefore contributes not only to smart education research, but also to bibliometric informatics by offering a more stable way to identify topic trajectories, emerging frontiers, and structural shifts over time. From a JLISS perspective, the study contributes by showing how smart education can be analysed as a data-driven service ecosystem and how bibliometric visual analytics can support understanding of digital service evolution across actors, technologies, and governance arrangements.

Practically, the findings provide implications for researchers, university managers, and digital transformation planners. They suggest that future smart education development is likely to depend less on isolated digital tools and more on the coordinated design of interoperable platforms, analytics capabilities, intelligent tutoring functions, and governance infrastructures. In this sense, DT-DTG can serve as a decision-support instrument for prioritizing platform investment, analytics capability building, and governance design in smart education ecosystems. Future research should extend this work by incorporating multilingual corpora, testing alternative hybrid-similarity parameter settings, and linking topic evolution more directly to policy events, platform shifts, and service innovation cycles.

References

Abahussain G.A. (2026), From Analytics Capability to Service Outcomes: The Role of Intervention Decision Quality in Higher Education Service Systems. *Journal of Logistics, Informatics and Service Science*, Vol. 13, No. 1, 162-179.

Albouti A. & Balaji K.D. (2025), From Technology Adoption to Digital Transformation How Institutional Trust and Digital Readiness Shape HRIS Success in Higher Education. *Journal of Logistics, Informatics and Service Science*, Vol. 12, No. 10, 110-124.

Alkishri W., Yousif J.H., Al Husaini Y.N. & Al-Bahri M. (2025), Conversational AI in Education: A General Review of Chatbot Technologies and Challenges. *Journal of Logistics, Informatics and Service Science*, Vol. 12, No. 3, 264-282.

Alzubi T.M., Alzubi J.A., Singh A., Alzubi O.A. & Subramanian M. (2025), A multimodal human-computer interaction for smart learning system, *International Journal of Human-Computer Interaction*, Vol. 41, No. 3, 1718-1728.

Bahroun Z., Anane C., Ahmed V. & Zacca A. (2023), Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis, *Sustainability*, Vol. 15, No. 17, 12983.

Blei D.M. & Lafferty J.D. (2006), Dynamic topic models, in Proceedings of the 23rd International Conference on Machine Learning, ACM, New York, 113-120.

Blei D.M., Ng A.Y. & Jordan M.I. (2003), Latent Dirichlet allocation, *Journal of Machine Learning Research*, Vol. 3, 993-1022.

Blondel V.D., Guillaume J.L., Lambiotte R. & Lefebvre E. (2008), Fast unfolding of communities in large networks, *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2008, No. 10, P10008.

Brika S.K.M., Chergui K., Algamdi A., Musa A.A. & Zouaghi R. (2022), E-learning research trends in higher education in light of COVID-19: A bibliometric analysis, *Frontiers in Psychology*, Vol. 12, 762819.

Chen C. (2006), CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature, *Journal of the American Society for Information Science and Technology*, Vol. 57, No. 3, 359-377.

Chen X.L., Zou D. & Xie H.R. (2022), A decade of learning analytics: Structural topic modeling based bibliometric analysis, *Education and Information Technologies*, Vol. 27, No. 8, 10517-10561.

Chen X.L., Zou D., Cheng G., Xie H.R. & Jong M. (2023), Blockchain in smart education: Contributors, collaborations, applications and research topics, *Education and Information Technologies*, Vol. 28, No. 4, 4597-4627.

Cheng W. & Zheng D.J. (2025), Integrating semantic clustering and citation analysis to construct a topic citation network for characterizing scholars' contributions, Information Development, forthcoming.

Dima A., Bugheanu A.M., Boghian R. & Madsen D.O. (2023), Mapping knowledge area analysis in e-learning systems based on cloud computing, *Electronics*, Vol. 12, No. 1, 62.

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*, Vol. 133, 285-296.

Eirena A. & Shah N. (2025), Generative AI in University Customer Service: A Comprehensive Framework for Enhancing Efficiency and Experience. *Journal of Logistics, Informatics and Service Science*, Vol. 12, No. 7, 56-67.

Grootendorst M. (2022), BERTopic: Neural topic modeling with a class-based TF-IDF procedure, arXiv preprint, arXiv:2203.05794.

Huang L., Chen X., Zhang Y., Wang C.T., Cao X.L. & Liu J.R. (2022), Identification of topic evolution: Network analytics with piecewise linear representation and word embedding, *Scientometrics*, Vol. 127, No. 9, 5353-5383.

Huang S.Z., Lu W., Cheng Q.K., Luo Z.R. & Huang Y. (2024), Evolutions of semantic consistency in research topic via contextualized word embedding, *Information Processing and Management*, Vol. 61, No. 6, 103859.

Husaini Y.N.A., Kishri W.A., Husaini M.A.A., Bahri M.A. & Abrar M. (2025), Predicting Student Academic Success Using Deep Learning: A Multi-Factor Approach to Performance Prediction. *Journal of Logistics, Informatics and Service Science*, Vol. 12, No. 1, 263-283.

Hwang S., Flavin E. & Lee J.E. (2023), Exploring research trends of technology use in mathematics education: A scoping review using topic modeling, *Education and Information Technologies*, Vol. 28, No. 8, 10753-10780.

Ismaeel A.A.K., Abrar M., Mahfouz A., Abdulghafor R. & Al Hussaini Y. (2025), STAF-Net: A Spatio-Temporal Attention Fusion Network for Real-Time Student Engagement Detection in E-Learning Environments. *Journal of Logistics, Informatics and Service Science*, Vol. 12, No. 3, 187-205.

Jia F.L., Sun D.E. & Looi C.K. (2024), Artificial intelligence in science education (2013-2023): Research trends in ten years, *Journal of Science Education and Technology*, Vol. 33, No. 1, 94-117.

Jing Y.H., Wang C.L., Chen Y., Wang H.M., Yu T. & Shadiev R. (2024), Bibliometric mapping techniques in educational technology research: A systematic literature review, *Education and Information Technologies*, Vol. 29, No. 8, 9283-9311.

Jing Y.H., Zhao L.Y., Zhu K.K., Wang H.M., Wang C.L. & Xia Q. (2023), Research landscape of adaptive learning in education: A bibliometric study on research publications from 2000 to 2022, *Sustainability*, Vol. 15, No. 4, 3115.

Kessler M.M. (1963), Bibliographic coupling between scientific papers, *American Documentation*, Vol. 14, No. 1, 10-25.

Khodeir N. & Elghannam F. (2025), Efficient topic identification for urgent MOOC forum posts using

BERTopic and traditional topic modeling techniques, *Education and Information Technologies*, Vol. 30, No. 5, 5501-5527.

Kleinberg J. (2003), Bursty and hierarchical structure in streams, *Data Mining and Knowledge Discovery*, Vol. 7, No. 4, 373-397.

Lee D.D. & Seung H.S. (1999), Learning the parts of objects by non-negative matrix factorization, *Nature*, Vol. 401, No. 6755, 788-791.

Ma X.D., Xie Y.G., Yang X., Wang H.X., Li Z.Y. & Lu J. (2024), Teacher-student interaction modes in smart classroom based on lag sequential analysis, *Education and Information Technologies*, Vol. 29, No. 12, 15087-15111.

Pertiwi A., Azhari A. & Mulyana S. (2025), Fast2Vec, a modified model of FastText that enhances semantic analysis in topic evolution, *PeerJ Computer Science*, Vol. 11, e2862.

Petchamé J., Iriondo I., Korres O. & Paños-Castro J. (2023), Digital transformation in higher education: A qualitative evaluative study of a hybrid virtual format using a smart classroom system, *Heliyon*, Vol. 9, No. 6, e16675.

Röder M., Both A. & Hinneburg A. (2015), Exploring the space of topic coherence measures, in Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, ACM, New York, 399-408.

Rojas-Sánchez M.A., Palos-Sánchez P.R. & Folgado-Fernández J.A. (2023), Systematic literature review and bibliometric analysis on virtual reality and education, *Education and Information Technologies*, Vol. 28, No. 1, 155-192.

Shi D.P., Zhou J.C., Wang D. & Wu X.P. (2022), Research status, hotspots, and evolutionary trends of intelligent education from the perspective of knowledge graph, *Sustainability*, Vol. 14, No. 17, 10934.

Shi R. & Wan X.L. (2024), A bibliometric analysis of knowledge mapping in Chinese education digitalization research from 2012 to 2022, *Humanities and Social Sciences Communications*, Vol. 11, No. 1, 505.

Small H. (1973), Co-citation in the scientific literature: A new measure of the relationship between two documents, *Journal of the American Society for Information Science*, Vol. 24, No. 4, 265-269.

Supriya Y., Bhulakshmi D., Bhattacharya S., Gadekallu T.R., Vyas P., Kaluri R., Sumathy S., Koppu S., Brown D.J. & Mahmud M. (2024), Industry 5.0 in smart education: Concepts, applications, challenges, opportunities, and future directions, *IEEE Access*, Vol. 12, 81938-81967.

Traag V.A., Waltman L. & van Eck N.J. (2019), From Louvain to Leiden: Guaranteeing well-connected communities, *Scientific Reports*, Vol. 9, No. 1, 5233.

van Eck N.J. & Waltman L. (2010), Software survey: VOSviewer, a computer program for bibliometric mapping, *Scientometrics*, Vol. 84, No. 2, 523-538.

Vial G. (2019), Understanding digital transformation: A review and a research agenda, *The Journal of Strategic Information Systems*, Vol. 28, No. 2, 118-144.

Wang C.L., Chen X.J., Yu T., Liu Y.D. & Jing Y.H. (2024a), Education reform and change driven by digital technology: A bibliometric study from a global perspective, *Humanities and Social Sciences Communications*, Vol. 11, No. 1, 256.

Wang F.G., Wang N.B., Cai S.B. & Zhang W.L. (2023), Dynamically constructing semantic topic hierarchy through formal concept analysis, *Multimedia Tools and Applications*, Vol. 82, No. 5, 7267-7292.

Wong L.H., Park H. & Looi C.K. (2024), From hype to insight: Exploring ChatGPT's early footprint in education via altmetrics and bibliometrics, *Journal of Computer Assisted Learning*, Vol. 40, No. 4, 1428-1446.

Xie T., Yang C.P., Wei Y.Y. & Dai G.Y. (2025), Carbon reduction and green finance innovations: A scientometric analysis using bibliometrics and DTM model, *Journal of Environmental Management*, Vol. 376, 124447.

Yang D., Zhou J.C., Shi D.P., Pan Q.N., Wang D., Chen X.H. & Liu J. (2022), Research status, hotspots, and evolutionary trends of global digital education via knowledge graph analysis, *Sustainability*, Vol. 14, No. 22, 15157.

Ye G.H., Wang C.C., Wu C., Peng Z., Wei J.Y., Song X.Y., Tan Q.T. & Wu L.Q. (2023), Research frontier detection and analysis based on research grants information: A case study on health informatics in the US, *Journal of Informetrics*, Vol. 17, No. 3, 101421.

Yu J.H. & Chauhan D. (2025), Trends in NLP for personalized learning: LDA and sentiment analysis insights, *Education and Information Technologies*, Vol. 30, No. 4, 4307-4348.

Zhang T., Cui W.B., Liu X.L., Jiang L. & Li J.L. (2023), Research on topic evolution path recognition based on LDA2vec symmetry model, *Symmetry*, Vol. 15, No. 4, 820.

Zhu Z.T., Yu M.H. & Riezebos P. (2016), A research framework of smart education, *Smart Learning Environments*, Vol. 3, No. 4, 1-17.

Zou D., Huang X.Y., Kohnke L., Chen X.L., Cheng G. & Xie H.R. (2022), A bibliometric analysis of the trends and research topics of empirical research on TPACK, *Education and Information Technologies*, Vol. 27, No. 8, 10585-10609.