Topham, L, Atherton, P, Reynolds, T, Hussain, Y, Hussain, A, Kolivand, H and Khan, W

Artificial Intelligence in Educational Technology: A Systematic Review of Datasets and Applications

Artificial Intelligence in Educational Technology: A Systematic Review of Datasets and Applications

School of Computer Science and Maths, Liverpool John Moores University, L3 3AF, Liverpool, UK School of Education, Liverpool John Moores University, L3 3AF, Liverpool, UK EdenFiftyOne, UKCollege of Engineering, University of Sharjah, UAE

LUKE TOPHAM, PETER ATHERTON, TOM REYNOLDS, YASIR HUSSAIN, ABIR HUSSAIN, HOSHANG KOLIVAND, and WASIQ KHAN

Artificial Intelligence (AI) has the potential to impact a diverse range of domains. For instance, AI for the education domain has received increasing interest with various applications, including predicting performance, curating learning materials, and automated assessment and feedback. Despite the developments, some imbalances appear in the literature; for example, traditional classrooms and non-scientific academic subjects received little attention. This survey provides a systematic review of the current trends in AI research for education, specifically addressing applications within secondary education (ages 11+) through to higher education (HE), and offers a detailed compilation of datasets and methods, facilitating a deeper understanding of the field and encouraging further investigation. It includes a thorough review of the datasets available to encourage and enable future research, development, and collaboration, as well as the establishment of performance benchmarks. Furthermore, this survey provides an overview of issues and problems arising from recent developments, which may aid policymakers in their decision-making and addressing ethical concerns and standards. For example, many AI in Education (AIEd) platforms are not grounded in educational theory. We also present several guidelines to aid future developments in AIEd, guiding long-term impactful projects and investments.

CCS Concepts: • **Applied computing → E-learning**; **Computer-assisted instruction**; *Interactive learning environments*; *Computer-managed instruction*; • **Computing methodologies → Machine learning approaches**; *Natural language processing*.

Additional Key Words and Phrases: Education, Virtual Learning, Artificial Intelligence, Machine Learning, Chatbots, Large Language Models, AIEd

ACM Reference Format:
Luke Topham, Peter Atherton, Tom Reynolds, Yasir Hussain, Abir Hussain, Hoshang Kolivand, and Wasiq Khan. 2024. Artificial Intelligence in Educational Technology: A Systematic Review of Datasets and Applications. 1, 1 (September 2024), 29 pages. https://doi.org/XXXXXXX.XXXXXXX

Authors' address: Luke Topham; Peter Atherton; Tom Reynolds; Yasir Hussain; Abir Hussain; Hoshang Kolivand; Wasiq KhanSchool of Computer Science and Maths, Liverpool John Moores University, L3 3AF, Liverpool, UK School of Education, Liverpool John Moores University, L3 3AF, Liverpool, UK EdenFiftyOne, UKCollege of Engineering, University of Sharjah, UAEmymainaddressPAaddressTRaddressPAHaddress.

Manuscript submitted to ACM

1 INTRODUCTION

Education has witnessed a variety of technological innovations over the past decades. For instance, in the 1950s and 1960s, pre-recorded lessons were aired on TV, bringing education to the masses in their own homes [119]. Similarly, in the early 2000s, Massively Open Online Courses (MOOCs) made university curricula available to students across the globe without the need to enrol at a university [119]. Despite such innovations' benefits, neither of these examples brings innovation into the traditional brick-and-mortar classroom. Recent AI developments may allow innovation in conventional and distance-learning classrooms. AI refers to computer programs which complete tasks that typically require human intelligence. For example, speech recognition is a task that usually requires human intelligence and is very difficult to achieve using standard programming concepts. However, many AI applications now solve such tasks [14]. Furthermore, AI has many applications which may benefit the education domain.

AIEd is the application of AI techniques in the educational domain. It may be implemented in aspects of education such as administration, learning material curation, assessment, prediction, and more. AIEd can benefit teachers, from school to Higher Education (HE), by reducing their workload and eliminating time-consuming, laborious tasks [24]. For instance, [2] suggests that up to 40% of lesson time is spent on activities that could otherwise be automated. Specifically, they highlight activities such as preparation, assessment and feedback, and administration [2]. The work also suggests some AIEd applications which may help teachers to reclaim some of this lost time, such as profiling and prediction, assessment and evaluation, adaptive systems and personalisation, and intelligent tutoring systems [2]. Thus, AIEd could free teachers to spend more time with their students, which may, in turn, help to improve learning outcomes.

1.1 Previous Survey Studies

Existing AIEd survey studies are presented in Table 1. These works highlight several common aspects of AIEd. For example, [15, 21] suggest that many applications of AIEd begin by building comprehensive learner models of students enrolled on the system, and [66] indicates that this may be used as a base for personalised learning. Moreover, [2, 15, 21] provide scenarios and aspects of education that may be addressed by AIEd, for example, administration, course material curation, prediction (e.g., achievement and dropout), instruction, grading, and feedback.

The subject categories in Table 1, Neutral, STEM, and Languages, reflect how prior reviews have grouped AIEd applications. "Neutral" refers to domain-agnostic tools such as dropout prediction or intelligent tutoring systems that apply across subjects. STEM and Languages are highlighted separately because they have been especially prominent in AIEd research: STEM due to its structured, problem-oriented content, and languages owing to the close link between language learning and Natural Language Processing (NLP). These categories, therefore, align with the focus of existing reviews while capturing the areas where AIEd has been most actively applied.

Table 1. An overview of existing surveys on AIEd-related topics, including the learning environment (i.e., classroom (C), online (O), or Massive Open Online Course (MOOC) (M)), academic subjects covered including neutral (non-subject specific), languages, Science, Technology, Engineering, and Mathematics (STEM), and languages, as well as brief descriptions of the content.

Ref	Year	Subject	Env.	Description
[76]	2024	Neutral	C, O	An overview of AIEd perspectives, learning theories, and suggested future directions. Highlights the need to reform educational policy and suggests an interdisciplinary approach. Limitations: no discussion of ethics or teacher perspectives.

Continued on next page

Table 1 – continued from previous page

Ref	Year	Subject	Env.	Description
[37]	2023	Neutral	C, O	Provides an overview of the controversial topic of ChatGPT in education. Opposes the complete ban of ChatGPT, comparing it to calls for the ban on calculators. Presents mitigation strategies relating to ChatGPT-misue but accepts that it is impossible to prevent all misuse. Limitation: does not consider other LLMs or related technologies.
[18]	2023	Neutral	C	A review of AI chatbots to support students, including benefits and potential challenges. It reports increased student engagement and student preference for conversational style output. It highlights the limitation of chatbots not understanding students' emotions. Limitations: non-technical participants, doesn't consider human traits (e.g., age, gender, etc.), and doesn't cover ethical considerations.
[86]	2022	Neutral	O	A review of AI in education with a focus on the AI algorithms commonly used. It reports increased student engagement and additional benefits such as accurate assessment and predictions (e.g., of performance). It highlights that advanced AI, such as deep learning and genetic algorithms, are rarely used. Limitations: limited search criteria are unlikely to provide sufficient breadth, and it does not consider any conference papers.
[73]	2022	STEM	O, M	A review of student modelling methods for adaptive online learning environments. It reports that Deep Knowledge Tracing (DKT) models produce the best prediction performance. It highlights that student models with problem characteristics perform better than models without. Limitations: It does not cover ethics, has limited model coverage, and mentions interpretable AI but does not cover the topic in depth.
[2]	2021	Neutral	C	An outline of challenges and potential of AI in education. It highlights that there are no standardised methods of integrating AI into existing educational models. Limitations: It does not cover a wide range of ethical issues, and although it mentions the acceptability of AI in general, it does not consider whether teachers would accept AI in the classroom.
[66]	2021	STEM	C, O, M	A review of machine learning for precision education and a description of the key findings. It reports that previous works focus mainly on prediction, and there are fewer works on other applications of AIEd. It highlights a convergence of AIEd and neuroscience, suggesting related methods to improve models. Similarly, it highlights the need to integrate AIEd with existing learning theories. It emphasises that AIEd makes the resource-intensive task of personalised learning achievable. Limitations: limited research synthesis, lack of consideration for conference papers, and some restrictions based on the journal ranking rather than paper quality.
[17]	2021	Languages	C, O	A survey of twenty years of precision language education focused on topic modelling and knowledge mapping for personalised learning. It highlights the use of AIEd for personalised learning, including feedback and assessment. Limitations: limited search databases and unoptimised search strings.

Continued on next page

Table 1 – continued from previous page

Ref	Year	Subject	Env.	Description
[15]	2020	Neutral	C, O	An assessment of AI's impact on education in administration, instruction, and learning. It highlights the use of AIEd to provide practical or experiential learning. It also promotes the use of AIEd-enabled robots in physical classrooms. Limitations: Little discussion of disadvantages or ethical considerations.
[16]	2020	Neutral	O	A systematic review of influential AIEd articles. It evaluates the definitions of AIEd and clarifies the relationships among AIEd, educational data mining, computer-based education, and learning analytics. It recommends further use of advanced AI, such as deep learning. Limitations: restricted search databases, conference papers were not considered, restricted articles based on citation count, and a limited discussion of ethical considerations.
[13]	2018	Neutral	O, M	A description of the impact of AIEd. It aims to identify the prospective impact of AIEd and to predict possible changes in the educational domain. It focuses on customised educational content, innovative teaching methods, technology-enhanced assessment, and communication between students and teachers. It highlights the use of AIEd for personalised learning. Limitations: limited search breadth and no discussion of ethics.
[21]	2017	Neutral	O, M	A survey of related topics to AI techniques for adaptive educational systems within e-learning. It highlights the need to ground AIEd in pedagogical theory. Limitations: no discussion of ethics.
[96]	2016	Neutral	C, O	Provides a review of the evolution of AIEd over the past 25 years, focusing on current strengths and future opportunities. It also suggests two parallel strands of future research. Limitations: technology evolves rapidly, so many older works are irrelevant now. It also suggests that AIEd may replace human teachers, which is unsupported.

Although these studies provide valuable information, they do have limitations. For example, [2, 21] do not offer a systematic review. Moreover, [2] does not provide a significant review of current applications or relevant current AI technologies. Similarly, [13] is relatively narrow in scope and does not analyse overall AIEd trends. None of the works significantly explores the ethical issues surrounding AIEd; in particular, none, other than [37], address the growing problem of students' misuse of AI. Similarly, none of the works significantly explore teacher or student acceptance or perceptions of AIEd, a key factor in their future adoption.

Moreover, to the best of the authors' knowledge, no existing survey provides a significant review of current datasets available for AIEd. Data is a vital aspect of AI and, therefore, must be carefully examined in the context of education. Moreover, AI and AIEd are fields that are evolving rapidly; consequently, it is vital to ensure that current issues and challenges are thoroughly examined as they occur. However, although [2, 13, 96] provide a review of challenges and potentials for AIEd, they ignore some critical current issues, for example, the protection of intellectual property (IP) when using tools such as ChatGPT. Similarly, to the best of the authors' knowledge, none of the existing surveys explore current educational theory. This is despite works such as [16] identifying that many AIEd platforms are not grounded in educational theory.

Furthermore, some studies severely limit the scope of articles they consider for review; for example, [17, 66, 86] do not consider conference papers or books. Similarly, [16] considers only articles with a minimum of 20 citations, which will likely produce a bias in favour of older articles. Moreover, [96] selects articles from three specific years only: 1994, 2004, and 2014.

We present a systematic overview of the recent AIEd literature to address these limitations. This ensures a thorough examination of relevant datasets and applications. Moreover, for the first time, we describe the emerging challenges and consequences of rapid advancements in AIEd. Unlike some existing works, we focus on AIEd and do not aim to provide a general exploration of the developments in AI beyond education. Readers interested in wider aspects of AI developments should consult works such as [80, 99].

1.2 Motivation

Education has the power to transform lives, empower individuals, and aid class mobility. Accordingly, highly motivated and competent teachers are central to delivering high-quality education. However, [71] reports that out of 830 teachers surveyed, 76.4% are considering quitting the profession. In particular, high pressure, increasing workloads, and low morale are reported as common reasons for the motivation to quit [71]. Furthermore, [71] suggests that a key reason for the increasing workloads and decreasing morale is the mass exodus of staff from the teaching profession. Therefore, it is necessary to identify methods of alleviating the pressure on teachers whilst maintaining the quality of education to help reduce the number of teachers quitting. AIEd can potentially lessen the pressure on teachers and make learning more effective.

AIEd presents the opportunity to automate or semi-automate many labour-intensive tasks such as material curation, assessment, and administration [120]. Thus freeing teachers to concentrate their efforts on delivering high-quality education without becoming overburdened [81]. AI and AIEd are rapidly evolving fields, and to make future contributions, it is necessary to understand educational theory [67], AI, and relevant datasets. Notably, the literature does not significantly review or describe datasets and modern issues. Therefore, a study surveying these topics could benefit future AIEd contributions. Such contributions may improve teachers' mental health and well-being, improve the quality of education for learners, and potentially bring high-quality education to individuals who may not otherwise receive it.

Despite the potential benefits for teachers and students, there are possible disadvantages. As AI technology rapidly evolves, policymakers struggle to create and maintain timely guidelines and policies [45] on what should and should not be allowed in classrooms and how best to maintain academic and ethical standards. For AIEd to be safely incorporated into classrooms, a concrete understanding of the policies and guidelines should be acquired [54]. However, to the best of the authors' knowledge, while several studies have addressed individual challenges in AIEd, none have systematically analysed and synthesised the full spectrum of problems, such as gaps in publicly available datasets, limitations of existing platforms and applications, and contemporary challenges like AI misuse in assessments. This lack hinders a holistic understanding and coordinated advancement in the field.

In order to investigate the aforementioned points, this manuscript is specifically interested in secondary education (ages 11+) through to HE (i.e., University). Students typically have higher digital literacy rates and are more independent at these levels. Similarly, the curriculum is usually more complex, with a broader variety of AIEd integration. To address the limitations of existing studies, this manuscript reviews the datasets available to implement AIEd, explores the challenges and ethical issues, and carefully considers the roles of teachers in the world of AIEd.

1.3 Contributions

For the first time, we present a thorough review of AIEd and the datasets necessary to develop AIEd solutions. Similarly, a common complaint of AIEd research is that many applications are not grounded in educational theory. Therefore, we present a concise survey of appropriate educational theory, which may form a basis for future AIEd research. Furthermore, the recent explosion in the popularity and availability of AI tools has meant that policymakers are struggling to develop effective policies around AI and AIEd in academic institutions. Therefore, we highlight issues and problems that have recently arisen due to such technology. These findings will aid policymakers in forming guidelines and policies relevant to the current and evolving situation. This survey contributes to the following aspects:

- A comprehensive systematic review that synthesises existing research, models, and technological platforms and applications in AIEd, highlighting key trends, gaps and future research directions.
- A thorough and detailed review of publicly available datasets pertinent to the development, training, and evaluation of AIEd solutions.
- An examination of contemporary challenges arising from recent advances and the growing use of AI tools, including concerns relating to academic misconduct and students' misuse of AI in assessments.

1.4 Organisation

The remainder of this article is organised as follows: Section 2 describes the methodology of this survey study, including the research aspects and search strategy. Section 3 presents the results of the AIEd survey and includes a general overview of the AI techniques and models which form the basis of AIEd technology. Section 4 offers the authors' recommendations for future AIEd research and applications. Finally, Section 5 provides the concluding remarks and suggestions for future AIEd surveys.

2 METHODOLOGY

This survey aims to systematically review the literature regarding AIEd to help guide future AIEd platforms and research. The scope of this article and the literature reviewed within it are guided by two elements: research aspects and search strategy. The following subsections describe the methodology for this review.

2.1 Research Aspects

The research aspects investigated in this survey include: (a) What is the general taxonomy of AIEd? (b) What datasets are available for AIEd research and development? (c) How does AIEd relate to established educational theory and practices? (d) What are the modern challenges and opportunities in AIEd?

2.2 Search Strategy

The PRISMA diagram in Figure 2 illustrates the selection and screening process for the articles included in this manuscript.. The articles reviewed for this work were identified using the following search engines and digital libraries: Google Scholar, IEEExplore, and Web of Science. These databases were used to provide a means of accessing a broad range of high-quality research. IEEExplore provides access to a plethora of peer-reviewed articles from computer science-related fields, whereas Google Scholar and Web of Science provide access to interdisciplinary research. Therefore, the authors could browse and search both technical and pedagogic research related to AIEd. Moreover, Table 2 presents the search strategy used to explore the literature for this work. Furthermore, articles must be written in English, peer-reviewed, and published within the past five years to be considered for inclusion in this work. The 5-year publication limit coincides with recent technological advances, such as the introduction of Large Language Models (LLMs) in 2018

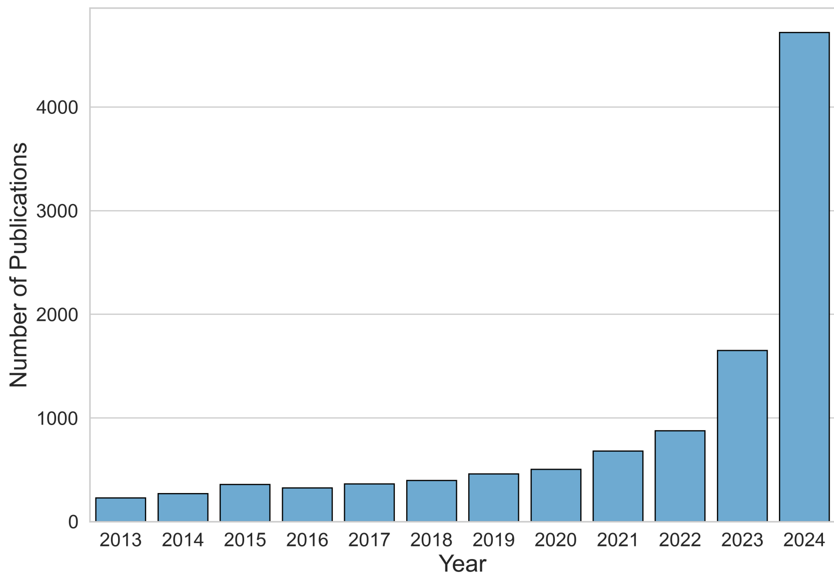

Fig. 1. The number of AI educational platform papers published per year.

[27]. Additionally, the authors reviewed the articles' abstracts and conclusions to eliminate inappropriate articles before performing a complete review. Consequently, a total of 56 articles were reviewed for this work.

Table 2. The search strategy used to explore the AIEd literature describing the permutations of context and objectives

Goal	Keywords
Context	Artificial Intelligence, AI, Machine Learning, Intelligent Systems, Large Language Models, Natural Language Processing, Chatbot, GPT, Data Mining
Objective	Education, Educational, Virtual Learning
Permutations	(Context AND Objective) AND ("Platform" OR "Environment" OR "AIEd")

Figure 1 displays the number of AI educational platform papers published yearly for the previous ten years, excluding the year of writing (2024). The number of such articles has increased from 227 in 2014 to 1,650 in 2023, and finally 4,720 in 2024. As of April 2025, 1,380 papers have been published, suggesting that the increasing trend may continue. Therefore, it is apparent that AIEd has proliferated in recent years, particularly since 2018, when LLMs were introduced.

2.3 Screening and Selection

Using the search strategy described in Section 2.2, 3,650 records were identified as potential articles to include in this study, as described in Figure 2. The screening stage resulted in 367 documents being identified as duplicates or records lacking full text, and were therefore removed. Moreover, based on the title, 2762 articles were removed. We excluded survey articles, letters to editors, papers without peer review, and irrelevant articles. During the eligibility phase, 82 out of 154 records were excluded based on their abstract, for example, due to a lack of relevance. Next, the full text of 72 articles was reviewed, and 14 were removed due to a lack of relevance, rigour, quality, etc. Finally, 58 articles were included in this study.

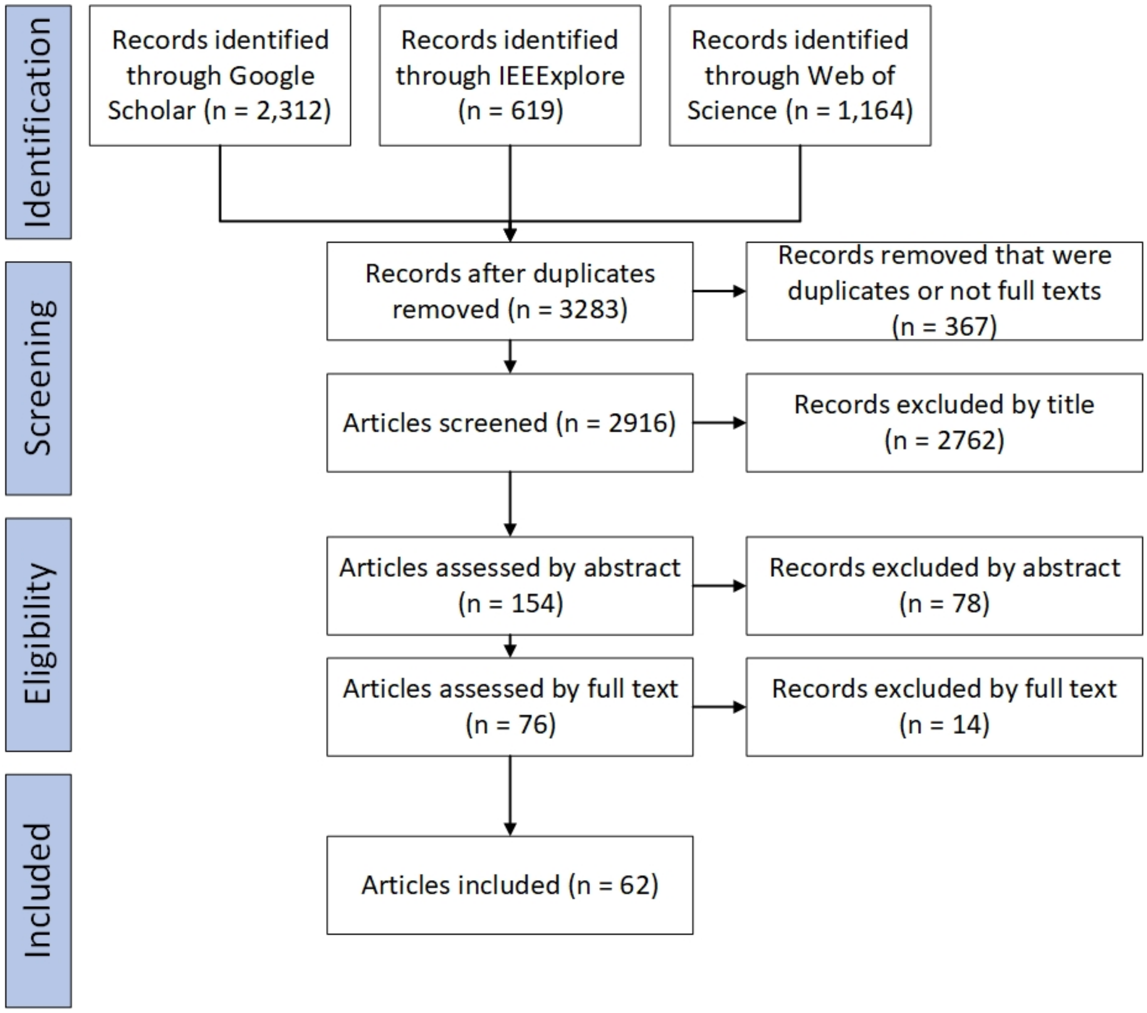

Fig. 2. Flow chart of the article selection process adapted from the PRISMA guidelines [77], showing the stages of selection through identification of sources, screening, eligibility, and resulting in the final included lists of articles.

2.4 Discussion

Our methodology aims to balance breadth and focus. By limiting the survey to peer-reviewed journals published in the past five years, we are ensuring quality and currency in a post-LLM landscape. Arguably, the trade-off of this approach is that some highly cited papers may be excluded if they are more than five years old; however, this constraint is necessary to ensure that the selected works reflect current developments in a post-LLM era. Similarly, the search strategy described in Table 2 was designed to capture the major trends and emerging topics in AIEd. Therefore, niche topics may not be fully represented in the search results. Future surveys may consider underrepresented niche subcategories, non-English language articles, and prominent pre-print articles. This could help capture a more globally inclusive and current view of the research landscape, especially in areas where innovation may be happening outside traditional academic publication channels or dominant academic languages (i.e., English).

3 ARTIFICIAL INTELLIGENCE IN EDUCATION

AIEd is an interdisciplinary field combining AI and education. AIEd produces intelligent software that may automate aspects of a teacher's role, help personalise students' educational journeys, or make predictions about student achievement. The following subsections provide some relevant background on AI, a review of the datasets which make AIEd possible, and a survey of relevant AIEd platforms.

3.1 Artificial Intelligence Background

AI refers to software that completes tasks usually deemed to require human intelligence. Moreover, machine learning is a subcategory of AI where software learns and adapts without being explicitly programmed. Machine learning can broadly be split into four categories: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. In supervised learning, machine learning models are provided with annotated input data and the desired output to enable them to learn the mappings between input and output. Alternatively, in unsupervised learning, models are provided with unlabeled data and no output labels; models then analyse and group (i.e., cluster) data to identify otherwise hidden patterns or groupings in data. Semi-supervised learning uses a combination of labelled and unlabelled data, whereby the desired output label is usually provided only for a small subset of the data. Figure 3 provides an overview of machine learning methods, problem types, and example algorithms relevant to AIEd.

The concepts explored in Figure 3 empower the implementation of many features in AIEd. For example, student performance may be treated as a classification problem, where students are grouped, or a prediction problem, where regression estimates grades. Furthermore, semi-supervised learning provides applications such as speech analysis, which can convert student or teacher speech to text for analysis. Generative methods can generate new content, such as images. Similarly, NLP and LLMs may be used to generate text, for instance, in response to a student's question. NLP is an approach to analysing natural language in the form of text or speech to gain an understanding of the content. Moreover, LLMs are based on transformer architecture and have been trained on massive datasets to enable them to produce human-like text.

In recent years, NLP and LLMs have grown in popularity and are implemented in many domains, including AIEd. NLP aims to understand natural language content such as text documents or speech. LLMs, such as the popular Generative Pre-trained Transformer (GPT) used in OpenAIs ChatGPT [35], aim to produce human-like text, for example, in response to a student's query.

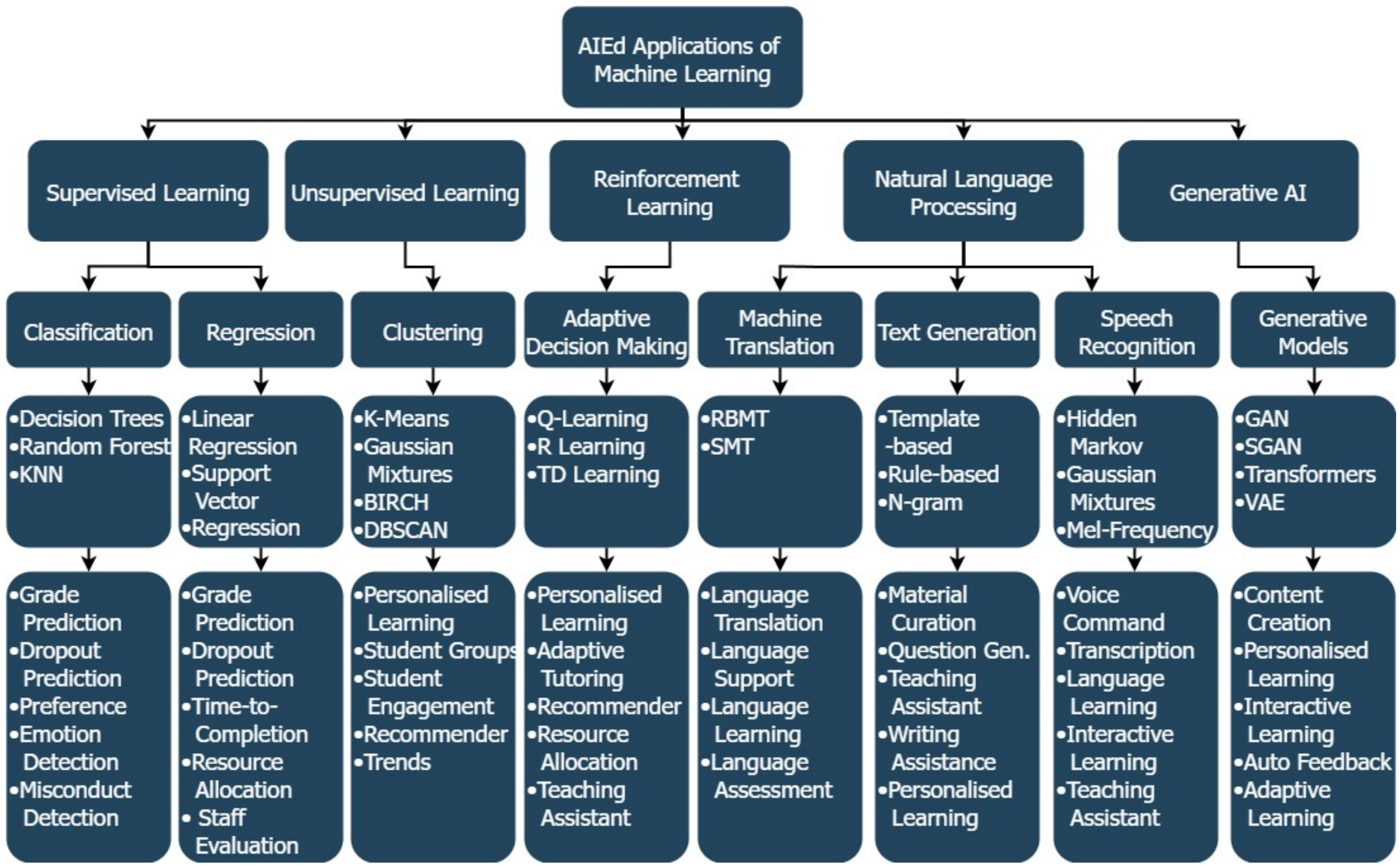

Fig. 3. A taxonomy of AIEd applications of Machine Learning with related concepts, algorithms and supported activities. Including a non-exhaustive list of recommended AIEd applications. Abbreviations: K-Nearest Neighbours (KNN), Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH), Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Generative Adversarial Network (GAN), Semi-Supervised Generative Adversarial Network (SGAN), Variational Autoencoder (VAE), Rule-Based Machine Translation (RBMT), Statistical Machine Translation (SMT).

Table 3. A description of available datasets appropriate for AIEd, including the year of publication, source, academic level, academic subject, the aspect of education targeted, the number of participants recorded in the dataset, and the total number of records.

Ref	Year	Source	Level	Subject	Aspect of Education	No. Participants	No. Records
[19]	2020	GitHub	MOOC	English as a second language	Prediction (multiple domains)	784,309	131,417,236
[97]	2014	Institute	MOOC	Psychology	Prediction (dropout)	20,828	3,475,485
[87]	2013	Institute	School	Mathematics	Prediction (performance)	70,808	7,178,761
[94]	2015	GitHub	MOOC	Mathematics	Knowledge mapping	4,000	200,000
[12]	2015	GitHub	MOOC	Mathematics	Knowledge mapping	247,606	25,925,922
[56]	2010	Institute	School, HE	Algebra, Chemistry, Chinese, English, geometry, Physics	Knowledge mapping	335	361,092
[107]	2015	Institute	HE	Computer Engineering	Prediction (performance)	115	230,318
[60]	2017	Institute	HE	Multiple (22 courses)	Prediction (performance)	32,593	10,655,280
[47]	2018	Institute	HE	Non-specific	Prediction (performance, dropout)	300	7,200
[52]	2020	Mendeley	MOOC	Multiple (200 courses)	Material curation (NLP)	N/A	12,032
[36]	2020	Institute	HE	Physics and Electricity	Prediction (performance)	1,233	1,233
[26]	2020	Mendeley	HE	Engineering	Prediction (performance)	12,411	546,084

3.2 Datasets for Education

Table 3 provides an overview of the available datasets that may be useful in the training and evaluation of AIEd models. It includes the curriculum subject from which the data was recorded. However, in the case of data for prediction, the subject may be irrelevant to the end user. Table 3 shows that 8 out of 12 datasets are attained from Science, Technology, Engineering, and Mathematics (STEM) based courses, suggesting an imbalance in the academic subjects targeted by such research. Moreover, the educational aspect, such as prediction (i.e., performance, dropout, etc.), knowledge mapping, and material curation, is stated. 8 out of 12 datasets are available for prediction, 3 for knowledge mapping, and 1 for material curation, suggesting a severe imbalance in terms of educational aspects. Table 3 also shows that from the 12 identified datasets, 5 contain data from MOOCs, 6 contain HE, and only 3 contain data collected from school students ([56] contained both school and HE data). Table 3 shows that the number of student participants in the datasets varies massively; for example, 115 students participated in [107], whereas 784,309 participated in [19]. Moreover, less than half of the datasets in Table 3 use external dataset repositories such as GitHub and the Mendeley data repository. Whereas, the majority host the data at an institutional level, which may pose accessibility issues in the long term.

More specifically, some prediction datasets are for prediction [26, 34, 36, 60, 87, 107], some such as [55] are for dropout, and some, such as [19, 47], enable various aspects to be predicted. However, it may be noted that significantly more datasets are provided for predicting performance than for dropout and other purposes. Such datasets provide various data aspects such as observed student affects (e.g., bored, engaged, etc.) [87], actions [60] (e.g., Learning environment logs and interactions), personal information (including academic information) [26], and more. Whereas knowledge mapping datasets such as [12, 56, 94] tend to provide data regarding tasks, lessons, and assessments related to knowledge areas. Moreover, [52] provides word embeddings and document topic distribution vectors to enable content curation. Despite the potential time-saving benefits to teachers of being able to automate some material curation tasks, very few datasets provide this ability. This is potentially linked to the fact that some teachers do not share materials they create [111].

Furthermore, datasets such as [19, 34, 55, 87] provide data regarding students' interactions with digital education platforms. Such data may be used to predict problems such as students' performance or dropout rates. Such features may include the number of page requests, active days, forum views, etc. Alternatively, [12, 56, 94] may be used for knowledge mapping, where students' knowledge is modelled as they interact with their work. Such work's results may help train models capable of customising education per student.

Moreover, the majority of the datasets in Table 3 use real-world data; however, [94] uses simulated data for one of their chosen datasets. Virtual students are simulated, providing example answers to exercises based on probabilities from a defined latent knowledge state [94]. Although additional non-simulated datasets are also used, no validation of the simulated data is provided (i.e., responses to the exercises from real students are not captured and compared). Simulated data may offer a convenient method to produce larger datasets and compensate for gaps in existing datasets. However, further work is required to assess the efficacy of simulated data in developing AIEd applications.

Table 3 suggests that there are significant imbalances in the academic subjects and educational aspects that are currently available. This will almost certainly lead to an imbalance in the types of research projects conducted. For example, very few datasets described in Table 3 are based on non-scientific subjects. Similarly, few datasets provide data relevant to aspects such as intelligent tutors, material curation, or automated assessment.

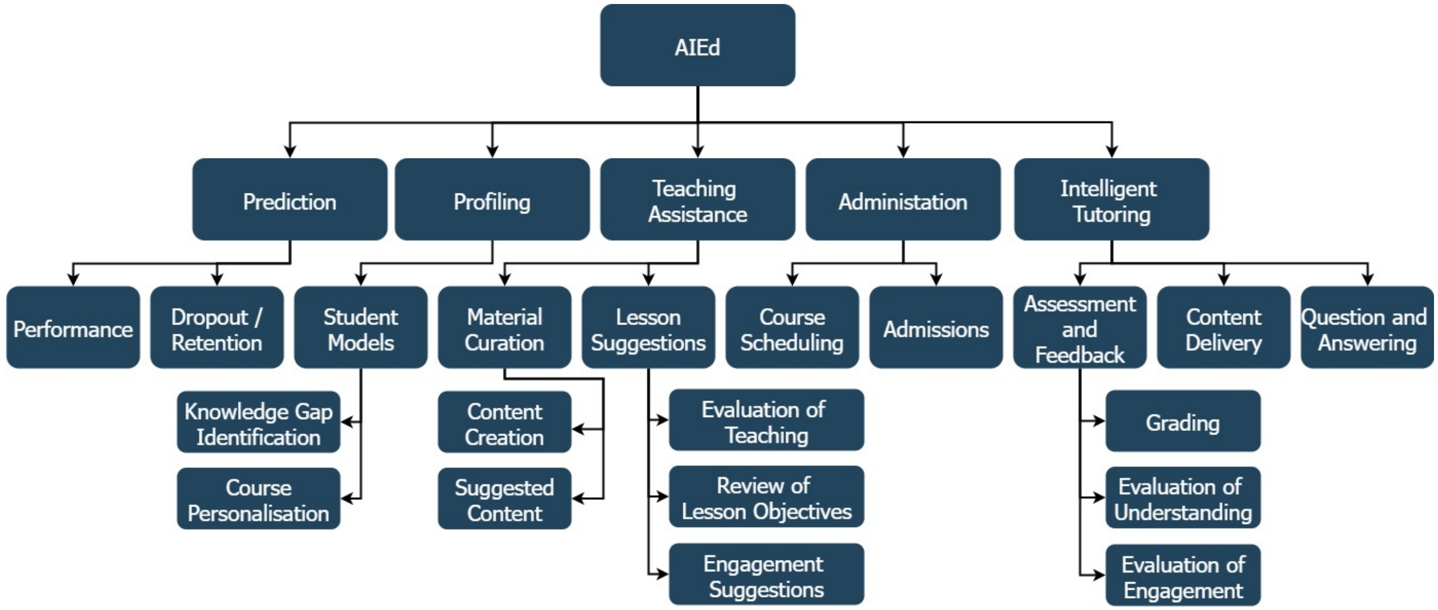

Fig. 4. The taxonomy of AIEd applications can broadly be categorised into application tasks regarding prediction, profiling, teaching assistance, administration, and intelligent tutoring.

3.3 AIEd Platforms

To the best of the authors' knowledge, there is not yet a widely accepted taxonomy of AIEd platforms and applications. Figure 4 presents a proposed taxonomy covering the range of AIEd applications currently identified. AIEd provides five main areas of application: prediction, knowledge mapping, teaching assistance, administration, and intelligent tutoring. Table 4 presents an overview of articles reviewed during this research. It shows the variety of educational aspects, learning environments, academic subjects, and AI models used to implement the application. In this context, traditional refers to classroom-based settings (e.g., schools or universities), online refers to digital or virtual learning environments, and both denote applications, such as performance monitoring, that are relevant to either setting." The remainder of this subsection describes these concepts and works in more detail.

Table 4. An overview of existing AI literature describing applications and studies related to education.

Ref	Year	Aspect of Education	Purpose	Learning Environment	Academic Subject	AI Model(s)
[102]	2023	Material Curation	Question Generation	Online	Medicine	LLM
[59]	2023	Material Curation	Question Generation	Online	Medicine	LLM
[85]	2023	Prediction	Performance	Online	Engineering	Genetic Program-ming
[61]	2023	Teaching Assistant	Performance and Feedback	Online	English	RF
[64]	2023	Teaching Assistant	Classification (DA)	Online	Various	NLP, LLM
[65]	2023	Teaching Assistant	Intelligent Tutoring	Traditional	English	NLP, LLM
[28]	2023	Material Curation	Narrative Fragments	Online	Various	NLP, LLM
[118]	2022	Teaching Assistant	Chatbot	Traditional	Various	NLP
[74]	2022	Knowledge Mapping	Estimate Student Knowledge	Online	Various	Naïve Baye
[105]	2021	Personalised Learning	Profiling (knowledge)	Online	English	Decision Tree
[90]	2021	Teaching Assistant	Chatbot	Online	Various	NLP
[20]	2020	Prediction	At Risk (Intervention)	Traditional	N/A	SVM

Continued on next page

Table 4 – continued from previous page

Ref	Year	Aspect of Education	Purpose	Learning Environment	Academic Subject	AI Model(s)
[42]	2020	Teaching Assistant	Chatbot	Online	Languages (Japanese)	NLP
[25]	2020	Prediction	Dropout	Classroom	Various	LDA, SVM Random Forrest
[43]	2020	Prediction	Performance	Both	Various	Random Forrest
[101]	2019	Prediction	Performance (test scores)	Both	Various	SVM, LSTM
[41]	2019	Teaching Assistant	Chatbot	Online	Languages (Japanese)	NLP
[88]	2019	Teaching Assistant	Chatbot	Online	Languages (Korean)	NLP
[91]	2019	Prediction	Dropout (Intervention)	Online	Programming	Decision Tree
[108]	2019	Knowledge Mapping	Estimate Student Knowledge	Online	Various	Factorisation Machines
[78]	2018	Teaching Assistant	Chatbot	Online	Various	NLP
[5]	2018	Personalised Learning	Profiling (learning styles)	Online	N/A	N/A

Firstly, AIEd can provide predictions, for example, the prediction of student performance, including identifying students at risk of not achieving, and it can predict future dropout or retention. Secondly, knowledge mapping builds an understanding of students' knowledge, understanding, and achievement and may aid the personalisation of their education. Thirdly, AIEd may directly benefit the teacher by assisting with curating materials and providing suggestions to improve planned classes. Similarly, AIEd can help reduce teachers' workload by assisting with or automating aspects of their administrative duties, such as course scheduling and admission. Finally, intelligent tutoring may automate some aspects of teachers' academic tasks, such as assessment and feedback, content delivery, and question and answer. Intelligent tutoring may be particularly advantageous to students who do not have significant access to a teacher, such as those on distance learning courses. The remaining subsections explore these aspects in more detail with examples from the literature.

3.3.1 Prediction. AIEd may support academics and administrators by providing prediction support in terms of student performance, admission, at-risk identification, dropout, and more. Predicting performance and at-risk students has important applications. For example, it may be used to target intervention to prevent students from underperforming or failing. Similarly, it may be used for early withdrawals if a student is unlikely to pass, allowing them to find a more suitable course or prevent wasted course fees.

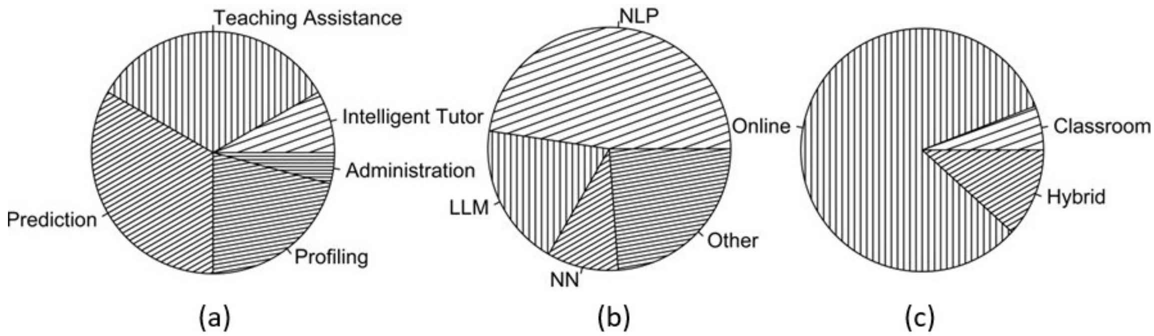

Fig. 5. The prominence of AIEd research aspects: (a) educational aspects, (b) AI techniques used, and (c) educational environment.

For example, in [20], the supervised machine learning algorithm Support Vector Machine (SVM) is used to classify students in terms of their predicted performance to identify students who may be at risk of not achieving. Specifically, a reduced-training SVM (RTV-SVM) is presented to reduce training time without impacting performance by removing redundant vectors, and the method still achieves classification accuracy between 92.2% and 93.8% while reducing training time by approximately 60% [20]. Similarly, [101] predicts and classifies student performance but does not identify at-risk students. For both experiments, several machine algorithms are considered. In the prediction experiment, Support Vector Regression (SVR) was reported to produce the highest accuracy [101]. Whereas in the classification experiment, Back Propagation (BP) was reported to produce the highest accuracy [101]. Furthermore, in [85], performance prediction was combined with learning analytics to target improvements in course delivery.

Moreover, [25] predicts future student dropouts and experiments with three machine learning algorithms: Linear Discriminant Analysis (LDA), SVM, and Random Forest (RF). The findings suggest that additional learning requirements and course credits are potent features for predicting dropout [25]. Furthermore, the best-performing predictive models were the SVM and RF models, which achieved 87% accuracy, followed by the LDA model, which achieved 85% [25]. Similarly, [91] used decision trees to predict student dropouts reporting an accuracy of 80%. Such predictive models can assist institutions in targeting support and help forecast course running costs and staffing requirements. While deep learning methods can enhance performance, they may compromise the explainability of the decisions made by the models.

Figure 5 presents the prominent research aspects identified from the literature collected. For example, Figure 5-a shows that teaching assistance, prediction, and profiling are more prominent aspects of AIEd in the identified literature than intelligent tutoring and administration. Moreover, Figure 5-b shows that NLP and LLM were identified in many of the works identified in the literature, and other AI techniques were much less prevalent. However, this is likely due to the limitations on publications (2018 to 2023). Furthermore, Figure 5-c shows that most works are implemented for online or hybrid courses, with relatively few implemented for the traditional classroom. Therefore, the lower prominence of administration-focused AIEd and AIEd for the classroom may suggest that the technology's benefits are not significantly integrated into the conventional classroom. Further, teachers and students may not fully benefit from a traditional educational setting.

3.3.2 Profiling. Student profiling and knowledge mapping can help to build a model of a student's understanding and may be used to predict their performance. This model may then be leveraged to personalise a student's learning journey, ensuring that they are directed to appropriate learning materials at a level and in the best format for that student [116]. A personalised approach, supported via AIEd, may help to alleviate the feeling of information overload for students [17] and help to target their efforts to receive the maximum results, for instance, in terms of their grades.

Unlike previous knowledge tracing works such as [112], which use deep learning to maximise knowledge tracing performance, instead [108] and [74] present interpretable knowledge tracing. In [108], Factorisation Machines (FMs), a supervised machine learning algorithm capable of estimating interactions in sparse settings, are implemented to estimate student knowledge. Similarly, [74] presents an interpretable method of knowledge tracing; however, it is based on a Tree-Augmented Naive Bayes Classifier (TAN). Although the works do not include adaptive learning modules, the interpretable nature of the works may allow helpful feedback to be provided to the student.

Although many existing solutions model students' knowledge via their interactions with materials, [5] also analyses inter-student interactions via discussion forums. Such approaches increase the number of domains from which valuable data may be extracted to build models of students and may encourage soft skills such as communication and teamwork.

Furthermore, in addition to modelling a student's knowledge, [5] determines their preferred learning style. Therefore, besides suggesting resources and learning paths for a student, this work may provide suggestions for how the resources may be presented to maximise their effectiveness for each student.

3.3.3 Teaching Assistance. Aspects of AIEd can support classroom teaching without directly delivering content, providing some of the support traditionally provided by teaching assistants in some institutes. For example, questioning, answering, reading, pronouncing words, and answering general course queries [15]. Moreover, it has been suggested that robotic teaching assistants could provide an additional physical presence in the classroom [15], such as in [57]. AIEd teaching assistant support may respond to the more straightforward issues and free the classroom teacher to focus on teaching and dealing with more complex problems. Additionally, it may also provide some assessment support; for instance, in [90], exam papers are automatically screened for evidence of plagiarism and reports evidencing the suspicion are generated in suspected cases.

NLP and LLMs are commonly used to build chatbots, which may be used to implement such assistance. For example, [90] implements a question-and-answer chatbot using time series analysis of cloud-based big data to find relevant information. The chatbot assists students, encourages them to collaborate, and provides teachers with an understanding of students' current knowledge. Likewise, in [95], personalised instructions are provided to students, and language writing and translation are provided using NLP. To support such tasks, Lin et al [64] use a combination of NLP and LLM methods to classify Dialogue Acts (DA), the function or intentions of a statement. This can then be used to improve intelligent tutoring systems or to support the automated labelling of datasets.

Similarly, AIEd has been used to help curate teaching and learning materials. For example, Diwan et al., [28] demonstrate an LLM-based approach to generating narrative fragments, small story-like content to create an interactive learning pathway. Similarly, in both [59] and [102], LLMS are used to generate questions and answers for students based on their learning materials. These approaches not only reduce the time burden on educators but also demonstrate increased student engagement, support elements of personalised learning, and facilitate more effective solo study.

Also using NLP, in [118], a chatbot is implemented to support teachers in learning to use educational software by answering questions and providing demonstrations on how to complete essential tasks. Furthermore, using school library resources and topics provided by the teacher, [118] can compile reading lists for their students.

3.3.4 Administration. AIEd has been implemented to improve the effectiveness and efficiency of educational administration tasks, such as providing feedback, grading work, secretarial queries, course selection and student admittance

[15]. Administration work can be time-consuming and labour-intensive, and likely contributes to teacher burnout [71]. Automating some administration tasks may reduce the pressure on teachers and free them to focus on content delivery and other important teaching activities. However, [106] reports that AI in educational administration receives significantly less focus than topics more directly related to teaching and learning. Despite this, there are significant examples of AIEd providing benefits in educational administration. For instance, automated feedback may reduce teaching workload and provide increased scalability due to the processing speed, which has been found to increase student participation in assessments [58]. However, [11] reports that less than 5% of the literature regarding student feedback considers the teacher's needs, and therefore, it may not provide the desired reduction in administration overhead for teachers. Similarly, AIEd has been used to speed up the return of assessment feedback. For instance, in [40], decision trees are reported as producing accurate predictions of student success, which may be used to aid decisions regarding student admission, course progression or course selection. More recently, in [43], random forests are reported to perform well on the same task.

Despite the benefits, serious issues must be considered. For example, if decisions such as student admittance and grading are automated using AIEd, the AI model must be fair [70]. It must not discriminate against individuals or groups while favouring others [70]. Such issues can threaten student diversity or cause students to avoid a particular institution.

3.3.5 Intelligent Tutoring. Intelligent tutoring refers to content delivery tasks typically performed by a teacher and may include content-specific questioning and answering. Intelligent tutors are often implemented as chatbots using NLP and LLMs. However, these differ from the administration chatbots, which deal only with administration queries, not course content. Intelligent tutoring may help provide teaching to those who otherwise have limited access to teachers, such as those on distance learning courses. Moreover, such applications may help overcome the issues arising from recent problems regarding low teacher-to-student ratios [18].

Chatbots have been used in several works to provide a space for students to interactively practice foreign languages two-way, which would traditionally require another competent speaker. For example, Gengobot [41] is a chatbot designed to support the practice of the Japanese language. Moreover, Gengobot is combined with a social media application [42], which reported increased student confidence and motivation compared to working with teachers. Similarly, [65] uses a combination of NLP and LLMs to help students practice their English reading through student-AI co-creation of questions, and was reported to improve student motivation and low-level understanding. This concept is an expansion on the Student Question Generation (SQB) pedagogical approach, where students create their own questions based on what they have learned. The concept of using intelligent tutors to practice languages is further extended in [88], where a chatbot is integrated with Virtual Reality (VR) to enable students to practice the Korean language.

LLMs are a relatively recent advancement and may offer the potential to generate educational content and provide automated interaction to increase student engagement [51]. Chat-GPT [35] is a particularly popular LLM that has gained significant interest recently. Furthermore, it has been successfully demonstrated in educational settings, providing interactive questions and answers for students [82], curating learning materials [48], producing assessment questions [9], assessing student answers [79], and providing explanations [68].

3.4 Contemporary and Future AIEd Issues and Challenges

Despite the advantages that AIEd brings, it also presents several issues. For instance, the recent introduction of LLMs and generative models may present issues with trust and misinformation.

3.4.1 Privacy and Ethical Concerns. Like other AI methods, AIEd collects and analyses large amounts of data. The large-scale data collection of students may present human rights and related privacy and ethical issues [46]. Furthermore, [46] highlights several additional concerns regarding AIEd, such as personal privacy, freedom of choice, long-term implications, and choices regarding students below the age of legal adulthood. For instance, inclusiveness should be protected, and decisions and predictions should not discriminate against students based on protected characteristics such as gender, ethnicity, and additional learning needs [6].

Similarly, AIEd should be transparent [46], meaning informed consent must be gained before collecting data. It should include an explanation of the data to be collected, the data processing methods, and how the data will be used [7]. Similarly, AIEd should not be allowed to become a "black box"; instead, any predictions produced should be explainable so that teachers, students, and parents can understand [22]. This should allow for decisions and predictions to be explained and justified.

Similarly, protecting personal data is a legal requirement. Once data has been collected, it must be protected; this includes controls on how the data is managed, where it is stored, and who has access to it [46]. Furthermore, security procedures must be in place to ensure safeguards against data breaches, for example, protection from cybercriminals [104]. Moreover, to ensure the trustworthiness of AIEd, UNESCO highlights six challenges that should be solved [89]: lack of comprehensive public policy, exclusion and inequity, teachers' unpreparedness for AI, the requirement of quality data systems, need for significant AIEd research, and the need for a holistic comprehension of ethics.

3.4.2 Generative Pre-trained Transformers and Large Language Models Trust and Validation. Unlike traditional chatbots, which use tree-based pre-generated responses, generative models have a significantly more extensive range of responses for broader topics. However, they are much more prone to grammatical and linguistic errors [78]. Furthermore, such technology presents additional issues such as copyright issues, bias, unfairness, difficulty distinguishing model-generated answers from human-generated answers, training costs, data privacy, content errors, and more [3, 51].

For example, as users are unlikely to be aware of the source of information provided by models, they may not be mindful of copyrighted content and their rights and responsibilities in using this information. Furthermore, copyright and intellectual property owners may find it difficult to control and enforce their rights if their materials are used within training datasets.

Similarly, the model's output may contain unchanged phrases or sentences from the training set. This further presents opportunities for copyright violations and the possibility of unintended plagiarism, for example, for students or authors [51]. To the best of the authors' knowledge, there is not yet a widespread consensus on using generative models in published articles. Currently, many journals allow such content if authors declare it. However, such policies do not explain the process of dealing with issues that arise from their use. Furthermore, it is often challenging to differentiate between model-generated and human-generated content, which presents related problems in authorship and plagiarism detection for students and authors [33]. Therefore, developing improved methods of detecting generated content would be of significant value.

Furthermore, it is also difficult to validate the correctness of the information provided by generative models, which are susceptible to an issue known as hallucination, whereby the models produce convincing information that is incorrect

[8]. It is, therefore, difficult to verify the correctness of the information produced. Methods of verifying the correctness and source of the information would help improve confidence in using such models. Moreover, as such methods are relatively new, the long-term implications are unknown or untested. For example, students and teachers may become too reliant on generative models. This may result in laziness and a lack of motivation for personal learning and development.

Alternatively, Small Language Models (SLMs) have recently been introduced as an alternative to LLMs. SLMs are trained using smaller datasets and usually produce simpler models that are faster to train [113]. Early results have suggested that SLMs may outperform LLMs in specific domains, such as medical exams [38]. However, due to the more simple model and limited training dataset, SLMs may produce less accurate language in terms of writing style and grammar [113]. Therefore, if the writing of SLMs could be improved in terms of grammar and style, they may be a promising direction for future work in developing AIEd for specific academic subjects.

3.5 The Evolving Landscape of Academic Integrity with Generative AI

The increasing reliance on Generative AI (GAI) has given rise to concerns about academic integrity, as students can misuse it to produce plagiarised or fabricated content. It poses immense concerns regarding the authenticity of educational assessments as well as the validity of learning / knowledge maps when it comes to outcomes. This section, in its consideration, explores two consequential problems. It examines current work surrounding how learners misuse generative AI. Second, we address proactive detection and prevention mechanisms that would help curb such misuse.

3.5.1 The Misuse of Generative AI by Students. Students have also been swift to adopt generative AI tools (i.e., ChatGPT, Bard) to complete academic work, raising widespread integrity concerns. Weber et al. [114] highlighted that students are increasingly using AI to complete tasks like essays and coding homework. Similarly, Bobula et al. [10] surveyed students and found that over a third of students used AI tools for assignments, despite acknowledging it was misconduct.

Educators emphasise that presenting AI-generated text as one's own is plagiarism [31, 49]. Further, Wach et al. [109] emphasised that AI misuse extends beyond the academic context, threatening the overall reliability of information environments. Moreover, Johnston et al. [50] demonstrate that students recognise AI's value as a learning tool but also view its misuse. The over-reliance on AI not only undermines authentic intellectual development but also lowers the value of the integrity of educational assessments [10]. The study by Liang et al. [62] examines GAI's potential to transform teaching and learning in HE, particularly in social entrepreneurship education. It identifies three key areas where GAI enables new forms of interaction: collaborative learning, knowledge connectivity, and theory-practice integration. While promising, the research notes limitations such as its qualitative nature, narrow scope, AI hallucination and bias inheritance. The existing literature highlights how large language models have paved new avenues for academic misconduct while rendering detection increasingly challenging.

As highlighted by Cotton et al. [23] and Dwivedi et al. [30], the need for new institutional policies has been made even more urgent due to AI's blurred ethical boundaries. The research identifies the possible benefits and inherent risks of using generative AI in academic institutions. Marchal et al. [69] taxonomy of actual misuse tactics suggests advanced means used by threat actors in various modalities. In contrast, Perkins et al. [93] study reveals that teaching staff tend to struggle in detecting AI-assisted submissions. Furthermore, Williams et al. [115] analysis centres on the severe ethical issues surrounding the deployment of generative chatbots in education, namely data privacy, algorithmic bias, and student agency. As per Akpan et al. [1], the mounting use of conversational AI tools in industries is both enhancing their potential for instruction and research and widening areas of concern regarding plagiarism, misinformation, and intrusion into privacy. Shepherd [103] further illustrated how students in cybersecurity education have strategically misused generative AI, highlighting the urgent need for both technical solutions and adaptive pedagogical frameworks. Preserving academic integrity in the era of generative AI demands a hybridised approach that harmonises technological advancement, behavioural insight, and visionary educational reform.

3.5.2 Detection and Prevention Mechanisms. A diverse array of AI-content detection systems [114] has been developed to address academic misconduct, yet empirical evaluations reveal significant performance limitations. Tools such as GPTZero [39], OpenAI's Text Classifier [83], and DetectGPT [75] have undergone systematic analysis, demonstrating only moderate efficacy. For instance, GPTZero exhibited approximately 65% sensitivity in identifying AI-generated medical texts [39], while DetectGPT's perturbation-based method showed diminished effectiveness when exposed to paraphrased outputs [75]. These findings align with broader critiques by Perkins et al. [92], who argue that detector reliability is often overestimated and that transparent benchmarking practices are urgently needed. As such, detection technologies alone remain inadequate, underscoring the imperative for integrated academic integrity strategies. Such limitations underscore the importance of preventative measures and alternative assessments over detection tools.

In response, educational institutions increasingly prioritise transparent governance frameworks and reconfiguration of assessment methodologies. Universities are revising honour codes and embedding explicit guidelines regarding AI utilisation within course syllabi [72, 110] to foster a culture of academic honesty. Concurrently, there is a pedagogical shift towards designing AI-resilient assessments, such as multi-phase research projects, oral examinations, and reflective writing assignments that emphasise higher-order critical thinking [32, 44]. For example, in one of our related works, [53], we propose "auto-assessment of assessment" whereby students are presented with multiple-choice questions based on the content that they have submitted to check authorship. Importantly, as Lim et al. [63] argue, absolute prohibitions on AI usage are neither feasible nor conducive to educational development. Instead, institutions should "accept, adapt, and integrate" AI technologies within curricula to enhance learning outcomes [44].

The increased utilization of generative AI tools within learning spaces brings short-term and long-term problems to schools. As students see these technologies as beneficial tools for learning, instructors have a complex landscape where positive potential must be weighed against dire academic integrity risks. Current detection mechanisms have repeatedly failed to identify AI-generated work, further emphasising the ineffectiveness of traditional approaches in maintaining academic standards. This technological challenge requires a more nuanced and integrated response from institutions of learning. The future of education demands an integrated response that acknowledges the certainty of technological advancement while ensuring authentic intellectual growth through innovative means of evaluation and healthy institutional policy. This congruence is crucial for protecting the dignity of scholarly endeavours while allowing meaningful learning experiences within an increasingly AI-enhanced setting.

3.6 Discussion

There is a consensus in the literature on the importance of establishing a foundation of AI knowledge and standard terminology for both educators and students. Therefore, this section began by providing foundational knowledge and terminology to support their understanding of subsequent sections. However, it is beyond the scope of this article to provide a complete representation of the field of AI, and readers are encouraged to consult more comprehensive sources such as [98].

Building on the foundational background in AI, the discussion next turned to the landscape of publicly available datasets that support research and development in AI-driven education, as evidenced in the literature. Table 3 provided an overview of such datasets that support various AIEd research. In particular, ample datasets are available for prediction and knowledge mapping activities. However, it is clear that there are gaps, for example, in the availability of datasets for alternative activities such as course administration, teaching assistance, intelligent tutoring, and more. Similarly, there is a clear focus on dataset availability in STEM-related subjects, particularly mathematics. In contrast, non-STEM disciplines—such as the humanities, social sciences, and arts—remain significantly underrepresented. This imbalance limits the applicability and generalisability of AIEd tools across the full spectrum of academic disciplines, potentially reinforcing subject-area inequities and marginalising learners and educators in non-STEM fields.

Despite the datasets lacking diversity in subject matter, the literature demonstrates more diversity in AIEd applications. For example, Table 4 provides an overview of AIEd applications, including material curation, performance prediction, intelligent tutoring, and more. Further suggesting the desire for more diverse datasets. However, significantly fewer studies have addressed AIEd in administrative domains, despite substantial opportunities to automate repetitive and time-consuming tasks. Leveraging AIEd in this context could reduce educators' administrative burden, allowing them to dedicate more time and attention to pedagogical activities and student engagement.

Granted, the advantages of AIEd are substantial and well documented; however, several contemporary issues and challenges have been identified. Such as privacy and ethical concerns, as well as trust in the feedback provided by AIEd models. Therefore, there is a clear and urgent need for reform in terms of academic policy as well as technological advancements to encourage trust in and the adoption of AIEd.

The increasing misuse of GAI by students in assessments raises complex and pressing concerns for educators, institutions, and policymakers alike. While AI offers substantial educational potential, its unregulated use in evaluative contexts challenges academic integrity, assessment validity, and learning outcomes. Although there are several detection methods available, the reliability and accuracy have been questioned in the literature. Therefore, there is a pressing need for technical solutions, adaptive pedagogical frameworks and reform of educational policy.

4 RECOMMENDATIONS FOR FUTURE AIED PLATFORMS AND APPLICATIONS

After reviewing the strengths and limitations of the literature, the authors propose several recommendations for future AIEd platforms. They suggest that AIEd should be:

- Grounded in relevant educational theory.
- Explainable, whereby users can see what sources, processes, and deductions have influenced the result received. Furthermore, this may include controls against issues regarding generative models creating false information.
- Fair, users should not be discriminated against due to their personal characteristics (i.e., ethnicity, gender, sexuality, etc.), for example, when decisions or predictions are made regarding their future performance or course admittance.
- Integrated with traditional classrooms where appropriate. The benefits of AIEd can benefit teachers and students if implemented appropriately in traditional educational platforms.

AIEd has the potential to support education and learning significantly. However, it must be grounded in educational theory to ensure that the technology has the maximum impact on learning [84]. It should also consider pedagogical, social, cultural, and economic aspects [84]. Furthermore, many existing applications explored in this article are not integrated with traditional classrooms, thus limiting their impact on mainstream education. Therefore, future AIEd should be developed in collaboration with teachers, other stakeholders, and policymakers to ensure its relevance in classrooms [100]. Further efforts should be made to integrate other classroom systems, such as attendance and attainment monitoring, to provide socially responsible data-driven approaches to AIEd.

Furthermore, to the best of the authors' knowledge, the current literature does not address how models are updated over time to ensure that a student's knowledge remains accurately mapped. For example, how would it deal with students forgetting information that they had previously learned? Moreover, [20] suggests that AIEd prediction could be expanded to include further essential applications, for instance, by predicting assignment deferments, late completions, dropouts, and narrow passes [20].

Another significant issue noted throughout this article is the current reliance on deep learning methods. Such methods are often described as "black box" because it is difficult for humans to understand how the models arrive at their conclusions. Where AIEd predicts important factors such as students' grades or the likelihood of them dropping out, it is crucial that humans can check and verify such decisions. However, this is not possible with current "black box" methods. Therefore, future efforts should be made to ensure that future AIEd applications are explainable. The literature describes various methods of producing explainable AI, for example, "white box" models such as linear models and decision trees [29]. Further methods of promoting explainability are described in [29].

A related issue is the lack of diversity of available AIEd datasets regarding the subjects and educational aspects provided. The majority of the datasets provide only science-based subjects. Similarly, the majority of datasets are designed for educational aspects such as performance prediction and knowledge mapping tasks. However, material curation and intelligent tutoring have received little interest in terms of datasets. The lack of such datasets will severely limit the projects that researchers can conduct. Therefore, efforts should be made to collect, prepare, and share datasets to fill such gaps.

5 SUMMARY

From the recent literature, several modern issues with AIEd and similar applications have been identified to help minimise their potential effects. Despite the plethora of helpful applications described, some aspects of the current research may present problems. For example, many recent works, such as [20, 25, 101, 113], are implemented using deep learning. Deep learning is a powerful concept shown in many domains to produce superior performance. However, such methods are not explainable, making it difficult to explain or justify the results generated [117]. Therefore, more explainable approaches may present an opportunity to provide early intervention. For example, if a reason for a student's dropout or poor performance is provided, the cause may be acted on to prevent it. Similarly, there is a growing trend in using LLMs, for example, in works such as [59, 88, 102]. Such works demonstrate the range of applications that LLMs can achieve; they are vulnerable to hallucinations and may present fabricated information in their answers [4]. An alternative approach which may present promising avenues for future research is SLMs. SLMs have been shown to produce high performance in limited domains; however, as they are trained on smaller datasets, they often produce poor grammar and writing [113]. Therefore, improving the written language level in SLMs could benefit AIEd. For example, they may be particularly beneficial for subject-specific applications, reducing the potential for hallucination and providing superior subject-specific performance compared to LLMs.

Moreover, this research highlighted several aspects of AIEd and education that have not received sufficient attention in recent years. For example, AIEd is not being developed and integrated into mainstream classrooms; instead, the majority of applications focus on online education. This is a missed opportunity for widespread impact. Similarly, it may help lessen the workload of teachers, many of whom are planning to leave the profession in the near future because of their workload. Similarly, applications focused on administration tasks and intelligent tutors also received little attention. Such applications may again lessen teachers' workloads but also provide opportunities to expand high-quality education to those who may not have access to it.

Similarly, we also addressed the growing and evolving issue of academic misconduct, particularly that related to student misuse of GAI in assessment. It is evident from the literature that current detection and prevention mechanisms are not adequate, and further measures are required to ensure academic integrity. This developing and multifaceted problem requires a cohesive methodology that blends technological advances, behavioural analysis, and forward-thinking pedagogical reform.

Furthermore, this work has also provided a range of recommendations for future applications to ensure fairness and to maximise the impact of such work in the future. However, despite these contributions, there is potential for future work. For example, as suggested in [66], a prospective meta-analysis of AIEd may provide more critical information, which may assist in shaping future AIEd platforms. Furthermore, to the best of the authors' knowledge, no current survey in the literature focuses on AIEd purely from a classroom-based teacher's perspective. Moreover, as many AIEd platforms are aimed at distance learners, a future study regarding the usefulness of avatars for intelligent tutors may be helpful for future AIEd research and platforms. As highlighted in this work, there is an imbalance in the focus of AIEd in traditional and online classrooms. Therefore, such a review explaining potential barriers to future development would be advantageous.

Despite the contributions of this research and the authors' best efforts, some potential limitations of the study remain. For example, as described in 2, only research published within the past five years was included for the reasons described and justified earlier. Despite the important justification for this range, it may have unforeseen consequences. For example, it may have affected the range of applications and topics within the identified literature and their relevant appearance frequencies. For example, the use of LLMs was prevalent in this period. Therefore, some readers may wish to consult reviews with more historical AIEd content. Additionally, this research does not consider whether such AIEd methods would be accepted by students, teachers, and other relevant stakeholders. If the technology is not accepted, then it cannot have an impact. Therefore, future work should investigate the needs and wants of potential users and identify methods for increasing the acceptance of AIEd.

6 CONCLUSION

This study has surveyed recent works in AIEd, providing a taxonomy of applications, an overview of available datasets, and insights into emerging trends, including the use of deep learning, LLMs, and SLMs. While these approaches offer significant potential, challenges such as a lack of explainability, risk of hallucinations, and limited integration into traditional classrooms remain. The review also highlights underexplored areas, including classroom-based teacher support, administrative applications, and strategies to mitigate academic misconduct. Future research should address these gaps, improve model interpretability and language quality, and consider user acceptance to ensure that AIEd applications are both effective and widely adopted, ultimately supporting equitable, high-quality education.

REFERENCES

[1] Ikpe Justice Akpan, Yawo M. Kobara, Josiah Owolabi, Asuama A. Akpan, and Onyebuchi Felix Offodile. 2025. Conversational and Generative Artificial Intelligence and Human–Chatbot Interaction in Education and Research. *International Transactions in Operational Research* 32, 3 (2025), 1251–1281. https://doi.org/10.1111/itor.13522

[2] Ashraf Alam. 2021. Possibilities and apprehensions in the landscape of artificial intelligence in education. In *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*. IEEE, 1–8.

[3] P Atherton, L Topham, and W Khan. 2024. AI and student feedback. *EDULEARN24 Proceedings* (2024).

[4] Razvan Azamfirei, Sapna R Kudchadkar, and James Fackler. 2023. Large language models and the perils of their hallucinations. *Critical Care* 27, 1 (2023), 120.

[5] Richa Bajaj and Vidushi Sharma. 2018. Smart Education with artificial intelligence based determination of learning styles. *Procedia computer science* 132 (2018), 834–842.

[6] Ryan S Baker and Aaron Hawn. 2022. Algorithmic bias in education. *International Journal of Artificial Intelligence in Education* (2022), 1–41.

[7] Elisa Bertino, Shawn Merrill, Alina Nesen, and Christine Utz. 2019. Redefining data transparency: A multidimensional approach. *Computer* 52, 1 (2019), 16–26.

[8] Gernot Beutel, Eline Geerits, and Jan T Kielstein. 2023. Artificial hallucination: GPT on LSD? *Critical Care* 27, 1 (2023), 148.

[9] Shravya Bhat, Huy A Nguyen, Steven Moore, John Stamper, Majd Sakr, and Eric Nyberg. 2022. Towards automated generation and evaluation of questions in educational domains. In *Proceedings of the 15th international conference on educational data mining*, Vol. 701.

[10] Michal Bobula. 2024. Generative Artificial Intelligence (AI) in Higher Education: A Comprehensive Review of Challenges, Opportunities, and Implications. *Journal of Learning Development in Higher Education* 30 (March 2024). https://doi.org/10.47408/jldhe.vi30.1137

[11] Anderson Pinheiro Cavalcanti, Arthur Barbosa, Ruan Carvalho, Fred Freitas, Yi-Shan Tsai, Dragan Gašević, and Rafael Ferreira Mello. 2021. Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence* 2 (2021), 100027.

[12] Haw-Shiuan Chang, Hwai-Jung Hsu, Kuan-Ta Chen, et al. 2015. Modeling Exercise Relationships in E-Learning: A Unified Approach.. In *EDM*. 532–535.

[13] Maud Chassignol, Aleksandr Khoroshavin, Alexandra Klimova, and Anna Bilyatdinova. 2018. Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science* 136 (2018), 16–24.

[14] Chih-Hung Chen, Chorng-Shiuh Koong, and Chien Liao. 2022. Influences of integrating dynamic assessment into a speech recognition learning design to support students' English speaking skills, learning anxiety and cognitive load. *Educational Technology & Society* 25, 1 (2022), 1–14.

[15] Lijia Chen, Pingping Chen, and Zhijian Lin. 2020. Artificial intelligence in education: A review. *Ieee Access* 8 (2020), 75264–75278.

[16] Xieling Chen, Haoran Xie, Di Zou, and Gwo-Jen Hwang. 2020. Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence* 1 (2020), 100002.

[17] Xieling Chen, Di Zou, Haoran Xie, and Gary Cheng. 2021. Twenty years of personalized language learning. *Educational Technology & Society* 24, 1 (2021), 205–222.

[18] Yu Chen, Scott Jensen, Leslie J Albert, Sambhav Gupta, and Terri Lee. 2023. Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers* 25, 1 (2023), 161–182.

[19] Youngduck Choi, Youngnam Lee, Dongmin Shin, Junghyun Cho, Seoyon Park, Seewoo Lee, Jineon Baek, Chan Bae, Byungsoo Kim, and Jaewe Heo. 2020. Ednet: A large-scale hierarchical dataset in education. In *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II 21*. Springer, 69–73.

[20] Kwok Tai Chui, Dennis Chun Lok Fung, Miltiadis D Lytras, and Tin Miu Lam. 2020. Predicting at-risk university students in a virtual learning environment via a machine learning algorithm. *Computers in Human behavior* 107 (2020), 105584.

[21] Khalid Colchester, Hani Hagras, Daniyal Alghazzawi, and Ghadah Aldabbagh. 2017. A survey of artificial intelligence techniques employed for adaptive educational systems within e-learning platforms. *Journal of Artificial Intelligence and Soft Computing Research* 7, 1 (2017), 47–64.

[22] Roberto Confalonieri, Ludovik Coba, Benedikt Wagner, and Tarek R Besold. 2021. A historical perspective of explainable Artificial Intelligence. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 11, 1 (2021), e1391.

[23] Debby R. E. Cotton, Cotton , Peter A., and J. Reuben and Shipway. 2024. Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT. *Innovations in Education and Teaching International* 61, 2 (March 2024), 228–239. https://doi.org/10.1080/14703297.2023.2190148

[24] Dale Crowe, Martin LaPierre, and Mansureh Kebritchi. 2017. Knowledge based artificial augmentation intelligence technology: Next step in academic instructional tools for distance learning. *TechTrends* 61, 5 (2017), 494–506.

[25] Francesca Del Bonifro, Maurizio Gabbrielli, Giuseppe Lisanti, and Stefano Pio Zingaro. 2020. Student dropout prediction. In *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part I 21*. Springer, 129–140.

[26] Enrique Delahoz-Dominguez, Rohemi Zuluaga, and Tomas Fontalvo-Herrera. 2020. Dataset of academic performance evolution for engineering students. *Data in brief* 30 (2020), 105537.

[27] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[28] Chaitali Diwan, Srinath Srinivasa, Gandharv Suri, Saksham Agarwal, and Prasad Ram. 2023. AI-based learning content generation and learning pathway augmentation to increase learner engagement. *Computers and Education: Artificial Intelligence* 4 (2023), 100110.

[29] Rudresh Dwivedi, Devam Dave, Het Naik, Smiti Singhal, Rana Omer, Pankesh Patel, Bin Qian, Zhenyu Wen, Tejal Shah, Graham Morgan, et al. 2023. Explainable AI (XAI): Core ideas, techniques, and solutions. *Comput. Surveys* 55, 9 (2023), 1–33.

[30] Yogesh K. Dwivedi, Nir Kshetri, Laurie Hughes, Emma Louise Slade, Anand Jeyaraj, Arpan Kumar Kar, Abdullah M. Baabdullah, Alex Koohang, Vishnupriya Raghavan, Manju Ahuja, Hanaa Albanna, Mousa Ahmad Albashrawi, Adil S. Al-Busaidi, Janarthanan Balakrishnan, Yves Barlette, Sriparna Basu, Indranil Bose, Laurence Brooks, Dimitrios Buhalis, Lemuria Carter, Soumyadeb Chowdhury, Tom Crick, Scott W. Cunningham, Gareth H. Davies, Robert M. Davison, Rahul Dé, Denis Dennehy, Yanqing Duan, Rameshwar Dubey, Rohita Dwivedi, John S. Edwards, Carlos Flavián, Robin Gauld, Varun Grover, Mei-Chih Hu, Marijn Janssen, Paul Jones, Iris Junglas, Sangeeta Khorana, Sascha Kraus, Kai R. Larsen, Paul Latreille, Sven Laumer, F. Tegwen Malik, Abbas Mardani, Marcello Mariani, Sunil Mithas, Emmanuel Mogaji, Jeretta Horn Nord, Siobhan O'Connor, Fevzi Okumus, Margherita Pagani, Neeraj Pandey, Savvas Papagiannidis, Ilias O. Pappas, Nishith Pathak, Jan Pries-Heje, Ramakrishnan

Raman, Nripendra P. Rana, Sven-Volker Rehm, Samuel Ribeiro-Navarrete, Alexander Richter, Frantz Rowe, Suprateek Sarker, Bernd Carsten Stahl, Manoj Kumar Tiwari, Wil van der Aalst, Viswanath Venkatesh, Giampaolo Viglia, Michael Wade, Paul Walton, Jochen Wirtz, and Ryan Wright. 2023. Opinion Paper: "So What If ChatGPT Wrote It?" Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research, Practice and Policy. *International Journal of Information Management* 71 (Aug. 2023), 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

[31] Damian Okaibedi Eke. 2023. ChatGPT and the Rise of Generative AI: Threat to Academic Integrity? *Journal of Responsible Technology* 13 (April 2023), 100060. https://doi.org/10.1016/j.jrt.2023.100060

[32] Ahmed M. Elkhatat, Khaled Elsaid, and Saeed Almeer. 2023. Evaluating the Efficacy of AI Content Detection Tools in Differentiating between Human and AI-generated Text. *International Journal for Educational Integrity* 19, 1 (Sept. 2023), 17. https://doi.org/10.1007/s40979-023-00140-5

[33] Katherine Elkins and Jon Chun. 2020. Can GPT-3 pass a writer's Turing test? *Journal of Cultural Analytics* 5, 2 (2020).

[34] Mingyu Feng, Neil Heffernan, and Kenneth Koedinger. 2009. Addressing the assessment challenge with an online system that tutors as it assesses. *User modeling and user-adapted interaction* 19 (2009), 243–266.

[35] Luciano Floridi and Massimo Chiriatti. 2020. GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines* 30 (2020), 681–694.

[36] José A Gómez-Tejedor, Ana Vidaurre, Isabel Tort-Ausina, José Molina-Mateo, María-Antonia Serrano, José M Meseguer-Dueñas, Rosa M Martínez Sala, Susana Quiles, and Jaime Riera. 2020. Effectiveness of flip teaching on engineering students' performance in the physics lab. *Computers & Education* 144 (2020), 103708.

[37] Simone Grassini. 2023. Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education sciences* 13, 7 (2023), 692.

[38] Zhen Guo, Peiqi Wang, Yanwei Wang, and Shangdi Yu. 2023. Dr. llama: Improving small language models on pubmedqa via generative data augmentation. *arXiv preprint arXiv:2305.07804* (2023).

[39] Farrokh Habibzadeh. 2023. GPTZero Performance in Identifying Artificial Intelligence-Generated Medical Texts: A Preliminary Study. *Journal of Korean Medical Science* 38, 38 (Sept. 2023), e319. https://doi.org/10.3346/jkms.2023.38.e319

[40] Alaa Hamoud, Ali Salah Hashim, and Wid Akeel Awadh. 2018. Predicting student performance in higher education institutions using decision tree analysis. *International Journal of Interactive Multimedia and Artificial Intelligence* 5 (2018), 26–31.

[41] NURIA Haristiani, AA Danuwijaya, MM Rifai, and H Sarila. 2019. Gengobot: A chatbot-based grammar application on mobile instant messaging as language learning medium. *Journal of Engineering Science and Technology* 14, 6 (2019), 3158–3173.

[42] Nuria Haristiani and Mumu Muhamad Rifa'i. 2020. Combining chatbot and social media: Enhancing personal learning environment (PLE) in language learning. *Indonesian Journal of Science and Technology* 5, 3 (2020), 487–506.

[43] Raza Hasan, Sellappan Palaniappan, Salman Mahmood, Ali Abbas, Kamal Uddin Sarker, and Mian Usman Sattar. 2020. Predicting student performance in higher educational institutions using video learning analytics and data mining techniques. *Applied Sciences* 10, 11 (2020), 3894.

[44] Charles B. Hodges and Paul A. Kirschner. 2024. Innovation of Instructional Design and Assessment in the Age of Generative Artificial Intelligence. *TechTrends* 68, 1 (Jan. 2024), 195–199. https://doi.org/10.1007/s11528-023-00926-x

[45] Wayne Holmes, Kaska Porayska-Pomsta, Ken Holstein, Emma Sutherland, Toby Baker, Simon Buckingham Shum, Olga C Santos, Mercedes T Rodrigo, Mutlu Cukurova, Ig Ibert Bittencourt, et al. 2022. Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education* (2022), 1–23.

[46] Yvonne Hong, Andy Nguyen, Belle Dang, and Bich-Phuong Thi Nguyen. 2022. Data Ethics Framework for Artificial Intelligence in Education (AIED). In *2022 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 297–301.

[47] Sadiq Hussain, Neama Abdulaziz Dahan, Fadl Mutaher Ba-Alwib, and Najoua Ribata. 2018. Educational data mining and analysis of students' academic performance using WEKA. *Indonesian Journal of Electrical Engineering and Computer Science* 9, 2 (2018), 447–459.

[48] Kevin Maik Jablonka, Philippe Schwaller, and Berend Smit. 2022. Is GPT-3 all you need for machine learning for chemistry?. In *AI for Accelerated Materials Design NeurIPS 2022 Workshop*.

[49] Adeeb M. Jarrah, Yousef Wardat, and Patricia Fidalgo. 2023. Using ChatGPT in Academic Writing Is (Not) a Form of Plagiarism: What Does the Literature Say? *Online Journal of Communication and Media Technologies* 13, 4 (Oct. 2023), e202346. https://doi.org/10.30935/ojcmt/13572

[50] Heather Johnston, Rebecca F. Wells, Elizabeth M. Shanks, Timothy Boey, and Bryony N. Parsons. 2024. Student Perspectives on the Use of Generative Artificial Intelligence Technologies in Higher Education. *International Journal for Educational Integrity* 20, 1 (Dec. 2024), 1–21. https://doi.org/10.1007/s40979-024-00149-4

[51] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences* 103 (2023), 102274.

[52] Zenun Kastrati, Arianit Kurti, and Ali Shariq Imran. 2020. WET: Word embedding-topic distribution vectors for MOOC video lectures dataset. *Data in brief* 28 (2020), 105090.

[53] Wasiq Khan, Luke K Topham, Peter Atherton, Raghad Al-Shabandar, Hoshang Kolivand, Iftikhar Khan, and Abir Hussain. 2024. Auto-assessment of assessment: A conceptual framework towards fulfilling the policy gaps in academic assessment practices. *arXiv preprint arXiv:2411.08892* (2024).

[54] Jinhee Kim. 2023. Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies* (2023), 1–32.

[55] Marius Kloft, Felix Stiehler, Zhilin Zheng, and Niels Pinkwart. 2014. Predicting MOOC dropout over weeks using machine learning methods. In *Proceedings of the EMNLP 2014 workshop on analysis of large scale social interaction in MOOCs*. 60–65.

[56] Kenneth R Koedinger, Ryan SJd Baker, Kyle Cunningham, Alida Skogsholm, Brett Leber, and John Stamper. 2010. A data repository for the EDM community: The PSLC DataShop. *Handbook of educational data mining* 43 (2010), 43–56.

[57] Elly A Konijn and Johan F Hoorn. 2020. Robot tutor and pupils' educational ability: Teaching the times tables. *Computers & Education* 157 (2020), 103970.

[58] Stephan Krusche and Andreas Seitz. 2018. Artemis: An automatic assessment management system for interactive learning. In *Proceedings of the 49th ACM technical symposium on computer science education*. 284–289.

[59] Tiffany H Kung, Morgan Cheatham, Arielle Medenilla, Czarina Sillos, Lorie De Leon, Camille Elepaño, Maria Madriaga, Rimel Aggabao, Giezel Diaz-Candido, James Maningo, et al. 2023. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLoS digital health* 2, 2 (2023), e0000198.

[60] Jakub Kuzilek, Martin Hlosta, and Zdenek Zdrahal. 2017. Open university learning analytics dataset. *Scientific data* 4, 1 (2017), 1–8.

[61] Alwyn Vwen Yen Lee. 2023. Supporting students' generation of feedback in large-scale online course with artificial intelligence-enabled evaluation. *Studies in Educational Evaluation* 77 (2023), 101250.

[62] Echo Shang Liang and Shurui and Bai. 2024. Generative AI and the Future of Connectivist Learning in Higher Education. *Journal of Asian Public Policy* 0, 0 (2024), 1–23. https://doi.org/10.1080/17516234.2024.2386085

[63] Weng Marc Lim, Asanka Gunasekara, Jessica Leigh Pallant, Jason Ian Pallant, and Ekaterina Pechenkina. 2023. Generative AI and the Future of Education: Ragnarök or Reformation? A Paradoxical Perspective from Management Educators. *The International Journal of Management Education* 21, 2 (July 2023), 100790. https://doi.org/10.1016/j.ijme.2023.100790

[64] Jionghao Lin, Wei Tan, Lan Du, Wray Buntine, David Lang, Dragan Gašević, and Guanliang Chen. 2023. Enhancing educational dialogue act classification with discourse context and sample informativeness. *IEEE Transactions on Learning Technologies* 17 (2023), 258–269.

[65] Ming Liu, Jingxu Zhang, Lucy Michael Nyagoga, and Li Liu. 2023. Student-AI question cocreation for enhancing reading comprehension. *IEEE Transactions on Learning Technologies* 17 (2023), 815–826.

[66] Hui Luan and Chin-Chung Tsai. 2021. A review of using machine learning approaches for precision education. *Educational Technology & Society* 24, 1 (2021), 250–266.

[67] Rosemary Luckin and Mutlu Cukurova. 2019. Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology* 50, 6 (2019), 2824–2838.

[68] Stephen MacNeil, Andrew Tran, Dan Mogil, Seth Bernstein, Erin Ross, and Ziheng Huang. 2022. Generating diverse code explanations using the gpt-3 large language model. In *Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 2*. 37–39.

[69] Nahema Marchal, Rachel Xu, Rasmi Elasmar, Iason Gabriel, Beth Goldberg, and William Isaac. 2024. Generative AI Misuse: A Taxonomy of Tactics and Insights from Real-World Data. https://arxiv.org/abs/2406.13843v2.

[70] Frank Marcinkowski, Kimon Kieslich, Christopher Starke, and Marco Lünich. 2020. Implications of AI (un-) fairness in higher education admissions: the effects of perceived AI (un-) fairness on exit, voice and organizational reputation. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 122–130.

[71] David T Marshall, Tim Pressley, Natalie M Neugebauer, and David M Shannon. 2022. Why teachers are leaving and what we can do about it. *Phi Delta Kappan* 104, 1 (2022), 6–11.

[72] Nora McDonald, Aditya Johri, Areej Ali, and Aayushi Hingle Collier. 2025. Generative Artificial Intelligence in Higher Education: Evidence from an Analysis of Institutional Policies and Guidelines. *Computers in Human Behavior: Artificial Humans* 3 (March 2025), 100121. https://doi.org/10.1016/j.chbah.2025.100121

[73] Sein Minn. 2022. AI-assisted knowledge assessment techniques for adaptive learning environments. *Computers and Education: Artificial Intelligence* 3 (2022), 100050.

[74] Sein Minn, Jill-Jênn Vie, Koh Takeuchi, Hisashi Kashima, and Feida Zhu. 2022. Interpretable knowledge tracing: Simple and efficient student modeling with causal relations. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 36. 12810–12818.

[75] Eric Mitchell, Yoonho Lee, Alexander Khazatsky, Christopher D. Manning, and Chelsea Finn. 2023. DetectGPT: Zero-Shot Machine-Generated Text Detection Using Probability Curvature. In *Proceedings of the 40th International Conference on Machine Learning (ICML'23, Vol. 202)*. JMLR.org, Honolulu, Hawaii, USA, 24950–24962.

[76] Ahmed Taiye Mohammed, Johanna Velander, and Marcelo Milrad. 2024. A Retrospective Analysis of Artificial Intelligence in Education (AIEd) Studies: Perspectives, Learning Theories, Challenges, and Emerging Opportunities. In *Radical Solutions for Artificial Intelligence and Digital Transformation in Education: Utilising Disruptive Technology for a Better Society*. Springer, 127–141.

[77] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, Prisma Group, et al. 2010. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International journal of surgery* 8, 5 (2010), 336–341.

[78] György Molnár and Zoltán Szüts. 2018. The role of chatbots in formal education. In *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)*. IEEE, 000197–000202.

[79] Steven Moore, Huy A Nguyen, Norman Bier, Tanvi Domadia, and John Stamper. 2022. Assessing the quality of student-generated short answer questions using GPT-3. In *European conference on technology enhanced learning*. Springer, 243–257.

[80] Neelesh Mungoli. 2023. Exploring the Advancements and Implications of Artificial Intelligence. (2023).

[81] Shuanghong Jenny Niu, Jiutong Luo, Hannele Niemi, Xiaoqing Li, and Yu Lu. 2022. Teachers' and students' views of using an AI-aided educational platform for supporting teaching and learning at Chinese schools. *Education Sciences* 12, 12 (2022), 858.

[82] Faruk Enes Oguz, Mahmut Nedim Ekersular, Kubilay Muhammed Sunnetci, and Ahmet Alkan. 2023. Can Chat GPT be Utilized in Scientific and Undergraduate Studies? *Annals of Biomedical Engineering* (2023), 1–3.

[83] OpenAI. 2024. New AI Classifier for Indicating AI-written Text. https://openai.com/index/new-ai-classifier-for-indicating-ai-written-text/.

[84] Fan Ouyang and Pengcheng Jiao. 2021. Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence* 2 (2021), 100020.

[85] Fan Ouyang, Mian Wu, Luyi Zheng, Liyin Zhang, and Pengcheng Jiao. 2023. Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education* 20, 1 (2023), 4.

[86] Fan Ouyang, Luyi Zheng, and Pengcheng Jiao. 2022. Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies* 27, 6 (2022), 7893–7925.

[87] Zachary A Pardos, Ryan SJD Baker, Maria OCZ San Pedro, Sujith M Gowda, and Supreeth M Gowda. 2013. Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the third international conference on learning analytics and knowledge*. 117–124.

[88] WooHyun Park, Dojin Park, ByungJune Ahn, SeokHun Kang, HaengYeong Kim, RaeHyeon Kim, and JaeHeum Na. 2019. Interactive AI for linguistic education built on VR environment using user generated contents. In *2019 21st International Conference on advanced communication technology (ICACT)*. IEEE, 385–389.

[89] European Parliament. 2022. *Artificial Intelligence in Education, Culture and the Audiovisual Sector*. https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_EN.html

[90] Ying Pei and Gang Li. 2021. Massive AI based cloud environment for smart online education with data mining. In *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, 1115–1118.

[91] Filipe D Pereira, Elaine Oliveira, Alexandra Cristea, David Fernandes, Luciano Silva, Gene Aguiar, Ahmed Alamri, and Mohammad Alshehri. 2019. Early dropout prediction for programming courses supported by online judges. In *Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part II 20*. Springer, 67–72.

[92] Mike Perkins, Leon Furze, Jasper Roe, and Jason MacVaugh. 2024. The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment. *Journal of University Teaching and Learning Practice* 21, 06 (April 2024). https://doi.org/10.53761/q3azde36

[93] Mike Perkins, Jasper Roe, Darius Postma, James McGaughran, and Don Hickerson. 2024. Detection of GPT-4 Generated Text in Higher Education: Combining Academic Judgement and Software to Identify Generative AI Tool Misuse. *Journal of Academic Ethics* 22, 1 (March 2024), 89–113. https://doi.org/10.1007/s10805-023-09492-6

[94] Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J Guibas, and Jascha Sohl-Dickstein. 2015. Deep knowledge tracing. *Advances in neural information processing systems* 28 (2015).

[95] Silvia Pokrivcakova. 2019. Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education* 7, 3 (2019), 135–153.

[96] Ido Roll and Ruth Wylie. 2016. Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education* 26 (2016), 582–599.

[97] Carolyn Rose and George Siemens. 2014. Shared task on prediction of dropout over time in massively open online courses. In *Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs*. 39–41.

[98] Stuart J Russell and Peter Norvig. 2016. *Artificial intelligence: a modern approach*. pearson.

[99] Muhammad Fakhrul Safitra, Muharman Lubis, Tien Fabrianti Kusumasari, and Deyana Prastika Putri. 2024. Advancements in Artificial Intelligence and Data Science: Models, Applications, and Challenges. *Procedia Computer Science* 234 (2024), 381–388.

[100] Daniel Schiff. 2021. Out of the laboratory and into the classroom: the future of artificial intelligence in education. *AI & society* 36, 1 (2021), 331–348.

[101] Boran Sekeroglu, Kamil Dimililer, and Kubra Tuncal. 2019. Student performance prediction and classification using machine learning algorithms. In *Proceedings of the 2019 8th international conference on educational and information technology*. 7–11.

[102] Umut Tan Sevgi, Gökberk Erol, Yücel Doğruel, Osman Fikret Sönmez, Richard Shane Tubbs, and Abuzer Güngor. 2023. The role of an open artificial intelligence platform in modern neurosurgical education: a preliminary study. *Neurosurgical review* 46, 1 (2023), 86.

[103] Carlton Shepherd. 2025. Generative AI Misuse Potential in Cyber Security Education: A Case Study of a UK Degree Program. https://doi.org/10.48550/arXiv.2501.12883 arXiv:2501.12883 [cs]

[104] Ramesh Subramanian. 2017. Emergent AI, social robots and the law: Security, privacy and policy issues. *Subramanian, Ramesh (2017)" Emergent AI, Social Robots and the Law: Security, Privacy and Policy Issues," Journal of International, Technology and Information Management* 26, 3 (2017).

[105] Zhuomin Sun, M Anbarasan, and DJCI Praveen Kumar. 2021. Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence* 37, 3 (2021), 1166–1180.

[106] André Ullrich, Gergana Vladova, Felix Eigelshoven, and André Renz. 2022. Data mining of scientific research on artificial intelligence in teaching and administration in higher education institutions: A bibliometrics analysis and recommendation for future research. *Discover Artificial Intelligence* 2, 1 (2022), 16.

[107] Mehrnoosh Vahdat, Luca Oneto, Davide Anguita, Mathias Funk, and Matthias Rauterberg. 2015. A learning analytics approach to correlate the academic achievements of students with interaction data from an educational simulator. In *Design for Teaching and Learning in a Networked World:*

10th European Conference on Technology Enhanced Learning, EC-TEL 2015, Toledo, Spain, September 15-18, 2015, Proceedings 10. Springer, 352–366.

[108] Jill-Jênn Vie and Hisashi Kashima. 2019. Knowledge tracing machines: Factorization machines for knowledge tracing. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 750–757.

[109] Krzysztof Wach, Cong Doanh Duong, Joanna Ejdys, Rūta Kazlauskaitė, Pawel Korzynski, Grzegorz Mazurek, Joanna Paliszkiewicz, and Ewa Ziemba. 2023. The Dark Side of Generative Artificial Intelligence: A Critical Analysis of Controversies and Risks of ChatGPT. *Entrepreneurial Business and Economics Review* 11, 2 (June 2023), 7–30. https://doi.org/10.15678/EBER.2023.110201

[110] Hui Wang, Anh Dang, Zihao Wu, and Son Mac. 2024. Generative AI in Higher Education: Seeing ChatGPT through Universities' Policies, Resources, and Guidelines. *Computers and Education: Artificial Intelligence* 7 (Dec. 2024), 100326. https://doi.org/10.1016/j.caeai.2024.100326

[111] Jingxian Wang, Dineke EH Tigelaar, and Wilfried Admiraal. 2021. Rural teachers' sharing of digital educational resources: From motivation to behavior. *Computers & Education* 161 (2021), 104055.

[112] Tianqi Wang, Fenglong Ma, and Jing Gao. 2019. Deep hierarchical knowledge tracing. In *Proceedings of the 12th international conference on educational data mining*.

[113] Yueqian Wang, Chang Liu, Kai Chen, Xi Wang, and Dongyan Zhao. 2022. SMASH: Improving SMAll Language Models' Few-SHot Ability with Prompt-Based Distillation. In *Findings of the Association for Computational Linguistics: EMNLP 2022*. 6608–6619.

[114] Debora Weber-Wulff, Alla Anohina-Naumeca, Sonja Bjelobaba, Tomáš Foltýnek, Jean Guerrero-Dib, Olumide Popoola, Petr Šigut, and Lorna Waddington. 2023. Testing of Detection Tools for AI-generated Text. *International Journal for Educational Integrity* 19, 1 (Dec. 2023), 1–39. https://doi.org/10.1007/s40979-023-00146-z

[115] Ryan Thomas Williams. 2024. The Ethical Implications of Using Generative Chatbots in Higher Education. *Frontiers in Education* 8 (Jan. 2024). https://doi.org/10.3389/feduc.2023.1331607

[116] Haoran Xie, Hui-Chun Chu, Gwo-Jen Hwang, and Chun-Chieh Wang. 2019. Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education* 140 (2019), 103599.

[117] Feiyu Xu, Hans Uszkoreit, Yangzhou Du, Wei Fan, Dongyan Zhao, and Jun Zhu. 2019. Explainable AI: A brief survey on history, research areas, approaches and challenges. In *Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II 8*. Springer, 563–574.

[118] Shanshan Yang and Kim Stansfield. 2022. AI Chatbot for Educational Service Improvement in the Post-Pandemic Era: A Case Study Prototype for Supporting Digital Reading List. In *Proceedings of the 2022 13th International Conference on E-Education, E-Business, E-Management, and E-Learning*. 24–29.

[119] Robert Zemsky. 2014. With a MOOC MOOC here and a MOOC MOOC there, here a MOOC, there a MOOC, everywhere a MOOC MOOC. *The Journal of General Education* 63, 4 (2014), 237–243.

[120] Ke Zhang and Ayse Begum Aslan. 2021. AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence* 2 (2021), 100025.

Received xyz; revised xyz; accepted xyz