Education and Information Technologies
https://doi.org/10.1007/s10639-024-12491-8

Artificial intelligence in education research during 2013–2023: A review based on bibliometric analysis

Shuchen Guo[1] · **Yuanyuan Zheng**[1] · **Xiaoming Zhai**[2]

Received: 19 October 2023 / Accepted: 16 January 2024

Abstract

Research on Artificial Intelligence in Education (AIED) has rapidly progressed in recent years, and understanding the research trends and development is essential for technological innovations and implementations in education. Using a bibliometric analysis of 6843 publications from Web of Science and Scopus, we found that China, US, India, Spain, and Germany led the research profuctivity. AIED research is concerned more with higher education compared to K-12 education. Fifteen research trends emerged from the analysis, such as Educational Robots and Large Data Mining. Research has primarily leveraged technologies of machine learning, decision trees, deep learning, speech recognition, and computer vision in AIED. The major implementations of AI include educational robots, automated grading, recommender systems, learning analytics, and intelligent tutoring systems. Among the implementations, a majority of AIED research was conducted in seven major subject domains, chief among them being science, technology, engineering and mathematics (STEM) and language disciplines, with a focus on computer science and English education.

Keywords Artificial intelligence (AI) · AI in education (AIED) · Machine learning (ML) · Bibliometric analysis · Citespace software

✉ Xiaoming Zhai
Xiaoming.Zhai@uga.edu

Shuchen Guo
gsc44@njnu.edu.cn

Yuanyuan Zheng
3048835050@qq.com

[1] School of Education, Nanjing Normal University, Nanjing, Jiangsu, China

[2] AI4STEM Education Center & Department of Mathematics, Science, and Social Studies Education, University of Georgia, 125M Aderhold Hall, 110 Carlton St, Athens, GA 30602, Georgia

Published online: 09 February 2024

1 Introduction

The Industrial Revolution 4.0 has promoted the rapid development of information technologies, of which Artificial Intelligence (AI) is the most important aspect. AI breakthroughs such as Large Language Models and ChatGPT have emerged and widely impacted all sectors of society, including education (Lee et al., 2023; Peters et al., 2023; Zhai, 2023). Research on AI in education (AIED), such as intelligent tutoring systems, automatic scoring, sentiment analysis, etc., has taken a substantial strike in recent years (Chen et al., 2022; Prahani et al., 2022). These developments in AI have provided growing opportunities to improve the quality and efficiency of education (Chen et al., 2022; Zhang & Aslan, 2021), with the potential to transform future education.

Although education has been regarded as one of the most fundamental and critical aspects of social development, technological implementations have always been lagging compared to other areas, such as science and medical treatment (Zhai et al., 2020). For example, when computers were first introduced to education, researchers and educators were excited about and embracing the new opportunity that might revolutionize how classroom learning and teaching happen. However, the actual impact took way more time compared with other areas, and indicating a long way to go. This may be because education is such a complex system, and technology itself hardly makes substantial changes without companion professional learning, development of materials, innovative pedagogies, etc. The practical impacts rely on educators and how they perceive and manifest the uses of technologies with matching pedagogies. In this sense, examining how innovative technologies such as AI evolve could inform the research and practices on developing new pedagogy and materials to meet educational needs.

This study thus employed a bibliography analysis approach focusing on 6843 publications in the past ten years to identify the trends of AIED research and how the field evolves. Bibliography analysis is a data-driven approach that allows for the large-scale analysis of publications, thereby offering a more comprehensive and global overview of existing literature. Unlike traditional reviews, which are often limited by subjective interpretation, bibliometric analysis employs quantitative metrics—such as citation counts, h-index, and impact factors—to provide an objective assessment of academic impact and relevance. It can serve as a robust complement to traditional descriptive literature review methods, offering several distinct advantages that enhance the rigor and breadth of scholarly inquiry. We focused on the past ten years due to several concerns. First, research on AIED experienced a sharp increase since 2013, as indicated in a review study by Zawacki-Richter et al. (2019) who tracked the publications of AIED from 2007 and visualized the trends in a diagram. Second, Li et al. (2020) reported a dramatic increase on the applications of key AI technologies such as deep learning in education since 2013. Given this information, we believe a bibliography analysis of the literature is timely and essential to synthesize the trends of research and identify the research gaps. In the research, we asked three questions: (1) How diverse has AIED research been in the past ten years in terms of users, subject domains, and author geographics? (2) What are the research trends of AIED regarding technology, applications, and subject domain outcomes?

1.1 AI in education

AIED has been established as a cohesive academic research field since the 1980s (Williamson et al., 2020). The original idea of this community was to foster interdisciplinary research by connecting education researchers and computer scientists and promote the development and use of AI applications for education. During the past decades, AI has been widely used in education, given the advancement of computing and information processing techniques, particularly with the rapid development with the emergence of deep learning technologies (Limna, 2022) and the availability of big data (Williamson et al., 2020). Most recently, the release of ChatGPT has brought the attention of AIED to more people (Adiguzel, 2023; Zhai, 2023).

There are various AI applications in education, and research has seen noticeable impacts and advantages. By using AI-powered tools, students can receive not only individualized learning experiences and materials that cater to their unique characteristics but also personalized and instant feedback on their learning performance. This affordance not only maximizes learning efficiency but also encourages self-reflection and self-directed learning and motivates learning (Adiguzel, 2023). Since AIED can provide specialized instruction and automated feedback, it can free instructors from daily menial tasks to some extent and allow them to focus on more complex skills like creativity and respond to students more effectively (Guan et al., 2020). Besides, for teachers and administrators, AIED can enhance insight into students' learning processes and performance by evaluating, tracking, and recording, which can help identify the most supportive way for student learning and educational choices (Adiguzel, 2023).

Recently, the growing demand in education and the introduction of certain national guiding reports (e.g., *Framework for K-12 Science Education*) have introduced new challenges, such as the cultivation of problem-solving skills, assessment of high-order thinking, and formative assessment. Traditional teaching methods, which were more theory-based and one-size-fits-all, however, may fail to cater to those needs (Limna, 2022; Zafari et al., 2022). To meet the challenges, using cutting-edge technology, such as AI, is a promising way, considering the recent rapid development of AI technology and the advantages of AIED. Therefore, AIED has been an emergent and thriving area (Zafari et al., 2022; Zhai, 2023) and is thought to have the potential to revolutionize education.

1.2 AIED literature review

To comprehensively elucidate the historical development and current state of the field, review studies have been undertaken within the domain of AIED. A majority of these studies used narrative synthesis or systematic review approaches (Chassignol et al., 2018; Zawacki-Richter et al., 2019; Zhang & Aslan, 2021). For example, Roll and Wylie (2016) analyzed 47 papers in three individual years, i.e., 1994, 2004, and 2014, as they represent early, middle, and recent AIED research. From the perspectives of type and focus, domain and breadth, interaction type and collaborative structure, technology used, learning setting, and learning goals, the study explored the evolutionary process and future opportunities of AIED. They found an increase in the level of evaluative rigor of the papers, and a growing number of papers presented empirical data. Studies were found to focus more on science, technology, engineering and mathematics (STEM), and the breadth of content that is covered and the time spent using the technological environments is also increasing. Moreover, research paid attention to not only domain-knowledge learning but also self-regulation, collaboration, motivation, etc., in education, AI is also used to facilitate these aspects in the studies. A total of 61 articles on adaptive learning were analysed by Martin et al. (2020) to figure out the publication trends, instructional context, research methodology components, research focus, adaptive strategies, and technologies used in the field. They found that the number of publications peaked in 2015, and the majority of the studies were conducted in higher education in the computer science discipline. The most investigated adaptive targets were adaptive feedback and adaptive navigation, while the most observed learner characteristic was the learning style. Zhai et al. (2020) reviewed 49 articles applying machine learning in science assessment systematically and found that most of the studies focused on providing evidence of automatic scoring accuracy, while only a smaller number of studies referred to the pedagogies or provided sufficient technical details of the AI methods. Ouyang et al. (2022) conducted a systematic review of 32 empirical studies from 2011 to 2020, focusing on the functions, the algorithms used, and the effects and implications generated in online higher education. It is found that the functions of AI include prediction, recommendation, automatic assessment, and improvement of learning experience. AI can make predictions, and high-quality recommendations and improve students' academic performance as well as engagement online. However, advanced techniques such as deep learning are rarely used compared to traditional ones. These reviews mainly focused on a specific aspect or application of AIED, and the analyses were based on a rather small sample. Thus, to capture a complete picture of the field, the traditional qualitative review method falls short in comprehensively analysing the trends and evolvements of the field, especially because of the rapid development of AI. Moreover, the descriptive review approach that highly relies on human coding may draw concerns about the objectivity of the conclusions.

To fill the gaps, this study employed a bibliometric analysis approach to delve into the insights of AIED research. Bibliometric tools facilitate the identification of emerging research trends, gaps in the existing body of knowledge, and the evolution of a research field over time. These capabilities are particularly beneficial for scholars seeking to understand the landscape of a research domain such as AIED quickly and objectively. Additionally, bibliometric analysis can identify key researchers and institutions, thereby serving as a valuable tool for academic networking and collaboration. While bibliometric methods are not without limitations, such as the potential for overlooking qualitative nuances and the influence of varying citation practices across disciplines, they offer a more systematic and scalable approach to literature review (Chen et al., 2012).

Several prior reviews on AIED have been conducted using this approach and have provided abundant information about the annual publication, document types, top publication sources, top countries and institutions, scientific collaborations, top-cited articles and research topics, and trends (Chen et al., 2022; Prahani et al., 2022). For example, Pu et al. (2021) conducted a bibliometric review of articles published from

2000 to 2020 with CiteSpace, a software specifically designed to facilitate the detection of emerging trends in the scientific literature (Chen et al., 2012), which can be used to conduct clustering based on information like keywords and citations. They concluded three topics of AIED: the influence of intelligent applications on student learning, the relationship between teachers and machines, and the contribution and risks of algorithms in education. Talan (2021), with 2686 publications, identified the most productive country, the most frequently published journals, and the most productive institutions and researchers and concluded that the most frequently used keywords were artificial intelligence, intelligent tutoring systems, machine learning, deep learning, and higher education. Based on 4519 articles from 2000 to 2019, Chen et al. (2022) found a rise in the publication, with the USA being the top-ranked country by the H-index and Germany being the closest partner. They also found five topics, including Educational Data Mining (EDM), intelligent tutoring for writing and reading, intelligent tutoring for K12 and special education, Artificial Neural Networks (ANNs), and graphical representation and knowledge connection, to be the topics that have been paid increasing attention. Computerized adaptive testing and diagnosis systems, ontology and knowledge management, problem-solving and example-based learning, and ITSs for authoring and scaffolding experienced a significant decreasing trend. An analysis of 457 documents by Prahani et al. (2022) revealed three top subject areas: Computer Science, Engineering, and Social Sciences. The study also identified the research trend, including its application to students, the subject of education in engineering education, teaching methods, e-learning-based education, education systems, and curricula, including AI. The prior bibliographic analyses yielded inconsistent research trends partially due to the methods employed, such as the eligibility of literature, the scope of the publications, and the programs used to conduct the analyses. More importantly, the analyses lacked an explicit framework to guide the interpretation of the analysis results (Hinojo-Lucena et al., 2019; Talan, 2021). In our research, we intend to analyze the research trends from three aspects: the technology, applications, and outcomes and the degrees of relevance to the educational objectives. The technology aspect is focused on the algorithmic development of AIED, which will unpack the technological advances. The applications indicate how AI has been integrated into education to support teaching and learning. The third aspect, outcomes, will review the impacts of AIED on teaching and learning. The three aspects present a comprehensive picture of AIED research trends that will inform future research and practice.

2 Methods

2.1 Data sources and screening

We used WoS (Web of Science) and Scopus database to obtain eligible literature. Given that AIED is rapidly developing, this review only focuses on research published from 2013 to 2023. To ensure the accuracy of searching, based on previous studies (Chen et al., 2022; Tahiru, 2021; Zafari et al., 2022), we used the first retrieval strategies shown in Table 1 at the beginning, which results in an amount of non-

Table 1 Eight retrieval strategies

Number	Search string	Number of articles	
		WoS	Scopus
1	("adaptive learning") AND ("education*" OR "educational*" OR "teaching*" OR "student*" OR "instruction*" OR "teacher*" OR "classroom*")	102	144
2	"intelligent tutoring system"	317	779
3	"automatic grading" OR "automated grading" OR "automatic scoring" OR "automated scoring" OR "automatic question generation" OR "automated question generation"	439	633
4	("facial expression*" OR "sentiment analysis*") AND ("education*" OR "educational*" OR "teaching*" OR "student*" OR "instruction*" OR "teacher*" OR "classroom*")	174	297
5	("prediction*" OR "administration*" OR "retention*") AND ("education*" OR "educational*" OR "teaching*" OR "student*" OR "instruction*" OR "teacher*" OR "classroom*") AND ("Artificial intelligence*" OR "machine learning*" OR "computer vision*" OR "Natural language processing*" OR "artificial neural network*")	114	210
6	("recommendation system*") AND ("education*" OR "educational*" OR "teaching*" OR "student*" OR "instruction*" OR "teacher*" OR "classroom*")	80	135
7	("learning analysis*") AND ("education*" OR "educational*" OR "teaching*" OR "student*" OR "instruction*" OR "teacher*" OR "classroom*")	37	50
8	("artificial intelligence*" OR "machine learning*" OR "computer vision*" OR "Natural language processing*" OR "artificial neural network*") AND ("Education" OR "educational*" OR "teaching*" OR "student*" OR "instruction*" OR "teacher*" OR "classroom*")	1967	3285

negligible publication missing. Therefore, according to several qualitative reviews in the AIED (Zhang & Aslan, 2021), we identified the other eight retrieval strategies (Table 1). Publications were mostly collected on May 31st, 2023, using the eight strategies and yielded literature written in English, including journal articles, conference proceedings, and preprints.

We then screened the titles and abstracts and removed the papers that were not eligible for AI in education ($n=1435$). Using Citespace software, we identified and removed duplicated literature ($n=485$). Eventually, 6843 records were retained, including 2585 records from the WoS and 4258 records from the Scopus database. Figure 1 shows the steps of data collection and screening.

2.2 Data analysis

This study mainly used Citespace 6.2.R2 Advanced software (The software can be obtained from https://citespace.podia.com/) to conduct bibliometric analysis. At first, descriptive analysis was carried out using Citespace, and data were further processed by using Microsoft Excel. Keyword frequency analysis, clustering mapping, and co-citation analysis were done to identify research topics of AIED and help understand each topic. Details of the analyses are shown in Fig. 2. Clustering mapping of keywords based on frequency was conducted to provide data for detecting research

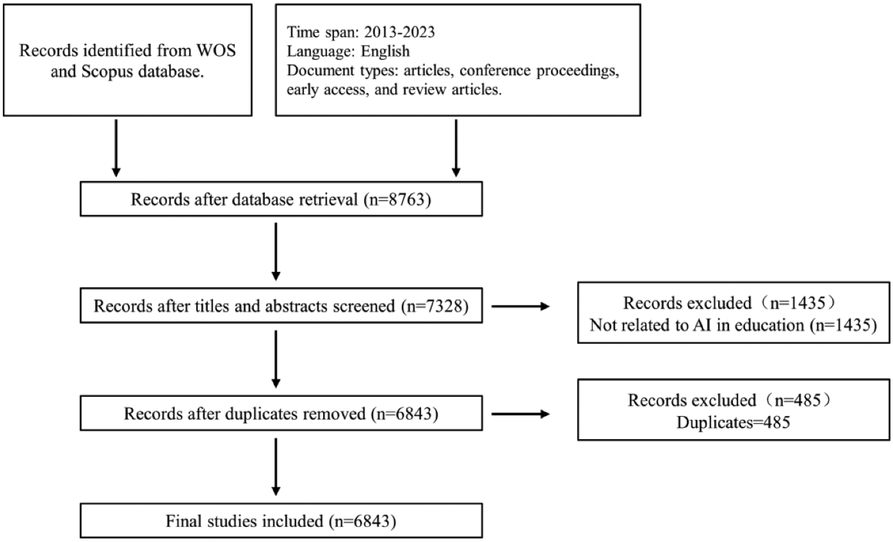

Fig. 1 Data searching and screening process

Keywords frequency analysis & Keyword clustering mapping

(1) Terms whose frequency was above 30 and meant the same were combined.
(2) Terms such as “Artificial Intelligence” and “Education” were excluded
(3) the node threshold of "TOPN=25" was set, and the network algorithm "Pathfinder" was used to trim the network

Co-citation analysis

(1) the node threshold of "TOPN=25" was set, and the network algorithm "Pathfinder" was used to trim the network

Fig. 2 Data analysis process

topics of AIED. Before the clustering was conducted, synonym keywords (e.g., intelligent tutor, intelligent tutoring, intelligent tutoring system, ITS, etc.) with frequency above 30 were merged. Keywords such as “Artificial Intelligence,” “Artificial Intelligence technology,” and “Education” were excluded because they form huge nodes but provide limited information useful to address our research questions. Next, using co-citation analysis, we identified the most influential articles to help delve into the research topics identified. In all analyses, the node threshold of “TOPN = 25” was set, and the network algorithm “Pathfinder” was used to trim the network.

3 Results

3.1 Descriptive information

The eligible articles come from 133 countries and areas. Figure 3 shows the distribution of publications in the top 10 countries and areas, among which China leads the productivity ($n=1284$), followed by the USA ($n=648$), then India ($n=249$), Spain ($n=132$), and Germany ($n=118$). However, the largest centrality value was found in the USA, indicating the greatest importance and contribution.

Among the studies reviewed, seven major subject domains stand out, with the most salient domains being STEM education ($n=464$) and language education ($n=192$) (see Table 2). In addition, we found that AIED research spanned from preschool to higher education, but significantly more research was conducted at the higher education level ($n=771$) compared to K-12 ($n=90$) and preschool ($n=15$) (see Table 3).

3.2 Research topics and trends of AIED

3.2.1 AIED research overview

Investigating keyword clusters permitted an overview of related topics on which AIED researchers have been working. Keyword clustering identified 16 significant clusters labeled by the log-likelihood ratio (LLR) test method (Fig. 4; Table 4). The quality of the clustering is indicated by modularity (Q value) and silhouette score. Modularity implies the strength of the division of a network into clusters (or "modules" or "communities"). It determines the degree to which nodes in the network can be divided into a number of sub-networks, within which the nodes are connected tighter (Chen et al., 2012). Q value ranges from 0 to 1. The higher the Q value is, the more robust the internal structure of the network is. Generally, a value of Q over 0.4 is acceptable (Hongqiang et al., 2020). Silhouette score indicates the separation distance between the resulting clusters (Chen et al., 2012). With the value ranging

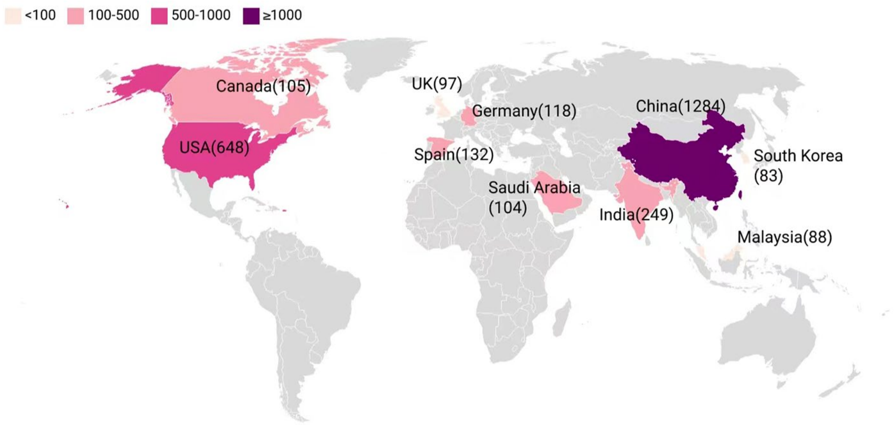

Fig. 3 Top 10 contributed countries in AIED research

Table 2 Distribution of literature in subject areas according to keyword frequency

Rank	Subject area	# of keywords	Related keywords (in descending order)
1	STEM Education	464	engineering education; Physical education; Mathematics education; STEM education; Computing education; Biology education; earth science
2	Language education	192	English teaching; college English; English educations; language learning; foreign language teaching; foreign language; language education; English languages; English; linguistics; computer assisted language learning; English speaking test; drama writing; creative drama education
3	Medical Education	142	medical education; medical students; medical student; major clinical study; medical curriculum; physiology; clinical competence; dentistry; health professions education; surgical education; breast cancer; nursing education
4	Music Education	56	music education; music; music teaching
5	Education Management	27	education management; human resource management; management systems
6	Vocational Education	25	vocational education; entrepreneurship education
7	Art Education	2	art education

Table 3 Distribution of literature in grade level according to keyword frequency

Rank	Grade Level	# of keywords	Related keywords (in descending order)
1	Higher Education	771	higher education; colleges and university; college students; college English; undergraduate students; college physical educations; higher vocational colleges; college student; college education; university student; higher school; higher educational institutions; college physical education; undergraduate courses; online higher education; college teaching
2	K-12 Education	90	high school students; k 12 education; secondary education; elementary education; primary and secondary schools; k-12 education; k 12; middle school; child; primary schools; middle school students
3	Preschool Education	15	preschool education; early childhood education

from 0 to 1, a higher silhouette score means a higher homogenous. A value over 0.5 is satisfying (Hongqiang et al., 2020). In the study, Q = 0.56 and mean Silhouette = 0.77 indicate a well-structured network and a robust separation between different clusters.

Researchers reviewed the 16 clusters and further identified three themes among them, including AIED technology, AIED application, and subject areas, which are interrelated to each other. Below, we describe each theme in detail.

3.2.2 AIED technology

Among various AI technologies, the popular cluster recognized *Large data mining* as the most frequently referred feature in AIED research (Cluster 2), followed by *Algorithms* (Cluster 6), *Deep learning* (Cluster 9), *Speech recognition* (Cluster 11), and *Computer vision* (Cluster 15).

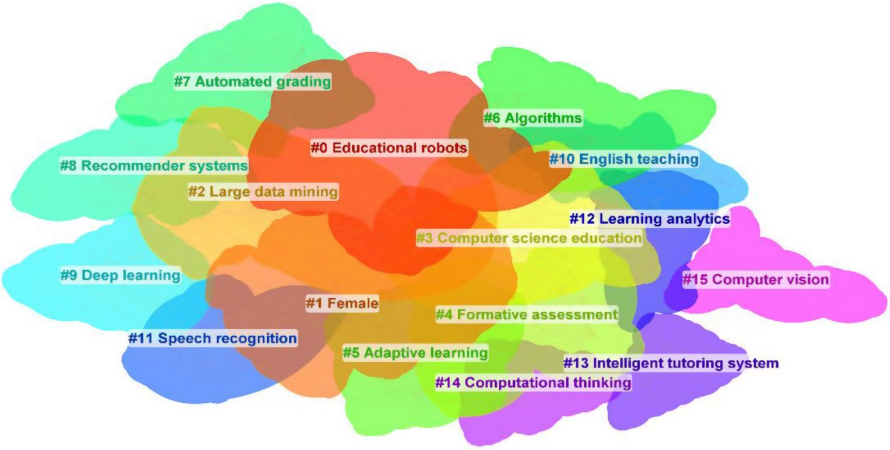

Fig. 4 Keyword clustering mapping

AI is increasingly applied in education research partially because of the availability of *large datasets* with various variables. AI technology, such as machine learning, could mine such complex data sets because it highlights the ability that computers learn from "experience" to develop algorithmic models and then implement the models developed to classify or predict. Machine learning has had a significant development in the past decade, and most other popular technologies, such as decision trees and deep learning, are subcategories of machine learning. Machine learning algorithms are classified mainly into supervised and unsupervised learning (Rastrollo-Guerrero et al., 2020). Supervised learning builds a mathematical model of training data that contains both the inputs and outputs to make predictions about future instances. In unsupervised learning, there are only inputs. The goal is to find commonalities and the structure in the data, based on which to classify or categorize the existing data or the new one. Besides, sentiment analysis and forecasting further the data mining approach and could provide more sentiment information about learning, which broadens the usability of AI in education.

Inevitably, AIED researchers are concerned with various innovative *algorithms* applied to solve education problems. For example, the decision tree is an algorithm broadly used to classify data into branch-like segments that construct a tree with a root node, internal nodes, and leaf nodes. The advantages of the decision tree are that the result is easy to understand and there is no need for data assumption or prior models, which makes it usable for non-linear relationships (Mauro et al., 2017). Decision tree is powerful in classification, prediction, interpretation, and data manipulation to deal with education problems. For example, Gobert et al. (2013) used the decision tree to build a machine-learned detector for scoring student science inquiry performance automatically. The detector was shown to be able to replicate human judgment and can be used in another physical science domain without modification or retraining of the algorithm. Kim et al. (2017) used a decision tree to automatically classify student answers to a novel intermediate constraint question. Alsalman et al. (2019) used the decision tree and artificial neural network (ANN) to build a classi-

Table 4 Results of keyword clustering

Cluster ID	Size	Cluster Label	Top Terms (LLR)
0	102	Educational robots	educational robots; augmented reality; multimedia technology; software engineering
1	97	Female	female; male; human; adult
2	81	Large data mining	machine learning; data mining; sentiment analysis; forecasting; learning algorithms
3	78	Computer science education	computer science education; a.i. literacy; a.i. education; artificial intelligence education; literature review
4	72	Formative assessment	formative assessment; knowledge base; music education; feedback; back propagation neural networks
5	68	Adaptive learning	adaptive learning; personalized learning; computer aided instruction; intelligent tutoring system; machine learning
6	65	Algorithms	decision trees; support vector machines; logistic regression; nearest neighbour search
7	63	Automated grading	automated grading; automation; grading; automatic scoring; automated scoring
8	58	Recommender systems	recommender systems; collaborative filtering; personalized recommendation; personalized recommendation systems; resource recommendation
9	58	Deep learning	deep learning; mental health; health; topic modeling; convolution
10	56	English teaching	English teaching; teaching; college English
11	52	Speech recognition	speech recognition; engineering education; data analytics; teaching systems; computer science
12	50	Learning analytics	learning analytics; higher education; learning analytic; MOOCs; dropout prediction
13	44	Intelligent tutoring system	intelligent tutoring system; intelligent tutoring systems; computer aided instruction; learning process; affective computing
14	35	Computational thinking	computational thinking; k-12 education; mathematics education; teachers; eye tracking
15	20	Computer vision	computer vision; high school; object tracking; applications; competitive exercise

Note LLR = Log-likelihood Ratio test method

fication model to predict university students' academic performance and found that decision trees and ANN perform better in different cases. Decision tree is also used in educational management. Yang (2022) used the decision tree in the university educational administration management system and found that it simplified the calculation method and improved the accuracy greatly.

Deep learning gained significant attention because of its breakthrough compared to traditional machine learning models. Deep learning is obtained by adding more depth and layers to machine learning models. By extracting higher-level features from the raw input using multiple layers, deep learning allows the solution of more complex problems with higher accuracy and less manual tuning. These advantages made it a spotlight for researchers to develop and implement in the educational field. Specifically, deep learning is more widely used in adaptive assessment and grading, predicting performance and student retention (Guan et al., 2020). Lin and Chen (2020) developed a deep learning recommendation system with augmented reality (AR) technology and found that students learning with this system performed better in terms of learning achievement and computational thinking ability than students learning with AR systems that do not embed deep learning recommendation. Anupama and Elayidom (2022) developed a recommendation system using deep learning to predict suitable academic paths for higher secondary students. Results showed that the recommendation accuracy is satisfying and can be used in practice.

Speech recognition can verbally input and then generate an output in forms such as text (Shadiev & Liu, 2023). Speech recognition technology is also known as automatic speech recognition (Yu, 2012). Traditionally, speech recognition mainly uses the hidden Markov modeling (HMM) approach combined with feedforward neural networks. Most recently, many aspects of speech recognition have been using deep learning methods, which can decrease word error rate by 30% and is a major innovation in the speech recognition field. Speech recognition is now widely used in education, especially in language learning (Shadiev & Liu, 2023). For example, Ahn and Lee (2016) employed speech recognition in a mobile-based learning system to improve learners' English speaking proficiency. They found that speech recognition made the speaking activity more interactive and motivating, and 57% of the participants thought that the application was helpful in practicing their oral English. Arcon et al. (2017) applied speech recognition to composition writing for elementary school English language learners (ELLs). They found the application led to higher holistic text quality, as well as less error and effort. Shadiev et al. (2017) investigated the effectiveness of speech recognition for English as foreign language leaners (EFLs) to learn when they attend lectures in English. They found that using speech recognition increased student learning performance, attention, and meditation. Students showed positive perceptions towards using speech recognition for learning. Besides, speech recognition can also benefit students with special needs, such as helping deaf students manage digital sources (Wald, 2005).

Computer vision can acquire, process, analyze, and understand digital images and extract high-dimensional data from the real world. The information extracted from digital images can eventually lead to a decision or execution of a suitable action (Sophokleous et al., 2021). The typical tasks of Computer vision include face recognition, motion analysis, human activity recognition, etc. With the use of deep learning, the accuracy of computer vision has increased in several applications. Lee et al. (2022) developed a behavior recognition system that combines deep learning and computer vision techniques for automatically analyzing the learning process of students in STEM education and found that the average precision is satisfying to keep track of the learning process. Many computer vision studies in education focus on tracking and analyzing classroom teaching and learning to provide useful information for teaching adjustment (Shenoy et al., 2022). Haar (2019) tried to integrate a model using computer vision methods that can capture video, pre-process it, and classify the students captured into eight emotion categories to create student emotion reports automatically, and the results showed that the derive of emotion can be done in near real-time while a class is being given. Bhavana et al. (2020) applied computer vision technology to develop a face recognition system, which was used as an automated attendance system in university classrooms for checking the daily attendance of students, and the studies showed the test accuracy reached 96.66%. There are many other studies that have integrated computer vision with other technology, like educational robots, augmented reality, etc., to build tools that can facilitate learning. Altin et al. (2014) developed a robotics educational system with computer vision technology, which can detect hand-written characters of new alphabet learners and help them write fast and clearly. Esquivel-Barboza et al. (2020) developed and tested three computer vision algorithms in a smartphone-based educational robot, with which students can solve challenging projects and experience more realistic robotics. Chursin & Semenov (2021) explored the use of computer vision in an educational game, which was developed to enhance student skills in mathematics, physics, and programming.

3.2.3 AIED application

As for AI applications, the largest cluster labeled is *Educational robots* (Cluster 0), followed by *Adaptive learning* (Cluster 5), *Automated grading* (Cluster 7), *Recommender systems* (Cluster 8), *Intelligent tutoring system* (Cluster 13), and *Learning analytics* (Cluster 8). These applications were widely used to enhance formative assessment and the cultivation of high-order thinking like *Computational thinking* (Cluster 14). A summary of these AIED applications, corresponding descriptions, and the applied AI technologies can be seen in Table 5.

Educational robots can accomplish specific teaching tasks by conveying knowledge and providing company and interaction. Kewalramani et al. (2021) explored the uses of interactive AI robotic toys on 4-5-year-old children, where children engage and interact with AI toys as a friend to foster their inquiry literacy. Based on the data

Table 5 AIED applications, descriptions, and applied AI technologies

AI Application	Description	AI Technologies
Educational robots	Achieve abundant functions like intelligent perception and reasoning, planning and decision-making, control, and interaction. Serve as tutors, tutees, or teaching aids, and allow students to interact with them orally and physically.	Machine Learning, Decision Trees, Computer Vision
Automated grading	Automatically evaluate students' performance (e.g., written explanations or drawn models) and provide feedback to ease the human burden.	Machine Learning, Deep Learning, Natural Language Processing, Image Processing, Speech Recognition
Recommender systems	Tools and techniques that suggest items that are most likely of interest to a particular user. In education, it supports personalized learning activities through enhanced information retrieval and suggesting suitable learning resources or educational choices.	Decision Trees, Deep Learning, Machine Learning
Intelligent tutoring systems	Presenting information, asking questions or assigning tasks, providing feedback, and choosing instructional activities and strategies in each step of the problem-solving process adapted to the characteristics and needs of students.	Machine Learning, Natural Language Processing), Data Mining,
Learning analytics	"the measurement, collection, analysis, and reporting of data about learners and their contexts" focuses on data related to learners' interactions with course content, other students, and instructors to provide personalized support to students in time.	Machine Learning, Data Mining, Educational Data Mining, Data Visualization

obtained from interviews, observations, and artifact analysis, the study concluded that playing with AI robots improved children's creative, emotional, and collaborative inquiry. Fernández-Llamas et al. (2020) compared the effectiveness of robotic and human teachers in teaching computational principles to pre-university students and found that older students get better scores with robotic teachers, and there is no significant difference between younger students whether they were taught by robots or human. In STEM education, with educational robots consisting of disassembled hardware and programmable software, students can also practice their design or programming better (W. Xu & Ouyang, 2022). For instance, to teach object-oriented computer languages, Rodriguez Corral et al. (2016) used a type of commercial ball-shaped robot with sensing, wireless communication, and output capabilities. They found that compared to students who learned through a standard introductory approach in the control group, students who learned with the robots got an overall better mark and showed higher interest.

Automated grading is used to evaluate types of student answers, including text, image, and speech (e.g., Zhai et al., 2022). Taghipour and Ng (2016) developed a system that can learn the features and relation between an essay and its score automatically based on recurrent neural networks and found the best system outperformed strongly. Zhai et al. (2023) identified three attributes of scientific argumentation, i.e., making claims, using evidence, and providing warrants, and developed machine learning algorithms to assess student written responses automatically. Results showed that the automatic grading could achieve a satisfying average machine–human agreement. In addition, automatic question generation is also growing, which utilizes AI technologies to automatically generate questions or tasks for instructors (W. Xu & Ouyang, 2022). For example, Aldabe and Maritxalar (2014) developed a system to automatically generate Multiple-Choice Questions (MCQ) using natural language processing (NLP) techniques and scientific terms as a starting point. The qualitative and quantitative analysis of the generated tests showed that the system can help teachers generate MCQ.

Recommender systems can help students find suitable learning resources they need to study or provide suggestions about educational choices, including selection of courses, programs, career paths, etc. (Rivera et al., 2018). Valdiviezo-Díaz et al. (2016) propose a recommender system to recommend learning resources in a smart classroom based on five types of knowledge: students, learning resources, topics, context, and criticism. Cabada et al. (2018) present a Web-based Environment, including a recommender and mining system, to provide real-time programming instruction for learners. The recommender system can recommend new exercises to a student based on the performance of previous learners. It has been found that students learning with this tool enjoy using it and perform better than students learning with the traditional method. Baskota and Ng (2018) developed a recommendation system that recommends appealing graduate programs to students based on their personal data and data of various graduate programs. They conducted an empirical study using data from current graduate schools and former graduate school applicants and verified the accuracy of the system. The main approaches used to generate recommendations are content-based, collaborative filtering, and hybrid (Rivera et al., 2018). In the content-based approach, the recommendation is assigned based on similarities in the items' properties. Collaborative filtering provides a recommendation based on the preferences of other users. These two approaches can be combined in various ways to form a hybrid approach (Deschênes, 2020).

As for *Intelligent tutoring systems* (ITS), many empirical research and systematic reviews found that it is widely used and effective in various subjects (Crow et al., 2018; Huang et al., 2023). For example, Walker et al. (2014) developed an ITS to help high students learn algebra. The research adopted a pre-and post-test design with a controlled group receiving non-adaptive support, while the experimental group received adaptive support using the ITS. It was found that ITS improved student learning by providing more relevant feedback. Kulik and Fletcher (2016) also conducted a meta-analysis of 50 controlled evaluation studies of ITS and found that 92% of students who received intelligent tutoring performed better than students from traditional classes, and the positive effect is robust and substantial. A systematic review of ITS also found that most of the empirical evaluations showed positive results in student learning gains (Paladines & Ramírez, 2020).

Learning analytics is thought to be beneficial in predicting the dropout of students, increasing engagement of students, improving learning outcomes, providing real-time feedback, and personalized learning (Banihashem et al., 2018; Mah, 2016). Freitas et al. (2014) proposed a learning analytics model using data from various sources to support personalized learning and services and strengthen student retention. Lu et al. (2017) conducted a controlled experiment applying learning analytics in a MOOC collaborative programming course. In the study, students in the experiment group received learning interventions from a teacher according to the result of learning analytics, and students in the control group received interventions according to the instructor's observation. Results showed that applying learning analytics improved students' engagement and learning outcomes. Lacave et al. (2018) used learning analytics to identify dropout factors of computer science studies through Bayesian networks. Reviews of learning analytics also found that a majority of studies use learning analytics to improve retention of students, while few are focused merely on improving the teaching/learning process or academic issues (Hernández-de-Menéndez et al., 2022). Also, learning analytics are integrated with intelligent tutoring systems. Yilmaz et al. (2022) present a design of an adaptive and dynamic intelligent tutoring system supported by learning analytics in order to make learning management systems like MOOCs smarter. Researchers have also used learning analytics to explore the performance of the Chinese mathematical intelligent tutoring system and found that a dialogue-based ITS with adaptive feedback is helpful for learning fraction multiplication and division (Sottilare & Schwarz, 2019).

3.2.4 Subject domain outcomes

Two topics of *Computer science education* (Cluster 3) and *English teaching* (Cluster 10) were formed in the cluster analysis. *Computer science education* is a subdomain of STEM education focusing on computing education. Thus, the result is consistent with the keyword count shown in Table 2.

AI was widely used in teaching *computer science*. One of the popular applications is intelligent tutoring systems. (Mousavinasab et al., 2021; Xu & Ouyang, 2022).

Nesbit (2014) found that the learning outcomes of ITS in computer science education were significantly better than teacher-led classroom instruction by conducting a meta-analysis of 22 studies with effect size presented. Particularly, a number of intelligent tutoring systems have been created for programming education, which can be referred to as Intelligent Programming Tutors (IPTs) (Crow et al., 2018). Most IPTs need environments to produce and run code. Due to the complexity of programming tasks, IPT can provide hints on the syntax and semantics of student-produced programs. According to previous literature, IPTs can help address difficulties for novice programmers (Crow et al., 2018). Besides, recommendation systems were used to personalize student learning by suitably suggesting exercises for students, which had a positive effect on students' programming learning abilities. Cabada et al. (2018) found that students enjoyed learning with systems, including recommender technology, and could achieve better learning achievement. Lin and Chen (2020) found that in programming teaching with a deep learning recommendation system, students' creativity, logical computing, critical thinking, and problem-solving skills are significantly increased. Moreover, automatic assessment is used to evaluate students' code and programming skills (Aldriye et al., 2019), which were shown to be beneficial. Based on the literature review, Pettit et al. (2015) concluded that automatic assessment tools are helpful in student learning, increasing assessment accuracy, and also supporting teachers by reducing their workload. Färnqvist and Heintz (2016) found that most students using automatic assessment tools considered it more objective and could positively influence their ways of working. However, studies that empirically examine the educational effectiveness of automated assessment in CS education are still not sufficient (Paiva et al., 2022). Additionally, considering the problem of student dropout in the computer science discipline, learning analytics are used to identify related factors and predict dropout (Lacave et al., 2018).

Likewise, various AI applications were applied in the field of *English education*. For example, automatic assessment, together with speech recognition, is mainly applied to writing and speaking contexts to facilitate skills such as pronunciation, vocabulary and writing skills, etc. The results of related studies mostly conclude that it can enhance student performance (Jiang, 2022; Shadiev & Liu, 2023). Ahn and Lee (2016) explored students' experience using a mobile-based learning system embedded with speech recognition technology in order to improve learners' English speaking skills. They found that automatic speech recognition enhances student speaking and pronunciation and makes the student feel more motivated to learn speaking. Bai and Hu (2017) explored the effectiveness of an automated writing evaluation (AWE) system for Chinese students in terms of the precision of the feedback and the students' uptake of such feedback. Studies showed that although the AWE feedback is not accurate enough, the student can critically use it to improve their writing. ITS was also widely used through various applications, with the main objective of enhancing students' learning of writing, reading, vocabulary, and grammar (Huang et al., 2023). For example, Ghali et al. (2018) introduced an ITS to help students learn English grammar. The system was evaluated with high school and university English learners and teachers who specialize in teaching English, in terms of the benefit, comprehensiveness of material, quality of system design, and quality of material and yield an acceptable result. Mohammadzadeh and Sarkhosh (2018) investigate the effects of

ITS on the improvement of students' English speaking ability. Results showed that students learning with ITS improved their speaking skills significantly better than the students who received the traditional method of teaching speaking. The effectiveness of using ITS was examined by studies in different contexts, with a majority of results being positive (Jiang, 2022). Xu et al. (2019) conducted a meta-analysis of 19 studies using ITS to improve reading comprehension for students in K-12 classrooms. The result yielded an overall random effect size of 0.60, indicating a medium effect of ITS on student reading comprehension. Other AIED technologies, including neural machine translation tools and affective computing, were also seen to be applied to English education and mostly yielded positive effects (Jiang, 2022).

4 Discussion and conclusions

This study conducted a bibliometric analysis of AIED to understand the status and major research trends in the past ten years. We found that a majority of publications come from China and the US, which is consistent with findings in other prior review studies about AIED (Efendi et al., 2022; Maphosa & Maphosa, 2021; Prahani et al., 2022). Aligned with prior studies, our descriptive analyses revealed that substantial research was conducted at higher educational levels. For example, Mousavinasab et al.'s (2021) systematic review of 53 studies from 2007 to 2017 using ITS reported that the major users were university students (75%). In a systematic review of using machine learning to predict student performance, although the research did not filter studies at the pre-university level in the data selection on purpose, the reviewed studies were mainly conducted at the higher educational level (Albreiki et al., 2021); Zafari et al., 2022).

In terms of the subject domain, STEM education and language education received major attention. In STEM education, more attention has been given to computer science education. And in language, English education accounts for the majority. In 2020, Paladines and Ramírez conducted a systematic review of 49 studies using ITS with dialogue systems and found that most ITSs aimed at STEM learners at the university level (Paladines & Ramírez, 2020). In a systematic review of ITS, Mousavinasab et al. (2021) found that the experts of AIED were mainly in computer sciences (37.73%), followed by mathematics and health/medical education. Apart from these STEM fields, they also found that ITS was more applied in language education, with a majority of which being English education. Based on AIED articles from 2010 to 2020, Zhai et al. (2021) also found that most research-sampled students majored in science and language.

Based on the results of keyword clustering, we divided the results into three aspects: AIED technology (i.e., large data mining, algorithms, deep learning, speech recognition, and computer vision), applications (i.e., educational robots, automated grading, recommender systems, intelligent tutoring system, and learning analytics) and subject domain outcomes (computer science education and English teaching). The results in the subject area are consistent with the descriptive analysis and many previous studies. As for AIED technology and applications, previous reviews have also found similar results. For example, Zawacki-Richter et al. (2019) conducted a systematic review of AIED in higher education and identified AIED applications as profiling and prediction, assessment and evaluation, adaptive systems and personalization, and intelligent tutoring systems. These four areas of applications were also mostly seen in the results of our study. Pu et al. (2021) found core topics in AIED by conducting a bibliometric analysis based on papers from 2000 to 2020, and the main clusters also included intelligent tutoring systems, deep learning, and decision trees. In the bibliometric study in the AIED field done by Talan (2021), it is also found artificial intelligence, intelligent tutoring systems, machine learning, deep learning, and higher education are located in the center of the keyword network map.

In all, the study contributes to the field by identifying the current status, the major research topics, and the educational outcomes in the last ten years in AIED. The findings provide researchers with a general picture of this promising research field and hints for future research topics and directions. With much of the existing research conducted in higher education that addressed specific issues at the educational level, K-12 and preschool education needs more attention. K-12 students are more diverse and are likely to need more customized learning support that AI may provide. Therefore, it is necessary to consider more AI technology in K-12 education accordingly. For research persistent to AIED subject domains, it is not clear why some subject domains are more productive than others, but it is necessary to apply AI technology in other disciplines that received less attention, such as arts, sports, special education, etc. In STEM education, apart from computer science, researchers in other disciplines, including physics, biology, chemistry, etc., also need to be involved in more AI-based research to address problems in the respective subject domains. Besides, other emerging technologies and applications with generative AI, such as ChatGPT (Latif & Zhai, 2023; Zhai, 2023), GPT4V (Lee & Zhai, 2023), Gemini (Lee, Latif et al., 2023), and affective computing (Jiang, 2022) also need more research to examine their potential to address education problems.

Acknowledgements The authors thank China Scholarship Council (CSC) for supporting the study. The study was also partially funded by the National Science Foundation (NSF) (Award # 2101104, PI: Zhai). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the CSC and NSF.

Data availability The data generated and analyzed during the current study are available in WoS (Web of Science) and Scopus database. The retrieval and screening strategies were shown in the method section of the study.

Declarations

Competing interests The authors declared that they have no conflicts of interest to this work.

References

Adiguzel, T. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Conteporary Educational Technology*, *15*(3), ep429.

Ahn, T., & Lee, S. M. (2016). User experience of a mobile speaking application with automatic speech recognition for EFL learning. *British Journal of Educational Technology*, *47*(4), 778–786. https://doi.org/10.1111/bjet.12354.

Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student' performance prediction using machine learning techniques. *Education Sciences*, *11*(9). https://doi.org/10.3390/educsci11090552.

Aldabe, I., & Maritxalar, M. (2014). Semantic similarity measures for the generation of science tests in basque. *IEEE Transactions on Learning Technologies*, *7*(4), 375–387. https://doi.org/10.1109/TLT.2014.2355831.

Aldriye, H., Alkhalaf, A., & Alkhalaf, M. (2019). Automated grading systems for programming assignments: A literature review. *International Journal of Advanced Computer Science and Applications*, *10*(3), 215–221. https://doi.org/10.14569/IJACSA.2019.0100328.

Alsalman, Y. S., Abu Halemah, K., Alnagi, N., E. S., & Salameh, W. (2019). Using decision Tree and Artificial neural network to Predict Students Academic performance. *2019 10th International Conference on Information and Communication Systems ICICS 2019*, *104-109*, https://doi.org/10.1109/IACS.2019.8809106.

Altin, H., Aabloo, A., & Anbarjafari, G. (2014). New Era for Educational Robotics: Replacing Teachers with a Robotic System to Teach Alphabet Writing. Proceedings of 4th International Workshop Teaching Robotics, Teaching with Robotics & 5th International Conference Robotics in Education, January 2015, 164–166.

Anupama, V., & Elayidom, M. S. (2022). Recommendation system using deep learning to predict suitable academic path for higher secondary students. 2022 International Conference on Innovative Trends in Information Technology, ICITIIT 2022, 1–7. https://doi.org/10.1109/ICITIIT54346.2022.9744245.

Arcon, N., Klein, P. D., & Dombroski, J. D. (2017). Effects of dictation, Speech to text, and handwriting on the Written composition of Elementary School English Language Learners. *Reading and Writing Quarterly*, *33*(6), 533–548. https://doi.org/10.1080/10573569.2016.1253513.

Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: how do students respond? Educational Psychology, 37(1), 67–81. https://doi.org/10.1080/01443410.2016.1223275.

Banihashem, S. K., Aliabadi, K., Pourroostaei Ardakani, S., Delaver, A., & Nili Ahmadabadi, M. (2018). Learning analytics: A systematic literature review. *Interdisciplinary Journal of Virtual Learning in Medical Sciences*, *9*(2). https://doi.org/10.5812/ijvlms.63024.

Baskota, A., & Ng, Y. K. (2018). A graduate school recommendation system using the multi-class support vector machine and KNN approaches. Proceedings –2018 IEEE 19th International Conference on Information Reuse and Integration for Data Science, IRI 2018, 277–284. https://doi.org/10.1109/IRI.2018.00050.

Bhavana, D., Kumar, K. K., Kaushik, N., Lokesh, G., Harish, P., Mounisha, E., & Tej, D. R. (2020). Computer vision based classroom attendance management system-with speech output using LBPH algorithm. *International Journal of Speech Technology*, *23*(4), 779–787. https://doi.org/10.1007/s10772-020-09739-2.

Cabada, R. Z., Estrada, M. L. B., Hernández, F. G., Bustillos, R. O., & Reyes-García, C. A. (2018). An affective and web 3.0-based learning environment for a programming language. *Telematics and Informatics*, *35*(3), 611–628. https://doi.org/10.1016/j.tele.2017.03.005.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. *Procedia Computer Science*, *136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233.

Chen, C., Hu, Z., Liu, S., & Tseng, H. (2012). Emerging trends in regenerative medicine: A scientometric analysis in CiteSpace. *Expert Opinion on Biological Therapy*, *12*(5), 593–608. https://doi.org/10.1517/14712598.2012.674507.

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of Artificial Intelligence in Education: Contributors, collaborations, Research Topics, challenges, and future directions. *Educational Technology and Society*, *25*(1), 28–47.

Chursin, G. (2021). M. Semenov (Ed.), Using computer vision in the gameplay of educational computer games. *Journal of Physics: Conference Series* 1989 1 https://doi.org/10.1088/1742-6596/1989/1/012011.

Crow, T., Luxton-reilly, A., & Wuensche, B. (2018). Intelligent Tutoring Systems for Programming Education: A systematic review. *Proceedings of the 20th Australasian Computing Education Conference*, 53–62. https://doi.org/10.1145/3160489.3160492.

Deschênes, M. (2020). Recommender systems to support learners' Agency in a learning context: A systematic review. *International Journal of Educational Technology in Higher Education*, *17*(1). https://doi.org/10.1186/s41239-020-00219-w.

Efendi, T., Lubis, F. F., Mutaqin, Putri, A., Waskita, D., Sulistyaningtyas, T., Rosmansyah, Y., & Sembiring, J. (2022). A Bibliometrics-based systematic review on Automated Essay Scoring in Education. *2022 International Conference on Information Technology Systems and Innovation ICITSI 2022 - Proceedings*, *275-282*, https://doi.org/10.1109/ICITSI56531.2022.9970982.

Esquivel-Barboza, E. A., Llamas, L. F., Vazquez, P., Bellas, F., & Arias-Mendez, E. (2020). Adapting Computer Vision Algorithms to Smartphone-based Robot for Education. 2020 4th International Conference on Multimedia Computing, Networking and Applications, MCNA 2020, 51–56. https://doi.org/10.1109/MCNA50957.2020.9264289.

Färnqvist, T., & Heintz, F. (2016). Competition and feedback through automated assessment in a data structures and algorithms course. Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, 11-13-July, 130–135. https://doi.org/10.1145/2899415.2899454.

Fernández-Llamas, C., Conde, M., Rodríguez-Sedano, F. J., Rodríguez-Lera, F. J., & Matellán-Olivera, V. (2020). Analysing the computational competences acquired by K-12 students when lectured by robotic and human teachers: Can a Robot teach computational principles to pre-university students? *International Journal of Social Robotics*, *12*(5), 1009–1019. https://doi.org/10.1007/s12369-017-0440-9.

Freitas, S., De, Gibson, D., Plessis, C., Du, Halloran, P., Williams, E., Ambrose, M., Dunwell, I., & Arnab, S. (2014). Foundations of dynamic learning analytics: Using university student data to increase retention. *British Journal of Educational Technology*, *46*(6), 1175–1188. https://doi.org/10.1111/bjet.12212.

Ghali, M. A., Ayyad, A. A., Abu-Naser, S. S., & Laban, M. A. (2018). An Intelligent Tutoring System for Teaching English Grammar. *International Journal of Academic Engineering Research*, *2*(2), 1–6. www.ijeais.org/ijaer.

Gobert, J. D., Pedro, S., Raziuddin, M., J., & Baker, R. S. (2013). From Log Files to Assessment Metrics: Measuring Students' Science Inquiry Skills Using Educational Data Mining. *Journal of the Learning Sciences*, 22(4), 521–563. https://doi.org/10.1080/10508406.2013.837391.

Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *International Journal of Innovation Studies*, *4*(4), 134–147. https://doi.org/10.1016/j.ijis.2020.09.001.

Hernández-de-Menéndez, M., Morales-Menendez, R., Escobar, C. A., & Ramírez Mendoza, R. A. (2022). Learning analytics: State of the art. *International Journal on Interactive Design and Manufacturing*, *16*(3), 1209–1230. https://doi.org/10.1007/s12008-022-00930-0.

Hinojo-Lucena, F. J., Aznar-Díaz, I., Cáceres-Reche, M. P., & Romero-Rodríguez, J. M. (2019). Artificial intelligence in higher education: A bibliometric study on its impact in the scientific literature. *Education Sciences*, *9*(1). https://doi.org/10.3390/educsci9010051.

Hongqiang, M., Hongyu, C., Jinliang, H., Jieya, G., & Wei, L. (2020). CiteSpace-based visualization analysis on mappings of alkali-activated cementitious materials in the field of engineering research. *Case Studies in Construction Materials*, *12*, e00350. https://doi.org/10.1016/j.cscm.2020.e00350.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, Research Issues and Applications of Artificial Intelligence in Language Education. *Educational Technology and Society*, *26*(1), 112–131. https://doi.org/10.30191/ETS.202301_26(1).0009.

Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review on artificial intelligence in the EFL context. *Frontiers in Psychology*, *13*(Ml). https://doi.org/10.3389/fpsyg.2022.1049401.

Kewalramani, S., Kidman, G., & Palaiologou, I. (2021). Using Artificial Intelligence (AI)-interfaced robotic toys in early childhood settings: A case for children's inquiry literacy. *European Early Childhood Education Research Journal*, *29*(5), 652–668. https://doi.org/10.1080/1350293X.2021.1968458.

Kim, K. J., Meir, E., Pope, D. S., & Wendel, D. (2017). WordBytes: Exploring an Intermediate Constraint Format for Rapid classification of student answers on constructed response assessments. *Journal of Educational Data Mining*, *9*(2), 45–71.

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring systems: A Meta-Analytic Review. *Review of Educational Research*, *86*(1), 42–78. https://doi.org/10.3102/0034654315581420.

Lacave, C., Molina, A. I., & Cruz-lemus, J. A. (2018). Learning analytics to identify dropout factors of Computer Science studies through bayesian networks. *Behaviour & Information Technology*, *37*(10–11), 993–1007. https://doi.org/10.1080/0144929X.2018.1485053.

Latif, E., & Zhai, X. (2023). Fine-tuning chatgpt for automatic scoring. https://arxiv.org/pdf/2310.10072.

Lee, G. G., Latif, E., Shi, L., & Zhai, X. (2023). Gemini Pro Defeated by GPT-4V: Evidence from Education. https://arxiv.org/html/2401.08660v1.

Lee, G. G., Shi, L., Latif, E., Gao, Y., Bewersdorf, A., Nyaaba, M., ... & Zhai, X. (2023). Multimodality of ai for education: Towards artificial general intelligence. https://arxiv.org/pdf/2312.06037.pdf.

Lee, G. G., & Zhai, X. (2023). Nerif: Gpt-4v for automatic scoring of drawn models. https://arxiv.org/pdf/2311.12990.pdf.

Lee, H. Y., Chang, W. C., & Huang, Y. M. (2022). Combining Deep Learning and Computer Vision Techniques for Automatic Analysis of the Learning Process in STEM Education. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 13449 LNCS. Springer International Publishing. https://doi.org/10.1007/978-3-031-15273-3_3.

Li, Y., Xu, Z., Wang, X., & Wang, X. (2020). A bibliometric analysis on deep learning during 2007 – 2019. *International Journal of Machine Learning and Cybernetics, 11*(12), 2807–2826. https://doi.org/10.1007/s13042-020-01152-0.

Limna, P. (2022). A review of Artificial Intelligence (AI) in education during the Digital era. *Advance Knowledge for Executives*, *1*(3), 1–9.

Lin, P. H., & Chen, S. Y. (2020). Design and evaluation of a deep learning recommendation based augmented reality system for teaching programming and computational thinking. *Ieee Access : Practical Innovations, Open Solutions*, *8*, 45689–45699. https://doi.org/10.1109/ACCESS.2020.2977679.

Lu, O. H. T., Huang, J. C. H., Huang, A. Y. Q., & Yang, S. J. H. (2017). Applying learning analytics for improving students engagement and learning outcomes in an MOOCs enabled collaborative programming course. *Interactive Learning Environments*, *25*(2), 220–234. https://doi.org/10.1080/10494820.2016.1278391.

Mah, D. K. (2016). Learning analytics and Digital badges: Potential impact on Student Retention in Higher Education. *Technology Knowledge and Learning*, *21*(3), 285–305. https://doi.org/10.1007/s10758-016-9286-8.

Maphosa, V., & Maphosa, M. (2021). The Trajectory of Artificial Intelligence Research in Higher Education: A Bibliometric Analysis and Visualisation. IcABCD 2021–4th International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems, Proceedings, 1–7. https://doi.org/10.1109/icABCD51485.2021.9519368.

Martin, F., Chen, Y., Moore, R. L., & Westine, C. D. (2020). Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. *Educational Technology Research and Development*, *68*(4), 1903–1929. https://doi.org/10.1007/s11423-020-09793-2.

Mauro, C., Gomes, A., Federal, U., & Gerais, D. M. (2017). Advocating the Broad Use of the Decision Tree Method in Advocating the Broad Use of the Decision Tree Method in Education. 22.

Mohammadzadeh, A., & Sarkhosh, M. (2018). The effects of self-regulatory learning through computer-assisted intelligent tutoring system on the improvement of EFL learners' speaking ability. *International Journal of Instruction*, *11*(2), 167–184. https://doi.org/10.12973/iji.2018.11212a.

Mousavinasab, E., Zarifsanaiey, N., Niakan Kalhori, R., Rakhshan, S., Keikha, M., L., & Ghazi Saeedi, M. (2021). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, *29*(1), 142–163. https://doi.org/10.1080/10494820.2018.1558257.

Nesbit, J. C. (2014). How Effective are Intelligent Tutoring Systems in Computer Science Education ? 2014 IEEE 14th International Conference on Advanced Learning Technologies, 99–103. https://doi.org/10.1109/ICALT.2014.38.

Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. In Education and Information Technologies. *Issue 6) Springer US*, *27*, https://doi.org/10.1007/s10639-022-10925-9.

Paiva, J. C., Leal, J. P., & Figueira, Á. (2022). Automated Assessment in Computer Science Education: A state-of-the-art review. *ACM Transactions on Computing Education*, *22*(3). https://doi.org/10.1145/3513140.

Paladines, J., & Ramírez, J. (2020). A systematic literature review of intelligent tutoring systems with dialogue in natural language. *Ieee Access : Practical Innovations, Open Solutions*, *8*, 164246–164267. https://doi.org/10.1109/ACCESS.2020.3021383.

Peters, M. A., Jackson, L., Papastephanou, M., Jandrić, P., Lazaroiu, G., Evers, C. W., Cope, B., Kalantzis, M., Araya, D., Tesar, M., Mika, C., Chen, L., Wang, C., Sturm, S., Rider, S., & Fuller, S. (2023). AI and the future of humanity: ChatGPT-4, philosophy and education–critical responses. *Educational Philosophy and Theory*, *0*(0), 1–35. https://doi.org/10.1080/00131857.2023.2213437.

Pettit, R. S., Homer, J. D., Holcomb, K. M., Simone, N., & Mengel, S. A. (2015). Are automated assessment tools helpful in programming courses? ASEE Annual Conference and Exposition, Conference Proceedings, 122nd ASEE(122nd ASEE Annual Conference and Exposition: Making Value for Society). https://doi.org/10.18260/p.23569.

Prahani, B. K., Rizki, I. A., Jatmiko, B., Suprapto, N., & Amelia, T. (2022). Artificial Intelligence in Education Research during the last ten years: A review and bibliometric study. *International Journal of Emerging Technologies in Learning*, *17*(8), 169–188. https://doi.org/10.3991/ijet.v17i08.29833.

Pu, S., Ahmad, N. A., Khambari, M. N. M., & Yap, N. K. (2021). Identification and analysis of core topics in educational artificial intelligence research: A bibliometric analysis. *Cypriot Journal of Educational Sciences*, *16*(3), 995–1009. https://doi.org/10.18844/CJES.V16I3.5782.

Rastrollo-Guerrero, J. L., Gómez-Pulido, J. A., & Durán-Domínguez, A. (2020). Analyzing and predicting students' performance by means of machine learning: A review. *Applied Sciences (Switzerland)*, *10*(3). https://doi.org/10.3390/app10031042.

Rivera, A. C., Tapia-Leon, M., & Lujan-Mora, S. (2018). Recommendation systems in education: A systematic mapping study. *Advances in Intelligent Systems and Computing*, *721*(Icits), 937–947. https://doi.org/10.1007/978-3-319-73450-7_89.

Rodriguez Corral, J. M., Morgado-Estevez, A., Molina Cabrera, D., Perez-Pena, F., Rodriguez, A., C. A., & Balcells, C., A (2016). Application of robot programming to the teaching of object-oriented computer languages. *International Journal of Engineering Education*, *32*(4), 1823–1832.

Roll, I., & Wylie, R. (2016). Evolution and revolution in Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*, *26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3.

Shadiev, R., & Liu, J. (2023). Review of research on applications of speech recognition technology to assist language learning. *ReCALL*, *35*(1), 74–88. https://doi.org/10.1017/S095834402200012X.

Shadiev, R., Huang, Y. M., & Hwang, J. P. (2017). Investigating the effectiveness of speech-to-text recognition applications on learning performance, attention, and meditation. *Educational Technology Research and Development*, *65*(5), 1239–1261. https://doi.org/10.1007/s11423-017-9516-3.

Shenoy, M. A., Santhosh, S., & Kumar, S. S. (2022). A Study on various Applications of Computer Vision for Teaching Learning in Classroom. 6th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2022 - Proceedings, Iceca, 896–900. https://doi.org/10.1109/ICECA55336.2022.10009136.

Sophokleous, A., Christodoulou, P., Doitsidis, L., & Chatzichristofis, S. A. (2021). *Computer Vision meets Educational Robotics* (pp. 1–24). Electronics.

Sottilare, R. A., & Schwarz, J. (2019). Adaptive Instructional Systems. First International Conference, AIS, 469–484. https://doi.org/10.4324/9780203880869-41.

Taghipour, K., & Ng, H. T. (2016). A neural approach to automated essay scoring. Conference on Empirical Methods in Natural Language Processing, Proceedings, 1882–1891. https://doi.org/10.18653/v1/d16-1193.

Tahiru, F. (2021). AI in education: A systematic literature review. *Journal of Cases on Information Technology*, *23*(1), 1–20. https://doi.org/10.4018/JCIT.2021010101.

Talan, T. (2021). Artificial Intelligence in Education: A bibliometric study. *International Journal of Research in Education and Science*, 822–837. https://doi.org/10.46328/ijres.2409.

Valdiviezo-Díaz, P., Aguilar, J., & Riofrio, G. (2016). A fuzzy cognitive map like recommender system of learning resources. 2016 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2016, 1539–1546. https://doi.org/10.1109/FUZZ-IEEE.2016.7737873.

van der Haar, D. (2019). Student Emotion Recognition Using Computer Vision as an Assistive Technology for Education. In Information Science and Applications (pp. 183–192). Springer Singapore. https://doi.org/10.1007/978-981-15-1465-4.

Wald, M. (2005). Using automatic speech recognition to enhance education for all students: Turning a vision into reality. Proceedings - Frontiers in Education Conference, FIE, 2005, S3G-S3G. https://doi.org/10.1109/fie.2005.1612286.

Walker, E., Rummel, N., & Koedinger, K. R. (2014). Adaptive intelligent support to improve peer tutoring in algebra. *International Journal of Artificial Intelligence in Education*, *24*(1), 33–61. https://doi.org/10.1007/s40593-013-0001-9.

Williamson, B., Eynon, R., & Williamson, B. (2020). Historical threads, missing links, and future directions in AI in education education. https://doi.org/10.1080/17439884.2020.1798995.

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education*, *9*(1). https://doi.org/10.1186/s40594-022-00377-5.

Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. *British Journal of Educational Technology*, *50*(6), 3119–3137. https://doi.org/10.1111/bjet.12758.

Yang, J. (2022). Application of Decision Tree Algorithm in University Educational Administration. In Innovative Computing (pp. 605–611). Springer Singapore. https://doi.org/10.1007/978-981-16-4258-6.

Yilmaz, R., Yurdugül, H., Karaoğlan Yilmaz, F. G., Şahin, M., Sulak, S., Aydin, F., Tepgeç, M., Müftüoğlu, C. T., & Ömer, O. R. A. L. (2022). Smart MOOC integrated with intelligent tutoring: A system architecture and framework model proposal. *Computers and Education: Artificial Intelligence*, *3*(April). https://doi.org/10.1016/j.caeai.2022.100092.

Yu, Y., & Research on Speech Recognition Technology and Its Application. (2012). 2012 International Conference on Computer Science and Electronics Engineering, 1, 306–309. https://doi.org/10.1109/ICCSEE.2012.359.

Zafari, M., Bazargani, J. S., Sadeghi-Niaraki, A., & Choi, S. M. (2022). Artificial Intelligence Applications in K-12 education: A systematic literature review. *Ieee Access : Practical Innovations, Open Solutions*, *10*, 61905–61921. https://doi.org/10.1109/ACCESS.2022.3179356.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1). https://doi.org/10.1186/s41239-019-0171-0.

Zhai, X. (2023). ChatGPT for Next Generation *Science Learning XRDS*, 29(3), 42–46. https://doi.org/10.1145/3589649.

Zhai, X., Yin, Y., Pellegrino, J. W., Haudek, K. C., & Shi, L. (2020). Applying machine learning in science assessment: A systematic review. *Studies in Science Education*, *56*(1), 111–151. https://doi.org/10.1080/03057267.2020.1735757.

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J. B., Yuan, J., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. Complexity, 2021. https://doi.org/10.1155/2021/8812542.

Zhai, X., He, P., & Krajcik, J. (2022). Applying machine learning to automatically assess scientific models. Journal of Research in Science Teaching, August 2021, 1765–1794. https://doi.org/10.1002/tea.21773.

Zhai, X., Haudek, K. C., & Ma, W. (2023). Assessing Argumentation using Machine Learning and Cognitive Diagnostic modeling. *Research in Science Education*, *53*(2), 405–424. https://doi.org/10.1007/s11165-022-10062-w.

Zhai, Xiaoming, C., Haudek, K., Shi, L., Nehm, H., R., & Urban-Lurain, M. (2020). From substitution to redefinition: A framework of machine learning-based science assessment. *Journal of Research in Science Teaching*, *57*(9), 1430–1459. https://doi.org/10.1002/tea.21658.

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, 100025. https://doi.org/10.1016/j.caeai.2021.100025.