Journal of Hypermedia & Technology-Enhanced Learning
Vol. 4, No. 2, June (2026), Pages 145–171
e-ISSN: 2985–9166 | 10.58536/j-hytel.224

RESEARCH ARTICLE

Mapping Three and a Half Decades of AI and Education Literature: Trends, Gaps, and Future Directions

Constance de Saint Laurent[1] *, Vlad Glaveanu[2,4], Ioana Literat[3], Ingunn Ness[4]

[1] Department of Psychology, Maynooth University, Maynooth, Co. Kildare, **Ireland**
[2] DCU Business School, Dublin City University, Dublin, **Ireland**
[3] Teachers College, Columbia University, New York, **United States**
[4] SLATE: Centre for the Science of Learning and Technology, University of Bergen, Bergen, **Norway**

**Corresponding Author: vlad.glaveanu@dcu.ie*

This article contributes to:

4 QUALITY EDUCATION

SUSTAINABLE DEVELOPMENT GOALS

ABSTRACT

Research on artificial intelligence (AI) in education has expanded at an unprecedented speed, yet its development remains fragmented, hype-driven, and unevenly distributed across disciplines and regions. This paper presents one of the most comprehensive mappings to date, analysing 16,564 publications (1990–2025) identified through Web of Science and refined to 4,626 core contributions using a hybrid method that combined computational topic modelling, AI-assisted filtering, and human validation. To complement this macro-level mapping, we conducted a qualitative overview of the most cited papers in each major research topic, providing interpretive depth on how influential works have shaped debates on adoption, engagement, personalisation, chatbots, evaluation, risks, monitoring, reviews, and teaching AI. The findings suggest exponential growth since 2020, with major surges during the COVID-19 pandemic and following the release of ChatGPT, but also enduring fragmentation, a relative decline in evaluation studies, and a saturation of attitudinal surveys. Promising areas such as personalisation, teaching AI, and engagement have lost visibility despite their long-term importance, while risks and ethical issues remain underexplored and poorly connected to technical research. Citation patterns further show systemic imbalances, with education papers disproportionately cited and engineering outputs under-recognised. We conclude with recommendations to strengthen interdisciplinarity, address neglected areas, include historical perspectives, and address systemic challenges in order to foster a more cumulative, integrative, and critical field.

KEYWORDS

Artificial intelligence; review; education; technology; ethics; evaluation

Received: Mar. 04, 2026; *Revised:* Mar. 25, 2026; *Accepted:* Jul. 14, 2026; *Published Online:* Jul. 25, 2026

How to Cite: de Saint Laurent, C., Glaveanu, V., Literat, I., & Ness, I. (2026). Mapping Three and a Half Decades of AI and Education Literature: Trends, Gaps, and Future Directions. *Journal of Hypermedia & Technology-Enhanced Learning, 4*(2), 145-171. https://doi.org/10.58536/j-hytel.224

Published by Sagamedia Teknologi Nusantara

1. INTRODUCTION

Although generative Artificial Intelligence (AI) has recently captured widespread attention, AI has been shaping educational technology for over three decades, evolving from early intelligent tutoring systems and adaptive learning environments in the 1990s [1], [2] to today's advanced large language models and generative AI tools. While the core aims – personalising learning experiences, automating administrative tasks, and supporting student assessment – have remained remarkably consistent [3], [4], the technological capabilities, implementation contexts and ethical challenges have evolved significantly, reflecting the rapid pace of technological innovation and its integration into everyday educational contexts [5], [6], [7]. Recent developments in generative AI, particularly following the release of ChatGPT in late 2022, have intensified attention to this field [8]. In the past couple of years, research on AI in education has expanded from a niche field, with fewer than 30 papers published in the entire decade of the 1990s, to over 1,700 publications in the first eight months of 2025 alone.

Yet, despite this surge in interest, research on AI in education remains fragmented and siloed. This fragmentation reflects both the field's rapid evolution and the diversity of approaches across different contexts and disciplines (see also [9]). The field's inherently multi-disciplinary nature – spanning computer science, engineering, education, psychology, linguistics, and more – has tended to reinforce disciplinary siloes rather than encourage researchers to engage in cross-disciplinary dialogue [10]. Additionally, regional differences in educational priorities, technological infrastructure, and research traditions have led to parallel but disconnected streams of work [11]. This siloing makes it difficult to build cumulative knowledge about what works in AI education, particularly because the same educational challenges are often addressed independently across disciplines and regions.

The field of AI in education now stands at a critical juncture. The rapid acceleration of research has created an urgent need for a comprehensive understanding of the field's development, current state, and future trajectory. Existing reviews, typically examining fewer than 200 papers (with the exception of some bibliometric analyses; see [12]), have provided valuable but limited insights into specific aspects of AI in education. At this point, there is a need for systematic, large-scale reviews that can reveal broader patterns and interconnections (or lack thereof), allowing us to map the evolution of research priorities, methodological approaches, and regional variations in ways that smaller-scale analyses cannot capture.

In line with this need, our paper presents an analysis of AI in education research based on a comprehensive corpus of 16,564 articles published between 1990 and 2025. Using a combination of computational and qualitative methods, we examine how research in this field has evolved across disciplines, regions, and methodological approaches. Through careful AI-assisted filtering, we identify 4,626 articles focused primarily on AI and education, enabling detailed analysis of the interplay between technical innovation and educational implementation, while mapping geographical and disciplinary variations in research priorities.

The scale of our analysis represents a significant advance over existing reviews. Our initial corpus of over 16,000 articles enables us to identify patterns that smaller-scale analyses miss, such as how research priorities shift during global events like the COVID-19 pandemic, and how different regions respond to technological developments in AI. Statistical analysis of this larger corpus also reveals previously undetected relationships between research topics and citation patterns across disciplines.

Our methodological approach carefully integrates computational tools with traditional research methods to enable systematic analysis of large-scale educational research. Building on pragmatist approaches to mixing methods [13], we combine established topic modelling techniques with AI-assisted coding and a qualitative zoom-in to process and analyse our extensive corpus. Our methodological integration ensures both breadth and depth, allowing us to trace historical developments while capturing emerging themes in AI-driven education.

This hybrid approach maintains methodological rigor while expanding the scope of systematic reviews through computational support. By employing a carefully calibrated combination of traditional topic modelling and Large Language Models, our method facilitates the processing of larger corpora while remaining grounded in human oversight and interpretation. Importantly, this approach addresses the epistemological tensions inherent in applying AI to analyse human learning and teaching practices [14], recognising that computational tools must be guided by human expertise and understanding of educational contexts. The resulting framework provides a replicable methodology for future large-scale studies in educational research, demonstrating how computational methods can productively complement established analytical approaches.

This paper proceeds as follows: After detailing our methodological approach, we present our findings organised around three main axes: temporal evolution of the field, geographical patterns, and thematic focus areas. To complement our quantitative analysis, we also wrote detailed narrative summaries of the most-cited papers within each of the major themes identified by the quantitative review, summaries that helped add depth to our discussion section. These summaries provide a fuller sense of the most prominent topics of research in each of these areas, as well as the diversity of methodological approaches that are being employed. Thus, by combining broad computational analysis with close reading of influential works, we offer both macro-level patterns and granular insights into how the field has developed. Finally, in our conclusion, we discuss the implications of these findings for research and practice and conclude with recommendations for future directions in AI and education research.

2. Methods

2.1. Data Collection

We searched the database Web of Science on September 26th, 2024, for journal articles and conference proceedings published in English with the following words in the abstract: ((EDUCAT* OR TEACH*) AND (AI OR 'ARTIFICIAL INTELLIGENCE')). The search returned 8252 articles.[1] The full records and reference lists were then exported in .csv format and analysed using Python, as described below.

The search was conducted a second time on August 15th 2025, to update the dataset. 8312 new results were found. The procedures described below were developed using the first dataset and were carried out again with the updated dataset. The results presented below correspond to the whole dataset of 16564 articles.

The search was limited to Web of Science for two main reasons. First, databases use different methods to attribute keywords to articles – given that the keywords proposed by authors are too often meaningful only within a limited field of research – as well as topics and areas of research. This makes comparisons difficult. We therefore decided to limit our search to a single database. Second, we wanted a dataset reflecting the diversity of approaches to AI and education, and the Web of Science

[1] The word 'learn' or 'learning' could not be used, as machine learning is a commonly used term in AI research and is most often not related to education at all.

is one of the few databases that provides nearly balanced coverage across disciplines ranging from engineering to the humanities.

2.2. Selecting Articles and Analysis

A quick scan of some of the abstracts revealed that a large proportion was only tangentially about AI and Education. Many of them mentioned education as one of the fields affected by AI, or, conversely, included AI in a list of changes likely to affect education, without making it a focal point of the article. To keep the articles that focused substantively on AI and Education, we used a two-step method assisted by ChatGPT 4.0. In the first step, we used a ChatGPT assistant to summarise the abstracts in around 50 words. ChatGPT is particularly apt at this type of tasks, and an assistant allows us to give clear and specific instructions (see supplemental materials for the exact prompt used). The instructions were tested on 50 abstracts and tweaked until sufficient quality was obtained. Because assistants are queried through an API – an application programming interface that allows us to send data directly from a Python script – they can handle large amounts of queries. They also avoid 'leaks' between queries – where the answer to a prompt is affected by previous answers – as the conversation can be reset each time.

Summarising the abstracts served a double aim. First, it helped us strip them of content not directly relevant to the topic of the article – such as descriptions of the research's background in broad terms or of future research directions. Instead, the summaries focused on what the article was mainly about. Second, it translated the abstracts into a unified plain language, facilitating topic modelling (see below).

In the second step, we used another ChatGPT assistant, this time to determine if the article summarised was about AI and Education. The exact prompt can be found in the supplemental materials, and it was tested with 50 abstracts that had been manually classified by the first author. The summaries could be classified as 'on topic', 'off topic', or 'maybe on topic'. The last category was added because it greatly improved the results. While none of the 'maybe' articles inspected by hand in the test batch were found to be on topic by the manual coder, they were indirectly relevant. Without the 'maybe' category, they were sometimes classified as on topic by ChatGPT. With the inclusion of a 'maybe' category, agreement between the AI-assisted classification and manual coding was substantially improved in the validation sample. While the model performed consistently well in distinguishing clearly on-topic from off-topic articles, some borderline cases remained, particularly for indirectly relevant contributions. We therefore treat the classification as highly reliable but not error-free.

3. RESULTS

All the analyses presented below were done on the 4626 articles classified as primarily on AI and Education by the ChatGPT assistant (2487 until September 2024, 2139 between September 2024 and August 2025). Unless specified, only descriptive statistics were used, as the aim is not to generalise results from the WoS database to all the academic articles published on the topic.

Publications on AI and Education began appearing in the 1990s, but remained rare until the late 2010s (see Table 1). They have been growing exponentially since then, with 1759 publications in the first eight months of 2025. To facilitate subsequent analyses, earlier years will be grouped together.

Table 1. Number of Papers Published on Education and AI Each Year

1990's		2000's		2010's		2020's	
Year	**n**	**Year**	**n**	**Year**	**n**	**Year**	**n**
1990	1	2000	3	2010	3	2020	86
1991	1	2001	0	2011	7	2021	211
1992	9	2002	3	2012	4	2022	401
1993	1	2003	6	2013	7	2023	569
1994	5	2004	2	2014	4	2024	1424
1995	2	2005	1	2015	5	2025	1759[2]
1996	0	2006	4	2016	8		
1997	2	2007	1	2017	25		
1998	4	2008	4	2018	8		
1999	2	2009	8	2019	46		
Total	**27**	**Total**	**32**	**Total**	**117**	**Total**	**4450**

3.1. Validation Procedures

To ensure the robustness of the AI-assisted filtering and classification procedures, several validation steps were undertaken. First, both the summarisation and classification prompts were iteratively tested on a subset of abstracts (n = 50) manually coded by the first author, allowing refinement of instructions and categories. Second, a validation sample was used to compare AI-generated classifications with human judgement, focusing particularly on borderline cases. The introduction of a 'maybe' category reduced false positives and improved alignment between automated and manual coding.

Third, manual inspection was conducted at multiple stages of the analysis, including during topic labelling and cluster interpretation, during which the first 30 papers in each cluster were qualitatively reviewed. This allowed us to identify overlaps between clusters and merge conceptually similar topics where appropriate.

Despite these precautions, the approach remains subject to limitations inherent in large-scale automated classification. In particular, reliance on summaries rather than full texts and the assignment of each paper to a single primary topic may obscure more nuanced or multidimensional contributions. The findings should therefore be interpreted as indicative of broad patterns rather than precise categorical boundaries.

3.2. Regions

To analyse the geographical repartition of the papers, the country of the corresponding author was extracted. The countries were then grouped by region. Figure 1 shows the number of papers published from each region over time, and Figure 2 shows, for each year, the percentage that came from each region. Because the number of publications on AI and Education has exploded in recent years, looking at frequencies provides a better sense of the evolution across different regions.

Europe, and to a lesser extent North America, were the most prolific regions until the mid-2010s, when other regions began to join in. Between 2020 and 2022, at the height of the pandemic, production exploded in East Asia, then slowed in 2023 and picked back up in 2024 and 2025. While researchers in East Asia seem to have temporarily decreased their attention to the topic after the pandemic, the rest of the world appears to have been activated by the launch of ChatGPT.

[2] Until August 15^{th}, 2025

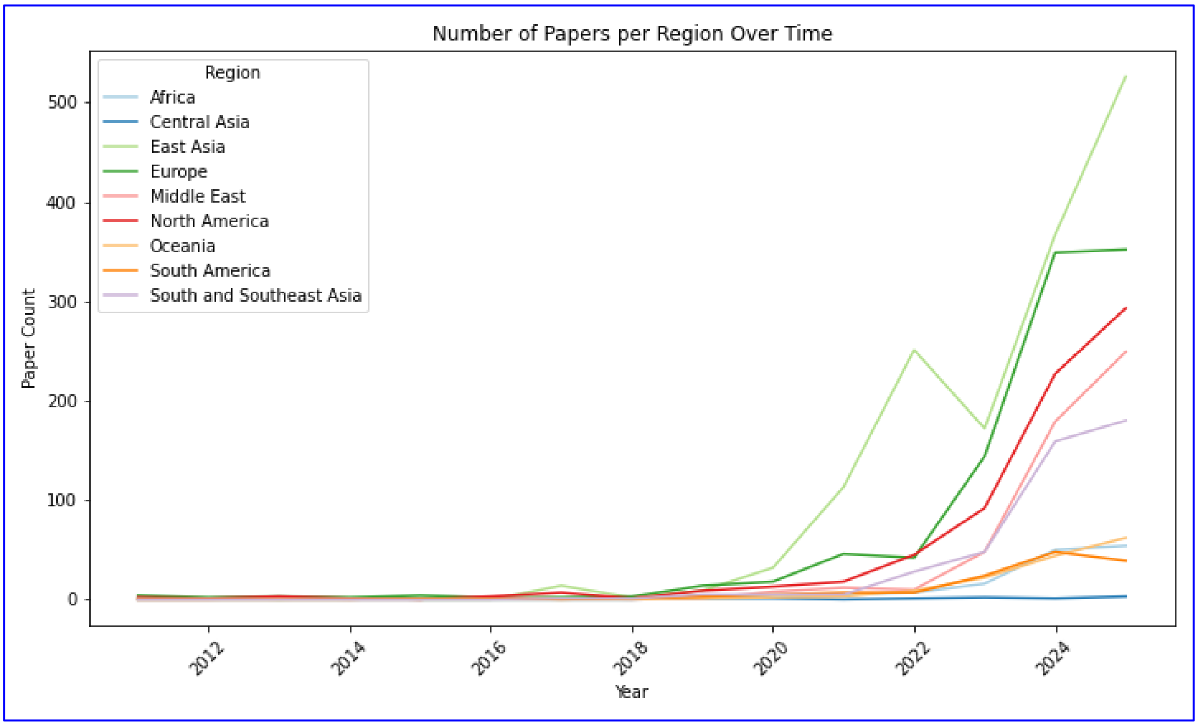

Figure 1. Number of Papers Published by Region per Year, 2011-2025

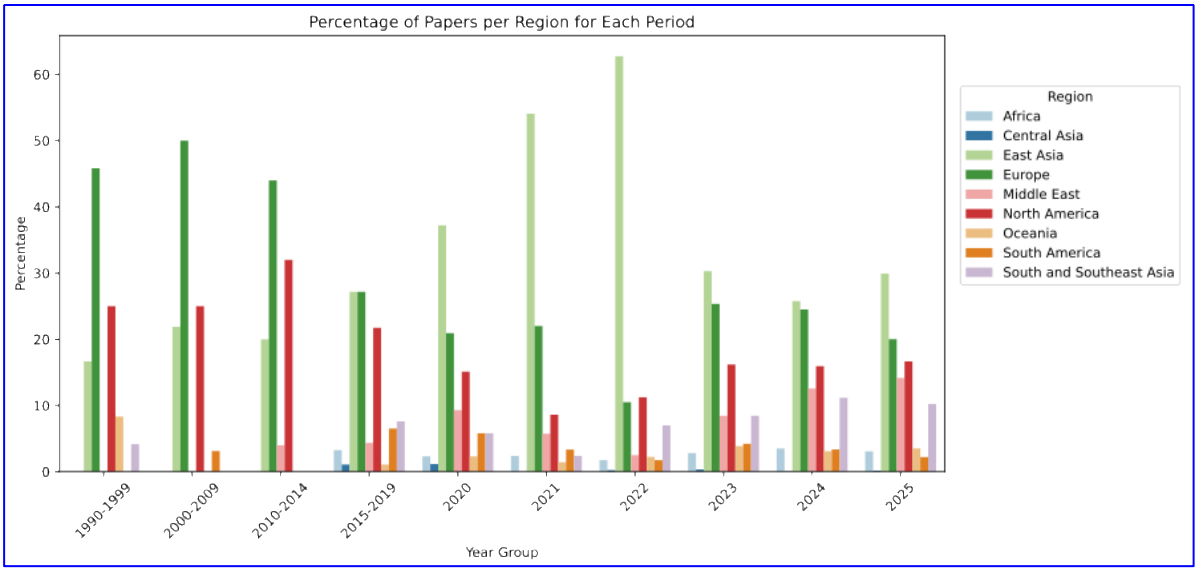

Figure 2. Percentage of Papers Published by Region Over Time

Indeed, there have been stark increases in the number of papers published in Europe, North America, the Middle East, and South and Southeast Asia (Figure 2 shows these regions slowing down in 2025, but there are only 8 months of data for that year). For the remaining analyses involving location, only the five most prolific regions have been kept.

3.3. Discipline

We next looked at which disciplines were most represented in the sample. We regrouped the disciplines identified by Web of Science into broader fields (see the conversion table in the supplemental materials). Given the importance of Education and Engineering in the corpus, they were kept as independent fields. Most of the papers fall within Education and Engineering, with 76.1% of papers (n=3519) in at least one of the two domains, as is to be expected for such a topic.

As shown in Table 2, each paper could be classified in multiple fields, although the vast majority fell into one area only (3808 papers, 82.3%). While papers outside of Education and Engineering were more likely to involve multiple disciplines, they rarely bridged the gap with these core fields of AI in Education. Instead, they tended to collaborate within clusters of related disciplines. For example, when Life Sciences papers were interdisciplinary, more than half collaborated with Physical Sciences rather than connecting with Education or Engineering expertise. This suggests that even as AI in Education attracts broader academic interest, new perspectives are developing alongside rather than in dialogue with core educational and technical expertise. The lack of interdisciplinarity has been decried in many areas, but it is particularly striking for an area of research that is, by definition, at the intersection between two disciplines.

Table 2. Co-Occurrences Between Disciplines and Number of Papers. Diagonal: Papers falling in only one discipline. Percentages in columns.

	Arts & Humanities	Education	Engineering	Health Sciences	Life Sciences	Physical Sciences	Social Sciences
Arts & Humanities	**30 (40.5%)**	12 (0.5)	2 (0.1%)	0	0	2 (0.5%)	30 (3.6%)
Education	12 (16.2%)	**1958 (80.5%)**	171 (13.6%)	58 (20.6%)	3 (1.7%)	26 (6.5%)	214 (25.5%)
Engineering	2 (2.7%)	171 (7.0%)	**918 (73.0%)**	10 (3.6%)	7 (4.0%)	133 (33.5%)	22 (2.6%)
Health Sciences	0	58 (2.4%)	10 (0.8%)	**177 (63.0%)**	32 (18.1%)	0	16 (1.9%)
Life Sciences	0	3 (0.1%)	7 (0.6%)	32 (11.4%)	**27 (15.3%)**	109 (27.5%)	0
Physical Sciences	2 (2.7%)	26 (1.1%)	133 (10.6%)	0	109 (61.6%)	**128 (32.2%)**	2 (0.2%)
Social Sciences	30 (40.5%)	214 (8.8%)	22 (1.7%)	16 (5.7%)	0	2 (0.5%)	**570 (68.0%)**
Total (100%)[3]	**74**	**2432**	**1258**	**281**	**177**	**397**	**839**

Looking at the evolution in time (Figure 3), we can see that up until 2020, the vast majority of papers on AI and Education were published in Engineering or Education, as expected. Engineering had a slight head start up until the mid 2010s, and completely took over during the pandemic, when other disciplines slowly started to join in. For the past three years, however, Education has been the discipline most represented, with over a thousand papers published in the first eight months of 2025. Regional analysis (Figure 4) helps explain this disciplinary divide. East Asia, which produced the majority of publications during the pandemic, often approaches AI in Education from an Engineering perspective. In contrast, in other regions, particularly North America and the Middle East, this research tends to originate almost exclusively from the field of Education.

Beyond those two disciplines, two things are worth noting. First, there has been an increased interest from the social sciences, disciplines particularly represented in North America and Europe. Second, the Arts and Humanities are notably underrepresented in the debate. Considering that these disciplines engage deeply with questions of ethics, cultural practices, and the very concept of

[3] Includes 13 articles with more than 2 disciplines. Arts & Humanities: 1; Education: 9; Engineering: 4; Health Sciences: 11; Life Sciences: 1; Physical Sciences: 2; Social Sciences: 12.

'generating' new artifacts, their absence from discussions on AI and education represents a significant missed opportunity to incorporate valuable perspectives.

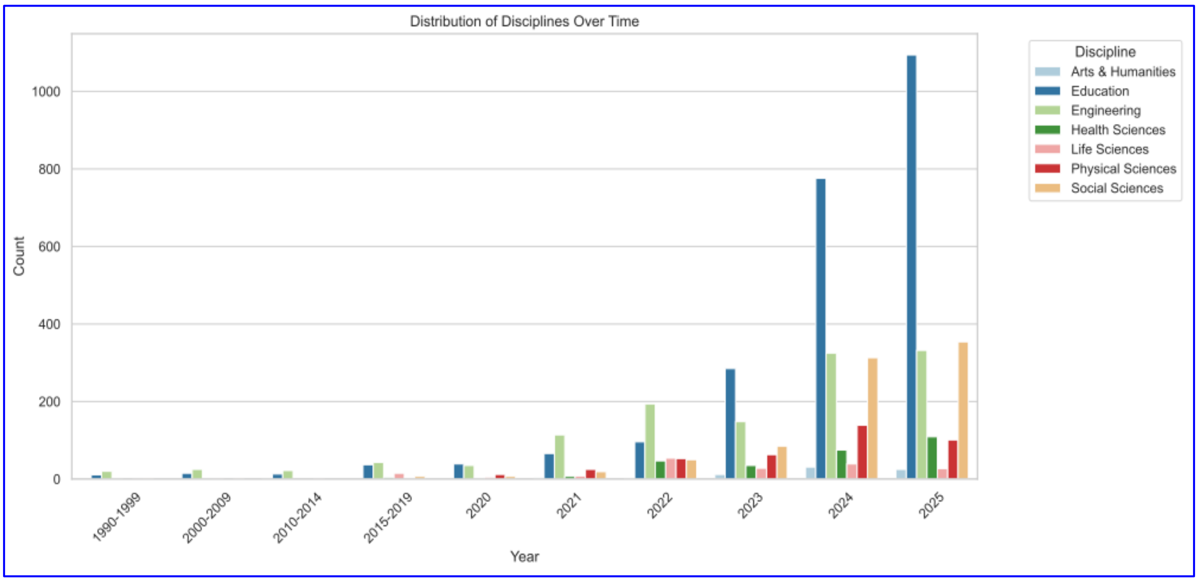

Figure 3. Number of Papers Published by Discipline per Year over Time

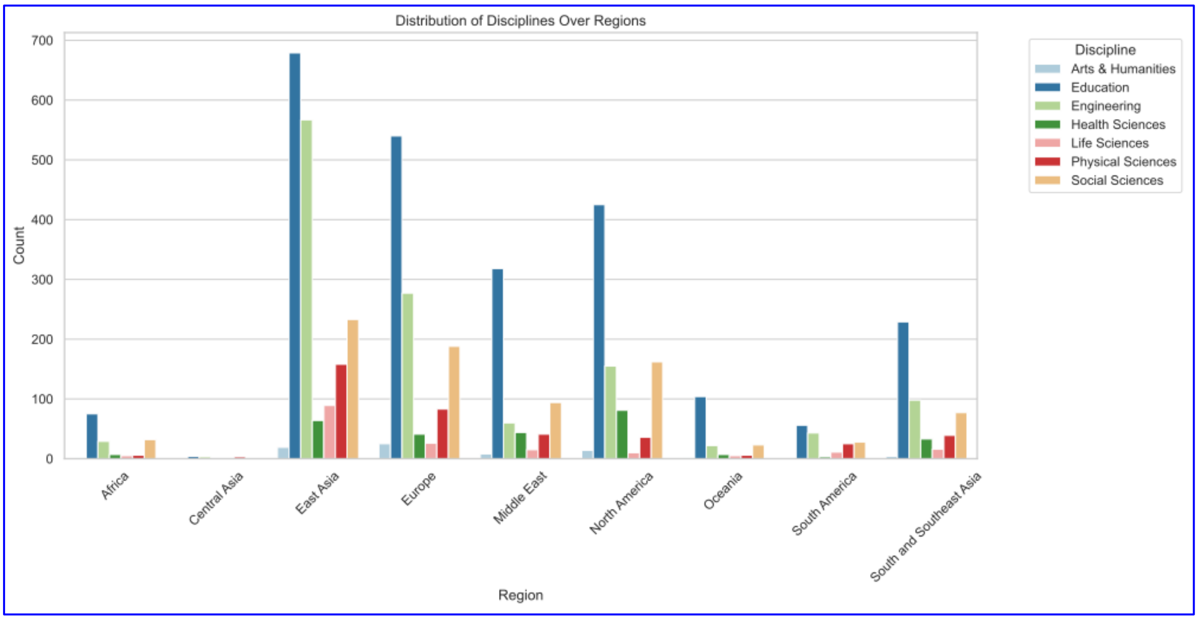

Figure 4. Number of Papers Published by Discipline in Each Region

3.4. Topics

To determine the topics covered in each of the papers, we used a four-step process, building on the capacities of Large Language Models (LLMs). We started with Topic Modelling, a Natural Language Processing method that automatically classifies documents into different topics, and expanded on it with ChatGPT.

In a first step, we used the BERT model to transform each abstract summary (uncased and truncated) into a vector in a high-dimensional space. BERT, a pre-trained language model, converts the whole document – here, each summary – into a vector approximation of its meaning. This approach is particularly efficient for clustering documents in which each word and sentence makes sense only in the broader context of the document, as is the case with article abstracts.

In the second step, we used K-Means to group together the abstracts, based on the BERT embeddings. It is an unsupervised learning algorithm that groups data points together based on their proximity in the embedding space. Since K-means requires the number of clusters to be specified a priori, we used the SSE (Sum of Squared Errors) – a measure of cluster quality – to determine the optimal number of clusters (see supplemental materials). There were multiple points of inflexion in the SSE (at 4, 9 and 16). The clusters were inspected by hand to select the version that lead to the most meaningful clustering (k=16). This combination of BERT embeddings and K-means clustering has been found to be the most efficient method to model the topics found in sets of abstracts from scientific articles [15], [16], [17].

In the third step, we looked at the first 30 papers for each cluster and proposed a name and definition for each. During this process four pairs of clusters were found to be similar in focus and were therefore grouped (see Table 3). While the Topic Modelling method used – with BERT embeddings – is considered the state of the art for this type of documents [15], [17], it suffers from two main flaws. First, the classification is made on the basis of the description made, not on what it says about the topic of the article. In other words, if two abstracts describe very different empirical studies but use a very similar structure and vocabulary, they are likely to be considered similar. Part of this issue was circumvented by using the summaries made by ChatGPT, but they were still partially influenced by the vocabulary used in the original abstracts. Second, each document can only be classified in one cluster, when in reality many papers fall in multiple categories. These two issues combined meant that we could not name the clusters in ways that were exhaustive – it accurately represented all the papers in the cluster – and exclusive – it couldn't be applied to papers in other clusters.

To resolve these issues, we decided to apply a fourth step. We created a ChatGPT assistant for each of the topics, tested with 20 summaries of 'on-topic' articles and 20 summaries of 'off-topic' articles (detailed prompts in supplemental materials). We then applied it to each of the summaries, using the same method as described in the 'Selecting Articles' section. This allowed us to build on traditional Topic Modelling methods – good at providing an overview of the topics that is truly representative of the content of the dataset – and improve them to have a quality of coding closer to what a human would do. The number of papers classified in each topic (Table 3) shows that factors influencing the use of AI (42.4% of papers) is by far the topic most frequently covered in the literature.

The topics were created using the dataset collected in September 2024 (n=2,487). While it was originally planned to update the topics once the final dataset was collected, we decided against it for two reasons. First, because of the large number of articles published since 2023 – and the release of ChatGPT – we would have ran the risk of over-weighting newer publications. In the original dataset, collected in 2024, papers published in 2023 and beyond represent 64.8% of the data. In the final dataset, it is 81% of the papers were published post ChatGPT. This means that the topics would have primarily been modelled on those. In other words, it would have represented well the very recent literature but risked missing important topical changes in the history of the field. Second, the results and trends – with a few minor exceptions related to chatbots – did not change following the addition of data, indicating that the original modelling remained robust.

To note, in order to further ensure robustness, we compared the distribution of topics before and after the inclusion of the 2025 data and found that the overall structure and relative prominence of topics remained stable, with only minor variations (primarily related to the rise of chatbot-related publications). Retaining the 2024 structure therefore allowed us to preserve historical continuity while avoiding an over-weighting of the most recent literature.

Table 3. Main Topics Covered in the Papers

Topic	Definition	Num. papers
Attitudes	Studies looking at teachers', students', or educational institutions' attitudes towards AI	1213 (26.2%)
Chatbots	Papers on the actual or potential uses of Chatbots such as ChatGPT in education	803 (17.4%)
Engagement	Papers on how AI could/can be used to improve or monitor engagement	831 (18.0%)
Evaluation	Studies empirically evaluating the effects of specific AI tools on education	1516 (32.8%)
Factors	Studies looking at the individual or contextual factors influencing AI adoption or effectiveness	1962 (42.4%)
Impact	Papers discussing or looking at the actual or potential impact of AI on education	789 (17.1%)
Monitoring	Papers on how AI can/could be used to monitor students' or teachers' performances or educational institutions	611 (13.2%)
Personalisation	Papers on how AI could/can be used to personalise educational contents or curricula	827 (17.9%)
Prediction	Papers on how AI can/could be used to predict the success of students, teachers, or programmes	332 (7.2%)
Reviews	Papers that offer a review of the literature or a framework to understand or study AI and education	194 (4.2%)
Risks	Papers discussing the potential risks of using AI in educational settings	671 (14.5%)
Teaching AI	Articles on how to teach AI	558 (12.1%)

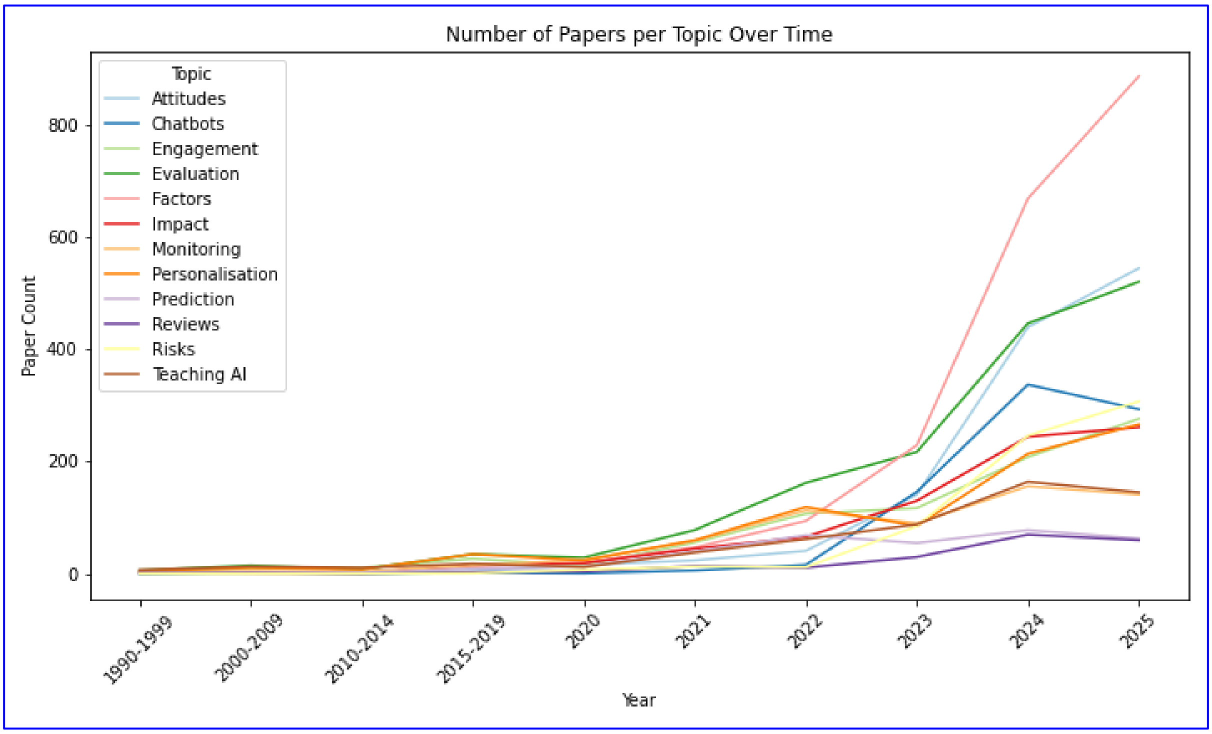

Figure 5. Number of Papers Published per Topic Over Time

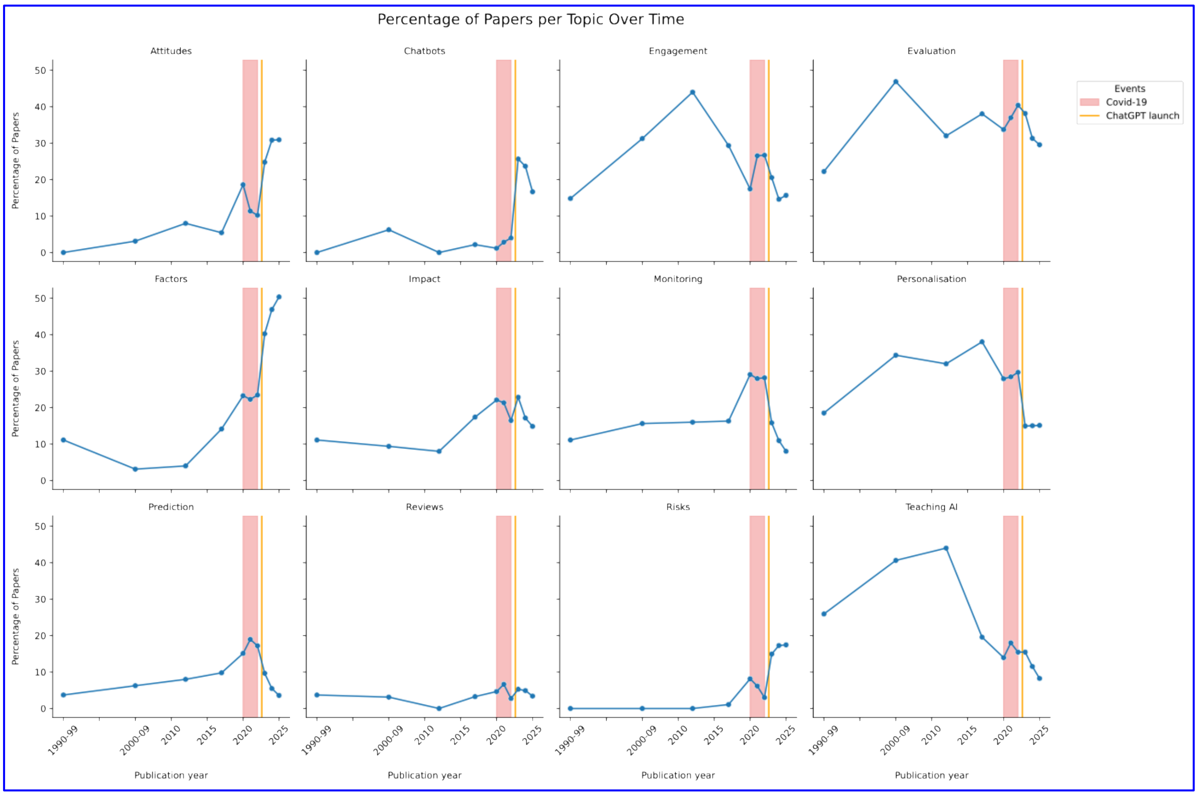

Figure 6. Percentage of Papers of Each Time Period Published per Topic

The importance of the topics has significantly changed over time (Figure 5). Early research on AI and Education mainly focused on four topics: how to teach AI, using AI to personalise content, using it to improve engagement, and evaluating the effectiveness of existing tools. A first change occurred around 2010, when interest in how to teach AI and improve engagement started to drop, in favour of research on which factors influence AI adoption, attitudes towards AI, and the impact of AI on education.

Things shifted again during the pandemic (Figure 6). Certain topics stalled or even dropped – such as research on factors, attitudes, impact or personalisation. Others registered a surge in interest, particularly visible for research on how AI can be used to monitor students, teachers or programmes, to predict educational outcomes, or to improve engagement. The shift to remote education during COVID-19 appears to have prompted efforts to enhance cohort monitoring and leverage AI to mitigate potential declines in student engagement in online learning environments. Notwithstanding the ethical issues raised by some of these uses, it is surprising how quickly they receded after the pandemic, given that some remain timely.

A final shift has occurred after the pandemic, with most topics returning to their pre-COVID trends – research on factors, for instance, resumed its ascension while work on predictions faded. The biggest change, however, seems to come from the launch of ChatGPT in late 2022: research on chatbots exploded in 2023 and 2024, before dropping in 2025. It also seems to have accelerated research on attitudes and related factors. Unfortunately, this has also led to a complete drop in research on evaluations. Given the hype around AI, the marked decrease in research assessing its effectiveness in education is particularly concerning.

Two topics have remained notably under-explored throughout. First, very few reviews or frameworks have been published. As a large proportion of the literature until 2023 included empirical evaluations of various AI tools (around 40%), reviews and integrations could really help the field take stock of what works, what doesn't, and why. Second, research discussing the risks posed by AI in education seems to be thoroughly lacking. It did increase following the release of ChatGPT, but it is already stagnating. For a field that often relies on large datasets collected on students and that could have tremendous implications on the future of many young people, it is rather shocking to see that even following the launch of ChatGPT, less than 15% of the literature focused on risks.

Finally, it is interesting to note that research on how to teach AI remains rare in the recent literature. Given the frequent calls for AI literacy – and the misconceptions surrounding AI amongst students (Authors, in press) – it is an area that would deserve further attention.

The disciplinary focus (Figure 7) on AI in education varies considerably across research fields. Social sciences, education, and humanities primarily engage with topics such as attitudes, risks, and factors influencing AI adoption and use, reflecting broader concerns about social, ethical, and pedagogical implications. In contrast, STEM disciplines, including computer science and life sciences, focus more on AI's technical applications, particularly in monitoring, personalisation, and prediction. Notably, while chatbots have attracted significant interest across multiple fields, engineering researchers appear largely disengaged from this trend, in contrast to education and social sciences – possibly because their technical expertise in the field protected them from the hype. However, evaluation studies – assessing the effectiveness of AI tools – and research on the factors that influence AI adoption and effectiveness show broad disciplinary interest, suggesting a shared concern across fields regarding the actual impact of AI on educational outcomes. But beyond those topics, a disciplinary divide underscores the fragmented nature of AI in education research, with technical and social considerations often developing in parallel rather than in dialogue.

Figure 8 presents the distribution of AI and education research topics across different regions over time. Early research showed varying levels of interest in each topic across regions – even as the number of papers increased. Yet one striking finding is that research focus seems to be moving in a similar direction worldwide, with similar trends everywhere. A few differences remain, however. Research on East Asia tends to focus more on engagement and personalisation and to pay less attention to risks and reviews. Studies in South and South-East Asia, as well as the Middle East, have increasingly examined Chatbots and related factors. Meanwhile, researchers from Europe are more likely to focus on risks and impact, and those from North America are more likely to propose reviews, examine impact, or investigate how to teach AI – while demonstrating less interest in evaluation and factors.

To analyse the overlap between AI and education research topics, Jaccard scores were used. These scores compare how frequently topics co-occur in the dataset with what would be expected by chance, while accounting for each topic's independent frequency. They are particularly useful for datasets where code frequencies can vary widely. Figure 9 depicts the percentage changes between the observed and expected scores: cells in red correspond to topics co-occurring more frequently than expected, cells in blue to those co-occurring less, and cells in white to non-significant differences.

Three clusters of topics can be observed. First, articles on prediction are more likely to also focus on monitoring, personalisation, and engagement, and to a certain extent, include an evaluation. Second, studies on factors often also investigate attitudes and are slightly more likely to relate to Chatbots. Finally, papers discussing the impact of AI on education often include a review and/or a discussion of potential risks.

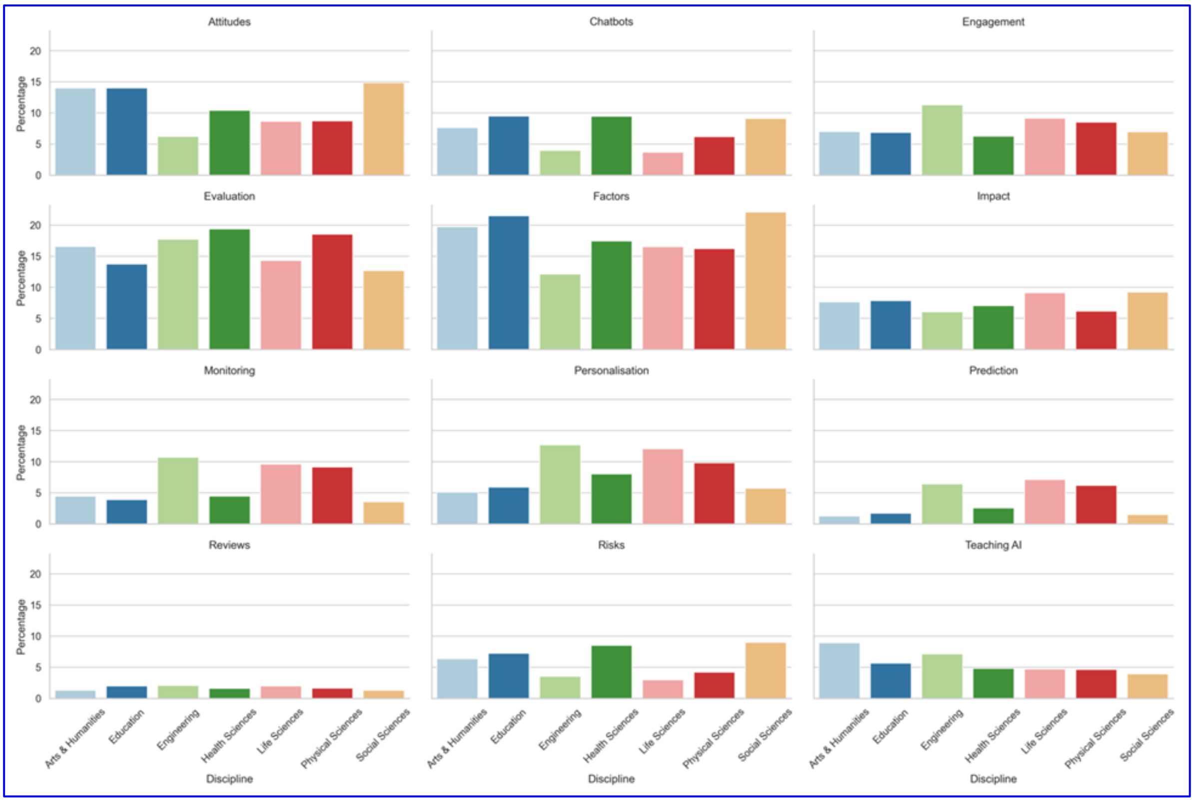

Figure 7. Percentage of Papers in Each Discipline Published per Topic

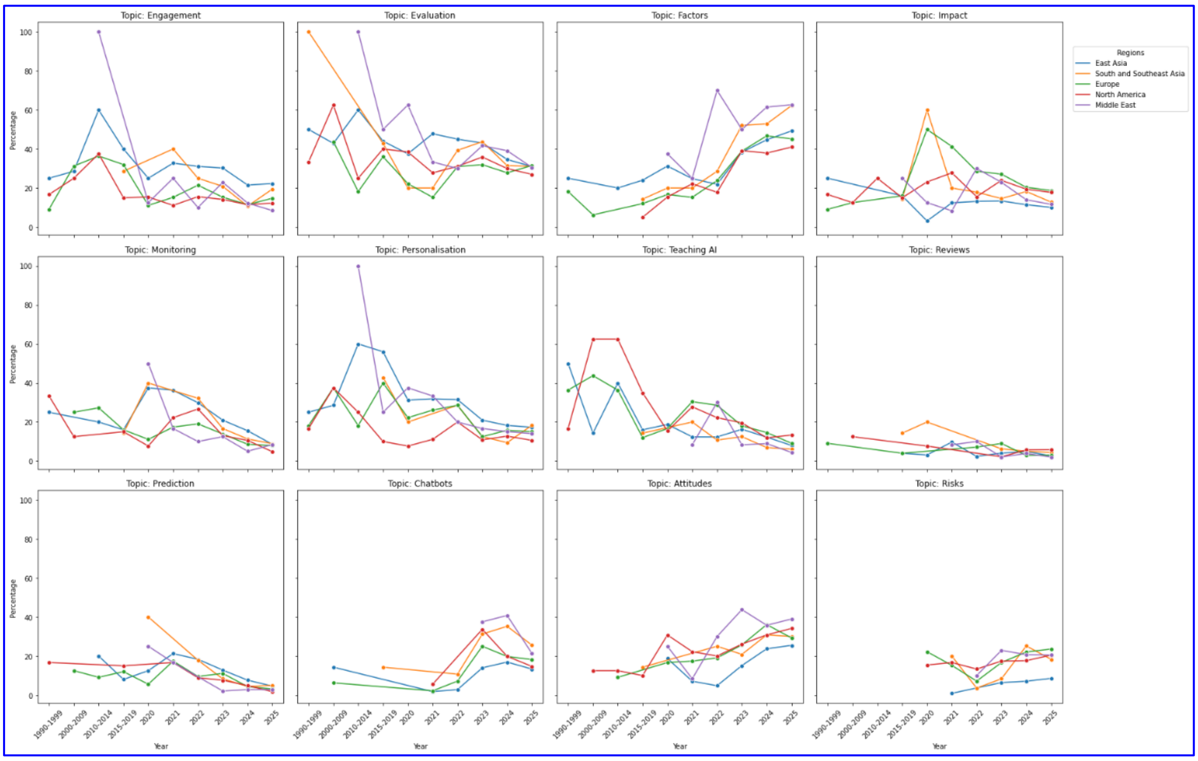

Figure 8. Number of Papers in Each Region Published per Topic

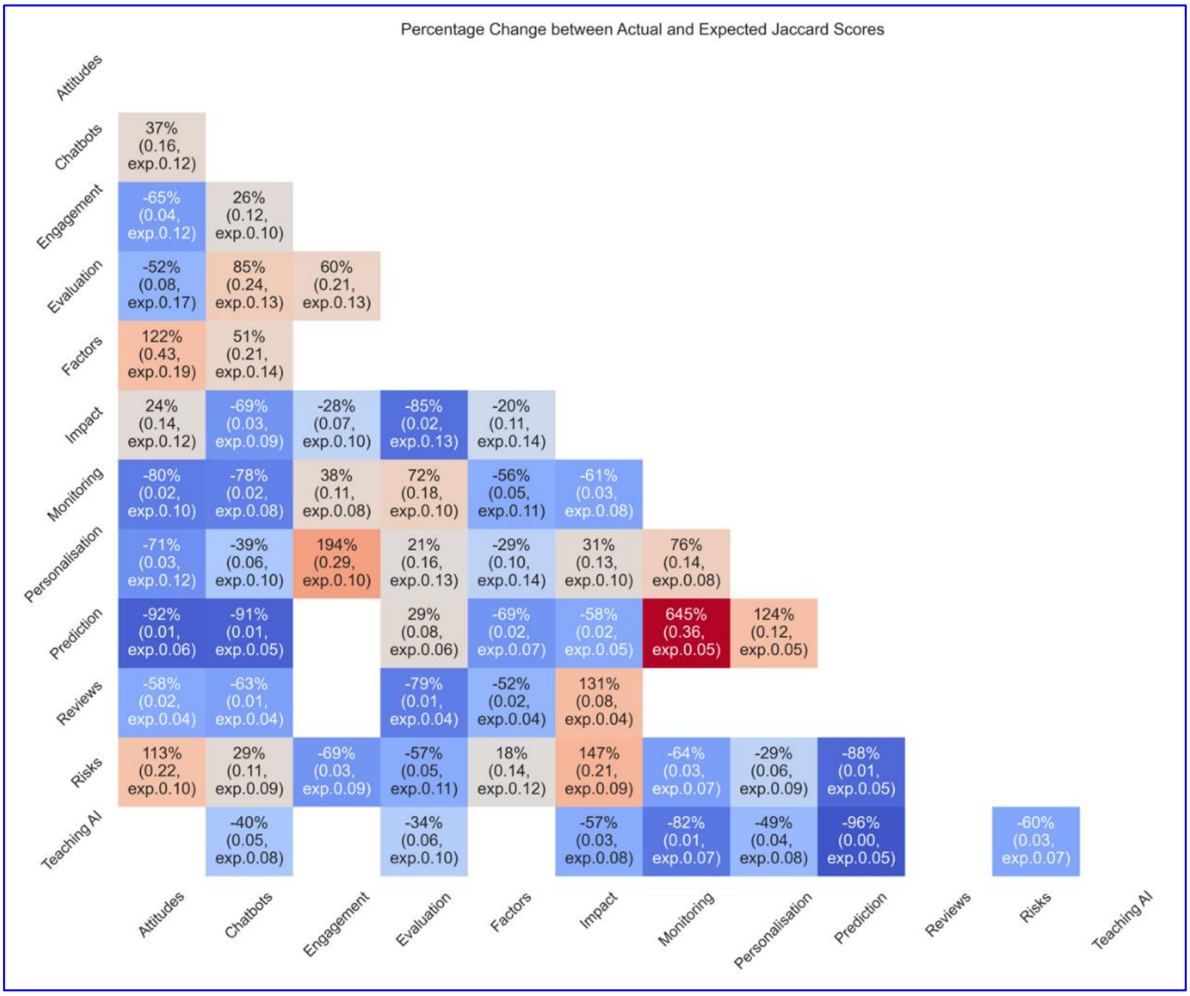

Figure 9. Percentage Change Between Actual and Expected Jaccard Scores

However, some critical gaps emerge: research on monitoring and prediction rarely intersects with studies on risks, despite the significant ethical concerns associated with AI-driven surveillance and decision-making in education. Similarly, impact studies seldom overlap with evaluations, suggesting that discussions of AI's broader effects may lack empirical grounding. Additionally, research on teaching AI remains largely separate from the rest of the literature, a division that made sense historically but now raises concerns about integrating AI literacy into the broader educational landscape. These patterns underscore the fragmentation of AI and education research, where technical and ethical considerations often develop in isolation rather than in dialogue.

Figures 10 and 11 highlight citation patterns across disciplines and topics in AI and education research. A large proportion of published papers get little to no citations, while a few papers are cited hundreds of times. This makes distribution plots difficult to interpret – with the bulk of the data around zero and a very long right tail. We therefore decided to focus on medians and means: in combination, they give an indication of where most papers fall while retaining some sensitivity to highly cited papers.

Surprisingly, papers in education receive the highest average citations, despite education traditionally having a lower citation culture compared to STEM fields. This is partly explained by the popularity of certain high-impact topics – such as risks and chatbots – that are frequently cited and predominantly studied within education. However, the pattern extends beyond topic effects alone, as physical sciences also receive substantial citations despite focusing less on trending topics. A likely explanation

Vol. 4, No. 2, June (2026), Pages 145–171 | e-ISSN: 2985–9166

is that engineering papers, which dominate AI tool development, are cited less often because they often present specific technological solutions that are difficult to build on without direct access to the tool. In contrast, papers offering broader theoretical perspectives, empirical evaluations, or literature reviews tend to be more widely referenced across disciplines. This pattern raises concerns about the current academic model's limitations in fostering high-tech educational tools, as impactful innovations may struggle to gain visibility and cross-disciplinary engagement due to restricted accessibility and applicability.

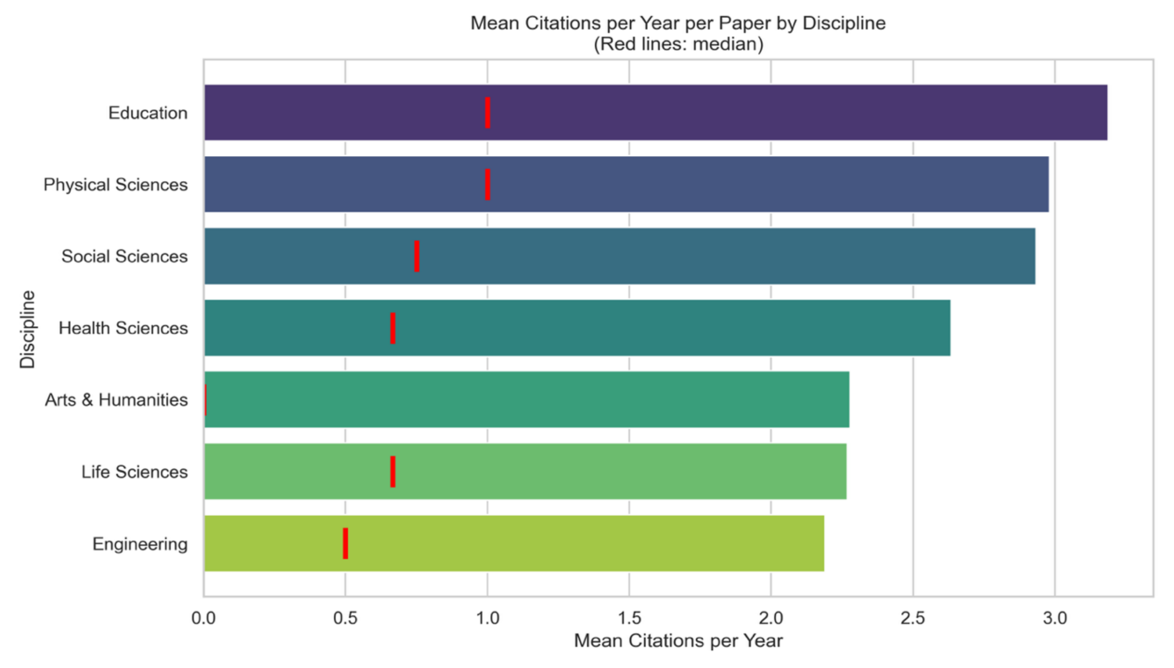

Figure 10. Mean Number of Citations per Year for Each Discipline

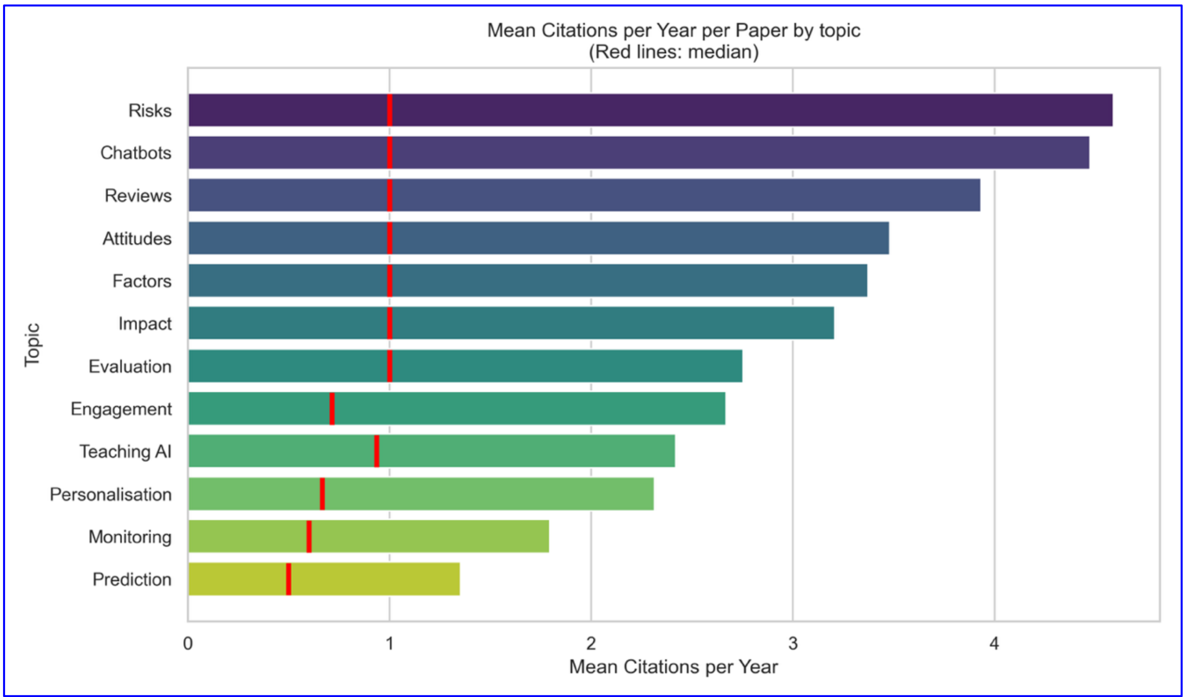

Figure 11. Mean Number of Citations per Year for Each Topic

It is also worth noting that, while papers on risks and reviews are relatively rare, they tend to be highly cited, showing the field's appetite for these topics.

Figure 12 presents a Sankey diagram illustrating citation flows between disciplines in AI and education research. To generate this figure, we used the reference lists for each paper as available on Web of Science. To assign a discipline to each cited article, we first matched journal titles to the journals already in the dataset and assigned the paper's discipline to the reference. For the references with no match, we then extracted the most frequent words in the article titles and manually coded those that univocally belong to a field (e.g., associating Learning to Education or Cells to Life Sciences, but ignoring words like Human, International, and so on). In case of multiple matches for a journal – either in the original dataset or in the keyword coding – the most frequent field was kept. In case of a tie, the reference was categorised in both areas. We then plotted, for each field, the proportion of cited references within that field.

The Sankey diagram reveals distinct citation patterns: Education research cites both itself and engineering at similar rates while also incorporating insights from social sciences. In contrast, engineering overwhelmingly cites its own field, engaging minimally with education literature – a pattern commonly observed in STEM research. Similarly, other STEM disciplines tend to reference engineering but rarely draw from education or social sciences. This disciplinary divide raises concerns about the development of AI-driven educational tools without sufficient input from educational research, potentially resulting in solutions that are poorly aligned with pedagogical needs. Addressing this gap requires greater interdisciplinary collaboration to ensure that AI applications in education are both technologically robust and pedagogically sound.

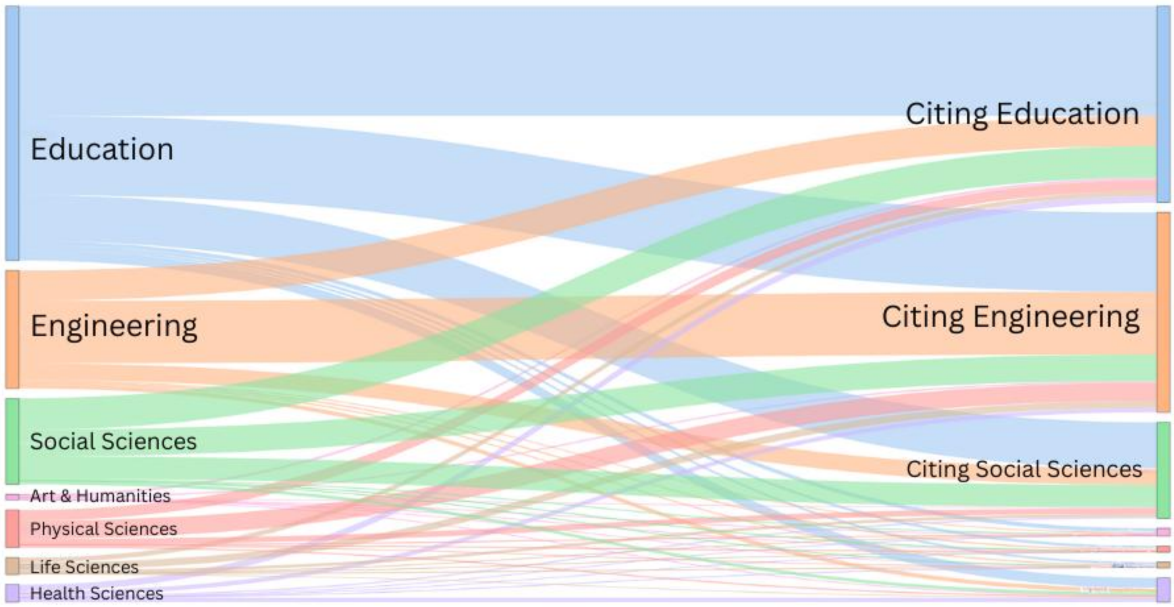

Figure 12. Sankey Diagram of the Flow of Citation from Citing Papers (left) to Cited Papers (right) by Field.

4. DISCUSSION

This study offers one of the most comprehensive mappings, to date, of research on artificial intelligence (AI) in education. In what follows, we distinguish between patterns directly observed in the data and interpretive claims that aim to explain these patterns, recognising that the latter remain provisional and open to further empirical testing. Our methodological approach – integrating computational tools, AI-assisted filtering, and human validation – allowed us to capture a corpus of 16,564 articles, of which 4,626 were identified as core contributions. This hybrid strategy not only ensured systematic coverage but also enabled us to explore key research themes, disciplinary orientations, regional trajectories, and citation patterns.

To complement this macro-level analysis, we also examined the twenty most-cited papers in each of the major research areas. This qualitative overview adds interpretive depth, showing how influential works have shaped debates on adoption, engagement, personalisation, chatbots, evaluation, risks, monitoring and prediction, reviews, and teaching AI. For instance, highly cited studies drawing on the Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT) have dominated discussions of adoption and attitudes (e.g., [18], [19]), while evaluations of generative AI tools such as ChatGPT [8], [20] illustrate both the promise and the risks of rapid adoption. These qualitative insights provide essential context for interpreting the broader patterns revealed by our quantitative analysis and are used to inform the present discussion.

Importantly, our findings align with and extend recent systematic reviews that have documented the exponential rise of AI in education since 2022. For example, Garzón, Patiño and Marulanda [21] identified a sharp post-ChatGPT surge, noting both motivational benefits and risks of over-reliance, while Lan and Zhou [22] highlighted how AI systems may support but also reshape self-regulated learning processes. Reviews of multimodal learning analytics [23] and ethical challenges in schooling [24] further underline the urgency of connecting technical innovation with pedagogical and ethical reflection. By combining large-scale mapping with qualitative synthesis, this paper provides a timely contribution to these ongoing conversations.

As following, we highlight the rapid growth and thematic shifts in the field, examine persistent disciplinary and regional divides, identify neglected areas and systemic challenges, and conclude with recommendations for consolidating and advancing AI and education research in more cumulative, integrative, and critical directions.

4.1. Growth of the Field and Hype Cycles

One of the most striking findings is the exponential growth of research on AI in education, particularly over the last decade. While fewer than thirty relevant articles appeared in the 1990s, the field has since accelerated dramatically, with 1,759 publications recorded in just the first eight months of 2025. This surge reflects not only the expansion of AI applications in educational contexts but also broader societal events that catalysed bursts of scholarly attention.

Two moments stand out in particular. The COVID-19 pandemic prompted an unprecedented wave of studies on monitoring, prediction, and engagement, reflecting urgent needs to sustain learning under conditions of disruption and distance. Shortly after, the release of ChatGPT in late 2022 triggered another explosion of research, this time centred on generative AI and chatbots. These thematic surges demonstrate the field's responsiveness to external shocks, but they also highlight a risk: scholarship may follow hype cycles rather than building cumulatively on prior knowledge. Recent systematic reviews confirm this tendency, showing that while generative AI and chatbots have stimulated impressive growth in output, much of the work is exploratory, descriptive, or focused on short-term applications [21], [25].

Our qualitative exploration confirmed this dynamic. For example, chatbot research has been dominated by studies emphasising their potential to enhance language learning, self-efficacy, and teacher workload management [26], [27], [28]. Yet, limitations have also been repeatedly documented, including difficulties in fostering higher-order thinking and ethical concerns around privacy and student dependency [29], [30]. Similarly, pandemic-driven monitoring and prediction studies showcased impressive computational advances, such as neural networks and multimodal analytics [31], [32], but critical analyses have cautioned against the reductive modelling of human cognition and the risks of algorithmic bias [14], [33].

Taken together, these findings suggest that while moments of crisis or technological novelty drive rapid innovation, they also risk narrowing the research agenda to immediate concerns or readily available tools. A more sustainable trajectory requires recognising these cycles and ensuring that new work builds on, rather than displaces, earlier insights.

4.2. Fragmentation and Interdisciplinary Gaps

Beyond its rapid growth, the field of AI in education is characterised by significant fragmentation across disciplines and regions. Our analysis shows that education and engineering together account for more than three-quarters of all publications, with the remaining disciplines, particularly the social sciences, life sciences, and arts and humanities, contributing only marginally. While the rise of education-led research since 2022 signals the field's increasing pedagogical orientation, engineering continues to dominate the technical development of tools and applications. However, the interaction between these two knowledge streams remains limited. More than 80% of publications are single-disciplinary, and citation flows reveal a sharp asymmetry: education scholars draw broadly on engineering and social sciences, but engineering researchers overwhelmingly cite within their own domain. This pattern suggests parallel trajectories of technical innovation and educational reflection that seldom intersect.

This dynamic was clear in the most cited papers we reviewed. Influential education and social science works often interrogate the ethical, cultural, and pedagogical implications of AI (e.g., [34], [35]), while highly cited engineering studies focus on algorithmic sophistication, monitoring, or predictive modelling (e.g., [31]). Few papers manage to bridge these orientations. For example, Perrotta and Selwyn's [14] critique of knowledge tracing algorithms explicitly engages both technical methods and epistemological assumptions, but such integrative approaches remain rare.

The consequence is a field at risk of reinforcing silos: technical systems are developed with insufficient consideration of pedagogical, ethical, or social contexts, while educational critiques often lack grounding in technical realities. This lack of integration undermines cumulative knowledge-building and increases the likelihood of duplication, with different teams "reinventing the wheel" rather than consolidating progress. It also helps explain citation disparities: education papers, especially those addressing broad issues like ethics and risks, achieve higher average citation counts, while engineering outputs remain under-cited despite their technical contributions.

Without stronger interdisciplinary dialogue, AI in education risks becoming a collection of isolated conversations, rather than a coherent field capable of shaping both technological development and educational practice. Recent contributions have underscored the urgency of this shift, advocating hybrid human-AI systems that explicitly connect technical innovation with pedagogical needs and ethical safeguards [36].

4.3. Thematic Shifts and Imbalances

The thematic distribution of research reveals both enduring concerns and striking imbalances. Adoption factors and attitudes dominate, representing over 40% of publications. Highly cited studies employ established frameworks such as the Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT) to explain behavioural intentions, trust, and perceived usefulness [18], [19]. Quantitative designs, including regression analysis and structural equation modelling [37], have been complemented by qualitative studies of experiences and perceptions [34]. Together, these studies have established adoption and attitudes as central themes. Yet the field risks becoming saturated with perception surveys, particularly of students and teachers, while neglecting design-oriented research that could directly inform more effective and ethical tool development.

Evaluation studies once formed another pillar of the literature, but their relative decline since 2022 is concerning. Influential evaluations of generative AI tools, including ChatGPT [8], [20], demonstrate both their potential to reduce workloads and enhance access, and their risks for academic integrity and critical thinking. Yet these contributions are increasingly overshadowed by attitudinal surveys or descriptive accounts of generative AI use. The reduced focus on systematic evaluations undermines evidence-based practice and leaves the field vulnerable to hype-driven conclusions.

Risks and ethics are similarly underrepresented, comprising less than 15% of the corpus. Nonetheless, when addressed, they attract high visibility. Highly cited works stress the importance of transparency, privacy, and algorithmic bias [33], [29], and call for governance structures to ensure equitable use [38]. However, these discussions rarely connect with technical domains such as monitoring and prediction, where ethical implications are most acute. Recent reviews confirm these concerns, noting that ethical and regulatory challenges of generative AI remain insufficiently addressed in both K-12 and higher education contexts [24], [25].

Other topics show uneven trajectories. Chatbot research surged after 2022, propelled by generative AI applications [26], [27], but these papers often report immediate benefits without long-term follow-up, and critiques warn against over-reliance [30]. Personalisation, historically a cornerstone of AI in education [1], [3], has declined in prominence despite its clear potential to support equity and autonomy. Likewise, teaching AI remains marginal, even though equipping learners with AI literacy is widely regarded as essential for future societies [39]. A further underexplored but promising area is the use of AI to support self-regulated learning, where recent reviews suggest that AI tools can scaffold autonomy but also risk reshaping learner agency in problematic ways [22]. The neglect of these promising areas illustrates the field's tendency to prioritise short-term interest over long-term educational transformation.

4.4. Historical Blind Spots and Cumulative Knowledge

A further pattern emerging from our analysis is the field's limited historical awareness. Research on AI in education often develops in sharp surges, triggered by external events or technological breakthroughs, but these spikes are rarely accompanied by sustained follow-up or systematic efforts to build cumulatively on prior work. The COVID-19 pandemic is emblematic: it prompted a rapid rise in studies of remote education, monitoring, and prediction, many of which documented promising applications for sustaining learning at scale. Yet once the immediate crisis subsided, this line of inquiry lost momentum, with few attempts to consolidate insights into enduring frameworks for digital resilience or hybrid learning. A similar pattern can be seen in the wake of ChatGPT's release, with research shifting dramatically toward chatbots and generative AI without consistently engaging earlier traditions of natural language processing, tutoring systems, or personalised learning.

The qualitative survey of the highly cited papers showed the risks of this short-termism. Early work on intelligent tutoring systems (ITS) traced the integration of AI with cognitive psychology and educational theory, emphasising the design of adaptive, exploratory, and dialogue-based learning environments [2]. These contributions offered conceptual foundations that remain highly relevant, yet are seldom cited in recent publications. Likewise, studies of personalisation and adaptive learning platforms [1], [3] demonstrated both the potential and the challenges of tailoring instruction to diverse learners. Instead of extending this knowledge base, much of the current literature seems to restart the conversation from scratch whenever a new tool or platform emerges. Recent systematic reviews of ITS and robot tutoring systems confirm this problem, showing that while foundational traditions continue to inform technical design, their pedagogical and theoretical contributions are rarely acknowledged in contemporary AI in education research [40].

The consequences of these historical blind spots are twofold. First, research risks reinventing the wheel, producing redundant findings that echo earlier insights but without advancing them. Second, it limits the field's ability to anticipate long-term implications, as each new surge is treated as unprecedented rather than part of a broader trajectory of AI's interaction with education. Without cumulative perspectives, there is little sense of how pedagogical theories, technical systems, and ethical debates evolve in relation to one another. Strengthening historical and theoretical continuity is therefore critical, both for consolidating what is already known and for avoiding cycles of hype-driven novelty that obscure deeper educational questions.

4.5. Systemic Issues in Academic Publishing and Impact

The patterns observed also reflect deeper systemic issues in how AI and education research is published, disseminated, and cited. One of the clearest findings is the relative under-citation of engineering papers compared to those in education and the social sciences. While engineering outputs can present sophisticated technical solutions, they rarely achieve the same visibility or influence. This is partly because of disciplinary silos but citations patterns – with education papers more likely to cite engineering papers than the other way around – points to the need for another explanation. One possible reason is that tool development is difficult to replicate and build upon in academic contexts. Without open access to code, data, or infrastructure, subsequent research cannot easily extend earlier contributions, which may help explain why such work is cited less. The danger is that technical innovation in education will increasingly shift to the private sector, where proprietary systems dominate but accountability and transparency are limited.

Conversely, education papers, especially those addressing broad themes such as ethics, risks, or adoption, are disproportionately well cited. This reflects the high demand for integrative and critical perspectives, but it also highlights the lack of effective mechanisms for bridging technical and pedagogical domains. Reviews and frameworks, which could serve this consolidating role, are scarce and rarely engage both ethical and educational questions with technical considerations. For example, new surveys of multimodal learning analytics map out technical advances in AI-driven data integration, but show little engagement with the broader pedagogical or ethical challenges these tools raise [23]. Similarly, while ethical guidelines for AI in higher education are beginning to emerge, including frameworks built around shared principles of accountability, transparency, and equity [41], these initiatives remain peripheral to the mainstream of technical development.

4.6. Limitations

A number of limitations should be acknowledged. First, the classification of articles relied on AI-assisted summaries rather than full-text analysis, which may have led to the loss of nuance in some cases. Second, although validation procedures were implemented, the categorisation of articles into "on-topic," "off-topic," and "maybe" remains subject to some degree of misclassification, particularly for interdisciplinary or indirectly relevant work. Third, the assignment of each article to a single primary topic simplifies what are often multi-dimensional contributions.

More broadly, the analyses presented here are descriptive and based on a specific database (Web of Science), and therefore do not aim to generalise to the entirety of AI and education research. Instead, they provide a structured overview of patterns within a large and diverse corpus, which should be interpreted as indicative rather than exhaustive.

5. Conclusion and Recommendations

This study has provided a comprehensive mapping of research on AI in education, combining large-scale quantitative analysis with a qualitative review of the most-cited works in each major topic area. The results reveal a rapidly expanding but seemingly fragmented field, marked by disciplinary silos, thematic imbalances, a tendency to short-termism, and systemic challenges in publishing and citation. While moments of crisis and technological innovation have catalysed rapid surges (most notably during the COVID-19 pandemic and following the release of ChatGPT), these waves have rarely been followed by cumulative knowledge-building. Instead, research agendas have often shifted abruptly, with neglected areas such as personalisation and teaching AI falling to the margins despite their clear long-term significance. At the same time, the field risks becoming dominated by quick, survey-based studies of adoption and attitudes, while systematic evaluations and integrative reviews decline. Recent systematic reviews confirm these tendencies, noting both the rapid post-2022 expansion of the field and the persistence of ethical, regulatory, and design-related gaps [21], [25], [24].

In light of these findings, several recommendations emerge for advancing the field in more sustainable and impactful directions. First, there is a need to *reinvigorate evaluation research*. Rigorous, empirical evaluations of AI tools are essential to counter hype-driven claims. Evaluations should be systematically integrated into reviews and frameworks to consolidate evidence. For instance, multimodal analytics and ITS traditions provide strong technical foundations that could inform more robust evaluation strategies [40], [23].

Second, it is crucial to *strengthen interdisciplinarity*. Greater collaboration is needed between technical and educational domains. This includes integrating research on monitoring and prediction with studies of risks and ethics, and moving beyond adoption surveys to design-focused inquiries. Calls for hybrid human-AI systems underline the need for such approaches [36].

Third, scholars should *revive neglected but promising areas.* Research on personalisation, teaching AI, and self-regulated learning should be re-prioritised, given their potential to enhance equity, autonomy, and AI/data literacy for future learners [22].

Fourth, the field *must adopt a historical and cumulative perspective*. Rather than treating each technological or societal disruption as unprecedented, future work should build explicitly on earlier traditions, consolidating knowledge and avoiding redundant cycles of hype. This requires revisiting and extending the conceptual foundations of tutoring systems, adaptive platforms, and evaluation research.

Finally, stakeholders must *address systemic issues in publishing*. Engineering outputs must be made more accessible and reusable through open practices, while reviews and frameworks should engage with technical as well as pedagogical domains. Encouragingly, new ethical guidelines for higher education provide models for more integrated frameworks [41]. Otherwise, academic research risks ceding tool development entirely to the private sector.

Together, these recommendations aim to move the field toward more integrated, rigorous, and forward-looking scholarship, ensuring that AI in education develops not as a collection of isolated responses to hype, but as a cumulative body of knowledge capable of informing both technological design and educational practice.

DECLARATIONS

Author Contributions

Constance de Saint Laurent: Conceptualization; Methodology; Investigation; Data Collection; Data Curation; Writing – Original Draft Preparation. **Vlad Glaveanu:** Theoretical Framework Development; Writing – Review & Editing. **Ioana Literat:** Theoretical Framework Development; Writing – Review & Editing. **Ingunn Ness:** Theoretical Framework Development; Writing – Review & Editing. All authors have read and approved the final version of the manuscript.

Acknowledgments

The authors would like to thank Natalie Robinson for her help formatting the final version of the paper.

Ethical Approval

This study was conducted in accordance with ethical standards.

Informed Consent

The study did not involve collecting empirical data from human participants.

Funding

The authors declare that this research did not receive specific funding from any public, commercial, or not-for-profit agencies.

Data Availability Statement

The data supporting this study's findings are not publicly available. Summary data are included in the manuscript, and more detailed information can be requested from the corresponding author under strict confidentiality agreements.

Competing Interests

The authors declare that they have no competing interests related to the content of this article.

Generative AI and AI-Assisted Technologies Statement

During the preparation of this manuscript, the authors used AI (ChatGPT 4.0) to assist as part of the methodology of summarizing data, as explained in the manuscript. Whenever AI was used, the authors carefully reviewed and edited the content to ensure accuracy and integrity, and they take full responsibility for the published work.

REFERENCES

[1] V. Caputi and A. Garrido, "Student-oriented planning of e-learning contents for Moodle," *Journal of Network and Computer Applications*, vol. 53, pp. 115–127, 2015, doi: 10.1016/j.jnca.2015.04.001.

[2] H. S. Nwana, "Intelligent tutoring systems: An overview," *Artificial Intelligence Review*, vol. 4, no. 4, pp. 251–277, 1990, doi: 10.1007/BF00168958.

[3] H. Yu, C. Miao, C. Leung, and T. J. White, "Towards AI-powered personalization in MOOC learning," *npj Science of Learning*, vol. 2, no. 1, Art. no. 15, 2017, doi: 10.1038/s41539-017-0016-3.

[4] B. Cope, M. Kalantzis, and D. Searsmith, "Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies," *Educational Philosophy and Theory*, vol. 53, no. 12, pp. 1229–1245, 2021, doi: 10.1080/00131857.2020.1728732.

[5] T. K. Chiu, Q. Xia, X. Zhou, C. S. Chai, and M. Cheng, "Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 4, Art. no. 100118, 2023, doi: 10.1016/j.caeai.2022.100118.

[6] S. J. Lee and K. Kwon, "A systematic review of AI education in K-12 classrooms from 2018 to 2023: Topics, strategies, and learning outcomes," *Computers and Education: Artificial Intelligence*, vol. 6, Art. no. 100211, 2024, doi: 10.1016/j.caeai.2024.100211.

[7] F. Martin, M. Zhuang, and D. Schaefer, "Systematic review of research on artificial intelligence in K-12 education (2017–2022)," *Computers and Education: Artificial Intelligence*, vol. 6, Art. no. 100195, 2024, doi: 10.1016/j.caeai.2023.100195.

[8] A. Tlili et al., "What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education," *Smart Learning Environments*, vol. 10, no. 1, Art. no. 15, 2023, doi: 10.1186/s40561-023-00237-x.

[9] F. J. Hinojo-Lucena, I. Aznar-Díaz, M. P. Cáceres-Reche, and J. M. Romero-Rodríguez, "Artificial intelligence in higher education: A bibliometric study on its impact in the scientific literature," *Education Sciences*, vol. 9, no. 1, Art. no. 51, 2019, doi: 10.3390/educsci9010051.

[10] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education—where are the educators?" *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, Art. no. 39, 2019, doi: 10.1186/s41239-019-0171-0.

[11] A. Bozkurt, A. Karadeniz, D. Baneres, A. E. Guerrero-Roldán, and M. E. Rodríguez, "Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century," *Sustainability*, vol. 13, no. 2, Art. no. 800, 2021, doi: 10.3390/su13020800.

[12] S. Guo, Y. Zheng, and X. Zhai, "Artificial intelligence in education research during 2013–2023: A review based on bibliometric analysis," *Education and Information Technologies*, vol. 29, no. 13, pp. 16387–16409, 2024, doi: 10.1007/s10639-024-12491-8.

[13] A. Gillespie, V. P. Glăveanu, and C. de Saint Laurent, *Pragmatism and Methodology: Doing Research That Matters with Mixed Methods*. Cambridge, U.K.: Cambridge University Press, 2024, doi: 10.1017/9781009031066.

[14] C. Perrotta and N. Selwyn, "Deep learning goes to school: Toward a relational understanding of AI in education," *Learning, Media and Technology*, vol. 45, no. 3, pp. 251–269, 2020, doi: 10.1080/17439884.2020.1686017.

[15] M. C. Wijanto, I. Widiastuti, and H. S. Yong, "Topic modeling for scientific articles: Exploring optimal hyperparameter tuning in BERT," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 14, no. 3, pp. 912–919, 2024, doi: 10.18517/ijaseit.14.3.19347.

[16] G. Piao, "Scholarly text classification with sentence BERT and entity embeddings," in *Trends and Applications in Knowledge Discovery and Data Mining*, Cham, Switzerland: Springer, 2021, pp. 79–87, doi: 10.1007/978-3-030-75015-2_8.

[17] S. Selva Birunda and R. Kanniga Devi, "A review on word embedding techniques for text classification," in *Innovative Data Communication Technologies and Application*, Singapore: Springer, 2021, pp. 267–281, doi: 10.1007/978-981-15-9651-3_23.

[18] R. Chocarro, M. Cortiñas, and G. Marcos-Matás, "Teachers' attitudes towards chatbots in education: A technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics," *Educational Studies*, vol. 49, no. 2, pp. 295–313, 2023, doi: 10.1080/03055698.2020.1850426.

[19] J. E. Raffaghelli, M. E. Rodríguez, A. E. Guerrero-Roldán, and D. Bañeres, "Applying the UTAUT model to explain the students' acceptance of an early warning system in higher education," *Computers & Education*, vol. 182, Art. no. 104468, 2022, doi: 10.1016/j.compedu.2022.104468.

[20] M. Farrokhnia, S. K. Banihashem, O. Noroozi, and A. Wals, "A SWOT analysis of ChatGPT: Implications for educational practice and research," *Innovations in Education and Teaching International*, vol. 61, no. 3, pp. 460–474, 2024, doi: 10.1080/14703297.2023.2195846.

[21] J. Garzón, E. Patiño, and C. Marulanda, "Systematic review of artificial intelligence in education: Trends, benefits, and challenges," *Multimodal Technologies and Interaction*, vol. 9, no. 8, Art. no. 84, 2025, doi: 10.3390/mti9080084.

[22] M. Lan and X. Zhou, "A qualitative systematic review on AI empowered self-regulated learning in higher education," *npj Science of Learning*, vol. 10, no. 1, Art. no. 21, 2025, doi: 10.1038/s41539-025-00319-0.

[23] M. Mohammadi, E. Tajik, R. Martinez-Maldonado, S. Sadiq, and W. Tomaszewski, "Artificial intelligence in multimodal learning analytics: A systematic literature review," *Computers and Education: Artificial Intelligence*, vol. 8, Art. no. 100426, 2025, doi: 10.1016/j.caeai.2025.100426.

[24] A. Gouseti, F. James, L. Fallin, and K. Burden, "The ethics of using AI in K-12 education: A systematic literature review," *Technology, Pedagogy and Education*, vol. 34, no. 2, pp. 161–182, 2025, doi: 10.1080/1475939X.2024.2428601.

[25] J. García-López and L. Trujillo-Liñán, "Ethical and regulatory challenges of generative AI in education: A systematic review," *Frontiers in Education*, vol. 10, Art. no. 1565938, 2025, doi: 10.3389/feduc.2025.1565938.

[26] C. Y. Chang, G. J. Hwang, and M. L. Gau, "Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training," *British Journal of Educational Technology*, vol. 53, no. 1, pp. 171–188, 2022, doi: 10.1111/bjet.13158.

[27] J. Jeon and S. Lee, "Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT," *Education and Information Technologies*, vol. 28, no. 12, pp. 15873–15892, 2023, doi: 10.1007/s10639-023-11834-1.

[28] S. Fergus, M. Botha, and M. Ostovar, "Evaluating academic answers generated using ChatGPT," *Journal of Chemical Education*, vol. 100, no. 4, pp. 1672–1675, 2023, doi: 10.1021/acs.jchemed.3c00087.

[29] C. Kooli, "Chatbots in education and research: A critical examination of ethical implications and solutions," *Sustainability*, vol. 15, no. 7, Art. no. 5614, 2023, doi: 10.3390/su15075614.

[30] D. Yan, "Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation," *Education and Information Technologies*, vol. 28, no. 11, pp. 13943–13967, 2023, doi: 10.1007/s10639-023-11742-4.

[31] J. Bernard, T. W. Chang, E. Popescu, and S. Graf, "Learning style identifier: Improving the precision of learning style identification through computational intelligence algorithms," *Expert Systems with Applications*, vol. 75, pp. 94–108, 2017, doi: 10.1016/j.eswa.2017.01.021.

[32] K. Sharma, Z. Papamitsiou, and M. Giannakos, "Building pipelines for educational data using AI and multimodal analytics: A 'grey-box' approach," *British Journal of Educational Technology*, vol. 50, no. 6, pp. 3004–3031, 2019, doi: 10.1111/bjet.12854.

[33] B. Berendt, A. Littlejohn, and M. Blakemore, "AI in education: Learner choice and fundamental rights," *Learning, Media and Technology*, vol. 45, no. 3, pp. 312–324, 2020, doi: 10.1080/17439884.2020.1786399.

[34] E. Dixon-Román, T. P. Nichols, and A. Nyame-Mensah, "The racializing forces of/in AI educational technologies," *Learning, Media and Technology*, vol. 45, no. 3, pp. 236–250, 2020, doi: 10.1080/17439884.2020.1667825.

[35] D. Schiff, "Out of the laboratory and into the classroom: The future of artificial intelligence in education," *AI & Society*, vol. 36, no. 1, pp. 331–348, 2021, doi: 10.1007/s00146-020-01033-8.

[36] D. Hooshyar et al., "Towards responsible AI for education: Hybrid human-AI to confront the elephant in the room," *Computers and Education: Artificial Intelligence*, vol. 9, Art. no. 100524, 2025, doi: 10.1016/j.caeai.2025.100524.

[37] C. S. Chai, X. Wang, and C. Xu, "An extended theory of planned behavior for the modelling of Chinese secondary school students' intention to learn artificial intelligence," *Mathematics*, vol. 8, no. 11, Art. no. 2089, 2020, doi: 10.3390/math8112089.

[38] P. Song and X. Wang, "A bibliometric analysis of worldwide educational artificial intelligence research development in recent twenty years," *Asia Pacific Education Review*, vol. 21, no. 3, pp. 473–486, 2020, doi: 10.1007/s12564-020-09640-2.

[39] Q. Xia, T. K. Chiu, M. Lee, I. T. Sanusi, Y. Dai, and C. S. Chai, "A self-determination theory (SDT) design approach for inclusive and diverse artificial intelligence (AI) education," *Computers & Education*, vol. 189, Art. no. 104582, 2022, doi: 10.1016/j.compedu.2022.104582.

[40] V. Liu, E. Latif, and X. Zhai, "Advancing education through tutoring systems: A systematic literature review," *arXiv*:2503.09748, 2025, doi: 10.48550/arXiv.2503.09748.. [Online]. Available: https://arxiv.org/abs/2503.09748

[41] M. Georgieva and J. Stuart, "Ethics is the edge: The future of AI in higher education," *EDUCAUSE Review*, Jun. 24, 2025. [Online]. Available: https://er.educause.edu/articles/2025/6/ethics-is-the-edge-the-future-of-ai-in-higher-education

Author Biographies

Constance de Saint Laurent is Assistant Professor of Sociotechnical Systems at the Department of Psychology at Maynooth University. She works on social thinking and the impact of technology and societal changes on people and organisations. She has previously researched on social media, artificial intelligence, misinformation, collective memory, and representations of alterity, some of which has been published in the open access book *Social Thinking and History: A Sociocultural Psychological Perspective on Representations of the Past* (2021). She has a BSc in Psychology from the University of Angers in France, an MSc in Social and Cultural Psychology from the London School of Economics, and a PhD in Social Sciences from the University of Neuchâtel in Switzerland.

Vlad Glaveanu is Full Professor of Psychology in the DCU Business School and Director of the DCU Centre for Possibility Studies at Dublin City University, and Professor II at the Centre for the Science of Learning and Technology, University of Bergen. He is the founder and president of the Possibility Studies Network (PSN). His work focuses on creativity, imagination, culture, collaboration, wonder, possibility, and societal challenges. He edited the *Palgrave Handbook of Creativity and Culture* (2016) and the *Oxford Creativity Reader* (2018), co-edited the *Cambridge Handbook of Creativity Across Domains* (2017) and the *Oxford Handbook of Imagination and Culture* (2017), authored *The Possible: A Sociocultural Theory* (Oxford University Press, 2020), *Creativity: A Very Short Introduction* (Oxford University Press, 2021), and *Wonder: The Extraordinary Power of an Ordinary Experience* (Bloomsbury, 2020), and authored or co-authored more than 200 articles and book chapters in these areas.

Ioana Literat is a Professor in the Communication, Media and Learning Technologies Design program at Teachers College, Columbia University, and the Director of the Media & Social Change Lab (MASCLab). Ioana's research centers on young people's creative and political expression online, with a particular focus on youth. Most recently, Ioana co-authored "Not Your Parents' Politics: The Unexpected Forms of Young People's Political Expression on Social Media" with Neta Kligler-Vilenchik, published by Oxford University Press in October 2024. Beyond academia, she frequently contributes insights on technology and society to

outlets like The New York Times, The Washington Post, The Guardian, Newsweek, The Atlantic, CNN, and Wired.

Ingunn Ness is an expert on learning, creativity and innovation in the field of education and business. She currently holds the position as a Senior Researcher and theme Leader at SLATE, the Centre for the Sciences of Learning & Technology at the University of Bergen, Norway. Ness has a particular interest for the sociocultural approach to innovative knowledge development and works with one of the world's leading environments on sociocultural theory, the OSAT group at the Department of Education, University of Oxford. Ness has done extensive empirical research on collaborative creativity in strategy and innovation contexts.

Publisher's and Journal's Note *Sagamedia Teknologi Nusantara, as the publisher and editor of the Journal of Hypermedia & Technology-Enhanced Learning (J-HyTEL), upholds the highest ethical standards in academic publishing. The journal remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Authors are fully responsible for the originality, accuracy, and integrity of their work. Post-publication ethical concerns will be addressed through corrections, clarifications, or retractions as necessary. The content of this publication has not been approved by the United Nations and does not reflect the views of the United Nations or its officials or Member States.* https://www.un.org/sustainabledevelopment