Computers and Education: Artificial Intelligence 6 (2024) 100195

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

Systematic review of research on artificial intelligence in K-12 education (2017–2022)

Florence Martin [a,*], Min Zhuang [a], Darlene Schaefer [b]

[a] North Carolina State University, 2310 Stinson Drive, Raleigh, NC, 27695, USA
[b] University of North Carolina Charlotte, 9201 University City Blvd, Charlotte, NC 28223, USA

ARTICLE INFO

Keywords:
Artificial intelligence
K-12 education
AI methods
AI technology applications
AI use

ABSTRACT

Background: The use of Artificial Intelligence (AI) has increased in all education sectors including K-12 settings where students can learn about AI and have an augmented learning experience using AI.
Purpose: The purpose of this systematic review is to provide a more complete and nuanced understanding of the role and impact of AI in K-12 education by synthesizing publication trends, AI research themes, AI methods and technology applications, and AI use by students and teachers in K-12 educational settings.
Methods: The systematic review searched Web of Science and six databases indexed in EBSCO host. A PRISMA flow chart was applied to search and screen for studies. Articles were screened at the title, abstract and full-text level and coded and analyzed.
Results: Themes in 66 AI studies include AI as a predictor and indicator of academic behavior or performance, AI curriculum design, integrating AI in various subjects, evaluation of AI in education, AI to enhance learning environments and school operations, AI ethics, and the equity and safety of AI. AI methods were grouped into Supervised Learning, Unsupervised Learning and Reinforcement Learning. AI technology applications were Machine Learning (ML) model building tools, intelligent tutors, chat bot, educational games, AI robots and virtual reality devices. AI applications were mostly used by teachers for ML model demonstration, academic performance prediction and behavior prediction. AI was used by students for scientific discovery learning, improving learning experience and data driven decisions.
Conclusion: This review has implications for K-12 school personnel and researchers. Practitioners can use the findings to implement AI in K-12 education. Researchers can benefit from the findings of the review but also build on the gap in research on AI K-12 education.

As the first quarter of the twenty-first century concludes, AI, or the ability for computer systems to achieve tasks associated with human behavior and intellect, continues to transform governments, businesses, industries, education, and healthcare. AI continues to gain momentum in education, but its true impact in K-12 learning environments is not realized yet (Tyson & Sauers, 2021). AI's significance in guiding daily instructional practice, curriculum design and leadership and policy are important considerations for K-12 leaders, educators, and researchers. Additionally, as AI rapidly accelerates, student readiness for AI supported careers becomes a critical component in curriculum planning. Educational researchers are examining student-readiness to use AI in careers ranging from music to medicine. Research is emerging regarding what specific core competencies must be integrated throughout the disciplines and the various subject areas that align with an AI supported world (Norouzi, Chaturvedi, & Rutledge, 2020). Additionally, researchers are investigating overall educator-readiness to teach practical skills that play a direct role in machine-learning including programming, game-design, statistical analysis, and computation (Zhang et al., 2021).

Whether AI is used as a tool and resource to improve educational environments (Cruz-Jesus et al., 2020) and outcomes (Harvey & Kumar, 2019) or as a vital part of a comprehensive K-12 curriculum (Norouzi et al., 2020), the discussion of the ethics and safety regarding AI continues to evoke discussion (Bilstrup, Kaspersen, & Petersen, 2020). Because of the nuances of AI anonymity and the interaction with sensitive student-data, legal and ethical concerns guide the decisions of how AI is used by schools and communities (Bilstrup et al., 2020). Conversations ignite in various directions regarding how AI can improve

* Corresponding author.
E-mail address: fmartin3@ncsu.edu (F. Martin).

https://doi.org/10.1016/j.caeai.2023.100195
Received 29 May 2023; Received in revised form 12 December 2023; Accepted 17 December 2023
Available online 23 December 2023
2666-920X/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

learning environments and the overall learning experience for students; AI's role in K-12 curriculum and how to negotiate AI's significance with safety and educational concerns. Ethics and safety become factors in the benefits and innovative possibilities (Adams, Pente, Lemermeyer, & Rockwell, 2023).

AI continues to permeate K-12 education, and AI's broad potential and concerns are still emerging in K-12 educational research (Wong, Ma, Dillenbourg, & Huan, 2020). Reviewing the publication trends of AI in K-12 education, including the research methodologies, themes, and findings regarding use and future direction can guide better understanding and drive future research and implementation. Prior systematic reviews of AI publications present, on the one hand, a specific aspect of AI in K-12 or, on the other hand, AI's broad use in educational settings, not specific to K-12. A systematic review focused on the comprehensive use and influence of AI in K-12 education for students and teachers may provide clarity on its past, present, and future impact for compulsory education.

1. Literature review

1.1. Previous systematic review

The existing Artificial Intelligence in Education (AIEd) systematic reviews have often focused on broad education settings. Guan, Mou, and Jiang (2020) examined the shifts of AI research themes from 2000 to 2019 and found the decline in conventional tech-enabled instructional design research to the flourishing of student profiling models and learning analytics. Zhai et al. (2021) summarized the AI research themes from 2010 to 2020 in the perspective of AI in knowledge representation, the themes were classified into development layer (classification, matching, recommendation, and deep learning), application layer (feedback, reasoning, and adaptive learning), and integration layer (affection computing, role-playing, immersive learning, and gamification). Zhang and Aslan (2021) focused on AI applications from 1993 to 2020, and categorized the applications into chatbots, expert systems, intelligent tutors, ML, personalized learning systems and visualization. Wang and Cheng (2022) reviewed research from 2001 to 2021 and identified three AI research themes including learning from AI, learning about AI and learning with AI. Chiu, Xia, Zhou, Chai, and Cheng (2023) examined the literature from 2012 to 2021 regarding the AI usage in student learning, teaching, assessment and administration and the relationship between AI and outcomes. They identified thirteen usages such as assigning tasks based on learners' individual preference, enhancing teachers' ability to teach, etc. , and seven outcomes such as leveraging teachers' teaching competency and improving students' motivation and engagement, etc.

The broad focus on education settings has provided some valuable insights of AIEd, while it has somewhat overlooked the crucial distinctions in the implementation of AI across different educational levels. On the one hand, the AI curriculum and pedagogical approaches required for higher education significantly differ from those suitable for K-12 settings. On the other hand, K-12 students have distinct cognitive levels compared to higher education students, necessitating customized AI applications and methodologies. Furthermore, K-12 education serves as the bedrock for nurturing the next generation of AI professionals and enthusiasts, therefore, there is a compelling need to conduct a more focused investigation into AIEd, differentiating between educational settings, emphasizing K-12 settings.

It's encouraging to see some systematic reviews focusing on AI in K-12 settings, as they focus on some individual aspects of AIEd. Celik, Dindar, Muukkonen, and Järvelä (2022) primarily focused on AI usage by teachers; they investigated the literatures from 2000 to 2020, and found AI offers teachers several opportunities for improved planning (e.g., by defining students' needs and familiarizing teachers with such needs), implementation (e.g., through immediate feedback and teacher intervention), and assessment (e.g., through automated essay scoring) of their teaching. Findings also identified AI methods like Artificial Neural Networks, Decision Trees, Bayesian and so on. Crompton and Burke (2022) examined research from 2010 to 2020 and focused on usage of AI for teachers and students, they found AI supports teachers in student monitoring, group management, automated grading and data-driven decisions. Additionally. AI supports students through AI tutors, student thinking, and Just-for-You learning. Marques, Gresse von Wangenheim, and Hauck (2020) focused on the AI curriculum design and investigated several instructional units in teaching ML concepts from 2009 to 2019 and found the instruction units mostly taught ML basics and neural networks, and increased students' understanding and interest of ML as well as contextualizing ML concepts through their social impact. Table 1 provides additional details on these systematic reviews.

1.2. Purpose of the study and research questions

When framing the research questions, we considered a few factors. Firstly, over the last five years, there has been a surge in research outputs in AI K-12 education, therefore, understanding publication trends during this period is crucial for identifying the most current shifts, and recognizing the most influential sources including countries leading in contributions. Secondly, analyzing research methodology components in AI K-12 education can provide a lens for researchers to understand the diversity and distribution of research practices in the field. Specifically, the identification of frequency of participants, subjects, school level and data collection techniques is useful to understand what is involved with and how the research is conducted in this field. Thirdly, identifying AI themes in AI K-12 education can help understand the overarching topics that researchers are exploring, and it can guide educators and policymakers to make decisions related to curriculum development, resource allocation, and policy formulation. Fourthly, examining the AI methods commonly used in K-12 education will shed light on the technological approaches employed in the educational settings and it is instrumental in understanding the technical landscape, assessing the feasibility and scalability of AI solutions, and guiding educators and policymakers in selecting appropriate technologies for K-12 contexts. Fifthly, investigating AI technology applications in K-12 education is crucial for understanding how AI is practically utilized, and it will inform educators about the potential benefits and challenges associated with specific applications, aiding in the design of effective educational interventions and the integration of AI tools into teaching practices. Lastly, exploring how AI is used by K-12 teachers and students is important for understanding the benefits gained from user perspectives, and it is essential for tailoring AI implementations to the needs of teachers and students, fostering successful integration and positive educational outcomes.

Previous systematic reviews in K-12 settings have provided valuable insights but had some specific areas of interest. Crompton and Burke (2022) only examined the usage of AI for teachers and students without technical facets. Celik et al. (2022) primarily focused on AI usage by teachers without student perspectives, they didn't include AI applications but presented AI methods based on specific algorithms without indicating taxonomy. Both Kim et al. (2020) and Marques et al. (2020) only focused on teaching AI that was just one of AI themes. We aim to provide a more complete and nuanced understanding of the role and impact of AI in K-12 education from both teacher and student perspectives, therefore in this study, we will address below research questions:

1. What are the publication trends of AI in K-12 education (publication years and countries)?
2. What research methodology components were used in AI in K-12 research (research methods, school Level, research participants, subject, and data collection)?
3. What AI themes are studied in K-12 research?
4. What AI methods are used in K-12 AI research?
5. What AI technology applications are studied in K-12 research?
6. How is AI used by teachers and students in K-12 research?

Table 1
Systematic reviews of AI in education.

Author	Years of Studies	AI Focus	Setting (K-12/ Higher Ed)	Number of Studies Included	Journal Published
Guan et al. (2020)	2000–2019	Paradigms/themes of technical applications, and the life cycles/trends	Education	425 articles	International Journal of Innovation Studies
Kim, Kim, Lee, and Kim (2020)	Not mentioned	The direction of teaching AI in K-12 in terms of the purpose, content, and methods	K-12	20 articles	The Journal of Korean association of computer education
Marques et al. (2020)	2009–2019	Overview of teaching ML concepts	K-12	33 articles	Informatics in Education
Zhai et al. (2021)	2010–2020	How AI has been applied to education and the research trends and challenges	Education	100 articles K-12 (n = 20) and Higher Ed (n = 23), not mentioned (n = 57)	Complexity
Tahiru (2021)	2010–2019	The opportunities, benefits, and challenges	Education	23 articles	Journal of Cases on Information Technology
Zhang and Aslan (2021)	1993–2020	Landscape of AI Ed research, and an overview of technology applications	Education	40 articles K-12 (n = 17) and Higher Ed (n = 21) Not specified (n = 2)	Computers and Education: Artificial Intelligence
Celik et al. (2022)	2000–2020	Overview of teachers' use of AI and ML methods	K-12	44 articles	TechTrends
Crompton and Burke (2022)	2010–2020	How AI specifically supports teaching	K-12	204 articles	SN Social Sciences
Wang and Cheng (2022)	2001–2021	Key themes and the summary of research trends, supported technologies, and role/purposes	Education	135 articles	Springer
Chiu et al. (2023)	2012–2021	How AI has been integrated into learning, teaching, assessment, and administration; the relationship between AI and learning outcomes	Education	92 articles K-12 (n = 21), Higher Ed (n = 61), and N/A (n = 10)	Computers and Education: Artificial Intelligence

1.3. *Conceptual framework for review codebook*

AI Methods. AI is a broader concept that encompasses the simulation of human intelligence by machines, especially computer systems. In other words, AI means machines have the capability to behave like humans in tasks such as communicating, memorizing, reasoning, and learning (Russell, 2010). Within the field of AI, ML is a specific subset that focuses on algorithms and models that allow machines to learn patterns and make predictions or decisions based on large amounts of data (Bishop, 2006). Throughout the systematic review, what we called "AI methods" refers to ML algorithms.

In terms of a given problem and available data, ML can be divided into supervised learning, unsupervised learning, and reinforcement learning (Janiesch, Zschech, & Heinrich, 2021; Sarker, 2021). Since the various categories can describe the variations in how ML is used in the broad sense, defining categories of ML can provide context to the various trends in AI research in K-12 (Fig. 1).

Supervised learning is a type of ML where the algorithm is trained on a labeled dataset, meaning each input sample is associated with a corresponding target or label (Han, Pei, & Tong, 2022). The goal of supervised learning is to learn mapping between the input features and the target labels, allowing algorithms to make accurate predictions on new, unseen data. Supervised learning can solve problems including regression (predicting numerical values) and classification (predicting categorical labels) (Han et al., 2022; Janiesch et al., 2021; Sarker, 2021). There are some specific algorithms that are commonly used in supervised learning such as Logistic Regression, Naive Bayes, K-Nearest Neighbors, Support Vector Machine, Decision Tree, Random Forest, Gradient Boosting (Sarker, 2021).

Unsupervised learning involves training the algorithm on an unlabeled dataset, where there are no corresponding target labels for the input samples (Han et al., 2022). The goal of unsupervised learning is to group in results, extract generative features, and identify meaningful structures. Unsupervised learning is often used for tasks like clustering (grouping similar data points together) and dimensionality reduction (reducing the number of features while preserving essential information) (Han et al., 2022; Janiesch et al., 2021). Some common algorithms are K-means clustering and Mean-shift clustering (Sarker, 2021).

Reinforcement learning is a type of ML which enables software agents and machines to automatically evaluate the optimal behavior in a

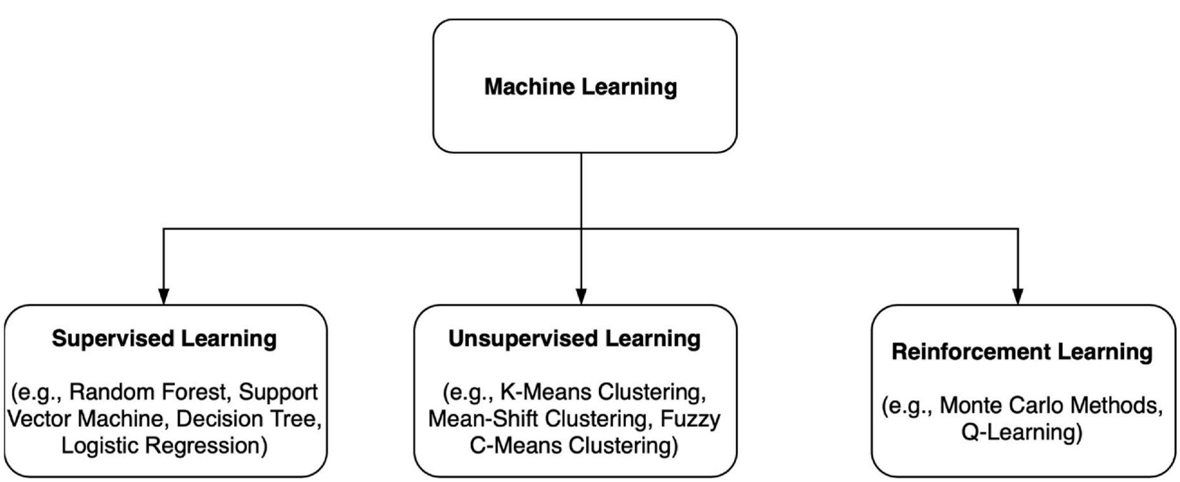

Fig. 1. The categorization of ML

particular context or environment to improve its efficiency (Kaelbling, 1996). The agent takes actions in the environment and receives feedback in the form of rewards or penalties based on its actions. The objective of the agent is to learn a policy, or a mapping from states to actions, that maximizes the total expected reward over time. Some specific algorithms include Monte Carlo methods and Q-learning (Sarker, 2021).

AI Technology Applications. For categorizing AI applications, we referred to Zhang and Aslan (2021), who presented a classification of AI Ed technology applications. Their framework encompasses various AI applications in education, including chatbots, expert systems, intelligent tutors/agents, ML, personalized learning systems, visualizations, and virtual learning environments. While utilizing these existing categorizations from the literature, we extended and adapted them as needed based on the emerging data from our review process.

AI Use. To examine the use of AI by teachers, we drew upon the categorization provided by Crompton and Burke (2022), which includes student monitoring, group management, automated grading, and data-driven decision-making. Additionally, Celik et al. (2022) offered a categorization of AI use by teachers, involving planning (e.g., providing information on student backgrounds, decision-making on learning content), implementation (e.g., timely monitoring, reducing teacher workload, providing immediate feedback), and assessment (e.g., automated assessment and evaluation, provision of feedback). Same as AI applications, those categories were used to develop codebooks, while we also made some adjustments based on the emerging data during the review process.

2. Methods

This systematic review of research on AI in K-12 education followed the process described by Higgins et al. (2019) in the Cochrane Handbook. The review (1) determined the research question, (2) defined eligibility criteria and methods for review, (3)

Searched for studies, (4) applied eligibility criteria, (5) collected data and critically appraise, (6)

analyze and present results, (7) interpreted results and form conclusions, and (8) completed a structured report.

2.1. Inclusion and exclusion criteria

The eligibility criterion and process for initial screening of the studies are described in Table 2.

2.2. Data sources and search strategies

Six databases indexed in EBSCO host Education Research Complete, Academic Search Complete, APA Psycinfo, Child Development and Adolescent Studies, ERIC, Library and Information Technology Abstracts and an additional database Web of Science were searched for this systematic review. Search keywords included ("artificial intelligence" or "ai" or "a.i". or "machine learning" or "deep learning") AND ("school" or "k-12"). Due to the recent technological advancements, pedagogical innovations, increased access, and the collective impact of AI on teaching and learning only five years (2017–2022) were used in this study. This five-year timeframe allows us to focus on the latest developments in AI relevant to K-12 education. The initial search did not utilize date filters and the team updated the inclusion and exclusion criteria indicating eligible publications should be from 2017 through May 2022.

2.3. Process flow

A PRISMA flow chart (Fig. 2) was drawn to depict the process flow showing the identification, screening, eligibility, and inclusion of articles. A faculty researcher and two doctoral students worked on screening and coding. One doctoral student researcher supported the screening

Table 2
Inclusion/exclusion criteria.

Criterion	Inclusion	Exclusion
Language	Article should be written in English	Articles written in other languages
Setting	Focus is on AI in K-12 Education	AI focus in non-education settings or on other educational settings such as higher education
Research Methods	Should be a primary research empirical article – either qualitative, quantitative, or mixed methods	Theoretical or conceptual pieces or review pieces were excluded
Publication Date	The search was conducted for articles published from 2017 to 2022. The publication years are until May 2022	Other years are excluded
Publication Type	Publications are peer reviewed journal articles or conference proceedings	Book chapters, posters and other news stories were excluded
Content Relevance (AI, Education)	AI-empowered tools for education, ML models for learning analytics, AI curriculum design/teaching AI and ML concepts	No mention of AI techniques used, AI techniques is not for teaching and learning (e.g., student health, teacher burnout syndrome, Covid testing), misleading AI concepts (e.g., deeper learning rather deep learning)

process while the other supported the coding process. The faculty researcher met weekly with the doctoral researchers to discuss the screening and coding process, and resolved any questions that arose during these meetings. At the identification stage, from the first six databases, 122 articles were imported to Zotero, and 207 articles were imported from Web of Science to Zotero, totaling 329 articles. A deduplication process to remove duplicate searches resulted in 276 articles. During the screening stage, 276 articles were initially screened for the titles and abstract; after applying the inclusion and exclusion criteria resulted in 137 articles. Among these 137 articles, full text was obtained for 133 articles and reviewed as four articles could not be accessed in full text form, even via interlibrary loan. When exclusion and inclusion criteria were applied during the full-text review, 66 articles were identified and included in the systematic review for the inclusion stage. Some of the reasons for exclusion at the full-text level included articles not being written in English, study not conducted in K-12 setting, not being an empirical study and not a peer-reviewed journal article or conference proceeding.

2.4. Data screening and coding

The research team included three team members. Article screening at the title and abstract level was done on a Google spreadsheet by the lead faculty researcher and one doctoral student. Full-text screening was conducted by the second doctoral student and was supported by the faculty researcher. Atleast, two researchers were involved in each step of the screening process. Afterwards, a Google Document coding form was created to code at the full-text level with discussions between two researchers. The codebook included the items shown in Appendix A. Research articles were primarily coded by one researcher but were discussed with the faculty researcher each week or when there were any challenges in coding.

2.5. Data analysis

A timeline chart was drawn for the publication years. Frequency and percentage are provided for all other variables. Open-ended data was analyzed to collapse for categories and identify themes. Different data collection methods were extracted. Through closed and open-coding

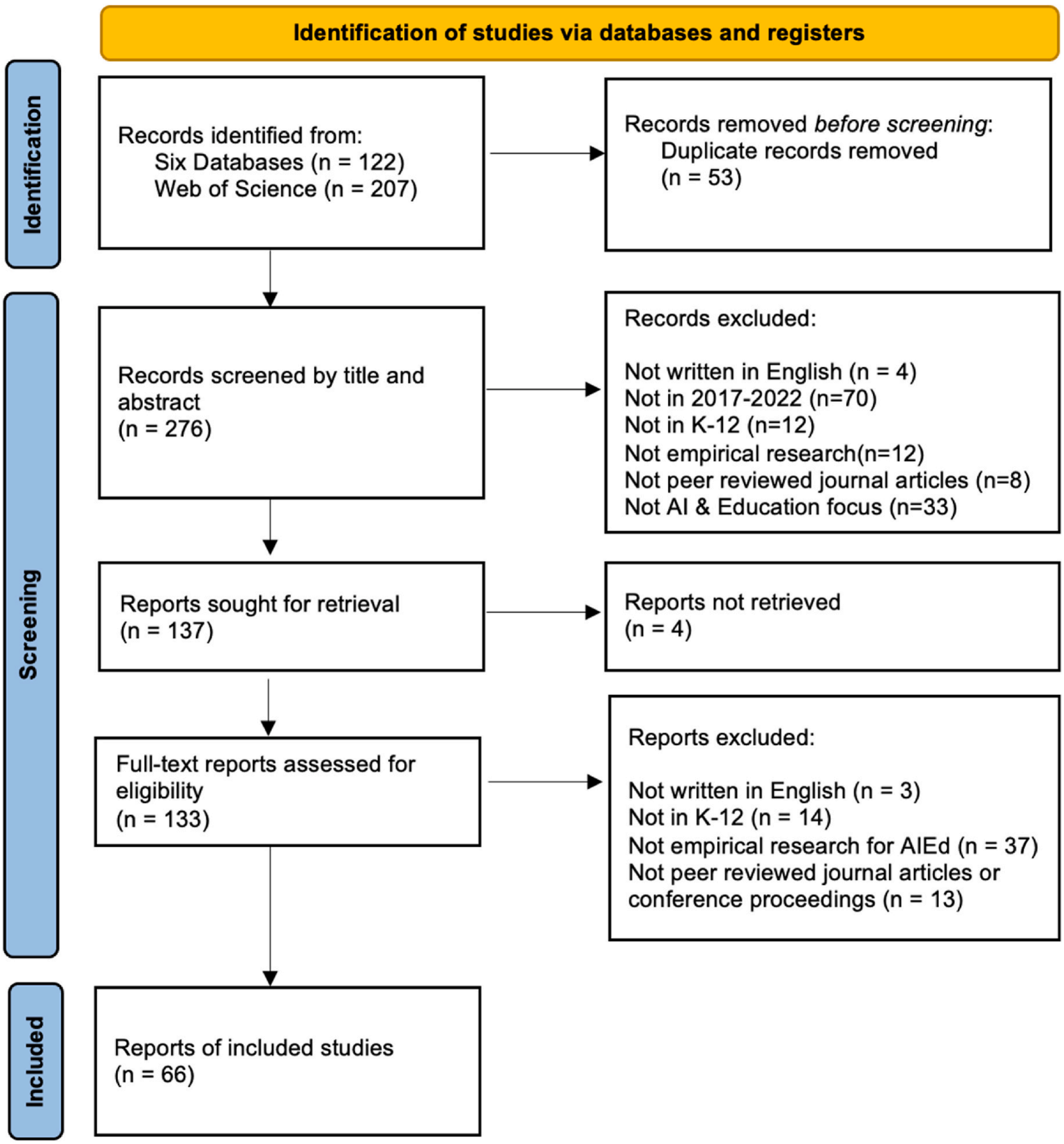

Fig. 2. Process Flow using PRISMA.

broad categories were identified from the AI studies in K-12 education. For data collection methods, research themes, technology applications, and AI use, narrative explanations with descriptions from example studies are included. The researchers met every week during the analysis process to discuss the analysis and resolve any questions.

Reliability is critical to minimize bias in the systematic review process. Some of the strategies that we used in this review were clearly defining inclusion and exclusion criteria, using a standardized coding form for data extraction, including multiple reviewer checks, and clearly documenting the methods to allow for replication. Validity is critical to provide meaningful results. Some of the strategies that were used in this systematic review to enhance validity included clearly defining research questions, being transparent in reporting of methods, and utilizing several databases to identify studies.

3. Results

The results section evaluates publication trends; research methodology components; AI research themes, research methods, and technology applications; and use of AI by teachers and students.

3.1. Publication trends

The first research question focused on publication trends of AI in K-12 education. Several trends were identified including research trends by year, publication type, and nation where the research was conducted. When examining publication years in this review, more articles were published on AI in K-12 education in 2019 and later with 2021 publishing 21 articles. For the year 2022, only a partial year of publications were used until May 2022 due to the time parameters of the search. Before 2018, only a few articles were found in AI in K-12 education. Fig. 3 shows a line graph with the publication years.

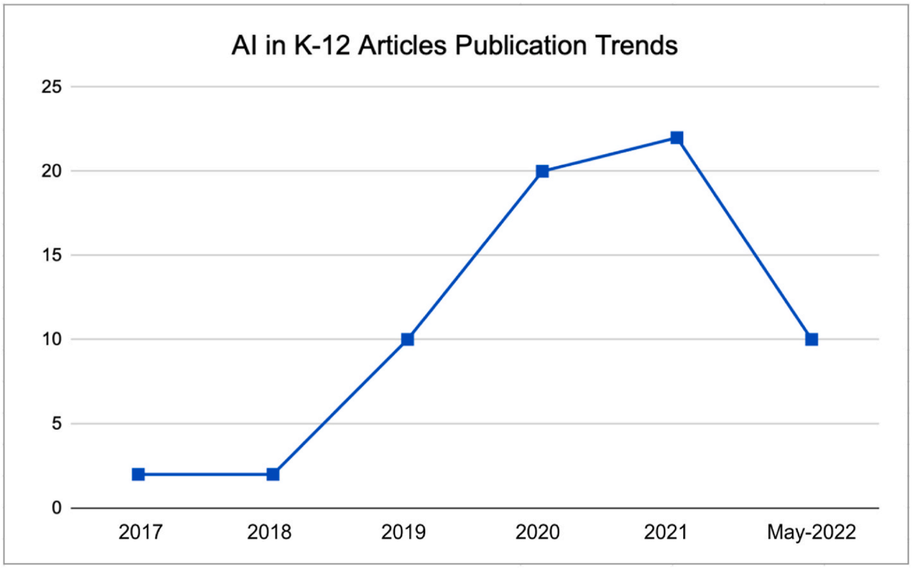

Fig. 3. Publication trends in years.

The top three publications that published research on AI in K12 were Proceedings of The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21) (n = 5), Institute of Electrical and Electronics Engineers (IEEE) Access (n = 2) and Künstliche Intelligenz (n = 2). The remaining 57 articles were in varied journals and conference proceedings. Among the total 66 publications, there were journal publications (n = 36), and conference proceedings (n = 30).

The research studies were conducted in various countries in the world. The United States led the list (n = 18) and was followed by China (n = 10). This was followed by Korea (n = 4), Brazil (n = 3), Spain (n = 3), Austria (n = 2), Denmark (n = 2), Finland (n = 2), and Turkey (n = 2). There were five studies that were also conducted in multiple countries. The remaining studies were conducted in varied countries.

3.2. Research methodology components

Research question 2 focused on the research methodology components used in AI in K-12 research. Various methodology components were identified in this review. Findings of research methods, research setting, research participants, research methods, and data collection are discussed. The majority of the studies used a qualitative method format (n = 28) and was followed by quantitative (n = 24), mixed methods (n = 14). In most situations, both qualitative and quantitative data were collected mostly through surveys, assessments, interviews, and observation. Qualitative Method and different types of Mixed-Method were mainly used to analyze the data.

3.2.1. School level, research participants and subject

Among the studies that were published, most of them were in high school (n = 45), followed by middle school (n = 32), and elementary school (n = 13). One study was conducted in a preschool setting (n = 1) and predicted student performance at age 8. Most of the studies in this review included K-12 students (n = 53) as the research participants. This was followed by studies that included in-service teachers (n = 15), and school administrators (n = 5). One study included counselors, employers, parents, and secondary data analysis (n = 1). Most of the studies on AI in K-12 researched Computer Science programs (n = 26); however, some studies (n = 10) were conducted for multiple subjects. Studies involving multiple subjects analyzed AI methods for predicting students' academic and general performance, including grades in various subject areas. Other studies were conducted in art, biology, biomedicine, engineering, foreign language, math, science, social studies, and writing.

3.2.2. Data collection methods

A variety of data collection methods were used in research studies focusing on AI in K-12 education. The most used data collection methods were survey, testing/assessment, and interviews. This was followed by observation, log data, video/recordings, discussion, photography, and EEG. Table 3 summarizes the frequency of data collection methods used.

Table 3
Data Collection methods.

Data collection method	Example studies	Frequency
Survey	● Xia et al. (2022) ● Lee (2022) ● El-Hajji, Obeid, Hammo, and Al Fayez (2021) ● Rebai et al. (2020)	47
Testing/Assessment	● Tuba and Pelin (2022) ● Harvey and Kumar (2019) ● Hsu, Abelson, and Van Brummelen (2022) ● Fernández-Martínez, Hernán-Losada, and Fernández (2021) ● Xie et al. (2020)	22
Interview	● Koren, Erell, and Sârb (2021) ● Sansone (2019) ● Southgate (2019) ● Chiu (2021) ● Lee et al. (2021) ● Tyson and Sauers (2021)	15
Observation	● Wan, Zhou, Ye, Mortensen, and Bai (2020) ● Bilstrup et al. (2020) ● Southgate (2019) ● Chai et al. (2021)	7
Log data	● Li, Ding, and Liu (2020) ● Hung, Rice, Kepka, and Yang (2020) ● Haendchen Filho, Concatto, do Prado, and Ferneda (2021). ● Zhuang and Gan (2017)	6
Video/Audio Recordings	● Bilstrup et al. (2020) ● Lin and Van Brummelen (2021) ● Kaspersen et al. (2021)	6
Discussion	● Ishikura, Takeda, and Iwashita (2020) ● Hung et al. (2020) ● Tedre et al. (2020)	5
Photography	● Bilstrup et al. (2020) ● Kaspersen et al. (2021)	2
Electroencephalogram	● Rasheed et al. (2021)	1

3.2.2.1. Survey. AI research studies evaluated general use of AI in K-12 settings using data collected through surveys. For instance, Xia et al. (2022) processed and modeled the dataset from the survey of Sixth National Sports Facility Census (NSFC, 2014) by a combinatorial unsupervised ML clustering approach (Fuzzy *C*-means algorithm) to evaluate sports facilities conditions in mainland China. Lee (2022) employed the micro-level data from the survey of 2018 Programme for International Student Assessment on a ML method to investigate what drives the academic performance of Chinese urban and rural secondary schools. Zafari, Sadeghi-Niaraki, Choi, and Esmaeily (2021) collected student data from one-by-one online surveys to evaluate high school student performance based on a ML-based system. El-Hajji et al. (2021) conducted an online survey to collect the student data for building a predictive ML model based on supervised learning algorithms to determine whether the international program is fit for students or not. Rebai, Yahia, and Essid (2020) extracted data from the survey of Program for International Student Assessment 2012 to predict secondary schools' performance in Tunisia using a graphically based ML approach.

3.2.2.2. Testing/assessment. AI research studies also used testing/assessment data to predict student performance or evaluating student learning on AI. Tuba and Pelin (2022) collected the testing data to estimate high school entrance examination success rates using a ML model - Beta Regression. Harvey and Kumar (2019) used assessments reported on student performance on state standardized tests to predict student performance using a predictive ML model based on three classifiers (Linear Regression, Decision Tree, and Naive Bayes). Hsu et al. (2022) conducted pre-test and post-test for measuring the learning effectiveness to investigate the effects on secondary school students of applying experiential learning to the Conversational AI Learning Curriculum. Fernández-Martínez et al. (2021) conducted quizzes and Situational Motivation Scale tests before and after participating in the AI workshop to evaluate students' motivation in learning AI. Xie et al. (2020) carried out tests (critical thinking scale, transfer application ability scale and problem-solving ability scale) before and after the mathematics teaching to investigate the effects of AI-supported intelligent environment in promoting math literacy, academic performance, and problem-solving ability.

3.2.2.3. Interview. AI research studies also used interviews to develop prediction models, developing curriculum, identifying how to support teachers, and investigate AI adoption and implementation in schools. Koren et al. (2021) interviewed the co-founders of Solfy programs to investigate the development and effects of AI didactic support for updating school music education. Sansone (2019) used the interview data from High School Longitudinal Study of 2009, which is a panel micro study interviewing around 21,440 students in 9th grade from about 940 participating schools to develop a more precise ML model in predicting high school dropouts (High School Longitudinal Study of 2009 (HSLS:09), 2009). Southgate (2019) deployed the AI technology - Virtual Reality in high school science and interviewed students and teachers about their experiences with VR to investigate how AI technology affects students' engagement and learning outcomes. Chiu (2021) incorporated the interview data from 24 teachers in developing a design model of AI K-12 curriculum. Lee et al. (2021) interviewed with elementary school teachers to understand how best to support them in integrating AI into their classrooms. Tyson and Sauers (2021) included structured interviews with seven school leaders to investigate their adoption and implementation of AI in school.

3.2.2.4. Observation. AI research studies also used student and classroom observational data. Wan et al. (2020) observed students' learning behaviors while they interacted with the SmileyCluster, a learning environment that assists students in learning an entry level ML technology. Bilstrup et al. (2020) recorded observations on field notes regarding students' designing process and conversations about ML during the card-based design workshop that was designated for teaching students AI ethical dilemmas. Southgate (2019) used the observations by the teacher to support and validate the students' perspective on their learning with AI technology on Virtual Reality for deeper learning in school science. Chai et al. (2021) observed classroom behaviors to interpret the perceptions of and behavioral intentions towards learning AI in primary school students.

3.2.2.5. Log data. AI research studies used log data for prediction and evaluating performance. Li et al. (2020) collected click-stream log data as the features of the ML model for effectively predicting dropouts. Hung et al. (2020) extracted data from the Blackboard activity accumulator and the server logs to analyze student online behaviors with ML and deep learning algorithms. Haendchen Filho et al. (2021) used the corpus log data extracted from Brasil Escola portal to evaluate the performance of automated essay scoring using different ML methods. Zhuang and Gan (2017) extracted most of the log data from CPS database - a private SQL server database to make enrollment predictions in Chicago Public School using a ML approach known as Conditional Logistic Regression.

3.2.2.6. Video/audio recordings. AI research studies used video/audio recordings in their analysis. Bilstrup et al. (2020) collected sound recordings and video recordings of the whole card-based design workshop designed for teaching students AI ethical dilemmas. Lin and Van Brummelen (2021)'s dataset consisted of the audio recordings of the entire workshop of engaging teachers to co-design the integrated AI curriculums. Kaspersen et al. (2021) collected audio recordings of students' intra group discussions during group work when using VotestratesML - a high school learning tool for exploring ML and its societal implications.

3.2.2.7. Discussion. A few researchers studying AI used discussion board data. Ishikura et al. (2020) extracted texts posed on discussion boards by school refusal students and classified their worry and concerns using a ML classifier. Hung et al. (2020) used the text on the online discussion forum to construct early warning models with ML and deep learning algorithms. Tedre et al. (2020) collected discussions during the students' co-creation of ML-based solutions on data intensive analysis practice.

3.2.2.8. Photography. Two researchers used photography as the data in their research studies on AI. Bilstrup et al. (2020) captured students' learning behavior data using photography during the card-based design workshop while designing a ML system and reflecting on the ethical considerations of implementing it into the world. Kaspersen et al. (2021) collected intervention data through photography during the students using VotestratesML - a high school learning tool for exploring ML and its societal implications.

3.2.2.9. Electroencephalogram (EEG). Finally, in the study by Rasheed et al. (2021), EEG data was collected from children who were four years old and was used in predicting academic achievement at age 8 among rural children in Pakistan by a ML classifier (K- Nearest Neighbor).

3.3. AI research themes

The third research question focused on identifying AI research themes in K-12 education focused research studies. Seven AI research themes were identified and are summarized in Table 4.

3.3.1. AI as a predictor and indicator of academic behavior or academic performance

AI is used to predict academic performance in a wide range of subject areas such as math and languages. AI can be used to predict possible

Table 4
AI research themes.

Themes	Frequency	Example studies
AI as a Predictor and Indicator of Academic Behavior or Academic Performance	25	● Hu, Dong, and Peng (2022) ● Tuba and Pelin (2022) ● Cruz-Jesus et al. (2020) ● Jarbou, Won, Gillis-Mattson, and Romanczyk (2022) ● Mnyawami, Maziku, and Mushi (2022) ● El-Hajji et al. (2021) ● Ishikura et al. (2020) ● Dubey and Mani (2019) ● Zhuang and Gan (2017)
AI Curriculum Design	20	● Van Brummelen, Heng, and Tabunshchyk (2021) ● Park and Shin (2021) ● Kaspersen et al. (2021) ● Wu, Li, Li, Zhang, and Wu (2021) ● Choi and Park (2021) ● Zhang et al. (2021) ● Henry, Hernalesteen, and Collard (2021) ● Hsu et al. (2022) ● Fernández-Martínez et al. (2021) ● Williams, Kaputsos, and Breazeal (2021) ● Chiu (2021)
Integrating AI in Various Subjects	11	● Koren et al. (2021) ● Ghareeb Ahmed Ali (2020) ● Southgate (2019) ● Oskotsky et al. (2022) ● Malach and Vicherková (2020) ● Xie et al. (2020)
Evaluation of AI Education (attitudes, intentions, perceptions)	5	● Bochniarz, Czerwiński, Sawicki, and Atroszko (2022) ● Demir and Güraksin (2022) ● Chai et al. (2021) ● Tyson and Sauers (2021) ● Chai, Wang, and Xu (2020)
Learning Environment & School Operations	5	● Xia et al. (2022) ● Burlig, Knittel, Rapson, Reguant, and Wolfram (2020) ● Muhamedyev et al. (2020) ● Bergeron et al. (2019)
Equity/Ethics/Safety of AI	1	● Bilstrup et al. (2020)

individual and school-wide success for specific academic programs such as a specialized CTE program or an advanced academic program such as International Baccalaureate. Academic behavior trends suggested through AI include variables such as attendance, chronic absenteeism, or risk of school withdrawal. AI can also predict behavior patterns including physical health, mental health, and incidents of violence among individuals and schools, districts, and regions overall. Several studies have examined the impact of AI as a predictor or indicator of academic behavior and academic performance. For instance, Hu et al. (2022) applied a ML model (Support Vector Machine) to discover key contextual factors that collectively affect students' reading performance, which reached a relatively high classification accuracy of 81.94%. Tuba and Pelin (2022) made estimations of high school entrance examination success rates using Beta Regression model and Random Forest model were found to be the best two models with the highest R^2 value in accuracy. Cruz-Jesus et al. (2020) explored the relationship between academic achievement and economic development, employment, and countries' well-being by using different AI techniques, Random Forest model performed the best with an accuracy rate of 87%, 4.65 higher than the average. Jarbou et al. (2022) predicted school attendance for autistic students using ML models - Long Short-Term Memory and Multilayer Perceptron algorithms, Long Short-Term Memory increased the accuracy and recall of short-term absence prediction by 20% and 13%, while the same scores of long-term SA prediction increased by 5% using Multilayer Perceptron. Mnyawami et al. (2022) predicted student dropouts in developing countries using different ML models, results showed that Decision Tree, K-Nearest Neighbors, and Multilayer Perceptron outperformed better than others with the prediction accuracy of Decision Tree at 99.8%, K-Nearest Neighbors at 99.6%, and Multilayer Perceptron at 99%. El-Hajji et al. (2021) built a predictive ML model to determine an international program for students, 80% were successfully predicted by the model. Ishikura et al. (2020) generated classification of worries and consultations with school refusal students using a ML classifier, the accuracy could reach more than 85%. Dubey and Mani (2019) predicted high school student employability using classification-based supervised ML models, the trained predictive models perform better with larger dataset with up to 93% accuracy. Zhuang and Gan (2017) predicted enrollment in Chicago Public School using a conditional Logistic Regression ML model.

3.3.2. AI curriculum design

AI curriculum design involves the development of educational materials and learning environments for teaching AI concepts and skills to students at various levels. These include web-based platforms, project-based workshops, board games, role-playing games and e-books. The curricula target different competencies and objectives, such as teaching ML models and applications, natural language processing applications, or AI ethics. The target audiences include elementary, middle, and secondary school students, as well as teachers who want to introduce AI to their students. Van Brummelen et al. (2021) designed a conversational agent interface and an AI workshop curriculum for teaching students with eight AI competencies. Kaspersen et al. (2021) built VotestratesML, a tool allowing students to build ML models and make predictions based on real world voting data. Wu et al. (2021) created a web-based platform, which provides essential AI learning and exercising components to both students and instructors. Choi and Park (2021) piloted a board game-based gamification curriculum aiming at getting elementary students a better understanding of AI. Zhang et al. (2021) designed an open and interactive e-book for teaching students AI. Henry et al. (2021) designed a role-playing game to let children discover the basic concepts of ML. Hsu et al. (2022) designed an AI curriculum that allows students to learn the application of conversational AI on a block-based programming platform. Fernández-Martínez et al. (2021) designed an AI workshop for middle schools for introducing AI at an early age. Williams et al. (2021) developed "How to Train Your Robot: AI and Ethics Curriculum" for middle school teachers who want to introduce AI to their students.

3.3.3. Integrating AI in various subjects

Integrating AI in various subjects has the potential to enhance teaching and learning by providing new tools, techniques, and perspectives. For instance, AI can be integrated in mathematics education through the development of an intelligent environment, in science education through virtual reality as an example, and in engineering education. AI can be also integrated in music education through an AI didactic support, in language education through an AI application. Koren et al. (2021) integrated an AI didactic support - Solfy for updating school music education. Ghareeb Ahmed Ali (2020) used an AI application for help in developing primary school pupils' oral language skills. Southgate (2019) integrated AI technique - Virtual Reality as an exemplar for deeper learning of high school science. Oskotsky et al. (2022) introduced AI in biomedicine through a virtual summer program for high school students. Malach and Vicherková (2020) integrated automation and robotization as components of AI in mechanical engineering production. Xie et al. (2020) developed an intelligent environment for primary school mathematics.

3.3.4. Evaluation of AI education

AI education has been evaluated by different stakeholders in education, to help identify the potential barriers and opportunities, and inform the design and implementation of effective AI education programs. Several studies have investigated the perceptions and attitudes of students and school leaders towards AI in education. Some studies researched the behavioral intentions of students towards learning AI. And some studies investigated the adoption and implementation of AI by school leaders. Bochniarz et al. (2022) investigated attitudes toward AI among high school students. Demir and Güraksin (2022) studied the perceptions of middle school students towards the concept of AI. Chai et al. (2021) researched on the perceptions of and behavioral intentions towards learning AI in primary school students. Tyson and Sauers (2021) investigated school leaders' adoption and implementation of AI in their schools. Chai et al. (2020) modeled Chinese secondary school students' intentions to learn AI.

3.3.5. Learning environment and school operations

Governments, districts, and communities around the world use AI to evaluate the impact of a school building's design, condition, and location. AI can analyze the geographic location and evaluate the impact a school's location has on its success and effectiveness. This data could inform necessary support for urban and rural schools. AI can also evaluate the conditions of a school building and its facilities and provide suggestions impacting budget and operations. For instance, AI has been used to evaluate a building's overall energy efficiency. Additionally, AI can evaluate the design of a school building itself and suggest how the design affects school performance. The use of AI can guide school operations in areas such as budget decision making, growth planning, and transportation routing. It can also be used as a resource for sports and extracurricular activities. For instance, AI has been utilized to live stream school sports. It can also enhance and enforce safety protocols. Xia et al. (2022) used an unsupervised ML model, Fuzzy *C*-means algorithm to evaluate sports facilities condition in primary school. Burlig et al. (2020) evaluated school energy efficiency using ML models. Muhamedyev et al. (2020) developed a multi-criteria decision support system based on AI techniques to improve the quality of school education. Bergeron et al. (2019) modeled high school sport concussion symptom resolve using ML models.

3.3.6. Equity/ethics/safety of AI

Several issues surround the equity, ethics, and safety of AI. On the one hand, AI can be used to improve equity through compiling and evaluating aligned with demographics. This data can help school leaders advance equity initiatives for historically underserved groups. On the other hand, determining whether there may be some level of machine-bias and whether this can contribute to further marginalization is a concern. There are also concerns of using AI in K-12 because of sensitive student data. Bilstrup et al. (2020) investigated the students' perceptions of AI ethics from their design decisions. Through workshops in four high school classrooms, they provided opportunities for students to reflect on ethical dilemmas related to technology and their design decisions when they were designing ML applications. This made the high school students be accountable in their choices related to ethical issues (Bilstrup et al., 2020).

3.4. AI methods

Research question 4 focused on AI methods used in K-12 AI research studies. The AI Methods used are grouped into supervised learning, unsupervised learning, and reinforcement learning. These methods are summarized in Table 5. In addition, the specific algorithms that were used were also coded.

Table 5
AI methods.

Machine Learning	*Specific Algorithms*	Frequency	Example studies
Supervised Learning (n = 79)	Random Forest	14	Rebai et al. (2020)
	Support Vector Machine	12	Hu et al. (2022)
	Decision Tree	10	El-Hajji et al. (2021)
	Logistic Regression	9	Zhuang and Gan (2017)
	K-Nearest Neighbors	6	Rasheed et al. (2021)
	Linear Regression	5	Lee (2022)
	Naive Bayes	5	Ishikura et al. (2020)
	Gradient Boosting	5	Costa-Mendes, Oliveira, Castelli, and Cruz-Jesus (2021)
	Beta Regression	1	Tuba and Pelin (2022)
	Neural Network (specific algorithms not indicated)	4	Ni et al. (2020)
	Multilayer Perceptron*	3	Mnyawami et al. (2022)
	Recurrent Neural Network*	2	Wang and Wang (2018)
	Long Short -Term Memory*	2	Jarbou et al. (2022)
	Radial Basis Function Network*	1	Bergeron et al. (2019)
Unsupervised Learning (n = 2)	Fuzzy *C*-Means Clustering	1	Xia et al. (2022)
	Specific algorithms not indicated	1	Oskotsky et al. (2022)
	K-Means Clustering	0	
	Mean-Shift Clustering	0	
Reinforcement Learning (n = 0)	Monte Carlo Methods	0	
	Q-learning	0	

Algorithms above indicated with an asterisk (*)s belong to the family of artificial neural networks, which is of significant interest in the field of ML. ANNs consist of mathematical representations of connected processing units called artificial neurons, which mimic the neurons in human brains, and the neurons are organized into networks with different layers which construct the architecture for signal transmits of the learning process (Janiesch, et al., 2021). Because of their flexibility structure, which enables them to be adapted and modified for a wide range of contexts across all three types of ML (supervised learning, unsupervised learning, and reinforcement learning) (Janiesch, et al., 2021). The artificial neural networks mentioned in Table 5 were open-coded from our review articles and are mainly used for supervised learning.

3.4.1. Supervised learning

Twenty-six articles utilized supervised learning approaches. As mentioned in section 1.1, supervised learning is primarily used for solving two kinds of problems, regression (predicting numerical values) and classification (predicting categorical labels) (Han et al., 2022; Janiesch et al., 2021; Sarker, 2021). There are ten articles dealing with regression problems, and 16 articles for classification. For regression problems, the numerical values predicted by specific algorithms included test scores (Costa-Mendes et al., 2021; Harvey & Kumar, 2019; Lee, 2022; Masci, Johnes, & Agasisti, 2018; Rebai et al., 2020), high school entrance success rate (Tuba & Pelin, 2022), enrollment numbers (Zhuang & Gan, 2017), risk value of future status in online learning environment (Li et al., 2020), quality score of education system (Muhamedyev et al., 2020), and specification choice in electric energy savings (Burlig et al., 2020).

Four articles use only one algorithm aiming to predict the specific value and/or identify the key factors that influence that value. For example, Burlig et al. (2020) found the electricity specification choice set with a central estimate of 60% would contribute a great energy savings by using a Linear Regression algorithm; Through Gradient Boosting algorithm, Muhamedyev et al. (2020) found location of the school, higher category teachers, more readers in the library, sports facilities, technical support and other clubs would contribute to schools of good quality. Masci et al. (2018) found that in almost all the countries, the most important variables that influenced math scores included students' self-reported anxiety toward tests, socioeconomic, and self-reported motivation.

Six articles compared many algorithms to identify the algorithm with the best performance. For example, Harvey and Kumar (2019) compared the performance among Linear Regression, Decision Tree, and Naive Bayes in predicting SAT math score, and Naive Bayes showed the highest accuracy; Costa-Mendes et al. (2021) made a comparison between traditional statistic model - multilinear regression model and a series of ML algorithms - Random Forest, Support Vector Machine, Artificial Neural Network, the results showed that ML algorithms attained a higher level of predictive ability.

For classification problems, many categorical labels was set and predicted such as high academic performance vs. low academic performance (Hu et al., 2022), pass vs. fail in academic test (Rasheed et al., 2021), dropout vs. no dropout (Chung & Lee, 2019; De Melo et al., 2017; Ni et al., 2020), employed vs not employed (Dubey & Mani, 2019), IB Program vs. Other International Program (El-Hajji et al., 2021), different school types in Dutch (Niemeijer, Feskens, Krempl, Koop's, & Brinkhuis, 2020), and the reasons for truancy - Friend, Teacher, Family, and Study (Ishikura et al., 2020).

Five articles used only one specific algorithm, the other eleven articles included multiple algorithms. Hu et al. (2022) found Support Vector Machines and was able to effectively differentiate high- and low-performing students with respect to reading performance and they identified gender, motivation, and the number of children's books at home were the key contextual factors impacting student reading performance; Ishikura et al. (2020) used Naive Bayesian in identifying the reasons for truancy and the results showed that the accuracy was more than 70%, and the highest was more than 85%; Chung and Lee (2019) utilized Random Forest in predicting dropout or not and achieved the accuracy of 95%. Dubey and Mani (2019) compared the performance among Logistic Regression, Random Forest, and Support Vector Machine classifiers in predicting student employment and yielded the highest accuracy up to 92–93%; Mnyawami et al. (2022) used Decision Tree, Naive Bayes, Random Forest, Support Vector Machines, Multilayer Perceptron, Logistic Regression and K-Nearest Neighbors to predict students dropout, and they found that student marks (57%), student age (18%), distance (7%) and number of children (5%) are most statistically significant to student dropout compared to father's education (3%), student gender (3%), and means to school (2.5%).

Artificial neural networks were found mostly when dealing with classification problems, such as in the study of Bergeron et al. (2019), they classified three distinct category thresholds of symptom resolution time by using Multilayer Perceptron and Radial Basis Function Network to make estimation of symptom resolve time in high school athletes who incurred a concussion during sport activity. Costa-Mendes et al. (2021) used an artificial neural network without indicating specific algorithms for solving a regression problem - the prediction of the student's high school final grades. Another important finding was that Recurrent Neural Network was found in articles when the data was associated with text and for the use of Natural Language Processing, such as in Wang and Wang (2018), they used Recurrent Neural Network to solve the task of question answering matching in QA system; Haendchen Filho et al. (2021) utilized Recurrent Neural Network in automated essay scoring.

3.4.2. Unsupervised learning

There were only two articles associated with unsupervised learning to do the tasks of clustering, no articles were found for handling dimensionality reduction problems. Xia et al. (2022) used Fuzzy *C*-means to evaluate sports facilities condition in primary school. The condition was clustered into three types. Through analyzing the interclass differences and characteristics among those three types, they found that the location, grade levels, school category, and square meters were considered as the key evaluation indicators of the condition of sport facilities. In the study of Oskotsky et al. (2022), students learned to apply AI in biomedicine curriculum, specifically using unsupervised learning methods to predict whether a patient is COVID positive or negative and predict the severity of the COVID infection.

3.4.3. Reinforcement learning

No articles were identified in this study that focused on AI as a reinforcement learning approach. Burgsteiner, Kandlhofer, and Steinbauer (2016) only introduced the concept of reinforcement learning in their iRobot curriculum rather than using it for solving problems, and therefore was not included in the count.

3.5. AI technology applications

Research question 5 focused on AI technology applications used in AI K-12 research. AI technology applications are summarized in Table 6 and description of studies for each application is included in the following table.

3.5.1. ML model building tools

ML model building tools are web tools that make creating ML models for projects fast and easy with no coding required. Through training a computer to recognize the images, sounds, and poses, users will learn how to build ML models and how it works. Mariescu-Istodor and Jormanainen (2019) built a ML tool for object recognition that was implemented by students to learn the ML mechanism in recognizing objects. Lin and Van Brummelen (2021) leveraged tools like Teachable Machine and ML4Kids for their transparent data training capabilities, which allowed students to go through the steps of the ML cycle. Kaspersen et al. (2021) built VotestratesML, a web application enabling students to collaborate in real time on iteratively building ML models for predicting voter behavior using voter profile data.

3.5.2. Intelligent tutors

An intelligent tutoring system is a computer system that aims to provide immediate and customized instruction or feedback to learners, usually without requiring intervention from a human teacher. Koren et al. (2021) researched on the Solfy, an intelligent tutor that designed as a study and self-practice didactic support for promoting singing and strengthening the music literacy process. Ma et al. (2020) proposed an intelligent system that marks assignments and automatically recommends peer tutors by considering students' performance and

Table 6
AI technology applications.

AI Technology Application	Frequency	Example studies
ML Model Building Tools	9	● Mariescu-Istodor and Jormanainen (2019) ● Wan et al. (2020) ● Lin and Van Brummelen (2021) ● Kaspersen et al. (2021)
Intelligent Tutors	5	● Koren et al. (2021) ● Ma, Hwang, and Shih (2020) ● Wang and Wang (2018) ● Xie et al. (2020)
Chat Bot	4	● Ghareeb Ahmed Ali (2020) ● Hsu et al. (2022) ● Van Brummelen et al. (2021)
Educational Games	3	● Lee et al. (2021) ● Choi and Park (2021)
AI Robots	1	● Chai et al. (2020)
Virtual Reality Devices	1	● Southgate (2019)

relationships. Wang and Wang (2018) created a Q&A system that allows students to directly ask questions through natural language and can directly return accurate answers to students. Xie et al. (2020) developed an intelligent application in primary school mathematics to carry out depth-based teaching practice and promote the development of primary school students' core math literacy.

3.5.3. Chat bot

A chatbot or chatterbot is a software application used to conduct an on-line chat conversation via text or text-to-speech, in lieu of providing direct contact with a live human agent. Ghareeb Ahmed Ali (2020) did research on the language lessons that built on varied oral activities and practices using chatbots, which aimed at developing and fostering students' listening and speaking skills. Hsu et al. (2022) used Conversational VoiceBot as both the learning tool and learning outcome of the conversational AI curriculum. Van Brummelen et al. (2021) built and learned about conversational agents in a remote workshop for enhancing students' understanding of AI and conversational AI competencies.

3.5.4. Educational games

Educational games are games explicitly designed with educational purposes, or which have incidental or secondary educational value. All types of games may be used in an educational environment, however educational games are games that are designed to help people learn about certain subjects, expand concepts, reinforce development, understand a historical event, or culture, or assist them in learning a skill as they play. Lee et al. (2021) developed and implemented PRIMARYAI - a game-based learning environment that centered on applying AI to solve life-science problems for upper elementary classrooms. Choi and Park (2021) deployed board games to promote elementary school students' motivation to learn the principles of deep learning algorithms on image recognition.

3.5.5. AI robots

AI Robots are the artificial agents acting in the real-world environment, which aim at manipulating the objects by perceiving, picking, moving, destroying it. Chai et al. (2020) trained primary school students to control the intelligent fire-fighting robots to search for and navigate to the spot of the fire and investigated their perceptions and intentions toward learning AI.

3.5.6. Virtual reality devices

Virtual Reality is a simulated experience that employs pose tracking and 3D near-eye displays to give the user an immersive feel of a virtual world. Southgate (2019) integrated AI technique - Virtual Reality in high school science classes to illustrate how AI technology can promote engagement and deeper learning for students.

3.6. Use of AI by teachers and students

The use of AI by both teachers and students has been examined in research. In this section, we address research question 6 by introducing how AI is used by teachers followed by how AI is used by students (Tables 7 and 8).

3.6.1. Use of AI by teachers

3.6.1.1. ML model demonstration. AI platforms are used by teachers to demonstrate how ML works. The ML tool in Mariescu-Istodor and Jormanainen (2019) helps teachers to explain ML mechanisms in recognizing objects. Conversational AI Applications of Hsu et al. (2022) facilitates teachers in explaining to students how the chatbot works and how to build a conversational AI interface. SmileyCluster designed by Wan et al. (2020) is a learning platform to assist teachers for

Table 7
AI use by teacher.

Use by Teacher	Frequency	Example studies
ML Model Demonstration	13	● Mariescu-Istodor and Jormanainen (2019) ● Hsu et al. (2022) ● Williams et al. (2021) ● Wan et al. (2020) ● Van Brummelen et al. (2021) ● Lin and Van Brummelen (2021)
Academic Performance Prediction and Intervention	10	● Harvey and Kumar (2019) ● Cruz-Jesus et al. (2020) ● Costa-Mendes et al. (2021) ● Tuba and Pelin (2022)
Academic Behavior Prediction and Intervention	10	● Mnyawami et al. (2022) ● Sansone (2019) ● Chung and Lee (2019) ● Zhuang and Gan (2017) ● Li et al. (2020) ● Rasheed et al. (2021)
Engaging Students	6	● Southgate (2019) ● Hsu et al. (2022) ● Lee et al. (2021) ● Oskotsky et al. (2022) ● Park and Shin (2021) ● Wu et al. (2021)
Automatic Support	4	● Koren et al. (2021) ● Wang and Wang (2018) ● Xie et al. (2020) ● Ghareeb Ahmed Ali (2020)
Advising	3	● El-Hajji et al. (2021) ● Niemeijer et al. (2020) ● Dubey and Mani (2019)
Learning Process Monitoring	2	● Koren et al. (2021) ● Xie et al. (2020)
Automated Grading/Assessment	2	● Haendchen Filho et al. (2021) ● Ma et al. (2020)

Table 8
AI use by students.

AI use by student	Frequency	Example studies
Scientific Discovery Learning	22	● Hsu et al. (2022) ● Choi and Park (2021) ● Fernández-Martínez et al. (2021) ● Williams et al. (2021) ● Zhang et al. (2021) ● Henry et al. (2021) ● Van Brummelen et al. (2021) ● Park and Shin (2021) ● Kaspersen et al. (2021) ● Wu et al. (2021)
Improve Learning Experience	12	● Koren et al. (2021) ● Ma et al. (2020) ● Ghareeb Ahmed Ali (2020) ● Southgate (2019) ● Wang and Wang (2018) ● Lee et al. (2021) ● Lin and Van Brummelen (2021) ● Malach and Vicherková (2020) ● Xie et al. (2020)
Data-Driven Decisions - Program Choosing/Major Choosing/School Type Choosing	3	● El-Hajji et al. (2021) ● Niemeijer et al. (2020) ● Dubey and Mani (2019)

demonstrating K-Means Clustering. Programmable robot in Williams et al. (2021) facilitates teachers in explaining to students the K Nearest Neighbor algorithm and how it is used in ML.

3.6.1.2. Academic performance prediction and intervention. ML and data analysis techniques have potential to support teachers in their efforts to improve student academic outcomes. The use of ML can also assist teachers in understanding the factors that contribute to academic achievement, and in identifying strategies that are most effective for individual students. Harvey and Kumar (2019) introduce teachers with the Naive Bayes techniques for predicting SAT Math scores for high school students. Cruz-Jesus et al. (2020) and Costa-Mendes et al. (2021) help teachers to understand students' academic achievement with various ML techniques such as Random Forest, Support Vector Regression, etc. Tuba and Pelin (2022) supports teachers to predict students' entrance examination success rates using Beta Regression model.

3.6.1.3. Academic behavior prediction and intervention. By leveraging the power of predictive modeling, teachers can identify students who are at risk of dropping out, falling behind, or experiencing violence, and intervene early to provide targeted support. Similarly, predicting enrollment numbers and entrance exam success rates can help schools plan resources and allocate funding more effectively. The study of Mnyawami et al. (2022), Sansone (2019), and Chung and Lee (2019) help teachers in predicting student dropouts using supervised ML. Zhuang and Gan (2017) support teachers in the prediction of enrollment numbers for the next year based on an ensemble ML model. Li et al. (2020) assists teachers in identifying at-risk students in multimodal online environments with a customized ML approach. Rasheed et al. (2021) offers teachers an objective and feasible screening measure to identify at-risk children in the early grades that helps to design appropriate interventions based on a supervised ML classifier (K-Nearest Neighbor). Ni et al. (2020) assists teachers to identify and evaluate risk protective factors for school violence through developing natural language processing and ML technologies to automate the risk assessment process.

3.6.1.4. Engaging students. AI applicants are designed and used by teachers to make the curriculum more interactive and engageable. In the study of Southgate (2019), AI technique - Virtual Reality is integrated in high school science classes for teachers to promote engagement and deeper learning for students. Conversational AI Applications of Hsu et al. (2022) helps teachers in engaging students with experiential learning. A game-based learning environment - PRIMARYAI in Lee et al. (2021) assists teachers to engage students to apply AI to solve life-science problems for upper elementary classrooms. In the study of Oskotsky et al. (2022), integrating AI in the biomedicine curriculum assists teachers to make the traditional science course more engaging for students.

3.6.1.5. Automatic support. AI is used by teachers to provide students with automatic support during their learning process. Deep Neural Networks that performs voice synthesis and analysis in Koren et al. (2021) helps teachers to provide students with didactic support in music education. The Q&A system developed in Wang and Wang (2018) assists teachers to directly return accurate answers to students. The intelligent application in Xie et al. (2020) facilitates teachers to provide immediate support at primary school mathematics to promote the development of students' core math literacy.

3.6.1.6. Advising. AI and ML assists teachers with providing students with evidence-based suggestions like student program selection, school advice, and career guidance. El-Hajji et al. (2021) generates a predictive model based on ML that determines whether the IB program is fit for the student or not, which is used by teachers in advising the suitable program for students according to his/her learning abilities, program requirements and characteristics. Niemeijer et al. (2020) constructs ML techniques in predicting school advice based on students' academic achievement, which is used by teachers to estimate students' abilities and thereby give an indication of whether their school type is suitable for them. Dubey and Mani (2019) use ML to predict high school student employability, which is used for teachers in advising students in choosing local businesses for part-time jobs.

3.6.1.7. Learning process monitoring. AI has the potential to facilitate teaching in monitoring and analyzing the learning progress of students. The Solfy of Koren et al. (2021) - an AI didactic support in music class helps teachers in monitoring students' music literacy process. In the study of Xie et al. (2020), the intelligent application facilitates teachers in analyzing and monitoring the development of primary school students' core math literacy.

3.6.1.8. Automated grading/assessment. AI applications can assist teachers in grading and assessing student work, which can save teachers time and improve the consistency of grading. Automated Essay Scoring systems of Haendchen Filho et al. (2021) assists teachers in classroom assessment both in low and large-scale participation and helps to reduce problems related to the proficiency of the Portuguese-Brazilian language in Brazil. Peer Tutor Recommender System of Ma et al. (2020) reduces teacher workloads and simplifies interactions between teachers and students by providing detailed feedback for every question and reasoning for incorrect answers.

3.6.2. Use of AI by students

3.6.2.1. Scientific discovery learning. AI can be used by students to enhance their scientific discovery learning. The use of AI can help young learners obtain basic ML concepts and technologies that support a diverse range of pattern recognition and inference activities, leading to the discovery and construction of new knowledge across different subject domains. The conversational AI curriculum developed in Hsu et al. (2022) allows young students to connect the application of audio interaction with the Internet of Things or simulative interaction in the block-based programming environment. Choi and Park (2021) proposed a board game-based gamification curriculum helping elementary students get a better understanding of AI before education.

3.6.2.2. Improve learning experience. AI can be a powerful tool for improving the learning experience of students by providing personalized feedback, guidance, and support. Automated Essay Scoring System of Haendchen Filho et al. (2021) provides students with timely feedback on their essays without the need to wait for teachers. In the study of Koren et al. (2021), Solfy is designed as a study and self-practice AI support for promoting singing and strengthening the literacy process, helping students in guided practice of Solfege, and reviewing the results obtained. Ma et al. (2020) proposed a ML based Peer Tutor Recommender System with automated assessment to enhance students' learning performance in computer application operating skills. The Advanced Automated Assessment System uses computer vision technology to evaluate student assignments and instantly return feedback. The AI-based class in Ghareeb Ahmed Ali (2020) provides students with an accurate personalization of their learning according to their requirements. In the study of Southgate (2019), Virtual Reality based on AI techniques in high school classes promoted engagement and deeper learning of sciences for students. The intelligent Q&A system of Wang and Wang (2018) allows students to directly ask questions through natural language and can directly get accurate answers.

3.6.2.3. Data-driven decisions - program choosing/major choosing/school type choosing. AI-powered recommendation systems and predictive models can be valuable tools for students to make data-driven decisions about their education and career paths, leading to more successful and fulfilling futures. El-Hajji et al. (2021) built a predictive ML model to determine an international program which can be used by students to choose the suitable international program for themselves. Niemeijer et al. (2020)' ML techniques can be used by students to estimate their abilities and thereby get an indication of whether their school type is suitable for them. Dubey and Mani (2019) provide an automated approach based on Supervised ML classifiers for high school students to see if they have enough employable skills and choose wisely in part-time job applications.

Fig. 4 includes a summary of the findings for AI themes, methods,

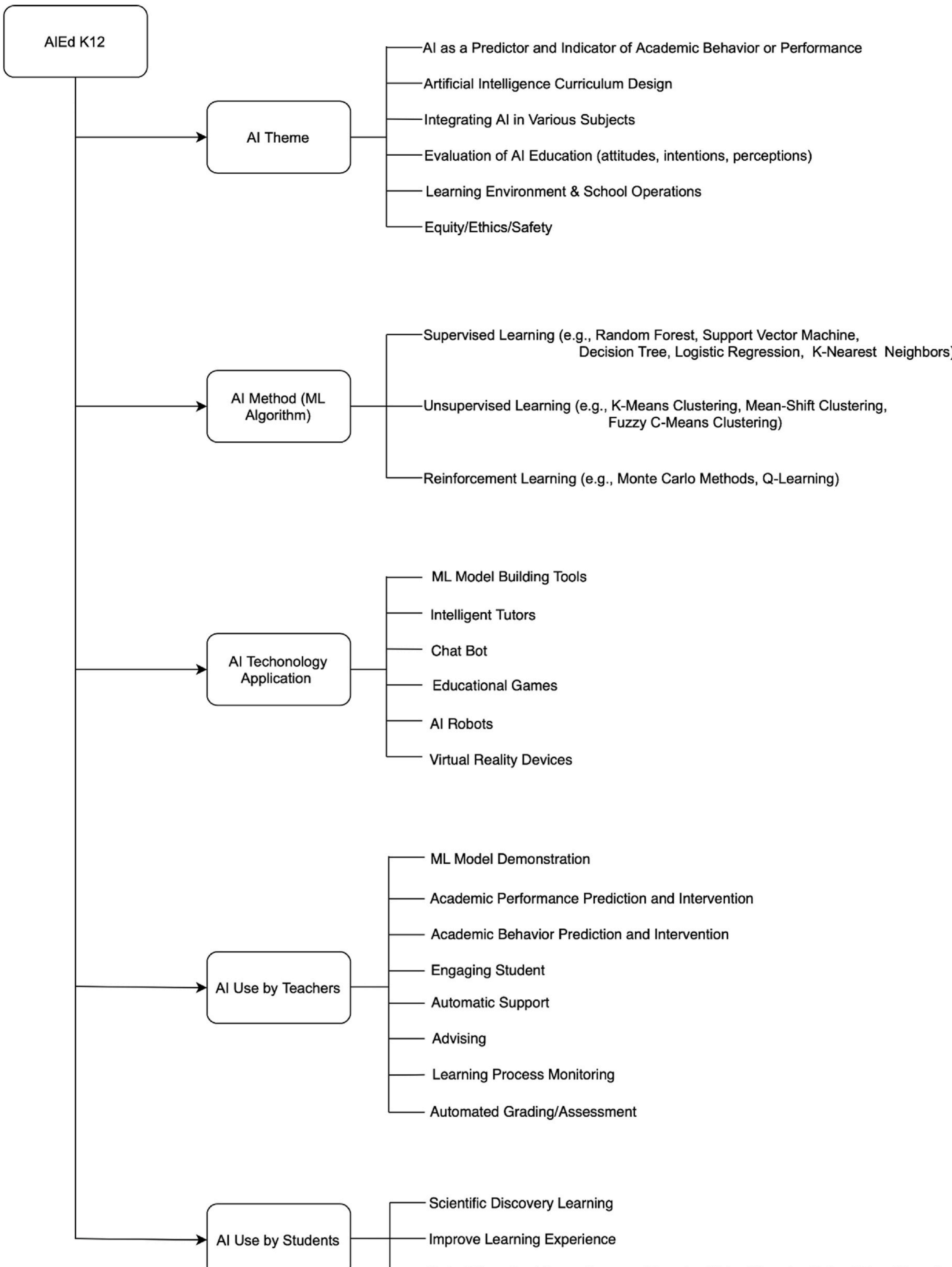

Fig. 4. Summary of findings.

applications, and AI use by teachers and students.

4. Discussion

In this section, we discuss the findings based on the research findings.

4.1. Publication trends and research methods

Research question 1 focused on overall research trends. In regards to research trends, this review saw an increase in AI in K-12 publications over time. The year 2022 did not have a full data set as publications were only reviewed through May 2022. Similar to this review, Wang and Cheng (2022) also report an increase in publications after 2018. The increase in published research on AI in K-12 after 2018 seems to align with Marquesas' study (2020) which revealed a sharp trend in AI Instructional Units starting in 2018. By 2019, Marquesas had noted 30 IUs in AI. Because AI in K-12 education research is still emerging as a publication focus, 30 of the 66 publications from this systematic review were conference proceedings. Given the increase in publications from 2018 onward, and given the late 2022 release of Microsoft's ChatGPT, it is anticipated journal publications will continue to increase in upcoming years. This study accounted for country trends in AI publication, like the Wang and Cheng study (2022) and Zhang and Aslan study (2021). All respective studies reported most of the publications were from the United States, China, and South Korea, specifically. To answer research question 1, this review found an increase in research in AI in K-12 with the three countries, United States, China and South Korea, leading the number of publications. This shows the growing importance of AI in K-12 education.

Research Question 2 focused on research methodology. Qualitative research presents as the leading form of research methodology. Comparable literature reviews including the Zhai et al. (2021) systematic review have focused on types of methodology, but one distinguishing feature of this systematic review is it not only includes study method, quantitative, qualitative, and mixed method, but it also coded studies according to data collection method. Comparable systematic reviews did not synthesize articles according to data collection methods. In this review survey and testing/assessment data were most commonly used data collection measures followed by interviews. Given the various data collection methods applicable to AI research, future studies are recommended to triangulate data sources to better inform AI integration and application in K-12 education. Research question 2 findings show the need for more mixed-methods studies, more studies in elementary school settings, and more studies focusing on the use of AI in other subject areas of research.

4.2. AI themes, research methods, and technology applications

Research question 3 focused on research themes in the literature. In the studies reviewed, the top three themes were AI as Academic Performance/Behavior Predictor or Indicator, AI Curriculum Design, Integrating AI in Various Subjects. AI was used to directly support learning for K-12 students, notably as a performance predictor or indicator. The capability to predict performance in specific programs, curriculum, and grade levels could guide vertical planning and suggest necessary scaffolds in learning to support student advancement. AI is used in K-12 settings to suggest a student's potential success in a specific academic program such as a CTE program track or advanced educational track like International Baccalaureate. AI has been used to assess student-readiness for subsequent grade level programs in subject areas such as math and computer science. The Crompton and Burke study (2022) focused specifically on articles that could support K-12 educators and students, primarily regarding student monitoring (comparable to performance/behavior), management, grading, data-driven decisions (Crompton and Burke, 2022), but the study did not distinguish articles focused on subject areas, curriculum design trends, or attitudes and intentions. The Kim et al. publication (2020) generalized AI use for teachers but specified students gaining competencies in AI in four areas: problem solving, reasoning, learning, and recognizing (Kim et al., 2021, May). Like the Kim et al. publication, this systematic review also identified articles focused on AI integration in specific student subject areas. This study also emphasized curriculum design trends and how students learn about AI in classroom settings and in various subject areas. Identifying AI curriculum as a part of diverse subject areas, besides computer science, was an area of research and focus that Moore et al. recommended (2022).

Least studied themes were evaluation of AI education, learning environment and school operations, and equity/ethics/safety indicating the need for additional research on these topics. A few studies examined how AI can promote the overall well-being of students and schools and forecast barriers to learning including incidents of school-violence as well as at-risk behaviors including chronic absenteeism or the potential of school withdrawal. It can also provide data and analysis on mental health issues or physical health issues to inform care measures to improve student outcomes. AI can also acquire information regarding teacher and school leader performance. AI has studied variables leading to teacher performance, success, and attrition, such as demographics and classroom practice. Research Question 3 findings show the most studied themes as AI as a predictor and indicator of academic behavior or academic performance, and AI curriculum design, and the least studied themes to be equity/ethics/safety of AI. The findings from the most studied themes can be used to integrate them into practice while the least studied themes show the gap in current research for future studies to address.

Research question 4 focused on AI methods used in K-12 research. The majority of the studies utilized supervised learning approaches in educational research. Supervised learning is well-suited for solving regression and classification problems in the context of education, such as predicting test scores and dropout. Similar results found in Wang and Cheng (2022), they indicated prediction was a prominent research focus of learning from AI. Many studies compared the predicting performance among different learning algorithms rather than merely using one specific algorithm. Some specific algorithms of supervised learning like Random Forest, Support Vector Machine, Decision Tree and Logistic Regression were mostly used. AI methods suggest how AI can support general teaching and learning processes for teachers and students (Crompton and Burke, 2022; Kim et al., 2021, May). Large number of ML algorithms provide educators with choices to select from in terms of different situations, and help to better monitor student behavior and performance to make predictions and guide intervention and support. Comparatively, unsupervised learning and reinforcement learning were rarely found in our review articles, which suggests that there is an opportunity for further exploration of its potential in this domain. Research question 4 shows that supervised learning was the most used ML algorithm and can be used by practitioners. However, it also shows the gap in using unsupervised learning and reinforced learning and provides recommendations for future studies.

Research question 5 focused on technology applications in AI in K-12. AI Model building tools, and intelligent tutors were the technology applications studied the most. This implies that AI was no longer a mere embedded technique in the intelligent tutors and expert systems that was highlighted in Guan et al. (2020), Zhang and Aslan (2021) and Wang and Cheng (2022). Instead, AI started to be designed and built for demonstrating students how ML learning model works and be used. Chatbot was still a popular AI technology for students to learn about and engage with, which was also found in Zhang and Aslan (2021), Wang and Cheng (2022) and Kim et al. (2020). Research question 4 found AI model building tools and intelligent tutors were most studied in AI K-12 research whereas AI robots and virtual reality devices are least studied. This shows the need for additional research to focus on AI robots and virtual reality devices to be examined in AI K-12 research.

4.3. AI use among teachers and students in K-12

Research question 6 identified various AI methods and technology applications used in K-12 learning environments to support teachers and students. Most AI applications were used by teachers to demonstrate ML models, and most of AI methods are for performance prediction and intervention which include academic performance like testing scores and academic behavior like dropouts. Celik et al. (2022) in their review found that teachers used AI for planning such as understanding student needs, for implementation such as providing immediate feedback, and for assessment such as automated essay scoring. The facilitate teaching finding and the assessment finding aligns with how teachers use AI based on their review. In addition, the Crompton and Burke (2022) review found Student Monitoring, Group Management, Automated Grading, and Data-Driven Decisions as four themes on how educators use AI. These findings also align in terms of learning process monitoring, automated grading and automatic support.

From the student perspective, AI was used for Scientific Discovery Learning and Improving Learning Experience. The Crompton and Burke (2022) review identified AI tutors, to extend student thinking and Just-for-You-Learning as student affordances from the use of AI. These slightly overlap with our findings though the exact names of themes were different. While Crompton and Burke (2022) review had data-driven decisions in the teacher use, in our review this was identified in student use specific to program, major or school selection. Research question 6 on use by teachers focused on ML model demonstration to the most studied which shows the importance of ML modeling in AI K-12 curriculum and has implications for designers and teachers. On the other hand, there were the least number of studies on learning process monitoring and automated grading/assessment and this is an area of need that future researchers could continue to research. For the students, scientific discovery of learning was the most use of AI, and data driven decision making was the least examined as student use of AI. Similarly, practitioners can continue to implement AI in K-12 for scientific discovery, and based on the gap researchers can frame studies on use of AI in K-12 for data driven decision making.

4.4. Implications for practice

Timely adoption of emerging trends. The overall examination of AI publication trends, themes, technical applications, and usage can equip practitioners with up-to-date insights, facilitating the adoption of innovative practices in the evolving technological landscape. This timely adoption is instrumental in enhancing teaching and learning experiences, ensuring that education remains aligned with the dynamic emergence of AI in K-12 settings.

AI integration and AI curriculum design. Educators can explore and incorporate the identified AI methods, AI applications into teaching practices, and design AI curriculum to foster innovation and prepare students for an AI-driven future. Insights into how AI is used by teachers and students guide the development of learning activities and training programs, ensuring effective utilization and promoting positive learning outcomes.

AI research design. Researchers can use the insights gained from the analysis of research methodology components to make informed decisions about the design of studies. Understanding the research methods, school levels, and participant demographics and the gaps can guide the selection of appropriate methodologies for investigating AI in K-12 education.

Policy development and resource allocation. Policymakers can use publication trends and AI focus on themes, technical applications, and usage to inform the development of policies related to AI in education. Understanding AI research and interest areas can guide resource allocation and policy initiatives to support the integration of AI technologies in K-12 classrooms.

4.5. Limitations and future directions for research

While this systematic review addresses a few important research questions, it also has some limitations. It analyzed only five years of data as research on AI is moving quickly. Since the introduction to ChatGPT, several research studies and special issues are in progress which have not been included as part of this review. Also, only specific search terms were used which may not have identified all research on AI in K-12. Future research studies are recommended to use broader keywords to include "chatbots", "intelligent tutors", intelligent agents", "conversational agent" for AI technology, and "primary school", "elementary school", "secondary school", "high school", "kindergarten", "middle school" and "junior high school" for the K-12 school setting. Also, only articles published in English and in databases that the authors had access to were analyzed. For example, databases such as Association for Computing Machinery (ACM) and IEEE were not used. This has excluded articles on AI in K-12 published in other languages and journals that are not indexed in these databases. Although three researchers were involved in this project with one doctoral student researcher supporting the screening process and another doctoral student researcher supporting coding and attending weekly meetings with the faculty researcher, the team did not calculate interrater reliability.

Educational researchers are able to review the research that has been published in the last five years, and also identify the gap and the themes that are not studied much. Researchers are also recommended to analyze findings by grade level. Additionally, more research regarding ethical and equity concerns with AI among diverse student populations, similar to the Bilstrup et al. (2020) article, can provide insights as AI use continues to scale and expand.

Statements on open data and ethics

The list of articles used in the systematic review is denoted using an asterisk in the reference list.

CRediT authorship contribution statement

Florence Martin: Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Conceptualization. **Min Zhuang:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis. **Darlene Schaefer:** Writing – review & editing, Writing – original draft, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

List of Acronyms

AI	Artificial Intelligence
AIEd	Artificial Intelligence in Education
ML	Machine Learning
EEG	Electroencephalogram

ACM Association for Computing Machinery
IEEE Institute of Electrical and Electronics Engineers

Appendix

Appendix A. Codebook with descriptions

Codebook Criteria	Description
Study ID	Each article was identified a unique ID during the screening process
Exclusion Rationale	If an article was not coded, the rationale was documented
Publication Year	Publication year was coded
Journal Name	Journal title was open coded
Author Name	Author names were included in the coding but not used for analysis
Country	Country was coded; both closed and open items
School Level	School level was coded as Preschool, Elementary School, Middle School, High School and Not Reported/Not Applicable, Other
Research Participants	Research participants were coded as K-12 Students, K-12 Special Need Students, In-service Teachers, Pre-service Teachers, Site Facilitators, District and State Level Administrators, School Administrators, Principals, Technology Facilitators/Coordinators, Librarians, Counselors, Employers, Parents, Secondary Data Analysis, Not Reported/Not Applicable, Other
K-12 Subject	Subject was coded as Reading, Writing, Foreign Language, Social Studies, History, Geography, Biology, Biomedicine, Chemistry, Math, Computer Science, Art, Science, Engineering, Multiple Subjects, Not Reported/Not Applicable, Other
Research Methodology	Research methodology was coded as Quantitative, Qualitative, Mixed-Method, Not Applicable
Data Collection Methods	Data collection methods were coded as Eye Tracking Device, Electroencephalogram, Accelerometer, Video/Recordings, Photography, Interview, Survey, Observation, Discussion, Log data, Testing/Assessment, Other
AI Methods in K-12 Research	AI Methods in K-12 Research were coded as Supervised Learning, Unsupervised Learning, Reinforcement Learning, and Not Applicable/Not Reported, Other
AI K-12 Research Themes	AI K-12 Research Themes were coded as AI Curriculum Design, Integrating AI in Various Subjects, AI as a Predictor and Indicator of Academic Behavior or Academic Performance, Evaluation of AI Education (attitudes, intentions, perceptions), Equity/Ethics/Safety of AI, Learning Environment & School Operations, and Not Applicable/Not Reported
AI Technology Applications	AI Technology Applications were coded as Chat Bot, Intelligent Tutors, Personalized/Adaptive Learning Systems, Virtual Reality Devices, Augmented Reality Devices, AI Robots, Educational Games, ML Model Building Tools, and Not Applicable/Not Reported, Other
AI Use by Teachers/ Administrators	AI Use by Teachers/Administrators was coded as Academic Performance Prediction and Intervention, Academic Behavior Prediction and Intervention, Learning Process Monitoring, Automatic Support, Automated Grading/Assessment, Engaging Students, Advising, ML Model Demonstration, and Not Applicable/Not Reported, Other
AI Use by Students	AI Use by Students was coded as Scientific Discovery Learning, Improve Learning Experience, Data-Driven Decisions - Program Choosing/Major Choosing/School, Type Choosing, and Not Applicable/Not Reported, Other

References

Denotes studies used in the systematic review and included in this manuscript

Adams, C., Pente, P., Lemermeyer, G., & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. *Computers & Education: Artificial Intelligence, 4*, Article 100131. https://doi.org/10.1016/j.caeai.2023.100131

* Bergeron, M. F., Landset, S. A. R. A., Maugans, T. A., Williams, V. B., Collins, C. L., Wasserman, E. B., et al. (2019). Machine Learning in modeling high school sport concussion symptom resolve. *Medicine & Science in Sports & Exercise, 51*(7), 1362–1371. https://doi.org/10.1249/MSS.0000000000001903.

* Bilstrup, K. E. K., Kaspersen, M. H., & Petersen, M. G. (2020). Staging reflections on ethical dilemmas in machine learning: A card-based design workshop for high school students. In *Proceedings of the 2020 ACM designing interactive systems conference* (pp. 1211–1222). https://doi.org/10.1145/3357236.3395558.

Bishop, C. M. (2006). *Pattern recognition and machine learning (information science and statistics)*. Springer-Verlag New York, Inc.

* Bochniarz, K. T., Czerwiński, S. K., Sawicki, A., & Atroszko, P. A. (2022). Attitudes to AI among high school students: Understanding distrust towards humans will not help us understand distrust towards AI. *Personality and Individual Differences, 185*, Article 111299. https://doi.org/10.1016/j.paid.2021.111299.

* Burgsteiner, H., Kandlhofer, M., & Steinbauer, G. (2016). Irobot: Teaching the basics of artificial intelligence in high schools. *Proceedings of the AAAI Conference on Artificial Intelligence, 30*(1). https://doi.org/10.1609/aaai.v30i1.9864.

* Burlig, F., Knittel, C., Rapson, D., Reguant, M., & Wolfram, C. (2020). Machine learning from schools about energy efficiency. *Journal of the Association of Environmental and Resource Economists, 7*(6), 1181–1217. https://doi.org/10.1086/710606.

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends, 66*, 1–15. https://doi.org/10.1007/s11528-022-00715-y

* Chai, C. S., Lin, P. Y., Jong, M. S. Y., Dai, Y., Chiu, T. K., & Qin, J. (2021). Perceptions of and behavioral intentions towards learning artificial intelligence in primary school students. *Educational Technology & Society, 24*(3), 89–101 https://www.jstor.org/stable/27032858.

* Chai, C. S., Wang, X., & Xu, C. (2020). An extended theory of planned behavior for the modelling of Chinese secondary school students' intention to learn artificial intelligence. *Mathematics, 8*(11), 2089. https://doi.org/10.3390/math8112089.

* Chiu, T. K. (2021). A holistic approach to the design of artificial intelligence (AI) education for K-12 schools. *TechTrends, 65*(5), 796–807. https://doi.org/10.1007/s11528-021-00637-1.

Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers & Education: Artificial Intelligence, 4*, Article 100118. https://doi.org/10.1016/j.caeai.2022.100118

* Choi, E., & Park, N. (2021). Demonstration of gamification in education for understanding artificial intelligence principles at elementary school level. *Ilköğretim Online, 20*(3), 709–715.

* Chung, J. Y., & Lee, S. (2019). Dropout early warning systems for high school students using machine learning. *Children and Youth Services Review, 96*, 346–353. https://doi.org/10.1016/j.childyouth.2018.11.030.

* Costa-Mendes, R., Oliveira, T., Castelli, M., & Cruz-Jesus, F. (2021). A machine learning approximation of the 2015 Portuguese high school student grades: A hybrid approach. *Education and Information Technologies, 26*(2), 1527–1547. https://doi.org/10.1007/s10639-020-10316-y.

Crompton, H., & Burke, D. (2022). Artificial intelligence in K-12 education. *SN Social Sciences, 2*(7), 1–14. https://doi.org/10.1007/s43545-022-00425-5

* Cruz-Jesus, F., Castelli, M., Oliveira, T., Mendes, R., Nunes, C., Sa-Velho, M., et al. (2020). Using artificial intelligence methods to assess academic achievement in public high schools of a European Union country. *Heliyon, 6*(6), Article e04081. https://doi.org/10.1016/j.heliyon.2020.e04081.

* De Melo, G., Oliveira, S. M., Ferreira, C. C., Vasconcelos Filho, E. P., Calixto, W. P., & Furriel, G. P. (2017). Evaluation techniques of machine learning in task of reprovation prediction of technical high school students. In *2017 CHILEAN conference on electrical, Electronics engineering, information and communication technologies (CHILECON), 1-7*. IEEE. https://doi.org/10.1109/CHILECON.2017.8229739.

* Demir, K., & Güraksin, G. E. (2022). Determining middle school students' perceptions of the concept of artificial intelligence: A metaphor analysis. *Participatory Educational Research, 9*(2), 297–312. https://doi.org/10.17275/per.22.41.9.2.

* Dubey, A., & Mani, M. (2019). Using machine learning to predict high school student employability–A case study. In *2019 IEEE international conference on data science and advanced analytics (DSAA)* (pp. 604–605). IEEE. https://doi.org/10.1109/DSAA.2019.00078.

* El-Hajji, H., Obeid, N., Hammo, B. H., & Al Fayez, R. Q. (2021). Building a predictive machine learning model to determine an international program: A case study from Jordan private schools for IB program. In *2021 IEEE Jordan international joint conference on electrical engineering and information technology (JEEIT)* (pp. 128–133). IEEE. https://doi.org/10.1109/JEEIT53412.2021.9634097.

* Fernández-Martínez, C., Hernán-Losada, I., & Fernández, A. (2021). Early introduction of AI in Spanish Middle Schools. A motivational study. *KI-Künstliche Intelligenz, 35*(2), 163–170. https://doi.org/10.1007/s13218-021-00735-5.

* Ghareeb Ahmed Ali, S. (2020). Using an artificial intelligence application for developing primary school pupils'. *Oral Language Skills, 75*(75), 67–110. https://doi.org/10.21608/edusohag.2020.97643.

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

* Haendchen Filho, A., Concatto, F., do Prado, H. A., & Ferneda, E. (2021). Comparing feature engineering and deep learning methods for automated essay scoring of Brazilian national high school examination. *ICEIS, 1*, 575–583. https://doi.org/10.5220/0010377505750583.

Han, J., Pei, J., & Tong, H. (2022). *Data mining: Concepts and techniques.* Morgan Kaufmann.

* Harvey, J. L., & Kumar, S. A. (2019). A practical model for educators to predict student performance in K-12 education using machine learning. In *2019 IEEE symposium series on computational intelligence (SSCI)* (pp. 3004–3011). IEEE. https://doi.org/10.1109/SSCI44817.2019.9003147. December.

* Henry, J., Hernalesteen, A., & Collard, A. S. (2021). Teaching artificial intelligence to K-12 through a role-playing game questioning the intelligence concept. *KI-Künstliche Intelligenz, 35*(2), 171–179. https://doi.org/10.1007/s13218-021-00733-7.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., et al. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions* (2nd ed.). John Wiley & Sons. https://doi.org/10.1002/9781119536604.

High School Longitudinal Study of 2009 (HSLS:09). (2009). *Base-year data file documentation.* U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences.

* Hsu, T. C., Abelson, H., & Van Brummelen, J. (2022). The effects on secondary school students of applying experiential learning to the Conversational AI Learning Curriculum. *International Review of Research in Open and Distance Learning, 23*(1), 82–103. https://doi.org/10.19173/irrodl.v22i4.5474.

* Hu, J., Dong, X., & Peng, Y. (2022). Discovery of the key contextual factors relevant to the reading performance of elementary school students from 61 countries/regions: Insight from a machine learning-based approach. *Reading and Writing, 35*(1), 93–127. https://doi.org/10.1007/s11145-021-10176-z.

* Hung, J. L., Rice, K., Kepka, J., & Yang, J. (2020). Improving predictive power through deep learning analysis of K-12 online student behaviors and discussion board content. *Information Discovery and Delivery, 48*(4), 199–212. https://doi.org/10.1108/IDD-02-2020-0019.

* Ishikura, R., Takeda, M., & Iwashita, S. (2020). Classification of worries and consultations with school refusal students using machine learning. In *2020 joint 11th international conference on soft computing and intelligent systems and 21st international symposium on advanced intelligent systems (SCIS-ISIS)* (pp. 1–3). IEEE. https://doi.org/10.1109/SCISISIS50064.2020.9322709.

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets, 31*, 685–695. https://doi.org/10.1007/s12525-021-00475-2

* Jarbou, M., Won, D., Gillis-Mattson, J., & Romanczyk, R. (2022). Deep learning-based school attendance prediction for autistic students. *Scientific Reports, 12*(1), 1431. https://doi.org/10.1038/s41598-022-05258-z.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research, 4*, 237–285. https://doi.org/10.1613/jair.301

* Kaspersen, M. H., Bilstrup, K. E. K., Van Mechelen, M., Hjorth, A., Bouvin, N. O., & Petersen, M. G. (2021). VotestratesML: A high school learning tool for exploring machine learning and its societal implications. In *FabLearn Europe/MakeEd 2021-An international conference on computing, design and making in education* (pp. 1–10). https://doi.org/10.1145/3466725.3466728.

* Kim, S., Jang, Y., Kim, W., Choi, S., Jung, H., Kim, S., et al. (2021). Why and what to teach: AI curriculum for elementary school. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(17), 15569–15576. https://doi.org/10.1609/aaai.v35i17.17833.

Kim, S., Kim, S., Lee, M., & Kim, H. (2020). Review on artificial intelligence education for K-12 students and teachers. *The Journal of Korean association of computer education, 23*(4), 1–11.

* Koren, M., Erell, A., & Sârb, D. E. (2021). *Solfy: An AI didactic support for updating school music education/Solfy: Un suport didactic IA pentru actualizarea educației muzicale·colare*.

* Lee, H. (2022). What drives the performance of Chinese urban and rural secondary schools: A machine learning approach using pisa 2018. *Cities, 123*, Article 103609. https://doi.org/10.1016/j.cities.2022.103609.

* Lee, S., Mott, B., Ottenbreit-Leftwich, A., Scribner, A., Taylor, S., Park, K., & Lester, J. (2021). AI-infused collaborative inquiry in upper elementary school: A game-based learning approach. *Proceedings of the AAAI Conference on Artificial Intelligence, 35* (17), 15591–15599. https://doi.org/10.1609/aaai.v35i17.17836.

* Li, H., Ding, W., & Liu, Z. (2020). *Identifying at-risk K-12 students in multimodal online environments: A machine learning approach.* https://doi.org/10.48550/arXiv.2003.09670. arXiv preprint arXiv:2003.09670.

* Lin, P., & Van Brummelen, J. (2021). Engaging teachers to co-design integrated AI curriculum for K-12 classrooms. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–12). https://doi.org/10.1145/3411764.3445377.

* Ma, Z. H., Hwang, W. Y., & Shih, T. K. (2020). Effects of a peer tutor recommender system (PTRS) with machine learning and automated assessment on vocational high school students' computer application operating skills. *Journal of Computers in Education, 7*, 435–462. https://doi.org/10.1007/s40692-020-00162-9.

* Malach, J., & Vicherková, D. (2020). Artificial intelligence in engineering as a factor influencing a choice of studies at vocational school and future career ideas of students. In *Proceedings of the 18th international conference on e-society* (pp. 89–98). https://doi.org/10.33965/es2020_202005L012.

* Mariescu-Istodor, R., & Jormanainen, I. (2019). Machine learning for high school students. In *Proceedings of the 19th Koli calling international conference on computing education research* (pp. 1–9). https://doi.org/10.1145/3364510.3364520.

Marques, L. S., Gresse von Wangenheim, C., & Hauck, J. C. (2020). Teaching machine learning in school: A systematic mapping of the state of the art. *Informatics in Education, 19*(2), 283–321. https://doi.org/10.15388/infedu.2020.14

* Masci, C., Johnes, G., & Agasisti, T. (2018). Student and school performance across countries: A machine learning approach. *European Journal of Operational Research, 269*(3), 1072–1085. https://doi.org/10.1016/j.ejor.2018.02.031.

* Mnyawami, Y. N., Maziku, H. H., & Mushi, J. C. (2022). Enhanced model for predicting student dropouts in developing countries using automated machine learning approach: A case of Tanzanian's secondary schools. *Applied Artificial Intelligence, 36* (1), Article 2071406. https://doi.org/10.1080/08839514.2022.2071406.

* Muhamedyev, R., Yakunin, K., Kuchin, Y. A., Symagulov, A., Buldybayev, T., Murzakhmetov, S., et al. (2020). The use of machine learning "black boxes'' explanation systems to improve the quality of school education. *Cogent Engineering, 7* (1), Article 1769349. https://doi.org/10.1080/23311916.2020.1769349.

* Ni, Y., Barzman, D., Bachtel, A., Griffey, M., Osborn, A., & Sorter, M. (2020). Finding warning markers: Leveraging natural language processing and machine learning technologies to detect risk of school violence. *International Journal of Medical Informatics, 139*, Article 104137. https://doi.org/10.1016/j.ijmedinf.2020.104137.

* Niemeijer, K., Feskens, R., Krempl, G., Koop's, J., & Brinkhuis, M. J. (2020). Constructing and predicting school advice for academic achievement: A comparison of item response theory and machine learning techniques. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 462–471). https://doi.org/10.1145/3375462.3375486.

* Norouzi, N., Chaturvedi, S., & Rutledge, M. (2020). Lessons learned from teaching machine learning and natural language processing to high school students. *Proceedings of the AAAI Conference on Artificial Intelligence, 34*(No. 09), 13397–13403. https://doi.org/10.1609/aaai.v34i09.7063.

* Oskotsky, T., Bajaj, R., Burchard, J., Cavazos, T., Chen, I., Connell, W. T., & Sirota, M. (2022). Nurturing diversity and inclusion in AI in Biomedicine through a virtual summer program for high school students. *PLoS Computational Biology, 18*(1), Article e1009719. https://doi.org/10.1371/journal.pcbi.1009719.

* Park, Y., & Shin, Y. (2021). Tooee: A novel scratch extension for K-12 big data and artificial intelligence education using text-based visual blocks. *IEEE Access, 9*, 149630–149646. https://doi.org/10.1109/ACCESS.2021.3125060.

* Rasheed, M. A., Chand, P., Ahmed, S., Sharif, H., Hoodbhoy, Z., Siddiqui, A., et al. (2021). Use of artificial intelligence on Electroencephalogram (EEG) waveforms to predict failure in early school grades in children from a rural cohort in Pakistan. *PLoS One, 16*(2), Article e0246236. https://doi.org/10.1371/journal.pone.0246236.

* Rebai, S., Yahia, F. B., & Essid, H. (2020). A graphically based machine learning approach to predict secondary schools' performance in Tunisia. *Socio-Economic Planning Sciences, 70*, Article 100724. https://doi.org/10.1016/j.seps.2019.06.009.

Russell, S. J. (2010). *Artificial intelligence a modern approach.* Pearson Education, Inc.

* Sansone, D. (2019). Beyond early warning indicators: High school dropout and machine learning. *Oxford Bulletin of Economics & Statistics, 81*(2), 456–485. https://doi.org/10.1111/obes.12277.

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN computer science, 2*(3), 160.

* Southgate, E. (2019). Virtual reality for Deeper Learning: An exemplar from high school science. In *2019 IEEE conference on virtual reality and 3D user interfaces (VR)* (pp. 1633–1639). IEEE. https://doi.org/10.1109/VR.2019.8797841.

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

* Tedre, M., Vartiainen, H., Kahila, J., Toivonen, T., Jormanainen, I., & Valtonen, T. (2020). Machine learning introduces new perspectives to data agency in K—12 computing education. In *2020 IEEE frontiers in education conference (FIE)* (pp. 1–8). IEEE. https://doi.org/10.1109/FIE44824.2020.9274138.

* Tuba, K. O. C., & Pelin, A. K. I. N. (2022). Estimation of high school entrance examination success rates using machine learning and Beta regression models. *Journal of Intelligent Systems: Theory and Applications, 5*(1), 9–15. https://doi.org/10.38016/jista.922663.

* Tyson, M. M., & Sauers, N. J. (2021). School leaders' adoption and implementation of artificial intelligence. *Journal of Educational Administration, 59*(3), 271–285. https://doi.org/10.1108/JEA-10-2020-0221.

* Van Brummelen, J., Heng, T., & Tabunshchyk, V. (2021). Teaching tech to talk: K-12 conversational artificial intelligence literacy curriculum and development tools. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(17), 15655–15663. https://doi.org/10.1609/aaai.v35i17.17844.

Wang, T., & Cheng, E. C. (2022). Towards a tripartite research agenda: A scoping review of artificial intelligence in education research. In *Artificial intelligence in education: Emerging technologies, models and applications* (pp. 3–24). Springer. https://doi.org/10.1007/978-981-16-7527-0_1.

* Wang, L., & Wang, W. (2018). Research and construction of junior high school subject Q&A system model based on deep learning. In *2018 international conference on information systems and computer aided education (ICISCAE)* (pp. 504–508). IEEE. https://doi.org/10.1109/ICISCAE.2018.8666853.

* Wan, X., Zhou, X., Ye, Z., Mortensen, C. K., & Bai, Z. (2020). SmileyCluster: Supporting accessible machine learning in K-12 scientific discovery. In *Proceedings of the interaction design and children conference* (pp. 23–35). https://doi.org/10.1145/3392063.3394440.

* Williams, R., Kaputsos, S. P., & Breazeal, C. (2021). Teacher perspectives on how to train your robot: A middle school AI and ethics curriculum. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(17), 15678–15686. https://doi.org/10.1609/aaai.v35i17.17847.

Wong, G. K., Ma, X., Dillenbourg, P., & Huan, J. (2020). Broadening artificial intelligence education in K-12: Where to start? *ACM Inroads, 11*(1), 20–29.
* Wu, C., Li, Y., Li, J., Zhang, Q., & Wu, F. (2021). Web-based platform for K-12 AI education in China. *Proceedings of the AAAI Conference on Artificial Intelligence, 35* (17), 15687–15694. https://doi.org/10.1609/aaai.v35i17.17848.
* Xia, J., Wang, J., Chen, H., Zhuang, J., Cao, Z., & Chen, P. (2022). An unsupervised machine learning approach to evaluate sports facilities condition in primary school. *PLoS One, 17*(4), Article e0267009. https://doi.org/10.1371/journal.pone.0267009.
* Xie, Y., Li, J., Ye, Z., Lin, X., Cao, L., & Huang, Y. (2020). The development and effect analysis of the deep-learning classroom model of primary school mathematics in the intelligent environment. In *2020 international symposium on educational technology (ISET)* (pp. 13–17). IEEE. https://doi.org/10.1109/ISET49818.2020.00013.
* Zafari, M., Sadeghi-Niaraki, A., Choi, S. M., & Esmaeily, A. (2021). A practical model for the evaluation of high school student performance based on machine learning. *Applied Sciences, 11*(23), Article 11534. https://doi.org/10.3390/app112311534.
Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., & Li, Y. (2021). A review of Artificial Intelligence (AI) in education from 2010 to 2020. *Complexity, 2021*, 1–18. https://doi.org/10.1155/2021/8812542
Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers & Education: Artificial Intelligence, 2*, Article 100025. https://doi.org/10.1016/j.caeai.2021.100025
* Zhang, X., Tlili, A., Shubeck, K., Hu, X., Huang, R., & Zhu, L. (2021). Teachers' adoption of an open and interactive e-book for teaching K-12 students artificial intelligence: A mixed methods inquiry. *Smart Learning Environments, 8*, 1–20. https://doi.org/10.1186/s40561-021-00176-5.
* Zhuang, Y., & Gan, Z. (2017). A machine learning approach to enrollment prediction in Chicago Public School. In *2017 8th IEEE international conference on software engineering and service science (ICSESS)* (pp. 194–198). IEEE. https://doi.org/10.1109/ICSESS.2017.8342895.