HiLL PUBLISHING

The Educational Review, USA, 2025, 9(2), 239-243
https://www.hillpublisher.com/Journals/er/
ISSN Online: 2575-7946
ISSN Print: 2575-7938

Ethical Issues and Risk Mitigation Strategies in the Application of Artificial Intelligence in Education

Siqi Guo, Xin Xie, Yanmin Niu*

School of Computer and Information Science, Chongqing Normal University, Chongqing 401333, China.

How to cite this paper: Siqi Guo, Xin Xie, Yanmin Niu. (2025). Ethical Issues and Risk Mitigation Strategies in the Application of Artificial Intelligence in Education. *The Educational Review, USA*, 9(2), 239-243.
DOI: 10.26855/er.2025.02.013

Received: January 24, 2025
Accepted: February 21, 2025
Published: March 18, 2025

Corresponding author: Yanmin Niu, School of Computer and Information Science, Chongqing Normal University, Chongqing 401333, China.

Abstract

In this artificial intelligence era, the continuous integration of artificial intelligence technologies with education offers significant opportunities to enhance the quality of education. However, it simultaneously gives rise to numerous complex ethical issues and challenges. Identifying the defining principles of ethics in AI educational applications, analyzing existing ethical problems, and developing feasible risk mitigation strategies have become critical research questions. This paper elucidates the concept and developmental trajectory of AI in education, analyzes ethical issues arising from its application, such as algorithmic hegemony, deviation from traditional teacher-student relationships, and loss of privacy rights. Furthermore, it proposes corresponding mitigation strategies for three types of risks: data security risks, technology dependence risks, and educational equity risks. The aim is to promote the practical application of AI technology in the education sector and provide guidance and support for the sustainable development of AI-powered education.

Keywords

Artificial Intelligence; AI in Education Applications; Ethical Issues; Risk Mitigation

1. Introduction

As a key driver of technological and industrial change, artificial intelligence (AI) is profoundly impacting education, ushering in a new era of digitally-enhanced learning. Since 2017, Chinese governmental bodies have increasingly emphasized responsible AI development and ethical guidelines, culminating in policies like the "Governance Principles for a New Generation of Artificial Intelligence" and the "Ethical Norms for a New Generation of Artificial Intelligence," which address privacy, fairness, and security. In September 2023, the Ministry of Science and Technology of China released the "Notice on Issuing the 'Measures for the Ethical Review of Science and Technology (Trial Version)'," which aims to improve the science and technology ethics system, enhance science and technology governance capabilities, and achieve scientific and technological self-reliance and self-improvement (Xiao, 2020).

However, the application of AI in education also raises significant concerns regarding student privacy, potential for discriminatory assessments, and data security vulnerabilities. Therefore, this paper examines the current state and future needs of AI in education, focusing on the ethical issues, their classifications, and risk mitigation strategies. By proposing practical governance solutions, we aim to promote the responsible use of AI in education and support its sustainable development.

2. Concept and Application of AI in Education

2.1 Artificial Intelligence in Education

The concept of Artificial Intelligence (AI) was first proposed by American computer scientist John McCarthy and other scholars in 1956, who defined AI as "the science and engineering of making intelligent machines." Throughout the more than 60 years of AI development, various "AI+" concepts have emerged. As AI technology has advanced, researchers have delved deeper into its meaning, resulting in related terms such as "intelligent education," "smart education," and "wisdom education."

The core of AI in education is to utilize AI technology to transform traditional teaching models. For example, through AI technology, teachers can provide personalized teaching plans based on each student's learning ability and interests. Furthermore, AI can assist teachers in understanding students' learning progress and challenges through data analysis, enabling targeted instruction.

2.2 Development of AI in Education Applications

Artificial Intelligence in Education (AIEd) involves the application of AI technologies to design, develop, and manage educational systems, with the goal of optimizing learning processes and achieving educational objectives (Cui & Ma, 2023). The evolution of AIEd has closely followed advancements in AI technology, progressing from early computer-assisted instruction to intelligent tutoring systems and, more recently, leveraging deep learning and big data for personalized learning.

AI applications in education can be categorized into three levels: the teaching and learning level, the management and evaluation level, and the learning and living level. At the teaching and learning level, AI technology can be used for personalized learning, intelligent assessment, and assisting teachers in lesson preparation, providing students with intelligent learning tools. At the management and evaluation level, AI technology can be used for collecting and analyzing teaching data, optimizing the allocation of educational resources, and providing decision support. At the learning and living level, AI technology can be used for intelligent tutoring, intelligent homework, and intelligent examinations (Cui & Ma, 2023). With the rapid development of this new educational model, while bringing numerous benefits and advantages to education, AIEd has gradually exposed many problems and challenges that urgently need to be addressed, with ethical concerns being particularly prominent.

3. Ethical Issues in AI in Education Applications

3.1 Definition and Principles of Ethical Issues

What is ethics? Ethics is a set of moral norms that people should abide by when living together. This norm consists of objective ethical relationships and external regulations, primarily referring to human relationships and their rules and norms (Cui, 2023). Only issues arising from the application of AI in education that possess ethical characteristics can be considered ethical issues, allowing for ethical analysis and the pursuit of ethical strategies. The so-called "ethicality" means that the behavior is subject to a certain moral consciousness, the subject consciously chooses and involves the interests of others or society, and has the meaning of good or evil, specifically manifested as self-awareness, voluntariness, and sociality. Defining whether the problems brought about by the application of AIEd in the educational context are ethical needs to be based on the three basic principles of self-awareness, voluntariness, and sociality of AI application behavior.

Self-awareness refers to the actor's conscious moral awareness of the behavior itself, and the ability to realize the essence, meaning, value, and impact of the behavior. The so-called "voluntary" means that the actor makes moral actions according to his own will under his voluntary circumstances. The principle of sociality means that the behavior itself has an impact on other people, groups, and society, and involves interests or interest relations. In other words, all behaviors that do not interfere with interests are neither moral nor immoral, that is, they are not ethical and cannot be judged as beneficial or harmful, or good or evil.

Usually, a behavior often contains complex content, bringing ethical problems, and other social problems will also arise. As long as the behavior that causes the problem has the characteristics of self-awareness, voluntariness and sociality at the same time, the problem can be called an ethical problem. In fact, the principle of self-awareness is the premise, the principle of voluntariness is the guarantee, and the principle of sociality is the basis. It is difficult to accurately define the ethical boundaries of behavior without complying with any of the principles, and it is impossible to determine whether the problems it brings are ethical problems.

3.2 Ethical Issues in the Application of AI in Education

3.2.1 Ethical Conflicts Between Deterministic Algorithmic Hegemony and Uncertain Development

Algorithmic hegemony, a term coined by data scientist Cathy O'Neil in her book "Weapons of Math Destruction," cautions against the pervasive influence of mathematical models in our lives, highlighting their potential threat to societal structures, particularly individual privacy and dignity. In education, the deterministic nature of algorithms, based on fixed formulas, may introduce cognitive biases regarding student development, conflicting with the complexity and creativity inherent in education. The ethical issues stemming from deterministic algorithmic hegemony can be broadly categorized into: algorithmic discrimination, which deprives educational subjects of equitable rights; algorithmic opacity ("black box" algorithms), which neglects the informed consent of educational subjects; and algorithmic recommendations, which limit the choices of educational subjects.

First, algorithmic discrimination arises from the limitations of AI systems in accurately analyzing individual information, leading to differential treatment of individuals and, ultimately, discrimination against specific groups. For example, if an assessment of equitable educational resource distribution only collects data from already-developed regions with high-quality resources, the results will inevitably show equitable distribution, which is inherently unfair to individuals in underdeveloped regions (Gao, 2022).

Second, algorithmic opacity (the "black box") refers to algorithms that are too complex for human understanding or observation. In deep learning, the "hidden layers" between input data and output answers are inscrutable, creating a "black box." This lack of transparency can lead to unfair data processing, gender bias, and a significant erosion of the user's right to be informed. For example, some schools may use mathematical models to assign teacher responsibilities that exceed their workload capacity, or automatically reassign students with low academic performance to different majors, thereby depriving them of their right to autonomous decision-making (Gao, 2022).

Third, algorithmic opacity (the "black box") refers to algorithms that are too complex for human understanding or observation. In deep learning, the "hidden layers" between input data and output answers are inscrutable, creating a "black box." This lack of transparency can lead to unfair data processing, gender bias, and a significant erosion of the user's right to be informed. For example, some schools may use mathematical models to assign teacher responsibilities that exceed their workload capacity, or automatically reassign students with low academic performance to different majors, thereby depriving them of their right to autonomous decision-making (Wang et al., 2023).

3.2.2 The Conflict Between Technological Neutrality and Traditional Teacher-Student Ethical Relationships

AI plays an increasingly significant role in education, but as a tool, it is not entirely "neutral." While AI can assist teachers in addressing complex pedagogical challenges, freeing them from monotonous tasks and fostering higher-order thinking skills in students, it also presents multifaceted challenges to education and teachers. For instance, intelligent systems, based on big data and machine learning, can surpass individual teachers' limited knowledge backgrounds and cognitive biases in diagnosing student learning and designing personalized instructional plans. This can lead to a dual threat to teachers' "disciplinary status" and "pedagogical status" (Wang, 2023).

3.2.3 Loss of Data Privacy, Security, and Legal Rights

Data are the cornerstone of AI. AIEd applications rely on the collection and analysis of large volumes of student data. However, blurred data boundaries, increased complexity, and potential misuse of data create significant ethical risks. For example, some AIEd applications analyze student learning data to predict academic performance and future development trends. While this can potentially enhance pedagogical effectiveness, inappropriate use of this data can negatively impact students' academic and personal lives. Moreover, the incompleteness of data and the development potential of learners may lead to data inaccuracy, which may affect the accuracy of educational decisions. Furthermore, when teachers base pedagogical decisions on data, it may infringe on students' legal rights, leading to issues such as educational inequality. Therefore, clear regulations regarding data usage, user permissions, and robust authentication methods are crucial to ensure data security (Wang, 2023).

4. Risk Mitigation and Strategies

4.1 Mitigation of Data Security Risks

AIEd systems require extensive data for learning and optimization, potentially including students' personal information. Data breaches can compromise student privacy. The complexity of AI also introduces the risk of data tampering. Altered data can lead to erroneous pedagogical decisions, impacting student learning outcomes. The following solutions address data security risks (Tian, 2020):

First, Establishing Strict Data Protection Policies: Educational institutions should implement rigorous data protection policies that clearly define rules for data collection, storage, usage, and sharing. A dedicated data protection unit should oversee and enforce these policies.

Second, Strengthening Data Encryption Technologies: Employing advanced data encryption techniques can effectively prevent data theft or tampering during transmission and storage. Blockchain technology, with its distributed ledger approach, can further ensure data integrity and immutability.

Third, Enhancing User Data Literacy: Educational institutions should provide data literacy education to students and teachers, emphasizing the importance of data security and best practices for protecting their data. Training should also cover the proper use of AIEd applications to avoid data breaches caused by operational errors.

4.2 Mitigation of Technology Dependence Risks

Technological dependence risk refers to the vulnerability of a system or process when it overly relies on a specific technology. If that technology fails or becomes obsolete, the system or process may be severely impacted. In AIEd, this risk manifests as technology update risk and technology dependence risk. While leveraging AI in instruction, it is crucial to cultivate students' independent thinking and problem-solving skills (Xie, 2023). Overreliance on AI can diminish these abilities. Teachers can design activities that require independent thought and problem-solving to mitigate this risk. Strategies to address technological dependence risks include.

First, Diversifying Pedagogical Approaches: Despite the unique advantages of AI in education, institutions should continue to utilize diverse teaching methods and approaches to ensure instructional quality. For example, a blended learning approach (combining online and offline instruction) can provide richer and more varied learning experiences.

Second, diversified teaching methods. Although AI technology has its unique advantages in education, educational institutions still need to use a variety of teaching methods and methods to ensure teaching quality. For example, a combination of online and offline methods can be used to provide a richer and more diverse teaching experience.

Third, Establishing Contingency Mechanisms: Institutions should establish comprehensive contingency plans to address potential technological issues. This includes developing emergency response plans, creating technical support teams, and conducting regular drills. For instance, if a school's network system fails, the plan should outline steps for rapid problem identification, service restoration, and prevention of recurrence. Technical support teams should comprise individuals with specialized technical knowledge and practical experience, capable of providing timely assistance and solutions. Regular drills allow institutions to assess the effectiveness of contingency plans, identify and address weaknesses, and enhance their ability to respond to technological issues swiftly and efficiently.

4.3 Mitigation of Educational Equity Risks

Educational equity ensures that all students have equal opportunities to receive high-quality education, regardless of their background, abilities, or socioeconomic status. However, inappropriate application of AIEd can exacerbate educational inequalities. For example, the high cost of tools or algorithmic bias can disadvantage marginalized student populations. When designing and optimizing AIEd systems, it is crucial to consider the needs and characteristics of all students, avoiding biases that favor specific groups. Regular evaluation and adjustment of AIEd systems are essential to ensure fairness. Strategies to mitigate educational equity risks include:

First, provide public funding support. Governments should allocate sufficient public funding to support the research, development, and implementation of AIEd applications, ensuring that all students can benefit from these tools. This may require legislative action or policy changes.

Second, designing equitable algorithms is a critical responsibility for AIEd developers. They must ensure that algorithms do not exhibit bias or prejudice against any specific student group during data processing and analysis. This means that regardless of gender, race, socioeconomic status, learning ability, or any other factor, the algorithm should treat every student equally. Achieving fair algorithm design requires extensive research and testing, including a deep understanding of factors that can influence algorithmic fairness, such as data collection and processing methods, algorithm design and implementation, and algorithm outputs and applications.

Third, provide training support and establish a feedback mechanism. Teachers and schools should provide necessary training and support to help students and parents understand and use AIEd applications effectively. This may include detailed user guides, workshops, and seminars. Schools and teachers should establish feedback mechanisms to collect input and suggestions from students and parents regarding AIEd applications. This feedback helps them understand the actual impact of these tools and identify areas for improvement to meet the needs of all students (Li, 2023).

5. Conclusion and Outlook

The AI era is full of opportunities and challenges. The famous physicist Stephen Hawking once said: "When AI develops to its extreme, we will face the best or worst thing in human history" (Xu, 2019). AI in education has made positive contributions to the development of education, but it has also brought severe and urgent challenges to mankind. The ethical risks of AI technology in the field of education are also ubiquitous. As users, designers, managers, and creators of AI technology, educators' ethical risks affect the reasonable, legal, and effective use of AI technology (Xu, 2019). In addition, in future research, educators also need to constantly reflect on the value orientation of AI technology in practice, and avoid or reduce the ethical risks that may be caused by AI educational applications by strengthening their own moral cultivation, improving teaching methods, and innovating classroom management.

References

Cui, J., & Ma, Y. (2023). Progress and prospect of artificial intelligence education research in China. *Journal of University Education Administration*, *17*(6), 31-39.

Gao, S., & Yang, D. (2022). Research on the ethical risks of artificial intelligence application in education and its countermeasures. *Exploration of Higher Education*, (1), 45-50.

Li, Z. (2023). Releasing imagination: Ethical review of the application of artificial intelligence in education and future education. *Audio-Visual Education Research*, *44*(8), 18-25.

Tian, X. (2020). Privacy protection and open sharing: Educational data governance change in the era of artificial intelligence. *Research on Audio-Visual Education*, *41*(5), 33-38.

Wang, Y., Luo, X., Wang, H., & Liu, C. (2023). Research progress of artificial intelligence ethics in education: An analytical framework based on activity theory. *Chinese Educational Informatization*, *29*(5), 53-64.

Xiao, R., Xiao, H., & Shang, J. (2020). Artificial intelligence and educational reform: Prospects, difficulties, and strategies. *China Audio-Visual Education*, (4), 75-86.

Xie, J. (2023). How to identify and judge ethical problems in the educational application of artificial intelligence: From the perspective of ethical strategies. *Research on Audio-Visual Education*, *44*(1), 36-41.

Xu, X. (2019). Ethical dilemmas and circumvention paths in the educational application of artificial intelligence. *Software Guide*, *22*(9), 243-247.