

Mapping Research Trends of Artificial Intelligence in Higher Education: A Topic Modeling Approach

ABSTRACT

Authors: Zhang Yuningjing¹, Zhang YingShui², Ko Jang Wan^{3*}

Affiliation: ¹²³Sungkyunkwan University, South Korea
*Corresponding author:
jakosu@skku.edu

Received: 9 December 2025 |
First revision: 4 March 2026 |
Second revision: 7 June 2026 |
Accepted: 10 June 2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Zhang, Y., Zhang, Y., & Ko, J. W. (2026). Mapping research trends of artificial intelligence in higher education: A topic modeling approach. *Journal of Institutional Research South East Asia*, 24(1), 78-96.

The rapid expansion of artificial intelligence (AI), particularly generative AI, has accelerated scholarly interest in higher education, yet the longitudinal evolution of research themes in this field remains insufficiently understood. This research aims to reveal the development trend and the topic evolution of research on AI in higher education from 2015 to 2024. A corpus of 475 open-access articles indexed in the Web of Science was analyzed using Latent Dirichlet Allocation (LDA) topic modeling. The temporal patterns identified through LDA were further interpreted through Diffusion of Innovations (DOI) theory and the Design and Development Research (DDR) framework. According to the research results, the literature on AI in higher education has seen a significant increase over the past 10 years. Meanwhile, we have found that the research on AI in higher education revolves around five core themes: (1) Generative AI Integration, (2) Empirical Evaluation Studies, (3) Contextual AI Application, (4) Digital Course Transformation, (5) Educational Process and Quality. The findings reveal substantial shifts in research attention over time and provide a longitudinal perspective on the development of AIEd scholarship. However, the findings are limited by the use of open-access WoS-indexed publications and the interpretive nature of topic labeling.

Keywords: Artificial Intelligence, Higher Education, Topic Modeling, LDA

1. Introduction

The rise of artificial intelligence (AI) has triggered profound changes across all sectors of society, including higher education. While AI has been studied for decades, the popularity of generative tools such as ChatGPT has raised urgent questions about how humans will interact, collaborate, and learn alongside intelligent systems (Gašević et al., 2023). This shift marks a fundamental challenge to traditional educational theories, assessment practices, and institutional structures. As higher education increasingly seeks to prepare learners for uncertain and complex futures, there is growing interest in using AI to support decision making, personalized learning, and enhance educational quality. These developments demand a comprehensive examination of how AI is transforming the approach and goals of education.

In the context of higher education, institutions and governments are under increasing pressure to accommodate diverse student populations, provide relevant and timely opportunities for education, and improve learning outcomes through scalable, data-driven approaches (Siemens et al., 2015). Technologies such as learning management systems and student information systems have laid a digital foundation that enables the application of AI-driven techniques. As educational data becomes more accessible, AI is no longer confined to understanding the learning process but can also be used to design intelligent systems that can respond to students' needs in real time. Consequently, AI is increasingly embedded in the practices of teaching, learning, and institutional management.

The growing interest in AI in higher education (AIEd) has led to the rapid rise of studies on AIEd in the last five years. Previous research reported that the number of AIEd publications have risen nearly 2-3 times the number of previous years in 2021 and 2022 (Crompton & Burke, 2023). This rapid growth is also reflected in the corpus analyzed in the present study. Based on the preliminary descriptive statistics of our WoS dataset, annual publication output increased from 10 articles in 2019 to 26 in 2020 and reached 43 in 2022, followed by a further increase to 81 in 2023 and 245 in 2024. These patterns provide additional evidence of the rapid expansion of AIEd research over the past several years. Nevertheless, several notable research gaps remained. First, existing systematic reviews concentrated on static snapshots rather than longitudinal perspectives on dynamic evolution of research topics (Chu et al., 2022) while few studies examining AIEd in the recent two to three years (Ouyang et al., 2022). Second, many previous studies rely on simple bibliometric analyses (e.g. co-citation network analysis or keyword co-occurrence) and therefore fail to capture the latent thematic structures within the AIEd literature (Chen et al., 2020; Hinojo-Lucena et al., 2019; Tang et al., 2023). Third, most AIEd studies are led by researchers from computer science and engineering backgrounds, with relatively few contributions from faculties of education (Zawacki-Richter et al., 2019). It limits our understanding of how AI is pedagogically and strategically integrated into higher education systems. In response to these research gaps, this study aimed to systematically analyze the evolution of AIEd research topics from 2015 to 2024, using Latent Dirichlet Allocation (LDA) to uncover the hidden topic patterns and trends.

This study made substantial contributions to the field of AIEd. First, by identifying the major latent themes within the AIEd literature, the study extends previous reviews that have relied largely on descriptive bibliometric techniques. Second, by tracing changes in these themes from 2015 to 2024, it provides a longitudinal perspective on the dynamic evolution of AIEd research. Building on these thematic and longitudinal analyses, the study further interprets the observed temporal and structural patterns through Diffusion of Innovations (DOI) theory and the Design and Development Research (DDR) framework, extending the discussion from technology-

centered applications to broader educational and institutional concerns. The longitudinal analysis also captures emerging thematic directions associated with the recent expansion of generative AI, an important period in the development of AIED. Accordingly, we have raised the following research questions:

- (1) What have been the main research topics in AIED during 2015-2024?
- (2) How have the key topics of AIED changed from 2015 to 2024?

2. Literature Review

Theoretical Framework

The diffusion of innovations (DOI) theory offers a robust analytical framework for understanding the evolution of research topics related to AIED. As Rogers (2003) stated, the adoption of innovation is a process that unfolds across five progressive stages: knowledge, persuasion, decision, implementation, and confirmation. It also categorizes adopters into five groups: innovators, early adopters, early majority, late majority, and laggards, and identifies five perceived attributes of innovations that influence the rate and extent of diffusion: relative advantage, compatibility, complexity, trialability, and observability. In the present study, DOI is used to interpret the temporal evolution of scholarly attention to AI-related topics rather than the actual adoption of innovations by individuals or institutions. As new AI technologies and practices emerge in higher education, related research topics may initially appear in a limited number of studies, subsequently attract broader scholarly attention, and eventually become more established within the field. From this perspective, changes in topic prevalence can reflect the emergence, expansion, and stabilization of scholarly attention, which resembles the broader diffusion process described by DOI, although publication frequency itself does not constitute direct evidence of innovation adoption. Accordingly, the five-stage DOI model is adapted into three broader phases (initial, integration, and maturation) to interpret the temporal patterns identified in the corpus. These phases therefore serve as an interpretive framework for topic evolution rather than as direct equivalents of individual or organizational adoption stages.

DOI has also been applied to the study of AI adoption in higher education. For example, Abdalla et al. (2024) employed DOI theory to examine university students' adoption of ChatGPT through innovation attributes such as relative advantage, compatibility, complexity, trialability, and observability. At the institutional level, Jin et al. (2025) applied DOI theory to analyze generative AI adoption policies and guidelines across universities in different global regions. These studies demonstrate the applicability of DOI to understanding AI-related adoption and diffusion across different levels of higher education.

To further clarify the paradigmatic distinctions across research themes related to AIED, we also drew on the Design and Development Research (DDR) framework proposed by Richey and Klein (2014). This framework emphasizes the dynamic integration between theory and practice, highlighting the importance of systematic research design in supporting process-oriented innovation in education. The DDR framework consists of two primary research paths. One focuses on the development, implementation, and empirical evaluation of instructional products and tools, while the other is model-oriented, aiming to construct and validate theoretical frameworks or strategies for instructional design and assessment processes. This approach emphasizes the reciprocal relationship between research and practice and serves as a basis for distinguishing the methodological orientation of different research topics. For instance,

studies focusing on learning assessment and instructional design are often associated with product or process-level empirical development, whereas research addressing AI ethics and institutional integration tends to reflect more systemic, model-driven exploration. In the present study, DDR is not used as a methodological framework that determines the research design. Instead, it provides a structural lens for interpreting differences among AIEd research themes. Themes emphasizing AI tools, instructional applications, assessment, implementation, and empirical evaluation can be understood as more closely aligned with product- and process-oriented research, whereas themes concerning institutional integration, quality, governance, and systemic transformation reflect broader model- and system-oriented inquiry. Changes in the relative prominence of these orientations therefore help characterize the structural evolution of AIEd research over time. DDR has also been applied to generative AI integration in higher education. For example, Lee et al. (2025) employed DDR to develop and validate an AI-TPACK-based instructional design for university instructors' use of generative AI. This application further demonstrates the relevance of DDR to the development and implementation of AI-supported educational practices in higher education.

Taken together, DOI and DDR provide complementary lenses for interpreting two dimensions of topic evolution. DOI offers a temporal lens for understanding the emergence, expansion, and stabilization of scholarly attention to AI-related topics, whereas DDR provides a structural lens for examining shifts between product/process-oriented and broader model/system-oriented research. Importantly, these frameworks are applied after the topic structure and temporal patterns have been empirically identified through LDA. They therefore do not determine topic formation or impose predetermined developmental stages on the data, but instead provide complementary concepts for interpreting the patterns observed in the corpus.

Systematic Review in AIEd

In recent years, the rapid development of AI technologies has led to the widespread examination in the context of higher education. Particularly under the accelerated expansion of online learning environments, AI has contributed to reshaping the technical infrastructure that underpins teaching and learning, and has influenced modes of instruction, assessment procedures, and institutional decision-making processes. Applications of AI in higher education span multiple domains, including personalized content recommendation, intelligent feedback during learning, academic performance prediction, and data-informed administrative planning (Hinojo-Lucena et al., 2019; Ouyang et al., 2022).

Although research in AIEd is still emerging, existing studies show a marked increase in scholarly output in recent years. Based on a bibliometric analysis of publications indexed in the WoS and Scopus, Hinojo-Lucena et al. (2019) reported a steady growth in AIEd-related research between 2007 and 2017, with a notable concentration in conference proceedings and engineering-focused journals. A more recent systematic review conducted by Ouyang et al. (2022) identified a surge in empirical studies after 2016, selecting 32 core publications that examine predictive analytics, recommendation systems, and automated assessment tools in online higher education contexts. In terms of thematic focus, previous research primarily addresses four functional domains including academic profiling and prediction, personalized content delivery, automated feedback and assessment, and the optimization of learning experiences (Chu et al., 2022; Ouyang et al., 2022; Zawacki-Richter et al., 2019). Chu et al. (2022) identified profiling and prediction as the most commonly studied application of AIEd, followed by intelligent tutoring systems (ITS) and assessment-related functions. Similarly, Zawacki-Richter et al. (2019) observed that many studies concentrate on academic

achievement prediction, learner modeling, automated grading, and the evaluation of teaching practices. However, it is notable that most studies are weakly connected to educational theory, and empirical analyses of teaching roles remain scarce. Additionally, while supervised machine learning methods (e.g., decision trees, support vector machines, and artificial neural networks) are frequently employed, the incorporation of advanced approaches such as deep learning and natural language processing is still in its early stages (Ouyang et al., 2022). More recent reviews have begun to capture the rapid transformation of AIED following the emergence of ChatGPT and generative AI. Bond et al. (2024), through a meta-systematic review of 66 secondary studies in higher education, highlighted growing concerns related to ethics, methodological rigor, and contextual diversity in AI research. Focusing specifically on ChatGPT, Baig and Yadegaridehkordi (2024) systematically reviewed 57 studies published in 2023–2024 and identified major research trends concerning adoption, applications, and challenges in higher education. Xia et al. (2024) further examined 32 empirical studies on generative AI and assessment in higher education, emphasizing changes in assessment practices across student, teacher, and institutional levels. More recently, Qian (2025) synthesized the first two years of pedagogical applications of generative AI in higher education and identified emerging attention to creativity, critical thinking, learning autonomy, and prompt literacy. Together, these reviews demonstrate that post-2022 AIED research has increasingly shifted toward generative AI, with growing attention to pedagogical integration, assessment, adoption, and institutional responses.

Despite these developments, several limitations remain in the existing review literature. Earlier reviews primarily cover AIED research before the emergence of generative AI, whereas recent reviews are largely concentrated on the post-2022 period, leaving the transition between these two stages insufficiently captured (Chu et al., 2022; Ouyang et al., 2022; Baig & Yadegaridehkordi, 2024; Qian, 2025). In addition, recent reviews tend to focus on specific aspects of generative AI rather than the broader thematic structure of AIED research (Baig & Yadegaridehkordi, 2024; Xia et al., 2024; Lee & Palmer, 2025; Qian, 2025). Taken together, these temporal and topical boundaries provide limited understanding of how major AIED research themes have evolved continuously from earlier AI applications to the generative AI era. To address these limitations, this study applies LDA topic modeling to AIED literature published from 2015 to 2024 to identify major latent themes and trace their longitudinal evolution.

Accordingly, there is a need for a systematically topic-based analysis of AIED literature spanning the period from 2015 to 2024. It should aim to identify dominant thematic structures, trace the development of topics over time, and highlight emerging areas of focus. By doing so, it can address the temporal and methodological limitations of existing reviews while providing valuable insight for educational policymakers and practitioners. A nuanced understanding of how research attention has shifted, and how AIED themes intersect with concerns of educational quality and institutional decision-making is vital for supporting more meaningful and sustainable integration of AI in higher education contexts.

3. Methodology

To answer the research questions, we utilized LDA that allows for the extraction of latent thematic structures from a large corpus of texts. Different from previous analysis methods that rely on keyword co-occurrence or co-citation network analysis, LDA can reveal the dynamic evolution of AIED research topics over time. We constructed a literature corpus by systematically collecting journal articles indexed in the Web of Science (WoS) database,

spanning the years 2015 to 2024. This temporal coverage allows for a comprehensive longitudinal analysis that encompasses both the early development and recent acceleration of AIED research.

The methodological procedure consisted of five sequential stages: literature identification and screening, text preprocessing, LDA implementation, topic-number selection, and longitudinal topic analysis.

Data Collection and Preprocessing

Once we identified the research goals and guiding questions, we established a research protocol and prepared for a systematic and rigorous data collection process. To align with our research aim of exploring the intersection of AI and higher education, we focused our initial search strategy on studies addressing both AI applications and their relevance to higher education or university settings.

Our initial dataset was obtained by targeting publications that combined the keywords "AI Education," "Higher Education," and "Universities," covering the publication period from 2015 to 2024. To enhance transparency and research reproducibility, only open-access articles were included in the initial search, yielding 1,780 records. This criterion ensured that the publications included in the corpus were publicly accessible and could be independently verified. Restricting the corpus in this way also facilitates reproducibility, as all selected publications are publicly accessible. Following the data cleaning process, which involved removing duplicates and records with missing essential information, we retained 1,456 entries. Further refinement was conducted by applying a targeted keyword filter across three key fields: Article Title, Author Keywords, and Abstract. Specifically, we retained articles where at least one of these fields contained both "AI" and "Higher education" or both "AI" and "Universities." During final data validation prior to repository deposition, 13 records indexed as 2025 publications were identified and excluded to ensure alignment with the predefined temporal scope (2015–2024). The final analytic corpus therefore consisted of 475 publications. Given the small proportion of excluded records, the overall thematic structure and interpretation were not materially affected. The detailed screening and selection process is illustrated in Figure 1 using a PRISMA-style flow diagram.

The dataset used for this study consisted of 475 academic publications retrieved from the WoS platform, representing scholarly contributions to the field of AIED. These publications reflect the work of a diverse range of researchers and institutions across the world. After data preprocessing and cleaning, a total of 1,545 unique terms were identified, forming the basis for the subsequent topic modeling analysis. We compiled a dataset of 475 selected papers for text mining analysis, including the text, journal names, and publication years.

In text mining research, it is necessary to preprocess the data to prepare the corpus for analysis. All text preprocessing was conducted in RStudio (version 4.3.3) using the packages *tm*, *topicmodels*, *ldatuning*, *textstem*, *tidytext*, and *stringr*. The corpus consisted of lemmatized text extracted from article titles and abstracts. First, texts were tokenized using whitespace-based segmentation implemented in RStudio. All tokens were converted to lowercase, and non-essential characters, including punctuation and numeric values, were removed during the initial cleaning stage. Following tokenization, lemmatization was conducted using the *textstem* package in RStudio, which applies a WordNet-based lemmatization procedure to normalize inflected word forms to their base forms.

To further reduce semantic redundancy, a rule-based synonym standardization procedure was implemented. A custom replacement dictionary was constructed to merge semantically equivalent terms (e.g., “institution,” “campus,” and “college” were standardized to “university”; “gpt” was standardized to “chatgpt”; “dataset” to “data”). This step improved topic interpretability by consolidating conceptually similar expressions.

Subsequently, stopwords removal was conducted using a manually curated stopwords list. In addition to conventional English stopwords, the list included domain-specific high-frequency terms (e.g., “ai,” “education,” “university”) and additional words identified through iterative inspection using TF-IDF and LDAvis diagnostics. This multi-stage stopwords refinement process ensured that overly generic terms did not dominate topic formation. Finally, a Document-Term Matrix was constructed using the tm package with a minimum word length of two characters. Sparse terms were removed using a sparsity threshold of 0.99, retaining only terms appearing in at least 1% of documents.

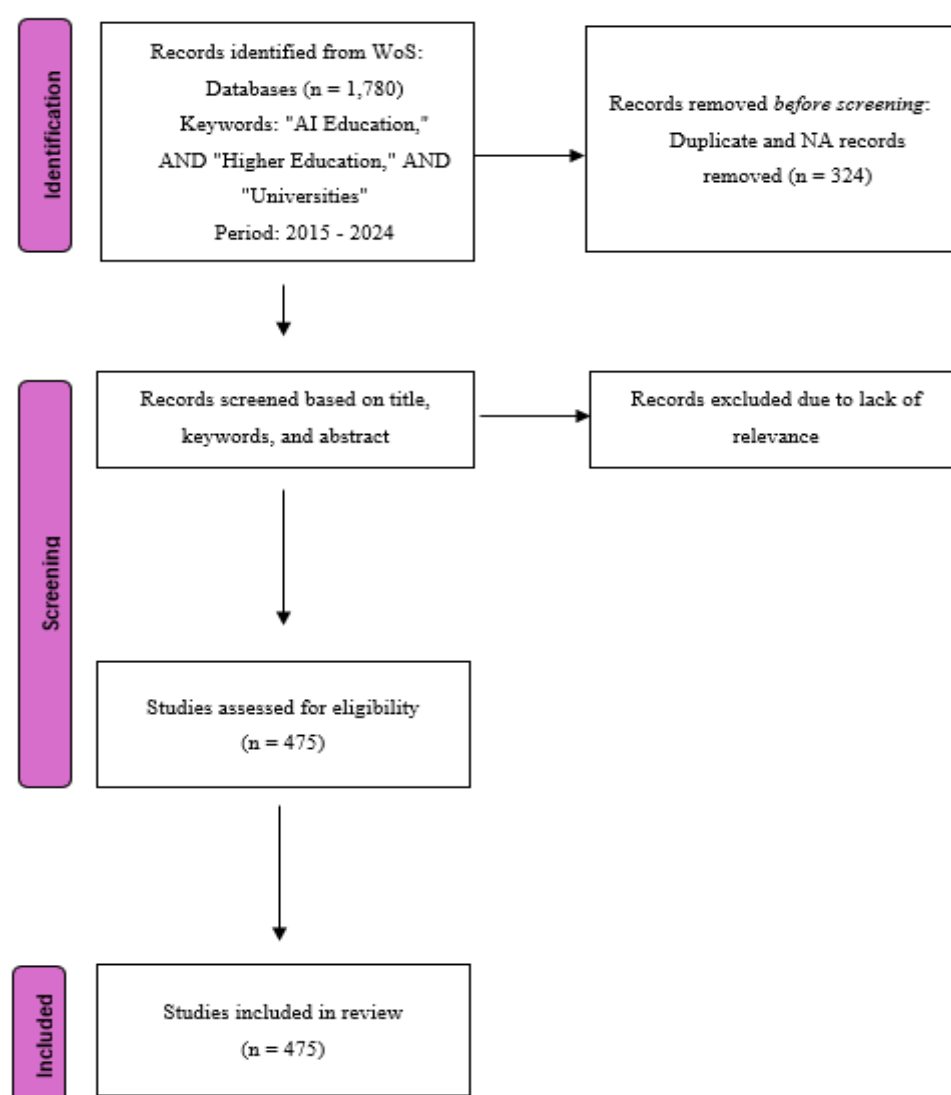


Figure 1: PRISMA Flow Diagram of the Literature Selection Process
Source: Based on the PRISMA 2020 flow diagram (Page et al., 2021)

LDA Implementation Details

Topic modeling was implemented in RStudio (version 4.3.3) using the topicmodels package. Latent Dirichlet Allocation (LDA) was estimated via Gibbs sampling with a fixed random seed (seed = 1234) to ensure reproducibility. The input to the model was a document–term matrix constructed using the tm package, with a minimum token length of two characters. Sparse terms were removed using a sparsity threshold of 0.99 to enhance model stability.

Model outputs included the topic–term distribution (β) and document–topic distribution (γ). The γ matrix was used to assign each document a dominant topic based on the highest posterior probability. Temporal variation in topic prevalence was then examined by aggregating dominant topic frequencies by publication year. This approach allowed for a longitudinal analysis of topic trends while maintaining a stable global topic structure across the entire corpus. The complete modeling script is provided as supplementary material to ensure transparency and reproducibility. LDA was selected because its probabilistic topic structure provides interpretable topic–term and document–topic distributions and enables the same global topic structure to be consistently compared across publication years. More recent embedding-based approaches, such as BERTopic and Top2Vec, can capture contextual semantic similarities through vector representations and may identify topic structures that are less dependent on word co-occurrence patterns. However, because the present study focuses on tracing the longitudinal prevalence of a stable set of topics across the 2015–2024 corpus, LDA provides a suitable and interpretable framework for the current analysis. Future studies may compare LDA with embedding-based approaches to assess the robustness of the identified thematic structures.

4. Results

TF-IDF Calculation and Visualization

To identify the most representative terms in the corpus, Term Frequency–Inverse Document Frequency (TF-IDF) was applied. TF-IDF is a widely used weighting technique in text mining and information retrieval. Its purpose is to evaluate how important a term is to a particular document while reducing the weight of terms that occur frequently across the entire corpus (Salton & Buckley, 1988). In this study, we calculated TF-IDF scores across the 475 abstracts to detect keywords that are both frequent and uniquely significant to specific documents. The top-ranked TF-IDF terms were then visualized using bar plots to show their relative importance. The specific data is shown as Figure 2. Minor adjustments to the corpus size during final validation did not materially alter the overall term distribution patterns.

The TF-IDF results provide an initial overview of the key semantic features in AI in higher education research. These high-weight terms also support the interpretation of the latent thematic structures identified through LDA topic modeling. Therefore, TF-IDF analysis serves as a complementary exploratory step by providing preliminary semantic evidence for the identification and interpretation of the five research themes.

and “level” indicate sustained interest in competency development, instructional implementation, and the practical deployment of AI-supported educational strategies. Overall, this visual representation provides a preliminary semantic overview of the corpus and served as an exploratory foundation prior to conducting LDA topic modeling.

Number of Topics Selection

This study selected 475 publications related to AIED and applied LDA topic modeling to identify key research themes. As LDA requires a predefined number of topics (k), we adopted a systematic iterative approach and estimated models with topic numbers ranging from 2 to 10. To determine the optimal number of topics, four widely used diagnostic metrics were employed for cross-validation: Griffiths et al. (2004), Cao et al. (2009), Deveaud et al. (2014), and Arun et al. (2010). Because these metrics evaluate model fit and topic redundancy from different perspectives and do not necessarily converge on a single solution, a multi-criteria evaluation strategy was adopted to avoid reliance on any single indicator.

As shown in Figure 4, model selection was guided by multiple diagnostic metrics rather than reliance on a single criterion. The Griffiths2004 metric increases markedly from $K = 2$ to $K = 5$ and then exhibits progressively smaller gains, suggesting diminishing improvements in model fit beyond five topics. Similarly, the Arun2010 metric declines sharply at lower values of K but begins to level off around $K = 5$, indicating increasing structural stability up to this point. Although the CaoJuan2009 and Deveaud2014 metrics show comparatively stronger performance at smaller topic numbers, these trends are not consistently supported across all criteria. Moreover, increasing the number of topics beyond $K = 5$ yields limited additional improvement while introducing greater model complexity and reduced interpretability.

Taken together, the stabilization patterns observed in Griffiths2004 and Arun2010, combined with considerations of parsimony and interpretive clarity, support the selection of a five-topic solution as an appropriate balance between statistical adequacy and conceptual coherence. To further assess the semantic interpretability of the selected model, topic coherence was evaluated using the C_v coherence measure based on the top 15 terms of each topic. The five-topic solution yielded an overall mean C_v coherence score of 0.451, with topic-level scores ranging from 0.284 to 0.601. These results indicated moderate overall semantic coherence, although some variation existed across individual topics. The coherence assessment was therefore treated as a complementary indicator alongside the four diagnostic metrics, parsimony, and interpretive clarity, rather than as a standalone criterion for model selection.

Beyond quantitative diagnostics, an inter-rater validation procedure was conducted to assess interpretability. For the selected five-topic model, the top 15 highest-probability terms and three to five representative articles for each topic were independently reviewed by two researchers to assess the interpretability of the topic labels. The inter-rater agreement was substantial, as both reviewers identified conceptually aligned labels for all five topics. Minor variations in wording did not affect thematic interpretation. The convergence of statistical diagnostics and interpretive validation further supported the selection of the five-topic model.

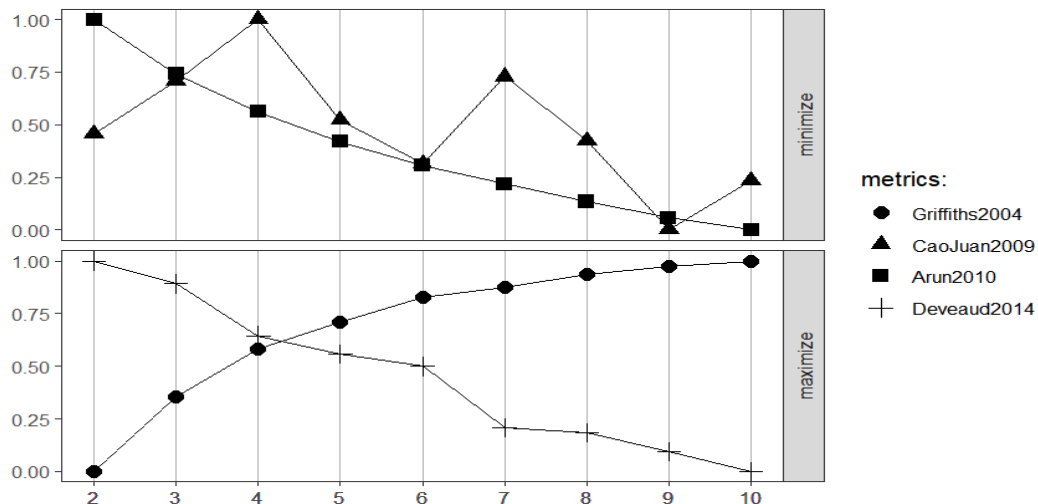
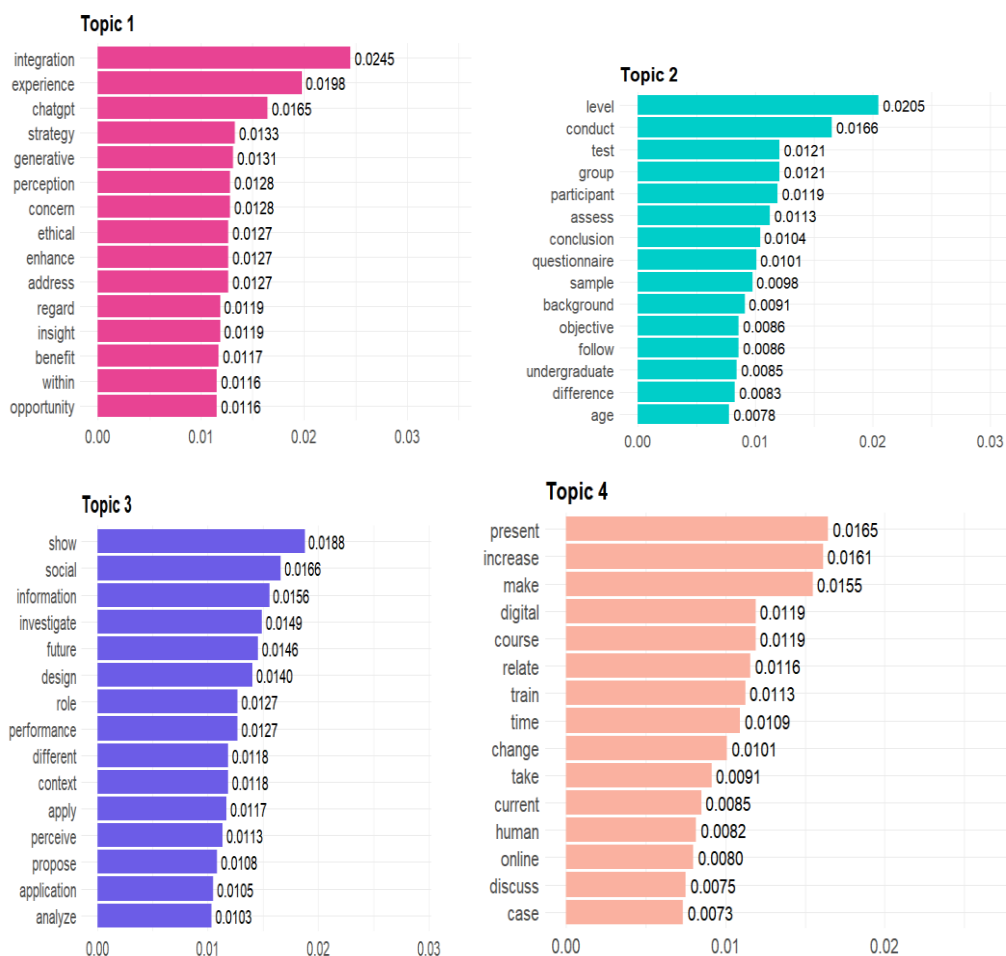


Figure 4: Number of Topics

Result of LDA Topic Modeling

Through LDA topic modeling of 475 papers related to AIED, five major topics were extracted in Figure 5. Each topic represents a key research direction in the field and is characterized by distinct keywords and educational implications. Details of each topic are as follows:



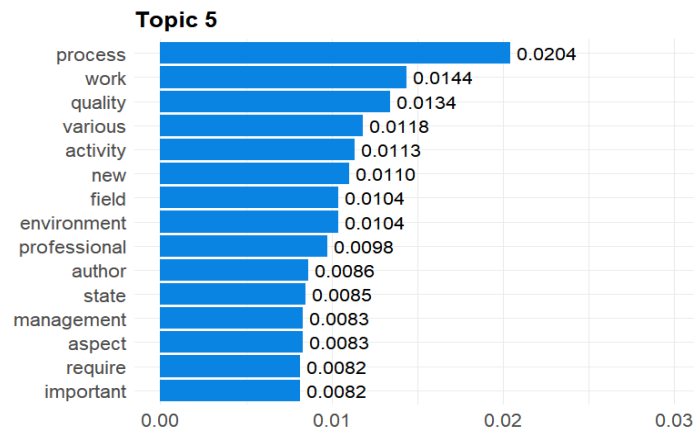


Figure 5: Keyword Probability Bar Chart of Topics

Topic 1 is labeled “Generative AI Integration” as it centers on the incorporation of generative technologies, particularly tools such as ChatGPT, into educational contexts. The prominence of terms such as integration, generative, ChatGPT, and strategy suggests a strong focus on how these tools are strategically embedded into teaching and learning practices. Additional keywords including experience, perception, concern, ethical, and benefit indicate that this line of research not only examines implementation strategies but also explores users’ experiences and evaluative perspectives. Studies within this theme frequently investigate how educators and students perceive generative AI, what opportunities it creates, and what concerns emerge in relation to its adoption. Overall, this topic reflects the rapid expansion of scholarship on generative AI integration in higher education.

Topic 2 is designated “Empirical Evaluation Studies” as its most salient terms. Conduct, participant, group, assess reflect a methodological focus on systematic investigation and participant-based assessment. These keywords point to research designs grounded in quantitative and survey-based evaluation approaches. The inclusion of terms such as level, undergraduate, background, and difference suggests comparative analyses across learner groups and contexts. Collectively, this topic represents empirical studies that measure outcomes, examine intervention effectiveness, and assess learner characteristics in AI-related educational research.

Topic 3 is labeled “Contextual AI Application” because its dominant terms such as social, information, context, apply, and analyze highlight the examination of AI technologies within broader environmental and situational frameworks. These words indicate that research in this cluster situates AI use within specific social and informational contexts rather than focusing solely on methodological evaluation. Additional terms such as role, performance, design, and future suggest forward-looking analyses of how AI applications evolve across domains. This topic thus encompasses studies that investigate contextual variation, implementation dynamics, and the functional roles of AI in diverse educational environment.

Topic 4 is titled “Digital Course Transformation” as its central keywords course, digital, online, present, train, and change point to the restructuring and adaptation of course delivery in digitally mediated settings. The clustering of these terms indicates a focus on instructional redesign and technological integration within online learning environments. Supplementary words such as increase, time, relate, and human further imply attention to interaction patterns, temporal organization, and evolving teaching practices. This thematic group therefore captures

research addressing how AI contributes to transformations in digital course formats and instructional processes.

Topic 5 is labeled “Educational Process and Quality” because its prominent terms like process, work, quality, professional, and management emphasize organizational workflows and quality-oriented dimensions of AI implementation. These keywords suggest a focus on institutional practices and professional standards rather than individual learner experiences. The presence of terms such as environment, field, state, and aspect indicates systemic perspectives on how AI technologies interact with broader educational structures. Research within this topic examines process optimization, professional practice, and structural considerations in AI-supported educational systems. For each topic, the five documents with the highest posterior topic probability (γ) were selected as representative articles. Short excerpts from the abstracts are provided in Appendix A to support semantic validation.

Transformation of Topics in 2015-2024

The overall distribution of the five themes across the entire study period was calculated to provide a structural overview of AI research in higher education. Based on the dominant topic assignment generated by the LDA model, Topic 1 represented the largest proportion of the research corpus, accounting for 117 publications (24.6%), followed by Topic 4 with 97 publications (20.4%), Topic 5 with 95 publications (20.0%), Topic 2 with 94 publications (19.8%), and Topic 3 with 72 publications (15.2%). Overall, the distribution indicates that research on AI in higher education was relatively balanced across the five thematic areas, although Topic 1 demonstrated a slightly greater presence within the literature corpus. These results provide an overview of the structural composition of AI-related higher education research and complement the subsequent analysis of topic evolution over time.

Although the five themes represent distinct research focuses, they are interconnected and collectively characterize the development of AI in higher education research. Generative AI Integration and Contextual AI Application show conceptual overlap in examining the adoption and implementation of AI technologies, while Digital Course Transformation and Educational Process and Quality reflect complementary perspectives on AI-driven pedagogical change and educational outcomes. Empirical Evaluation Studies serves as a cross-cutting theme by providing evidence-based assessments of AI applications and their educational impacts across different contexts. Overall, these themes represent interconnected dimensions ranging from technological adoption and application to pedagogical transformation and quality evaluation.

Figure 6 illustrates the annual distribution of the five identified AIED research themes across the global WoS-indexed corpus between 2015 and 2024. Rather than imposing a priori theoretical stages, phase boundaries were identified based on structural shifts in total annual publication volume. Total output remained below 20 publications per year prior to 2019, increased substantially in 2020 (from 10 to 26 publications), and accelerated sharply after 2022 (from 43 in 2022 to 81 in 2023 and 245 in 2024). These inflection points provide an empirical basis for distinguishing three periods: Initial Phase (2015–2019), Integration Phase (2020–2022), and Maturation Phase (2023–2024).

During the Initial Phase (2015–2019), overall publication volumes remained limited and thematically dispersed. Topic 5 (Educational Process and Quality) appeared earliest and maintained relative prominence during this period, reaching 8 publications in 2018. This indicates that early AIED research focused on institutional processes, professional practices,

and quality-related considerations rather than advanced technological applications. During the same period, Topics 2 and 4 began to surface incrementally. Topic 2 (Empirical Evaluation Studies) increased from 0 in 2015 to 7 in 2018, while Topic 4 (Digital Course Transformation) remained modest but visible. The absence of a dominant theme suggests that early-stage research was exploratory and conceptually open-ended, corresponding to the initial diffusion stages in DOI where knowledge accumulation precedes large-scale implementation. This phase primarily reflects a model-oriented orientation within DDR, where structural and pedagogical models were conceptualized before large-scale technological product diffusion.

In the Integration Phase (2020–2022), a noticeable inflection point emerges in 2020. Topic 4 rose sharply from 1 publication in 2019 to 10 in 2020 and 11 in 2021, indicating intensified attention to digital course restructuring and online learning environments. This surge unfolded within a broader policy context marked by the Ministry of Education’s 2020 Education Informatization Implementation Plan, which articulated the creation of a “human-centered future intelligent education environment.” The policy emphasized next-generation digital textbooks, AI-driven personalized learning systems, learning analytics platforms, and expanded K-MOOC offerings (Ministry of Education, 2020). The temporal alignment between the policy agenda and the observed increase in digital course transformation research suggests that institutional priorities may have reinforced scholarly attention to technology-enabled instructional redesign.

Simultaneously, Topic 2 experienced sustained growth, peaking at 15 publications in 2022. The increasing prevalence of participant-based assessment and evaluation-oriented research reflects a methodological consolidation phase, in which AI applications were subjected to systematic empirical scrutiny. Topic 3 (Contextual AI Application) also expanded gradually during this period, rising from 1 publication in 2019 to 4 in 2022. This pattern suggests growing awareness of the social and contextual dimensions of AI deployment, particularly as large-scale implementation required adaptation across diverse educational settings. Taken together, these shifts indicate a transition from exploratory inquiry toward structured application and evaluation. Rather than representing isolated academic developments, these trends appear embedded within a broader national strategy promoting digital transformation and AI-enabled education. This stage reflects a transitional configuration in DDR, where model-oriented instructional restructuring and product-oriented technological deployment advanced simultaneously.

By the Maturation Phase (2023–2024), research themes exhibited both acceleration and diversification. The most pronounced structural shift occurs after 2022. Topic 1 (Generative AI Integration) increased dramatically from 1 publication in 2022 to 23 in 2023 and 89 in 2024, representing the most substantial single-topic expansion in the entire dataset. This surge coincides with the rapid global diffusion of generative AI technologies, particularly ChatGPT, and signals a paradigmatic reorientation in research focus. This acceleration also unfolded within a significant policy context. The 2023 Digital-Based Education Innovation Plan formally incorporated generative AI into the national education reform agenda, advocating the development of a “personalized AI education system for all,” strengthening teacher AI competencies, and promoting instructional model innovation (Ministry of Education, 2023). The temporal correspondence between the release of this policy framework and the sharp increase in Topic 1 publications suggests that institutional endorsement of generative AI may have reinforced scholarly engagement with integration strategies, pedagogical applications, and user-centered adaptation.

Topic 3 also exhibited substantial growth, reaching 51 publications in 2024, indicating heightened attention to contextual and societal implications of AI deployment. As generative technologies became more widely adopted, research expanded beyond implementation toward examining broader environmental, informational, and systemic dynamics. Meanwhile, Topics 2 and 4 remained robust (40 and 39 publications in 2024, respectively), demonstrating that empirical evaluation and digital course transformation continued to develop alongside generative innovation rather than being displaced by it. Importantly, Topic 5 maintained a steady presence without comparable acceleration, suggesting that process- and quality-oriented discussions remained foundational but were not the primary drivers of the post-2022 expansion. Taken together, the 2023 policy initiative appears to have provided institutional framing and strategic direction that coincided with, and potentially amplified, the rapid scholarly pivot toward generative AI integration. This phase demonstrates a strong product-oriented acceleration, driven by generative AI technologies, while model-oriented stabilization mechanisms remained active through evaluation and contextual adaptation research.

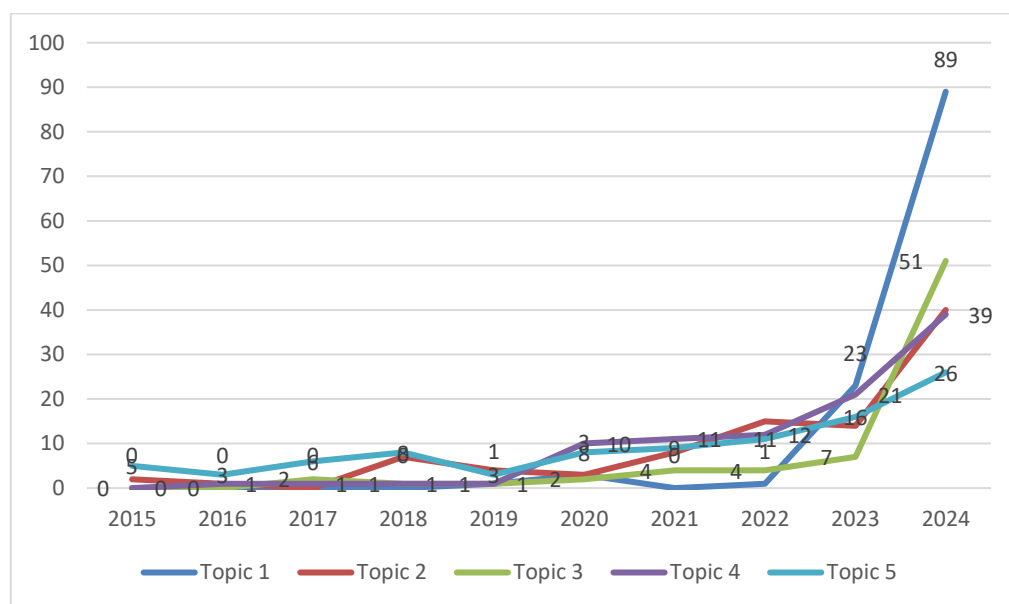


Figure 6: Frequency of Topic Occurrences by Year

To substantiate stage assignments, publication counts were aggregated by period and compound annual growth rates (CAGR) were calculated for each topic over the full 2015–2024 timeframe (Table 1). The results reveal differentiated diffusion velocity across themes. Topic 1 (Generative AI Integration) exhibits the highest overall growth rate (CAGR = 89%), with the majority of output concentrated in the final period. Topics 4 (84%) and 3 (71%) also demonstrate strong acceleration, particularly after 2022. In contrast, Topic 5 shows comparatively moderate long-term growth (31%), reflecting early prominence followed by stabilization. Topic 2 presents steady expansion (40%) across phases without abrupt escalation.

These quantitative indicators provide empirical grounding for diffusion-based interpretation. The first period is characterized by low publication intensity and limited technological emphasis, the second by structured expansion and institutional consolidation, and the third by accelerated growth and thematic concentration in generative AI research. Stage classification is therefore derived from observable publication dynamics rather than theoretical assumption.

Table 1: Stage-Level Publication Counts and Growth Rates of AIEd Topics

Topic	Initial Phase (2015–19)	Integration Phase (2020–22)	Maturation Phase (2023–24)	CAGR (2015–24)
T1: Generative AI Integration	1	4	112	89%
T2: Empirical Evaluation Studies	14	26	54	40%
T3: Contextual AI Application	4	10	58	71%
T4: Digital Course Transformation	4	33	60	84%
T5: Educational Process and Quality	25	28	42	31%

While the topic analysis reveals empirical patterns of thematic shifts, it does not by itself explain the underlying developmental logic. Therefore, Table 2 advances a conceptual integration that connects research themes with DOI stages and DDR pathways, offering a theoretically grounded interpretation of how technological experimentation evolves into systemic educational reform.

Table 2: Mapping AI Research Topics onto DOI Phases and DDR Dual Research Paths

T5: Educational Process & Quality	Initial Phase (2015–2019)	Model-Oriented
T4: Digital Course Transformation	Integration Phase (2020–2022)	Product-Oriented
T2: Empirical Evaluation Studies	Integration Phase (2020–2022) ~ Maturation Phase (2023–2024)	Product-Oriented
T3: Contextual AI Application	Maturation Phase (2023–2024)	Model-Oriented
T1: Generative AI Integration	Maturation Phase (2023–2024)	Model-Oriented & Product-Oriented

5. Conclusion

This study applied LDA topic modeling to examine the thematic evolution of AIEd research published between 2015 and 2024. Five dominant themes were identified: Generative AI Integration, Empirical Evaluation Studies, Contextual AI Application, Digital Course Transformation, and Educational Process and Quality. Their temporal dynamics reveal a structurally differentiated trajectory. Quantitative indicators, including annual publication counts and compound annual growth rates, show three empirically distinguishable phases: Initial Phase (2015–2019) marked by modest output and emphasis on process and evaluation; Integration Phase (2020–2022) characterized by expansion in digital transformation and empirical assessment; and Maturation Phase (2023–2024) driven primarily by generative AI-related research. These shifts indicate that AIEd scholarship has moved from foundational process-oriented inquiry toward large-scale technological embedding.

Interpreted through DOI as an analytical lens, this trajectory resembles a diffusion-like pattern in scholarly attention. Early research remained relatively dispersed, followed by broader attention to AI-supported instructional applications and, after 2022, a sharp expansion of generative AI-related scholarship. These patterns should not be interpreted as direct evidence of institutional adoption; rather, they indicate changes in the emergence, expansion, and stabilization of research attention. The DDR framework further complements this temporal interpretation by highlighting shifts between product/process-oriented and broader model/system-oriented research.

The findings also provide practical and policy implications for higher education stakeholders. For higher education institutions, the identified themes suggest the need to develop adaptive strategies for generative AI integration, contextual AI applications, and digital course transformation, supported by evidence-based evaluation practices to enhance educational quality. For education regulators, the evolving AI landscape highlights the importance of establishing flexible policy frameworks, ethical guidelines, and quality assurance mechanisms to support responsible AI adoption across diverse higher education contexts.

Despite its contributions, this study has several limitations. First, this study only included open-access articles retrieved from the Web of Science database to ensure data accessibility and research reproducibility. However, this approach may have introduced selection bias by excluding relevant subscription-based publications, which may limit the representativeness of the findings. Future studies should consider integrating multiple databases and including both open-access and non-open-access publications to provide a more comprehensive understanding of AI in higher education research. Second, as a statistical method, topic modeling captures patterns of lexical co-occurrence rather than the theoretical depth or methodological rigor of individual studies. Third, although the search strategy required the presence of both AI-related and higher-education-related keywords, some clusters reflect AI applications to university student populations (e.g., mental health prediction or behavioral profiling) rather than direct pedagogical integration. This indicates interdisciplinary spillover within the broader AIED discourse and suggests that the boundaries of AIED research are conceptually porous. Fourth, the adoption of only Web of Science as the single data source narrows disciplinary and regional coverage of AI higher education research; subsequent research may combine Scopus and Dimensions to mitigate database bias. Finally, topic modeling involves several methodological considerations. Although LDA effectively identifies latent thematic structures based on lexical co-occurrence patterns, topic interpretation and labeling inevitably involve analytical judgment, and thematic overlaps may persist despite inter-rater validation. The labels adopted in this study therefore represent one plausible interpretation of the underlying keyword distributions, and some topics may also support alternative but closely related interpretations. In addition, comparisons between LDA and emerging embedding-based approaches, such as BERTopic or Top2Vec, may provide further insights into the robustness and consistency of identified research themes. Future research may also extend the present analysis through cross-national and disciplinary comparisons and by integrating citation-network analysis with topic modeling to provide a more comprehensive understanding of the intellectual and thematic structure of AIED research.

References

- Abdalla, A. A., Bhat, M. A., Tiwari, C. K., Khan, S. T., & Wedajo, A. D. (2024). Exploring ChatGPT adoption among business and management students through the lens of diffusion of innovation theory. *Computers and Education: Artificial Intelligence*, 7, 100257.
- Arun, R., Suresh, V., Madhavan, C. E. V., & Murthy, M. N. (2010). On finding the natural number of topics with latent Dirichlet allocation: Some observations. In M. J. Zaki, J. X. Yu, B. Ravindran, & V. Pudi (Eds.), *Advances in knowledge discovery and data mining* (pp. 391–402). Springer.
- Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in the higher education: A systematic literature review and research challenges. *International Journal of Educational Research*, 127, 102411.
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21, Article 4.
- Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, 72(7–9), 1775–1781.
- Chen, X., Xie, H., & Hwang, G.-J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence*, 1, 100005.
- Chu, H.-C., Hwang, G.-H., Tu, Y.-F., & Yang, K.-H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, 38(3), 22–42.
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), 22.
- Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Document numérique*, 17(1), 61–84.
- Gašević, D., Siemens, G., & Sadiq, S. (2023). Empowering learners for the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, 4, 100130.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(suppl 1), 5228–5235.
- Hinojo-Lucena, F.-J., Aznar-Díaz, I., Cáceres-Reche, M.-P., & Romero-Rodríguez, J.-M. (2019). Artificial intelligence in higher education: A bibliometric study on its impact in the scientific literature. *Education Sciences*, 9(1), 51.
- Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, 8, 100348.

Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: A systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education*, 22, Article 7.

Lee, H. E., Kim, H., & Kim, J. H. (2025). A study on AI-TPACK-based instructional design for instructors' use of generative AI in higher education. *The Journal of Korean Education*, 52(2), 249–274.

Ministry of Education. (2020). *2020 Education informatization implementation plan*. Ministry of Education.

Ministry of Education. (2022). *AI ethics principles for the education sector*. Ministry of Education.

Ministry of Education. (2023). *Digital-based education innovation plan*. Ministry of Education.

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.

Richey, R. C., & Klein, J. D. (2014). *Design and development research: Methods, strategies, and issues*. Taylor & Francis.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513–523.

Siemens, G., Gašević, D., & Dawson, S. (Eds.). (2015). *Preparing for the digital university: A review of the history and current state of distance, blended, and online learning*. MOOC Research Initiative.

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.

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21, Article 40.

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.