Ethical Implications and Principles of Using Artificial Intelligence Models in the Classroom: A Systematic Literature Review

Lin Tang[1], Yu-Sheng Su[2] *

[1] Nanjing Normal University, (China)
[2] Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi county (Taiwan)

Abstract

The increasing use of artificial intelligence (AI) models in the classroom not only brings a large number of benefits, but also has a variety of ethical implications. To provide effective education, it is now necessary to understand the ethical implications of using AI models in the classroom, and the principles for avoiding and addressing these ethical implications. However, existing research on the ethical implications of using AI models in the classroom is rather sparse, and a holistic overview is lacking. Therefore, this study seeks to offer an overview of research on the ethical implications, ethical principles and the future research directions and practices of using AI models in the classroom through a systematic literature review. Out of 1,445 initially identified publications between 2013 and 2023, 32 articles were included for final coding analysis, identified using explicit inclusion and exclusion criteria. The findings revealed five main ethical implications, namely algorithmic bias and discrimination, data privacy leakage, lack of transparency, decreased autonomy, and academic misconduct, with algorithmic bias being the most prominent (i.e., the number of existing studies is the most), followed by privacy leakage, whereas decreased autonomy and academic misconduct were relatively understudied; and six main ethical principles, namely fairness, privacy, transparency, accountability, autonomy and beneficence, with fairness being the most prominent ethical principle (i.e., the number of existing studies is the most), followed by privacy, while autonomy and beneficence were relatively understudied. Future directions of research are given, and guidelines for future practice are provided: (1) further substantive discussion, understanding and solution of ethical implications are required; (2) the precise mechanism of ethical principles of using AI models in the classroom remains to be elucidated and extended to the implementation phase; and (3) the ethical implications of the use of AI models in the classroom require accurate assessment.

Keywords

I. Introduction

ARTIFICIAL intelligence (AI) is defined as a branch of computer science that simulates intelligent behavior in computers, in an attempt to develop human-like intelligence machines [1]. In recent years, AI has become an indispensable part of people's lives with its powerful functions, and it deeply affects all areas of human activities, including education [2]. In order to achieve Sustainable Development Goal 4 (SDG4) of UNESCO's Agenda 2023 on quality and inclusive education [3], many AI models have been applied in real classrooms to promote instruction and learning, such as Google Classroom (which was applied to online teaching management) [4], Google Dialogflow (which was used as a virtual education assistant) [5], and GPT (which was applied to automatic question generation and essay scoring) [6].

These AI models are based on powerful algorithms and generating capabilities to support personalized learning systems and automated assessment systems that facilitate students' learning and teachers' teaching [7]. They contribute to students' learning and can also free teachers from heavy work [8]. Compared with traditional computer-based models, these AI models can provide more dynamic and realistic learning experiences [9].

However, despite the use of AI models in the classroom having many undeniable benefits, it also raises potentially extensive ethical implications, for instance, leakage of personal private data caused by the collection of large amounts of data, discrimination and unfairness caused by algorithmic bias, and lack of integrity caused by the abuse of technology [10], [11]. Thus, using AI models in the classroom is seen as a complex and highly controversial issue [12]. Actually, however, we just need to assume what our response speed will be, rather than ignoring or banning AI, thus avoiding extremism [13]. SDG4 emphasizes that AI technologies must be applied to ensure equitable and inclusive access to education [3]. Hence, AI should be used to

* Corresponding author.

E-mail addresses: 200602093@njnu.edu.cn (L. Tang), ccucssu@gmail.com (Y. S. Su).

enhance and amplify the ability of teachers to teach and students to learn, instead of being replaced by them. In fact, both discriminative AI models and generative AI models are the result of data-driven model training, not a mystical magic [14]. Therefore, AI models are not a panacea, and this understanding of AI will help address the ethical implications explored in this study.

In order to avoid and address the ethical implications resulting from the use of AI models in the classroom, more ethical principles need to be considered, such as privacy, fairness, transparency and accountability [15]. Recently, some researchers and international organizations have specifically studied the ethical principles when applying AI in the field of education [16]. It is worth noting that some ethical principles overlap in these reports, but few studies have systematically examined the global consensus on the ethical principles of using AI in the classroom [17]. At a more formal and legal level, some countries and organizations have developed or are developing general laws about AI, such as the United States' AI Bill of Rights [18] and Canada's Artificial Intelligence and Data Act [19], which also cover the ethical aspects of using AI. Most notably, the European Union approved the Artificial Intelligence Act in December 2023 [20], which is the first global comprehensive regulation of the field of AI. Some rules for AI development and use have also been developed, such as human oversight, security, privacy, transparency, non-discrimination, and social and environmental well-being [21]. These efforts aimed to achieve a consensus on the rational and regulated use of AI through ethical and legal constraints.

In general, using AI models has brought about some ethical implications while improving the quality of instruction and learning. However, the ethical implications, ethical principles, and related research directions of using AI models in the classroom still need to be clarified. Previous systematic review work has provided some substantial insights into AI in education, including theoretical paradigms, applications, benefits, challenges and trends [8], [22]-[24]. However, literature reviews on the ethical implications about the use of AI models in the classroom are limited, and there is no research involving a systematic literature review, resulting in the lack of a holistic view. Additionally, the reviewing which principles are required to avoid and address the ethical implications of AI model use in the classroom remains inadequate and has only been macroscopically articulated in a few studies [11],[17]. Meanwhile, the future research and practice directions of related research could be clearer. Further research is urgently needed to clarify the ethical implications, ethical principles, and future research directions of using AI models in the classroom. Compared with the general literature review method, the systematic literature review method based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) emphasizes following strict steps to extract valid information from existing literature, to draw comprehensive conclusions [25], which are conducive to providing evidence for solving the research questions of this study. Therefore, to make up for this lack of research, this systematic literature review collected, reviewed, and summarized the research on the ethical implications, principles and future research directions of using AI models in the classroom.

The rest of this paper is structured as follows. Section II critically reviews the ethical implications and principles of using AI models in the classroom and raises the research questions. Section III describes the PRISMA method used in this study. Next, the results and findings of the literature analysis regarding the research questions are presented in Section IV. Based on the literature analysis, Section V provides further discussion of the results and findings. In Section VI, the implications, limitations and future work of the current study are illustrated.

II. Literature Review

A. The Ethical Implications of Using AI Models in the Classroom

In recent years, AI models have been widely used in the classroom. For example, Zhang et al. [26] introduced AI models into flipped classrooms to digitize and visualize course material preparation, supporting AI-assisted interactive classroom learning. Suresh et al. [27] discussed the application of several deep learning models for promoting fair classroom discussion. Recently, some studies have reported on generative language models such as chatbots being applied in the classroom to support student learning [28], [29]. However, despite the use of AI technology in teaching and learning bringing huge benefits to revolutionize education, the integrating AI models into the classroom could have significant ethical implications [7]. Some typical ethical implications have been identified. For instance, powerful algorithm-based AI predictive models can indeed provide personalized learning for students [30], assist teachers in instructional design [31], and provide references for administrators in making educational decisions [32]. However, one potential ethical issue is that AI algorithms can be biased. Verma [33] argued that if the data are biased, the AI models perpetuate those biases, thereby exacerbating existing discrimination in educational systems. In addition, data-driven AI models require collecting and storing large amounts of sensitive student data. This could raise another potential ethical issue, namely that these data could be used for unintended purposes or be accessed by unauthorized individuals, leading to students' privacy disclosure [16], [34]. Moreover, lack of accountability and transparency are also major ethical implications of using AI models in the classroom, which leads to the question of who is responsible for the accuracy of educational decisions based on AI models and how they are made [35]. Recently, Naidu and Sevnarayan [36] reported on the potential crisis of academic integrity arising from using ChatGPT, an emerging large AI language model, for online assessment in distance education.

In addition, the systematic literature review approach has provided comprehensive views of other aspects of using AI in education, such as paradigms, applications, benefits, challenges, and trends [8], [22]-[24]. For instance, Tahiru systematically reviewed the challenges of implementing AI in education, including ethics, privacy, and trust [24]. These are echoed in Murphy's research. He systematically summarized the major applications of AI models in education, such as rule-based expert systems, intelligent tutor systems (including a student model and a teacher model), and machine learning (including automated scoring systems and early warning systems). Meanwhile, he pointed out that these models are error-prone when used in different scenarios, which can have ethical implications including bias, transparency and trust [8]. Based on the e-learning background, Tang et al. [23] systematically summarized the future research trends of the use of AI in education, especially emphasizing the assessment of the environment and its implications.

However, in these studies, the ethical implications were not a major part of the review and were only briefly summarized. Moreover, previous research has not specifically elaborated the ethical implications about the use of AI models from a teaching and learning perspective, and there has been a lack of attention to the future direction of related research. Hence, it is not clear what the main aspects and concrete content of the ethical implications are, and what future research directions will be. Additionally, in contrast to other review methods, the approach of PRISMA emphasizes the search and selection of literature guided by the research question [25], which has the potential to provide a complete picture of ethical implications of AI models use in the classroom, as demonstrated by research on the ethical implications about AI in other fields [37]. Therefore, there is an urgent need for a systematic review to clarify the main ethical implications using AI models in the classroom, to guide instruction practice.

B. The Ethical Principles of Using AI Models in the Classroom

The ethical principles are the guidelines that should be followed for the ethical use of AI models in the classroom to avoid and address ethical implications [15]. Hagendorff [38] emphasized that ethical principles for using AI are necessary and must be aligned with societal values. Some international organizations (e.g., UNESCO Education & AI and the European Commission) have reported the general ethical principles that should be followed from AI design and development to its use, such as security, privacy, transparency, accountability, inclusiveness, sustainability, and human-centeredness [39], [40]. Likewise, these ethical principles have been further discussed in education. For example, in terms of the principle of privacy, Miao et al. [39] considered that to protect the privacy of teachers and students, it is necessary to collect and analyze the points of teachers and students before using AI models to decide how to deploy AI in the classroom. Additionally, the large collection of student and teacher data highlighted the need for transparency in using AI models [41]. The principle of transparency refers to the detailed explanation of using AI models, including what the data are, how they are collected, how they are interpreted, and how they are used [15]. Slimi and Carballido [42] emphasized that the principle of transparency is critical for teachers and students because data visualization can be used to analyze student learning behaviors and trajectories and to provide additional support for teachers' instruction. Moreover, the principle of accountability ethics has also been called for in some studies. For example, Klimova et al. [11] highlighted the primary responsibility for clarifying the use of AI-driven mobile apps in education. Hong et al. [43] pointed out that when AI is applied in education, it should be determined who is responsible for the consequences of the data use. These studies required clear subject responsibility for educational decision making based on AI models.

However, review work on ethical principles of using AI models in the classroom is still insufficient, and only a few reviews have been conducted [11], [17]. Specifically, Klimova et al. [11] synthesized eight articles on the ethical principles of using AI in education, and concluded four major principles, namely beneficence, accountability, justice and human values. Regrettably, this study reviewed only a few articles, and needed to provide further analysis of these ethical principles. Additionally, in Memarian and Doleck's research [17], they examined the fairness, accountability, transparency, and ethics in AI in the context of higher education, but they did not define ethical principles as the primary research focus. Importantly, previous studies did not explore the principles that should be followed to avoid and address the ethical implications of using AI models in the context of the classroom. More perspectives have focused on the macro context of education. However, some of the ethical implications arising from the current use of AI models in teaching urgently require a research perspective focused on the classroom. Further, future directions for related research have not been specifically discussed in previous studies. Therefore, it is still unclear what the main aspects of the ethical principles of using AI models in the classroom are, and what the future research directions are. Since the PRISMA method can extract and interpret data more accurately than the general review method [25], it is conducive to providing more accurate answers to the questions in this study. All in all, based on clarifying the ethical implications of using AI models in the classroom, this study further systematically reviewed the ethical principles that should be considered.

C. Research Questions

To further understand the ethical implications when applying AI models in the classroom, this systematic review examined the ethical implications and ethical principles from the teaching and learning perspective. Additionally, the future research directions of related research were also investigated. Specifically, the following research questions were proposed in this study:

RQ1: What are the ethical implications of using AI models in the classroom?

RQ2: What are the required principles of using AI models in the classroom to avoid and address the ethical implications?

RQ3: What are future directions of research and practice regarding the ethical implications and principles of using AI models in the classroom?

III. Method

The systematic review method was adopted in this study based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) principles, comprising a total of four phases: identification, screening, eligibility, and included [25]. This methodology was designed to answer specific research questions through clear, systematic, and repeatable search strategies [44]. Next, the procedure of systematic literature review in this study will be described.

A. Literature Search Process

Two international publication databases were selected to search the full-text archives, including the Web of Science and Scopus, which are the most comprehensive databases of academic literature [45]. In both databases, most journals are predominantly in the English language [46]. Burnham [47] insisted that the Web of Science and Scopus can complement each other to improve the coverage rate of related articles. Importantly, the works contained in these databases are seen to be of high quality and to have significant impact in social science, and the databases cover a wide range of educational journals [48]. Compared with other databases, these two provide a variety of search methods and browsing options, including standard, basic and advanced methods, which is more conducive to the accuracy of literature searches [46]. Additionally, the high accessibility of their journals in the academic community is more conducive to the conduct of this research. Furthermore, the literature searched for in this study covers articles published between 2013 and July 2023, because AI began to make significant progress in education from around 2013 [45]. To ensure the quality of the review, the selected articles were only from peer-reviewed papers, because they have a high degree of credibility and have undergone a rigorous review process [49]. In addition, conference proceedings (if available) were included in this study to obtain up-to-date information on the ethical aspects of the use of AI models in the field of education. Based on the Cochrane Handbook of Systematic Reviews of Interventions, searching for conference proceedings is considered a highly reliable practice, because it is beneficial to capture as many studies as possible, and can greatly reduce the risk of publication bias [50]. After the full text was filtered, according to the guidelines [51], the snowball method was applied to find further papers which were not retrieved through the search strings.

The structured search strategy was adopted in this study to search the databases. To find the most relevant literature in this field, the PICOC (Population, Intervention, Comparison, Outcome and Context) framework, proposed by Petticrew and Roberts [52], was adopted to define the search string and the scope of this study (see the details below):

a) Population: this study deals with terms, keywords, or some variation of the same meaning related to AI models, classrooms, and ethics. Therefore, the search string was defined according to these criteria.

b) Intervention: to implement this theme, some exclusion criteria were designed, as shown in Table I. Articles that did not meet these specific requirements were excluded.

c) Comparison: emphasis was on the specific ethical implications and principles of using AI models, rather than on a broader picture of their use.

d) Outcome: this step determined which outcomes were the most relevant to answering the research questions [52]. Hence, in addition to the ethical implications and principles of using AI models in the classroom, future research directions related to the topic were also included as outcomes.

e) Context: the last step is the "context" that defines the boundaries of the questions, which was defined as classroom teaching and education.

Ultimately, the following search string and Boolean operators AND/OR were utilized: ("Artificial intelligence" OR "AI" OR "AI model") AND ("classroom" OR "educat*") AND ("ethic*" OR "moral*"). This literature search was conducted in August 2023 and initially identified 1,445 records (1,063 from WOS, 382 from SCOPUS).

B. Inclusion and Exclusion Process

To improve the pertinence of the literature in the analysis of the research questions, a set of inclusion and exclusion criteria was designed to identify better the papers that focused on the ethical implications of the use of AI models in the classroom (shown in Table I). Specifically, these criteria were mainly based on the following considerations: (a) published between 2013 and 2023, as AI has made significant progress in the field of education since 2013 [45]; (b) written in English, not only because English is the internationally recognized language in the field of science, but also the same language is more conducive to text-mining analysis; (c) research from articles or conference proceedings were chosen, because they are highly scholarly; (d) sourced from peer-reviewed scientific papers, as these papers are typically evaluated by experts in their subject area, thus ensuring some form of quality check; (e) conducted in the field of education, because this was in line with the background of this study, for example, research in the field of medicine was excluded, but research in the field of medical education was included; (f) focus on the use of AI models in education, rather than the design and development of AI models; and (g) focus on the ethical implications of using AI models, rather than simply mentioning them, and discussion of the ethical implications as an important part of the research.

TABLE I. Inclusion and Exclusion Criteria

Inclusion criteria	Exclusion criteria
Research must be published from 2013 to 2023.	Research published before 2013.
Research must be written in English.	Research written in any other languages.
Research from articles or conference proceedings.	Research from book chapters, magazines, news, and posters.
Research must be sourced from peer-reviewed scientific papers.	Research not sourced from peer-reviewed scientific papers.
Research must be carried out in the field of education.	Research conducted in fields other than education.
Research must focus on using AI models in education.	Research not focused on using AI models in education.
Research must focus on the ethical implications of using AI models.	Research not focused on the ethical implications of using AI models.

After deleting 151 duplicates, the remaining 1,294 articles were screened according to the inclusion and exclusion criteria. The number of articles that did not meet the criteria by reviewing the titles and abstracts was 121. Subsequently, 79 articles that were inconsistent with the research purpose were further excluded by full text reading, and 28 relevant articles were identified. According to Webster and Watson's suggestions [53], a forward and backward reference search was carried out for these articles to identify further relevant records. In the backward search, references for 28 articles were analyzed, and in the forward reference, Google Scholar was used to analyze and identify articles that cited reservations; as a result, four articles were added after review according to the inclusion and exclusion criteria. Eventually, 32 eligible articles were identified for systematic review. The PRISMA flow diagram of the study is summarized in Fig. 1.
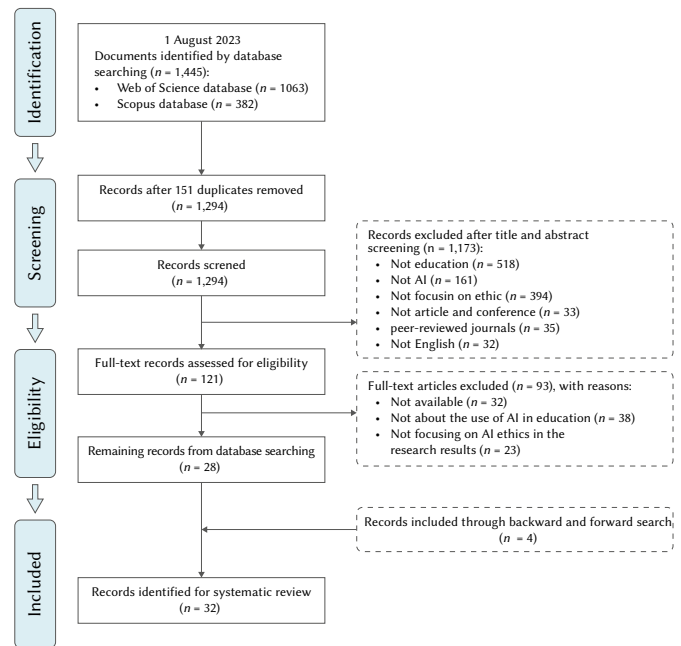

Fig. 1. PRISMA flow diagram of the study.

To ensure rigorous evaluation of articles included in the review, the following quality criteria were developed: (a) the article clearly defined the purpose of the research; (b) the article disclosed the research methodology used; (c) the article clearly presented one or more ethical implications or principles for the use of AI models in the classroom; and (d) a comprehensive description of the results was provided in the study.

As shown in Table II, among the 32 included articles, 23 were journal papers and nine were conference papers. In terms of regional distribution of the literature, the most prolific region for relevant literature was Europe (N = 17, 53%), followed by North America (N = 11, 35%), Asia with less (N = 3, 9%), and Oceania with the least (N = 1, 3%). In terms of the distribution of educational stages, except for discussion in which the scope was not specified (N = 21, 66%), the research was mainly concentrated on higher education (N = 8, 25%). In addition, the major research methods for the 32 studies were also identified, including literature study, quantitative survey, interview and observations, exploratory research, perspective, case study and mixed methods. Among them, literature study was the most commonly used method (N = 12, 38%), followed by perspective (N = 9, 28%).

C. Data Analysis

The inductive grounded method was applied to analyze and classify the information in the 32 eligible articles relevant to the research question [54]. This classification method identifies and refines topics through data rather than pre-determined categories or theories,

TABLE II. The General Information of the 32 Included Articles

General information	Category	N	%
Article type	Journal	23	72
	Conference paper	9	28
Country/Region	Europe	17	53
	North America	11	35
	Asia	3	9
	Oceania	1	3
Educational sector	Higher education	8	25
	K-12	3	9
	Unspecified field	21	66
Method	Literature study	12	38
	Quantitative survey	3	9
	Interview and observations	4	13
	Exploratory research	2	6
	Perspective	9	28
	Case study	1	3
	Mixed methods	1	3

TABLE III. The Ethical Implications of Using AI Models in the Classroom

Category	Example	Na	Sample studies
Algorithmic bias and discrimination	Gender discrimination.	7	Ghotbi and Ho [68]
	Racial or ethnic discrimination.	6	Ghotbi et al. [69]
	Class discrimination.	2	Matias and Zipitria [70]
	Cultural bias.	2	Masters [71]
Loss of privacy	Personal information is leaked.	6	Köbis and Mehner [72]
	Increasing culture of algorithmic surveillance.	5	Reiss [35]
	The absence of the right to be forgotten and to give informed consent.	3	Adams et al. [73]
Lack of transparency	Teachers and students may have difficulty understanding predictions related to learning performance.	5	Hong et al. [43]
	The explanation of potential or actual disadvantages or risks of using AI models in the classroom is not apparent.	3	Slimi and Carballido [42]
Decreased autonomy	The ability of students and teachers to manage their own lives is reduced.	6	Han et al. [74]
Academic misconduct	Cause cheating and plagiarism issues.	5	Adams et al. [75]

[a] The number of studies added up to more than 32 because multiple ethical implications of using AI models in the classroom are described in several studies.

which facilitates the extraction of new findings and protects the richness of the data [55]. To help answer the research questions, the following extraction framework was identified to extract data from the 32 included articles: study objectives, study design, study type, educational topic, AI application, main findings, ethical implications, ethical principles and the considerations of future work.

To answer RQ1—What are the ethical implications of using AI models in the classroom?—inductive analysis was performed to extract information about the ethical implications by manually mapping each article. These descriptive data were reconstructed through the process of coding, conceptualization, and classification. Specifically, coding was used to identify sentences or paragraphs from the data that were relevant to the ethical implications, and to describe them with short phrases. Then, the constant comparison method was applied to combine the similar codes to form categories about ethical implications [55]. To answer RQ2—What are the ethical principles of using AI models in the classroom?—the researchers read through the full text and adopted the above analytical steps to form categories about ethical principles. Meanwhile, the frequencies of each category coded were calculated in this study. To answer RQ3—What are future directions of research and practice regarding the ethical implications and principles of using AI models in the classroom?—the recommendations made by each researcher in the discussion and conclusion sections were read manually.

Additionally, three strategies were adopted in this study to ensure the reliability of the literature analysis. Firstly, two trained researchers carried out constant discussions until agreement was reached to verify the categories [56]. Secondly, according to Hsieh and Shannon [57], the Results section of this study explains in detail the categories of the literature findings for each research question. Finally, some examples are presented within each category to prove how well categories represent the data in response to the research questions [58].

IV. Results

A. What Are the Ethical Implications of Using AI Models in the Classroom?

Content analysis of the current literature revealed five categories of ethical implications about the use of AI in the classroom: data privacy leakage, algorithmic bias and discrimination, lack of transparency, decreased autonomy, and academic misconduct (shown in Table III).

1. Algorithmic Bias and Discrimination

Algorithmic bias and discrimination is the most studied ethical implication in the reviewed studies ($N = 17$). Due to the training data, AI models have been shown to exhibit bias and discrimination, which reinforce inherent stereotypes. The existing literature focuses on bias around databases and algorithms of AI models against certain groups of students (often underrepresented minorities), involving aspects such as gender, race or ethnicity, class and cultural background. Specifically, gender discrimination is one of the most apparent forms of this issue. Akgun and Greenhow revealed the gender stereotypes of using AI models in language learning classrooms. When students learn the translation of sentences using generative language models, such as those about doctors and soldiers, they are often translated as male, which exacerbates some social prejudices and gender stereotypes [10]. Additionally, Holmes et al. noted that the application of AI models could be influenced by "cultural imperialism," leading to ethical issues of cultural discrimination in the classroom [16]. Furthermore, biased decision algorithms have been shown in AI models, such as personalized learning, automated assessment, facial recognition systems and predictive systems in education [58]. Kooli considered that AI models, such as chatbots, can produce inaccurate results or misleading information that can lead to decisions being made against specific groups of students [59].

2. Data Privacy Leakage

Data privacy leakage is a critical ethical issue in debates of using AI models in the classroom ($N = 14$). AI models are used to analyze, assess and predict students' learning performance by accumulating large amounts of diverse data, such as personal background information, academic performance, facial expressions, and verbal records [35]. While these models can optimize the learning experience in the

classroom, they also raise some ethical problems about data privacy, including personal information leakage, surveillance and student tracking, the absence of informed consent and the right to be forgotten. Leakage of students' personal information is a frequently reported problem [60]. Kowch posed that the long-term tracking of students by AI has led to privacy disclosure, and AI surveillance is a difficult ethical issue that has long been considered [61]. Holmes et al. considered that students still lack a real opportunity to choose whether to opt in or out of educational AI systems, and that the right to informed consent and to be forgotten are important [62]. Therefore, enhancing privacy protection is a must for using AI models in the classroom.

3. Lack of Transparency

Lack of transparency is the third most ethical concern in the studies reviewed (*N* = 8). It is worth noting that transparency is also directly called explainability in some studies, such as Jang et al.'s [63] and Farrow's research [64]; it refers to the detailed explanation of algorithmic decisions or the collection and processing of data. There is a general decline in transparency around the use of AI models [17]. Kooli considered that the current application of AI models in the classroom still has a lack of explanation, that is, most teachers and students do not understand the process of AI decision making, and how and under what conditions to use these data [59]. Importantly, the former risks and practical downsides of using AI models in the classroom are not spelled out in detail [41]. Chen et al. [29] demonstrated that when a chatbot was designed for use in the classroom to support students' learning, it was not always able to identify spelling mistakes or understand colloquial speech. Further, the chatbot lacked a deeper understanding of the emotions expressed by students, such as sarcasm. As noted in Hong et al., a lack of transparency has led some teachers and students to question the results of AI algorithm-based learning predictions and decision models [35].

4. Decreased Autonomy

Decreased autonomy is also a serious ethical issue discussed in the reviewed studies (*N* = 6). It is worth noting that autonomy is also directly called agency in some studies, such as Tuomi [65] and Holmes et al. [16]; it refers to individuals being free to pursue goals and values that they deem important. Schiff suggested that in the case of using AI models in the classroom, inappropriate decision-making empowerment could potentially decrease and even undermine the autonomy of teachers, students, and parents, and he further emphasized that such problems have already arisen [66]. For example, Chen et al. [29] found that when chatbots were used in the classroom, some students only skimmed the learning content superficially rather than constructing their own thoughtful answers, and others even tended to engage in "smart loafing" in the classroom, handing the responsibilities for collaborative learning to the AI virtual assistant. Similarly, Akgun and Greenhow [10] and Klimova et al. [11] considered that algorithm-driven prediction systems and AI-driven mobile apps for education decrease the ability of students and teachers to manage their own lives, which may even lead to their conforming to norms in specific "data points."

5. Academic Misconduct

Academic misconduct is the least ethical concern in the reviewed studies *(N* = 5). However, the misuse of AI technology has led to some academic misconduct issues. The automatic generation function of AI models, such as ChatGPT, may be used by students to cheat and plagiarize while completing assignments and participating in assessments, which devalues the efforts of others and thus produces unfairness [59]. For example, Adams et al. mentioned that with regard to student writing, the use of AI models has caused the boundaries of who is writing to begin to blur: the student or AI [67]. Therefore, the issue

of academic integrity caused by the misuse of technology must be paid special attention to, because it not only involves the ethical use of AI during teaching and learning, but may also lead to educational inequity.

B. What Are the Ethical Principles of Using AI Models in the Classroom?

A count summary of ethical principle terms from the reviewed studies is presented in Table IV. Through content analysis, six ethical principles of using AI models in the classroom were summarized: fairness, privacy, transparency, accountability, autonomy, and beneficence.

TABLE IV. The Ethical Principles of Using AI Models in the Classroom

Category	Example	Na	Sample studies
Principle of fairness	Ensure that educational opportunities are equal among the students recommended by AI algorithms.	8	Matias and Zipitria [70]
	Ensure the accessibility of (digital) educational resources.	6	Köbis and Mehner [72]
	Inclusive of students from diverse backgrounds.	5	Schiff [66]
Principle of privacy	Keep the data provided by the students confidential.	9	Nguyen et al. [15]
	Acquire the students' active and full consent to access and use their personal data.	8	Masters [71]
Principle of transparency	Ensure that the educational decision-making process of AI models is explainable and understandable.	7	Chaudhry et al. [80]
	Specify the benefits, actual drawbacks, and possible risks of using AI models in the classroom.	4	Memarian and Doleck [17]
	Protect students' data ownership.	2	Nguyen et al. [15]
	Open communication regarding the expectations of using AI models in the classroom.	2	Köbis and Mehner [72]
Principle of accountability	Be responsible for the actions and decisions of using AI models.	9	Mouta et al. [81]
	Ensure teachers and students the right to access data.	4	Jang et al. [63]
Principle of autonomy	Teachers and students always maintain self-determination in deciding whether and how to adopt AI models.	8	Schiff [66]
Principle of beneficence	Provide comprehensive training before using AI models to enhance AI literacy.	4	Busch et al. [78]
	Support students' development and teacher well-being.	2	Adams et al. [75]

ᵃ The number of studies added up to more than 32 because multiple ethical principles of using AI models in the classroom are described in several studies.

1. Principle of Fairness

Fairness (or justice, or equity) is the most mentioned ethical principle in the reviewed studies (*N* = 19). According to the results of the systematic literature review, fairness generally subsumes representation, accessibility, and inclusiveness. First, AI models must be designed, developed and deployed with non-discriminatory and unbiased data and algorithms to ensure representation and equality between different educational groups. For example, when AI models are applied to student services, such as admissions and financial aid, they should ensure that they do not exacerbate existing biases and discrimination based on race, class, gender, or socioeconomic status [41]. In addition, Nguyen et al. posited that infrastructure, skills and social acceptance should be taken into account when using AI models, allowing equitable access and use by all teachers and students [15]. Schiff considered that the AI tutoring system needs to fit the background of students, such as their local customs, cultural background and learning styles [66].

2. Principle of Privacy

Privacy is the second ethical principle of using AI models in the reviewed studies (*N* = 17). First, the use of the AI-assisted tutor system in the classroom should protect students' personal information, such as gender, age, family address and mobile phone number, to avoid information leakage and personal harassment [76]. Moreover, Jang et al. further pointed out that students' data and privacy should be protected throughout the life cycle of using AI models, both in terms of raw data provided by students and new data generated about students during the interactions with AI systems (such as learning outcome analysis and recommendations) [63]. On the other hand, when collecting data about students, for whatever reason, it should be ensured that the student is giving active and not passive consent to the collection of personal data [71]. Meanwhile, Hong et al. considered that the use of AI models should also obtain full informed consent on how personal information and data are collected, shared and used [43].

3. Principle of Transparency

Transparency is the third ethical principle in the reviewed studies (*N* = 15). According to the results of the systematic literature review, the principle of transparency mainly includes interpretability, traceability, data ownership and communication. First, Holmes et al. emphasized that teachers and students should be provided with detailed explanations of the rationale, operational processes and outcomes of using AI models, so that they can better understand and apply the results [16]. For instance, when using AI models to make teacher ratings, student evaluations, and other educational decisions, the process, results and application condition of AI algorithm decisions must be explained in detail. Moreover, in addition to displaying the benefits of AI models, teachers and students must be informed of the actual drawbacks and potential risks of using AI models in the classroom, and even remedial suggestions [41]. Additionally, Nguyen et al. argued that data ownership, which relates to who owns and has access to students' personal data, is an important aspect of the principle of transparency [15]. The open communication regarding the expectations of using AI models in the classroom is also considered essential to promote trust [72].

4. Principle of Accountability

Accountability is the fourth ethical principle in the studies reviewed (*N* = 13). This principle requires responsibility for the actions and decisions of using AI models in the classroom and ensures that teachers and students have the right to access their data. Celik considered that teachers need to understand who the developers responsible for the design and decision-making of AI models are [77]. In addition, Hong et al. suggested that the principle of accountability can also be considered as the capacity to verify actions and decisions, so teachers and students must be provided with the right to own and control how AI models are used to facilitate their own teaching and learning [43]. Therefore, it is necessary to clearly state the acknowledgment and responsibility of the actions of every relevant person involved in using AI models.

5. Principle of Autonomy

Autonomy is the fifth ethical principle in the studies reviewed (*N* = 8). According to the literature reviewed, the principle of autonomy is generally associated with these key words, such as freedom, self-determination, independence, and empowerment. For example, Busch et al. emphasized that AI models should be considered as an addition to teaching and learning, rather than completely replacing traditional teaching materials and approaches, so that teachers and students can decide at any time whether or not to apply AI models [78]. Köbis and Mehner believed that it is essential to ensure the decisions made when using AI models in the classroom are aligned with human values and prevent compromising human independence [72]. Therefore, learner-centered use of AI models must be cultivated to strengthen students' authority and autonomy over their own learning.

6. Principle of Beneficence

Beneficence is the sixth ethical principle in the reviewed studies (*N* = 6). In the context of using AI models in the classroom, the principle of beneficence is always described in terms of providing appropriate training about AI applications, benefiting the development of students, and promoting the well-being of teachers. First, training courses on using AI models should cover knowledge, skills, and ethical considerations to improve the AI literacy of teachers and students [79]. Busch et al. considered that proper education and training on using AI models can not only effectively integrate AI into the classroom, but can also foster the AI literacy of students and teachers, and enhance autonomy and justice [78]. In addition, the use of AI models must meet the developmental needs of students and stay consistent with the educational goals [66]. Similarly, teacher well-being was also considered an important principle in Adams et al. [75] and Adams et al. [73]; it refers to the needs and the physical and mental health of teachers faced with the challenge of using AI models in the classroom.

C. What Are Future Directions of Research and Practice of Related Research?

Continuous discussions are required to comprehensively understand, prevent and overcome the ethical implications of AI model's use in the classroom. Table V displays the proposed future research and practice directions regarding the ethical considerations about the use of AI in the classroom.

First, while AI models have the capability to revolutionize education, it also raises a number of ethical implications. The results of this literature analysis show that the ethical implications of the use of AI in the classroom are not limited to bias and data privacy disclosure, but are also related to the ethical implications of reduced autonomy and academic integrity (see Table 2). Although these implications are mentioned or discussed in the existing literature, many of them have not been studied in detail. Hence, further substantive discussion, understanding and solution of these implications are required (see Table 5). On the one hand, educational decisions made by algorithm-based AI predictive models lead to bias and discrimination, but it is not clear what algorithmic features and attributes are needed to reduce such data bias. Future work should continue to optimize educational AI predictive models by training them using unbiased data. On the other hand, when AI is applied in the classroom, the algorithm-driven education prediction systems and AI-driven mobile learning apps decrease students' autonomy. In the future, learner-centered use of

TABLE V. The Proposed Future Directions of Research and Practice of Related Research

Implications	Direction for research	Guideline for practice
The ethical implications of using AI models in the classroom need to be addressed further.	What algorithmic features and attributes are needed to reduce data bias in AI prediction models?	Continuously optimize educational AI prediction models by training them with unbiased data.
	How can student autonomy in the use of AI models be maintained?	Cultivate learner-centered use of AI models.
		Teaching AI and ethics lessons in educational contexts.
The ethical principles of using AI models in the classroom lack elucidation of the precise mechanism.	How do the ethical principles lead to ethical functioning of using AI models?	Strong policy guidance for educators is needed.
	How can the ethical principles be integrated into teachers' teaching practice?	Adopt more pedagogical responsive and context-sensitive ethical approaches in the use of AI models.
The ethical implications of using AI models in the classroom lack accurate assessment.	How are the ethical implications of AI models use investigated in the classroom?	Various approaches, such as case studies or interviews, and more clear-cut empirical research is required.
		Clarify the definition of ethics of using AI models in the classroom.
	What should be considered to evaluate the ethical implications of AI models use in the classroom?	Positive and negative impacts should both be considered.
		The precise needs of the stakeholders should be taken into account.

AI models should be cultivated, and AI ethics courses are required in educational settings to learn about the ethical use of AI [82].

Second, in fact, most of the ethical principles discussed in the literature are more applicable to general AI systems or computing and design environments, and there is a lack of research on ethical principles for specific use cases in the classroom [15]. Thus, the precise mechanism of ethical principles of using AI models remains to be elucidated and extended to the implementation phase in the classroom. For example, how ethical principles lead to the ethical function of using AI models in practice remains ambiguous, and more robust policy guidance for educators is needed. Additionally, how to integrate the ethical principles of using AI models into the teaching practice of teachers remains to be further explored. More pedagogical responsive and context-sensitive ethical approaches should be designed and adopted in the use of AI models, to avoid and address these ethical implications [73].

Third, the ethical implications need to be more accurate assessment. On the one hand, how to investigate the ethical implications when applying AI model's during teaching and learning is a direction that needs further research. The existing discussion on the ethical implications about the AI models use is mostly descriptive research, and more clear empirical research is required [63],[11]. Moreover, future research needs to clarify the ethical definition of using AI models in the classroom to help identify the ethical implications. On the other hand, future work will be necessary to develop or customize ethical implication assessments for specific AI models use cases in classroom contexts. Thus, not only both positive and negative effects, but also the precise needs of the relevant stakeholders should be considered when assessing the ethical implications of AI models use in the classroom.

V. Discussion

A. Five Ethical Implications of Using AI Models in the Classroom

The first research question identifies the five major ethical implications of the use of AI models in the classroom, namely algorithmic bias and discrimination, data privacy leakage, lack of transparency, decreased autonomy, and academic misconduct. First, in terms of algorithmic bias and discrimination, although the main promise of AI models is to improve the objectivity and accuracy of instruction, the fact is that when AI is applied in the classroom, these inherent social biases, discrimination, and power structures are naturally embedded in them, and are even further perpetuated and exacerbated [10]. Masters emphasized that there is no such thing as ethically neutral AI, as all AI models react and make decisions that favor specific groups, leading to bias and discrimination in the classroom [71]. Second, another ethical implication surrounding the use of AI models in the classroom is data privacy leakage. The disclosure of personal information, surveillance and student tracking, lack of informed consent, and the right to be forgotten were often considered in the use of AI models. Previous review work has also identified the ethical implications of data privacy leakage [10], but in this study, the absence of students' right to be forgotten was further reviewed. Through the review, this study found that students lack the chance to choose whether to enter the educational AI system or not, but also lack the chance to opt out of the system. Third, in terms of the lack of transparency, when using AI models in the classroom, there is not only no clear explanation of the process and results, but also no detailed explanation of the actual shortcomings and potential risks. The latter, in particular, has not been discussed in great detail, but it does in fact exist [41]. In particular, while AI models provide personalized learning for students, they also have the problem of not always being able to understand the open-ended needs of students. Fourth, in terms of reduced autonomy for teachers and students, algorithm-based forecasting and decision-making systems and inappropriate delegation of authority have led to this ethical implication. Therefore, it is essential to consider the long-term consequences of using AI models for students' learning and cognitive abilities. Fifth, academic misconduct also emerges when AI technologies are misused by students. However, a previous review study has paid less attention to this ethical implication [10]. Therefore, this study extended the previous review work.

B. Six Ethical Principles of Using AI Models in the Classroom

The second research question revealed six ethical principles of using AI models in the classroom, namely fairness, privacy, transparency, accountability, autonomy and beneficence.

First, the principle of fairness, as the most mentioned ethical principle in the review study, requires representation, accessibility, and inclusiveness of using AI models, in order to achieve algorithmic processes and results without discrimination or bias for students and teachers [41]. As noted in the previous section, when AI models fail to

understand the needs of underrepresented students, such as minority students, this group of students may already feel marginalized. Hence, unbiased data training for AI models is recommended. Surprisingly, however, the AI Act recently approved by the EU did not even mention the principle of "fairness," but explicitly mentioned the term "non-discrimination." Actually, the "fairness" in the Act is relatively hidden, and the expression "non-discrimination" is intended to reflect specific regulatory objectives, because it has a more specific measure than the concept of fairness [83]. Therefore, the principle of fairness in this study is essentially consistent with the term "non-discrimination" in the Act.

Second, the principle of privacy calls for the protection of personal data and information of teachers and students in the use of AI models. It is worth noting that, based on the perspective of the AI life cycle, Jang et al. further pointed out the need to protect new data generated in the use of AI models [63]. This was not examined in the previous literature review work [10]. In addition to highlighting the full life cycle of AI, at the legal level, the EU's AI Act protects personal privacy by assessing the categories of AI risks. For example, the Act classifies the use of "real-time" remote biometrics in public places for law enforcement purposes as high risk. The practice is prohibited because it poses a great risk to an individual's private life [84]. Therefore, it is suggested that in the classroom, both aspects of the full life cycle of AI models and risk categories need to be considered to protect personal privacy.

Third, the principle of transparency mainly includes interpretability, traceability, data ownership and communication. Similar to the definition of transparency in the AI Act proposed by the EU [85], transparency here is meant not just as an algorithmic attribute, but as a means of supporting broader and different values. This act further distinguishes among technical, enabling and protective transparency. In particular, in addition to presenting the conditions, process, and results of using AI models in detail, the actual shortcomings and potential risks of using AI models in the classroom should be clearly stated, and even relevant remedial suggestions should be made [16]. It is worth noting that although some of the complex AI models, such as deep learning neural networks, have techniques for interpreting and proving results, there is still a need to customize different interpretations for different audiences [29]. In the field of education, it is necessary to provide detailed explanations for the use of AI models to teachers and students.

Fourth, the principle of accountability is closely related to the previous principle of transparency, and both principles are mentioned simultaneously in multiple studies [11] [17]. Canada's AI and Data Act also created a strong link between accountability, transparency and privacy provisions [86]. This means that individuals who use AI models responsibly also have an obligation to be transparent and provide data subjects with an explanation of the information intended or actually used by the AI model. In this Act, the principle of accountability specifically emphasizes responsible anonymization of data. In fact, accountability focuses more on requiring the establishment of mechanisms to ensure responsibility and accountability for AI models before and after their use [63]. The EU's AI Act adopted different regulatory measures and accountability based on classifying different risk levels of using AI [87]. Therefore, this study recommends that systems of responsibility for the possible consequences of using AI models should be developed and implemented, to clarify the obligations of teachers and students in the classroom, and especially to protect their privacy rights.

Fifth, the principle of autonomy emphasizes that teachers and students have the ability and right to act in accordance with their own interests and values, despite being under the monitoring system of AI. In previous studies, it is included in the human-centered principle, but

these studies all emphasize human values in the use of AI [11] [15]. From a legal point of view, the US AI Bill of Rights also mentioned that the use of AI must follow this principle: where appropriate, individuals can voluntarily opt out of the automated system and choose a human alternative [18]. However, the expression of the term "where appropriate" is vague and subject to different interpretations. Therefore, the boundary and degree of autonomy should be clearly defined. When teachers apply AI models to assist classroom teaching, the key is to reasonably design learning materials and tasks, and to consider in what dimensions and to what extent students' autonomy can be guaranteed, so as to avoid reducing students' learning efforts and their learning autonomy.

Finally, the principle of beneficence calls for attention to the sustainable development of teachers and students when using AI models. Importantly, this principle emphasizes appropriate education and training on AI for teachers and students, which would help students critically understand AI and promote the development of teachers' intellectual competence [65]. Hence, specialized AI ethics courses and lectures on improving AI literacy for teachers and students are suggested. This is similar to the "social and environmental well-being" mentioned in the EU's AI Act, which refers to the idea that AI should be developed and used in a sustainable and environmentally friendly way, while monitoring and assessing the long-term impacts on individuals and society [88]. Slightly different, training to enhance individual AI literacy is not mentioned in the Act, which focuses more on sustainable considerations in the development and use of AI.

C. Future Directions of Research and Practice of Related Research

The third research question concerns the main future directions of research and practice regarding the ethical implications and principles of AI models use in the classroom. Firstly, due to the lack of research on the specific solutions to the ethical implications about the use of AI in the existing literature, most of the discussions remain at the macro level, and so further exploration is needed in the future. From a technical point of view, how to train unbiased algorithms and what characteristics they should have needs to be explored. From a teaching point of view, how teachers can maintain students' autonomy when using AI models, and how teaching materials should be properly designed should be examined. On the one hand, it is recommended that AI ethics courses and lectures support students' autonomous development when using AI models. This has also been considered in previous studies [82]. However, it is further suggested that when teachers design teaching materials, they should allow ample opportunity for students' autonomous development, and some traditional classroom teaching is still valuable, such as class discussion, rather than relying entirely on AI models. Secondly, future research needs to further elucidate the precise mechanisms of ethical principles of using AI models in the classroom, and extend them to the implementation phase because the ethical principles discussed in the existing literature lack research on specific classroom use cases. This includes how ethical principles are translated into ethical functions and how they are integrated into teachers' instruction practices, all of which are unclear. In fact, more ethical principles in the context of teaching situations should be explored, which have not been mentioned in previous studies. Thus, future research should be based on different classroom types, such as online classes and flipped classes, to conduct different specific discussions. Finally, the accurate assessment of the ethical implications of AI use in the classroom is required, because what and how to assess it remains unclear, and more empirical research is called for. This finding echoes Memarian and Doleck's research [17], which reviewed the existing investigation methods of ethical implications and revealed the deficiency of quantitative research methods. However, this study further reported that ethical implication assessments for the future focus on teaching

and learning in the classroom, and developing or customizing ethical implication assessments for specific AI models use cases in classroom contexts. Due to the complexity of the situations presented by the real classroom, the evaluation of the ethical implications of using AI models should consider various factors, especially when it comes to sensitive topics such as student privacy aspects.

VI. Conclusions

A. Implications

The main theoretical contribution of this study is to outline the five ethical implications (including algorithmic bias and discrimination, data privacy leakage, lack of transparency, decreased autonomy, and academic misconduct), six ethical principles (including fairness, privacy, transparency, accountability, autonomy and beneficence), and the main future research directions and practices of the related research. This structure stems from a systematic review that helps to understand and conceptualize practice and research of using AI models ethically in the classroom. Additionally, this review is conducive to validating some less explored areas to help researchers determine the direction of future research efforts on the ethical implications of the AI models use in the classroom, for example, the strategic and evaluation study of the ethical implications, which still remain less researched. Meanwhile, some specific guidance schemes are provided in this study. On a practical level, this research helps educators and learners to understand which behaviors are ethical when using AI models for education-related purposes, which could lead to the implementation of appropriate regulation. Importantly, the study provides a detailed elaboration of ethical principles and practical recommendations to better promote the ethical use of AI models in the classroom.

B. Limitations and Future Work

However, several limitations of this systematic review must be acknowledged. First, the literature reviewed in this study mainly comes from two databases. Future research can consider other databases, such as Science Direct and Google Scholar, to retrieve suitable papers. In addition, during the eligibility phase of this systematic review, 32 articles were excluded because the full text was not available. Finally, since the articles reviewed in this study are mainly from Europe and North America, most represent Western perspectives. Therefore, there should be further reviews of the research from other continents or in other languages to gain a broader understanding of the ethical implications of AI model use in the classroom.

Acknowledgment

References

[1] S. K. Das, A. Dey, A. Pal, N. B. Roy, "Applications of Artificial Intelligence in Machine Learning: Review and Prospect," *International Journal of Computer Applications,* vol. 115, pp. 31-41, 2015, doi:10.5120/20182-2402.

[2] W. Holmes, M. Bialik, C. Fadel, "*Artificial Intelligence in Education: Promises and Implications for Teaching and Learning,*" Boston, MA: Center for Curriculum Redesign, 2019.

[3] J. M. Flores-Vivar, F. J. García-Peñalvo, "Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4)," *Comunicar,* vol. 31, no. 74, pp. 37-47, 2023, doi:10.3916/C74-2023-03.

[4] J. A. Kumar, B. Bervell, S. Osman, "Google classroom: insights from

Malaysian higher education students' and instructors' experiences. *Education and Information Technologies*, vol. 25, no. 5, pp. 4175-4195, 2020, doi:10.1007/s10639-020-10163-x.

[5] R. Reyes, D. Garza, L. Garrido, V. De la Cueva, J. Ramirez, "Methodology for the Implementation of Virtual Assistants for Education Using Google Dialgflow," In: Martínez-Villaseñor, L., Batyrshin, I., Marín-Hernández, A. (eds) *Advances in Soft Computing. MICAI 2019.* Lecture Notes in Computer Science, Springer, Cham, vol. 11835, pp. 440-451, 2019, doi:10.1007/978-3-030-33749-0_35.

[6] K. Naidu, K. Sevnarayan, "ChatGPT: An ever-increasing encroachment of artificial intelligence in online assessment in distance education," *Online Journal of Communication and Media Technologies*, vol. 13, no. 3, 2023, doi:10.30935/ojcmt/13291.

[7] M. Zafari, J. S. Zafari, A. Sadeghi-Niaraki, C. M. Choi, "Artificial Intelligence Applications in K-12 Education: A Systematic Literature Review,"*IEEE Access,* vol. 10, pp. 61905-61921, 2022, doi: 10.1109/ACCESS.2022.3179356.

[8] R. F. Murphy, "Artificial intelligence applications to support k-12 teachers and teaching: a review of promising applications, challenges, and risks," *Perspective,* pp. 1-20, 2019, doi:10.7249/PE315.

[9] M. Karabacak, B. B. Ozkara, K. Margetis, M. Wintermark, S. Bisdas, "The Advent of Generative Language Models in Medical Education," *JMIR Medical Education*, vol. 9, 2023, doi: 10.2196/48163.

[10] S. Akgun, C. Greenhow, "Artificial intelligence in education: Addressing ethical challenges in K-12 settings," *AI and Ethics,* vol. 2, no. 3, pp. 431-440, 2022, doi:10.1007/s43681-021-00096-7.

[11] B. Klimova, M. Pikhart, J. Kacetl, "Ethical issues of the use of AI-driven mobile apps for education," *Frontiers in Public Health*, v0l. 10, no. 1118116, 2023, doi:10.3389/fpubh.2022.1118116.

[12] N. Selwyn, "Debería los robots sustituir al profesorado? La IA y el futuro de la educación,"*Ediciones Morata, 2*019. Available: https://bit.ly/3zxyPmO.

[13] F. J. García Peñalvo, F. Llorens-Largo, J. Vidal, J. "The new reality of education in the face of advances in generative artificial intelligence," *ITEN - Ibero-American Journal of Distance Education*, vol. 27, no. 1, pp. 9-39, doi:10.5944/ried.27.1.37716.

[14] F. J. García-Peñalvo, A. Vázquez-Ingelmo, "What Do We Mean by GenAI? A Systematic Mapping of The Evolution, Trends, and Techniques Involved in Generative AI," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 4, pp. 7-16, doi:10.9781/ijimai.2023.07.006.

[15] A. Nguyen, H. N. Ngo, Y. Hong, Dang, Belle, B. P. T. Nguyen, "Ethical principles for artificial intelligence in education," *Education and Information Technologies,* vol. 28, pp. 4221-4241, 2023, doi:10.1007/s10639-022-11316-w.

[16] W. Holmes, K. Porayska-Pomsta, K. Holstein, E. Sutherland, T. Baker, S. B. Shum, O. C. Santos, M. T. Rodrigo, M. Cukurova, I. I. Bittencourt, K. R. Koedinger, "Ethics of AI in Education: Towards a Community-Wide Framework," *International Journal of Artificial Intelligence in Education,* vol. 32, no. 3, pp. 504-526, 2021, doi:10.1007/s40593-021-00239-1.

[17] B. Memarian, T. Doleck, "Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review," *Computers and Education: Artificial Intelligence,* vol. 5, no. 100152, 2023, doi:10.1016/j.caeai.2023.100152.

[18] E. Hine, L. Floridi, "The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction," *Minds & Machines*, vol. 33, no. 2, pp. 285-292, 2023, doi:10.1007/s11023-023-09625-1.

[19] A. E. Muhammad, K. -C. Yow, "Demystifying Canada's Artificial Intelligence and Data Act (AIDA): The good, the bad and the unclear elements," *2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE),* Regina, SK, Canada, pp. 510-515, 2023, doi: 10.1109/CCECE58730.2023.10288878.

[20] "EU AI Act: First regulation on artificial intelligence," European Parliament: News, 8 June 2023; www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

[21] European Parliament (2023) MEPs ready to negotiate first-ever rules for safe and transparent AI. European Parliament. 14.07.2023. Available at: https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai

[22] F. Ouyang, P. Jiao, "Artificial Intelligence in Education: The Three Paradigms," *Computers & Education: Artificial Intelligence,* vol. 100020, 2021, doi:10.1016/j.caeai.2021.100020.

[23] K. Y. Tang, C. Y. Chang, G. J. Hwang, "Trends in artificial intelligence supported e-learning: A systematic review and co-citation network analysis (1998-2019)," *Interactive Learning Environments,* vol. 31, no. 4, pp. 2134-2152, 2021, doi.org/10.1080/10494820.2021.1875001.

[24] F. Tahiru, "AI in Education: A Systematic Literature Review," *Journal of Cases on Information Technology,* vol. 23, no. 1, pp. 1-20, 2021, doi:10.4018/JCIT.2021010101.

[25] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, Prisma Group, "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement,"*PLoS medicine,* vol. 6, no. 7, pp. e1000097, 2009, doi:10.1371/journal.pmed.1000097.t001.

[26] L. Zhang, P. P. Wei, Y. W, Zhang, N. Wang, "Artificial Intelligence and Edge Computing Technology Promote the Design and Optimization of Flipped Classroom Teaching Models for Higher Vocational, Ideological, and Political Courses,"*Mobile Information Systems,* vol. 2022, doi:10.1155/2022/5385386.

[27] A. Suresh, J. Jacobs, C. Clevenger, V. Lai, C. H. Tan, J. H. Martin, T. Sumner, "Using AI to Promote Equitable Classroom Discussions: The TalkMoves Application," *22nd International Conference on Artificial Intelligence in Education (AIED),* vol. 12749, pp. 344-348, 2021, doi:10.1007/978-3-030-78270-2_61.

[28] M. Karabacak, B. B. Ozkara, K. Margetis, M. Wintermark, S. Bisdas, "The Advent of Generative Language Models in Medical Education," *JMIR Medical Education*, vol. 9, 2023, doi:10.2196/48163.

[29] Y. Chen. S. Jensen, L. J. Jensen, et al. "Artificial Intelligence (AI) Student Assistants in the Classroom: Designing Chatbots to Support Student Success," *Information Systems Frontiers*, vol. 25, no. 1, pp. 161-182, 2023, doi:10.1007/s10796-022-10291-4.

[30] X. Chen, H. Xie, G. J. Hwang, "A multi-perspective study on artificial intelligence in education: grants, conferences, journals, software tools, institutions, and researchers," *Computers and Education: Artificial Intelligence,* vol. 1, no. 100005, 2020, doi:10.1016/j.caeai.2020.100005.

[31] B. C. L. Christudas, E. Kirubakaran, P. R. J. Thangaiah, "An evolutionary approach for personalization of content delivery in e-learning systems based on learner behavior forcing compatibility of learning materials," *Telematics and Informatics,* vol. 35, no. 3, pp. 520-533, 2018, doi:10.1016/j.tele.2017.02.004.

[32] G. George, A. M. Lal, "Review of ontology-based recommender systems in e-learning," *Computers & Education,* vol. 142, no. 7, pp. 103642, 2019, doi:10.1016/j.compedu.2019.103642.

[33] S. Verma, "Weapons of math destruction: how big data increases inequality and threatens democracy," *Vikalpa,* vol. 44, no. 2, pp. 97-8, 2019, doi:10.1177/0256090919853933.

[34] P. M. Regan, J. Jesse, "Ethical challenges of edtech, big data and personalized learning: twenty-first century student sorting and tracking," *Ethics and Information Technology,* vol. 21, pp. 167-179, 2019, doi:10.1007/s10676-018-9492-2.

[35] M. J. Reiss, "The use of AI in education: Practicalities and ethical considerations," *London Review of Education,* vol. 19, no. 15, pp. 1-14, 2021, doi:10.14324/LRE.19.1.05.

[36] K., Naidu, K. Sevnarayan,"ChatGPT: An ever-increasing encroachment of artificial intelligence in online assessment in distance education,"*Online Journal of Communication and Media Technologies,* vo. 13, no. 3, 2023, doi: 10.30935/ojcmt/13291.

[37] N. R. Möllmann, M., Mirbabaie, S. Stieglitz, "Is it alright to use artificial intelligence in digital health? A systematic literature review on ethical considerations," *Health Informatics Journal,* vol. 27, no. 4, pp. 1-17, 2021, doi:10.1177/14604582211052391.

[38] T. Hagendorff, "The ethics of AI ethics: An evaluation of guidelines," *Minds and Machines,* vol. 30, no. 1, pp. 99-120, 2020, doi:10.1007/s11023-020-09517-8.

[39] F, Miao, W, Holmes, R. Huang, H. Zhang, "AI and education: Guidance for policy-makers," *United Nations Educational, Scientific and Cultural Organization,* 2021. Available: https://unesdoc.unesco.org/ark:/48223/pf0000376709.

[40] European Commission, "The European Commission's high-level expert group on artificial intelligence: Ethics guidelines for trustworthy AI," *European Union Publications Office,* 2019 Available: https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai

[41] S. Larsson, F. Heintz, "Transparency in artificial intelligence," *Internet Policy Review,* vol. 9, no. 2, pp. 1-16, 2020, doi:10.14763/2020.2.1469.

[42] Z. Slimi, B. V. Carballido, "Navigating the Ethical Challenges of Artificial Intelligence in Higher Education: An Analysis of Seven Global AI Ethics Policies," *TEM Journal-Technology Education Management Informatics,* vol. 12, no. 2, pp. 590-602, 2023, doi:10.18421/TEM122-02.

[43] Y. Hong, A. Nguyen, B. Dang, B. P. T. Nguyen, "Data Ethics Framework for Artificial Intelligence in Education (AIED)," *International Conference on Advanced Learning Technologies (ICALT),* pp. 297-301, 2022, doi:10.1109/ICALT55010.2022.00095.

[44] D. Gough, S. Oliver, J. Thomas, "An introduction to systematic reviews, (2nd ed.,)," Los Angeles: SAGE, 2017.

[45] C. Guan, J. Mou, Z. Jiang, "Artificial intelligence innovation in education: A twenty-year data-driven historical analysis," *International Journal of Innovation Studies,* vol. 4, no. 4, pp. 134-147, 2020, doi:10.1016/j.ijis.2020.09.001.

[46] A. A. Chadegani, H. Salehi, M. M. Yunus, H. Farhadi, M. Fooladi, M. Farhadi, N. A. Ebrahim, "A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases," *Asian Social Science,* vol. 9, no. 5, pp. 18-26, 2013, doi:10.5539/ass.v9n5p18.

[47] J. F. Burnham, "Scopus database: a review," *Biomedical digital libraries,* vol. 3, pp. 1-8, 2006, doi:10.1186/1742-5581-3-1.

[48] D. T. K. Ng, M. Lee, R. J. Y. Tan, et al. "A review of AI teaching and learning from 2000 to 2020," *Education and Information Technologies,* vol. 28, pp. 8445-8501, 2023, doi:10.1007/s10639-022-11491-w.

[49] D. Nicholas, A. Watkinson, H. R. Jamali, E. Herman, C. Tenopir, R. Volentine, et al. "Peer review: still king in the digital age," *Learned Publishing,* vol. 28, no. 1, pp. 15-21, 2015, doi:10.1087/20150104.

[50] C. Lefebvre, J. Glanville, S. Briscoe, A. Littlewood, C. Marshall, M.-I. Metzendorf, A. Noel-Storr, T. Rader, F. Shokraneh, J. Thomas, L. S. Wieland, *"Chapter 4: Searching for and selecting studies,"* In H. JPT, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), Cochrane Handbook for Systematic Reviews of Interventions (Version 6.2), 2021. Available:https://www.training.cochrane.org/handbook.

[51] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," *In Proceedings of the 18th international conference on evaluation and assessment in software engineering,* pp. 1-10, 2014, doi:10.1145/2601248.2601268.

[52] M. Petticrew, H. Roberts, *"Systematic reviews in the social sciences,"* 11, Blackwell Publishing Ltd, Oxford, UK, 2006, doi:10.1002/9780470754887.

[53] J. Webster, R. T. Watson, "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterly,* vol. 26, no. 2, pp. xiii-xxiii, 2022, doi:10.2307/4132319.

[54] V. Braun, V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology,* vol. 3, no. 2, pp. 77-101, 2006, doi:10.1191/1478088706qp063oa.

[55] A. Strauss, J. Corbin, "Grounded theory methodology: An overview," In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273-285). 1995, SAGE

[56] U. H. Graneheim, B. Lundman, "Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness," *Nurse Education Today*, vol. 24, no. 2, pp. 105e112, 2004, doi:10.1016/j.nedt.2003.10.001.

[57] H. F. Hsieh, S. E. Shannon, "Three approaches to qualitative content analysis," *Qualitative Health Research,* vol. 15, no. 9, pp. 1277-1288, 2005, doi:10.1177/1049732305276687.

[58] E. Dieterle, C. Dede, M. Walker, "The cyclical ethical effects of using artificial intelligence in education,"*AI & Society,* 2022, doi:10.1007/s00146-022-01497-w.

[59] C. Kooli, "Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions," *Sustainability*, vol. 15, no. 7, pp. 5614, 2023, doi:10.3390/su15075614.

[60] A. Latham, S. Goltz, "A Survey of the General Public's Views on the Ethics of Using AI in Education," *International Conference on Artificial Intelligence in Education (AIED),* vol. 11625, pp. 194-296, 2019, doi:10.1007/978-3-030-23204-7_17.

[61] E. Kowch, "Ethics to Prepare teachers for Professional Service Robots in Classrooms," *International Joint Conference on Information, Media and*

Engineering (IJCIME), 2021, doi:10.1109/IJCIME49369.2019.0010.

[62] W. Holmes, F. Iniesto, S. Anastopoulou, J. Boticario, "Stakeholder Perspectives on the Ethics of AI in Distance-Based Higher Education," *The International Review of Research in Open and Distributed Learning,* vol. 24, pp. 96-117, 2023, doi:10.19173/irrodl.v24i2.6089.

[63] Y. Jang, S. Choi, H. Kim, "Development and validation of an instrument to measure undergraduate students' attitudes toward the ethics of artificial intelligence (AT-EAI) and analysis of its difference by gender and experience of AI education,"*Education and Information Technologies,* vol. 27, pp. 11635-11667, 2022, doi:10.1007/s10639-022-11086-5.

[64] R. Farrow, "The possibilities and limits of XAI in education: a socio-technical perspective," *Learning, Media and Technology,* vol. 48, no. 2, pp. 266-279, 2023, doi: 10.1080/17439884.2023.2185630.

[65] I. Tuomi, "A Framework for Socio-Developmental Ethics in Educational AI," *In the Proceedings of the 56th Hawaii International Conference on System Sciences,* 2022, doi:10.13140/RG.2.2.36133.58089.

[66] D. Schiff, "Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies," *International Journal of Artificial Intelligence in Education,* vol. 32, pp. 527-563, 2022, doi:10.1007/s40593-021-00270-2.

[67] C. Adams, P. Pente, G. Lemermeyer, J. Turville, G. Rockwell, "Artificial Intelligence and Teachers' New Ethical Obligations," *The International Review of Information Ethics,* vol. 31, no. 1, 2022, doi:10.29173/irie483.

[68] N. Ghotbi, M. T. Ho, "Moral Awareness of College Students Regarding Artificial Intelligence," *Asian Bioethics Review,* vol. 13, pp. 421-433, 2021, doi:10.1007/s41649-021-00182-2.

[69] N. Ghotbi, M. T. Ho, P. Mantello, "Attitude of college students towards ethical issues of artificial intelligence in an international university in Japan," *AI & Society,* vol. 37, pp. 283-290, 2022, doi:10.1007/s00146-021-01168-2.

[70] A. Matias, I. Zipitria, "Promoting Ethical Uses in Artificial Intelligence Applied to Education. In: Frasson, C., Mylonas, P., Troussas, C. (eds) Augmented Intelligence and Intelligent Tutoring Systems," *International Conference on Intelligent Tutoring Systems,* vol. 13891, pp. 604-615, 2023, doi:10.1007/978-3-031-32883-1_53.

[71] K. Masters, "Ethical use of Artificial Intelligence in Health Professions Education: AMEE Guide No. 158," *Medical Teacher,* vol. 45, no. 6, pp. 574-584, 2023, doi: 10.1080/0142159X.2023.2186203.

[72] L. Köbis, C. Mehner, "Ethical Questions Raised by AI-Supported Mentoring in Higher Education," *Frontiers in Artificial Intelligence,* vol. 4, no. 624050, 2021, doi: 10.3389/frai.2021.624050.

[73] C. Adams, P. Pente, G. Lemermeyer, G. Rockwell, "Ethical principles for artificial intelligence in K-12 education.," *Computers and Education: Artificial Intelligence,* vol. 4, no. 100131, 2023, doi:10.1016/j.caeai.2023.100131.

[74] B. Han, S. Nawaz, G. Buchanan, and D. McKay, "Ethical and Pedagogical Impacts of AI in Education,"*International Conference on Artificial Intelligence in Education (AIED),* vol. 13916, pp. 667-673, 2023, doi:10.1007/978-3-031-36272-9_54.

[75] C. Adams, P. Pente, G. Lemermeyer, G. Rockwell, "Artificial Intelligence Ethics Guidelines for K-12 Education: A Review of the Global Landscape,"*International Conference on Artificial Intelligence in Education (AIED),* vol. 12749, pp. 24-28, 2021, doi.org/10.1007/978-3-030-78270-2_4.

[76] L. H. Yu, Z. G. Yu, "Qualitative and quantitative analyses of artificial intelligence ethics in education using VOSviewer and CitNetExplorer," *Frontiers in Psychology,* vol. 14, no. 1061778, 2023, doi: 10.3389/fpsyg.2023.1061778.

[77] I. Celik, "Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education," *Computers in Human Behavior,* vol 138, no. 107468, 2023, doi:10.1016/j.chb.2022.107468.

[78] F. Busch, L. C. Adams, K. K. Bressem, "Biomedical Ethical Aspects Towards the Implementation of Artificial Intelligence in Medical Education," *Medical Science Educator,* vol. 33, no. 1007-1012, 2023, doi.org/10.1007/s40670-023-01815-x.

[79] J. M. Flores-Vivar, F. J. García-Peñalvo, "Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4)," *Comunicar, v*ol. 31, no.74, pp. 37-47, 2023, doi:10.3916/C74-2023-03.

[80] M. Chaudhry, M. Cukurova, R. Luckin,"A Transparency Index Framework for AI in Education,"*23rd International Conference on Artificial Intelligence in Education (AIED)*, Durham Univ, Durham, ENGLAND, vol. 13356, pp. 195-198, 2022, doi:10.1007/978-3-031-11647-6_33.

[81] A. Mouta, A. M. Pinto-Llorente, E. M. Torrecilla-Sánchez, "Blending machines, learning, sense of agency, and ethics: Designing an in-depth framework with Experts using the Delphi Method approach," *In Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'21)* , Association for Computing Machinery, New York, NY, USA, pp. 665-670, 2021, doi:10.1145/3486011.3486545.

[82] T. Foltynek, S. Bjelobaba, I. Glendinning, et al. "ENAI Recommendations on the ethical use of Artificial Intelligence in Education," *International Journal for Educational Integrity,* vol. 19, no. 1, 12, 2023, doi:10.1007/s40979-023-00133-4.

[83] J. Pfeiffer, J. Gutschow, C. Haas, et al, "Algorithmic Fairness in AI," *Business & Information Systems Engineering,* vol. 65, no. 2, pp. 209-222, 2023, doi:10.1007/s12599-023-00787-x.

[84] D. Svantesson, "The European Union Artificial Intelligence Act: Potential implications for Australia," *Alternative Law Journal,* vol. 47, no. 1, pp. 4-9, doi:10.1177/1037969X211052339.

[85] B. Gyevnara, N. Fergusona, B. Schaferb, "Bridging the Transparency Gap:What Can Explainable AI Learn From the AI Act?," European Conference on Artificial Intelligence, vol. 372, pp. 964-971, doi:10.3233/FAIA230367.

[86] A. E. Muhammad, K. -C. Yow, "Demystifying Canada's Artificial Intelligence and Data Act (AIDA): The good, the bad and the unclear elements," *2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE),* Regina, SK, Canada, 2023, pp. 510-515, doi: 10.1109/CCECE58730.2023.10288878.

[87] R. Paul, "European artificial intelligence 'trusted throughout the world': Risk-based regulation and the fashioning of a competitive common AI market," *Regulation & Governance,* 2023, doi:10.1111/rego.12563.

[88] [K. Kalodanis, P. Rizomiliotis, D. Anagnostopoulos, "European Artificial Intelligence Act: an AI security approach," Information and Computer Security, vol. ahead-of-print, no. ahead-of-print, 2023, doi:10.1108/ICS-10-2022-0165.

Lin Tang

Lin Tang is a graduate student at Nanjing Normal University in China. She has experience in data mining, data analysis and text analysis. Her research interests are now focused on deep learning, modern educational technology, technical education, and other artificial intelligence applications in education.

Yu-Sheng Su

Dr. Yu-Sheng Su is currently an associate professor of the Department of Computer Science and Information Engineering at National Chung Cheng University, Taiwan. His research interests include Cloud Computing, Big Data Analytics, Intelligent Systems, and the Metaverse.