Journal of Contemporary Educational Research, 2025, Volume 9, Issue 3
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

The Double-Edged Sword of Artificial Intelligence: Ethics and Privacy Protection in Future Exams

Haidong Sun*, Huarui Lu, Diexiang Zhao

Inner Mongolia Autonomous Region Education Examinations Authority, Hohhot 010000, Inner Mongolia, China

***Corresponding author:** Haidong Sun, 491741892@qq.com

Abstract: The widespread application of artificial intelligence (AI) technology in exams has significantly improved the efficiency and fairness of exams; it has also brought challenges of ethics and privacy protection. The article analyzes the fairness, transparency, and privacy protection issues caused by AI in exams and proposes strategic solutions. This article aims to provide guidance for the rational application of AI technology in exams, ensuring a balance between technological progress and ethical protection by strengthening laws and regulations, enhancing technological transparency, strengthening candidates' privacy rights, and improving the management measures of educational examination institutions.

Keywords: Artificial intelligence; Examination; Ethics; Privacy protection

Online publication: March 28, 2025

1. Introduction

Artificial intelligence (AI) technology is profoundly changing the field of education, especially in exam management, where the application of AI technology has improved efficiency and fairness. However, the rapid development of technology has also brought ethical and privacy challenges, such as algorithmic bias, insufficient decision transparency, and privacy risks for candidates. The article aims to discuss these issues and propose specific strategies to balance technological progress and candidates' right protection, to provide reference for educational examination institutions.

2. Current application status of AI in exams

2.1. Intelligent paper marking inspection assistance and automatic test question generation

The application of intelligent paper marking inspection assistance significantly improves the overall quality of grading and reduces the possibility of human errors. However, the accuracy and fairness of this technology still receive much attention. Although AI can achieve efficient grading through big data analysis and pattern recognition, it cannot completely replace the judgment and experience of human examiners when dealing with complex subjective questions. In addition, biases and imbalanced training data that may exist in algorithms may also lead to unfairness in the grading results. Therefore, how to ensure fairness and impartiality in the grading results while improving efficiency remains an urgent problem to be solved.

In terms of intelligent question generation, AI systems can generate adaptive exam questions based on the candidate's knowledge level and exam targets. The application of this technology makes the exam content more personalized and diversified, breaking the limitations of the traditional question-setting mode. However, it is still faced with challenges on how to ensure exam fairness. If AI-generated test questions fail to consider the background differences of candidates, it may lead to biased exam results, thereby affecting the fairness of the exam [1]. Therefore, when developing and using intelligent question systems, it is necessary to fully consider the adaptability and fairness to avoid exam unfairness caused by technology.

2.2. Candidate behavior analysis and cheating detection

The introduction of AI invigilation systems provides a new means to combat exam cheating. Through real-time monitoring and analysis of candidate behavior, AI systems can quickly identify abnormal behavior, which greatly improves the fairness and security of exams. However, the widespread application of AI proctoring systems also comes with potential risks. Candidates are monitored throughout the exam process, which may raise concerns about their privacy rights. Meanwhile, the misjudgment or technical malfunction of AI systems may also lead to innocent candidates being identified for cheating, thereby affecting their exam scores and personal reputation.

The collection and analysis of candidate data is the foundation for the effective operation of AI invigilation systems, but it also brings serious ethical issues. A large amount of personal data faces the risk of data leakage and abuse during collection, storage, and analysis [2]. Do candidates have the right to know the specific use of their data? How do we find a balance between technological development and personal privacy protection? These issues still need to be explored and resolved in practice.

2.3. Application practice of Education Examination Institute

As an important institution for exam management, the Education Examination Institute has introduced AI technology to a certain extent to improve the efficiency and quality of exam management. For example, some provinces have attempted to apply intelligent marking to assist quality inspection in large-scale exams such as the college entrance examination, middle school entrance examination, and academic examination, in order to improve the quality of marking and ensure consistency in grading. At the same time, some regions are also exploring the application of intelligent question systems to achieve more scientific and fair exam evaluation. However, the application of these technologies also poses higher requirements for the management capabilities and policy-making of educational examination institutions.

Regarding the application of AI in exams, the Education Examination Institute needs to develop and improve relevant policies and management measures to ensure the compliance and rationality of technology applications, which includes establishing and improving data protection mechanisms, strengthening the evaluation and supervision of AI systems, and fully considering the rights and privacy protection of candidates in technological applications. The Education Examination Institute also needs to actively engage in social communication to enhance public understanding and trust in the application of AI technology in exams and seek the best balance between technological innovation and ethical protection.

3. Ethical issues caused by AI in exams

3.1. Fairness and prejudice

The design and training process of AI algorithms relies on large-scale datasets, and the imbalance and bias in these datasets may lead to unfair results in the application of AI systems. Specifically, AI algorithms may tend to favor a specific group during grading, question setting, and other processes, leading to the unfair treatment of some candidates. For example, the cultural background or language habits of certain groups may not be fully reflected in the algorithm training data, which may lead to poor performance of AI systems while processing exam answers from such candidates. To ensure the fairness of the exam, measures must be taken to identify and eliminate biases in AI algorithms, ensuring that candidates from different groups have equal opportunities in the exam. This requires full consideration of social diversity in algorithm design, dataset selection, and system testing processes, and multi-party review mechanism must be taken to monitor the fairness of AI systems [3].

3.2. Transparency and interpretability of decision-making

The “black box” problem in the AI decision-making process is a major challenge in current technological applications. Due to the complexity of AI algorithms, the decision-making process often lacks transparency and is difficult for non-technical personnel to understand and explain. In the examination environment, candidates, parents, and teachers may know very little about how AI systems make scoring or judgment results, which may lead to trust crises due to a lack of transparency. Lack of clear explanation path when candidates question AI decisions will lead to distrust and dissatisfaction with exam results. Therefore, the Education Examination Institute needs to ensure the transparency of the AI system and guarantee the candidates’ right to know and appeal by publicly disclosing algorithm decision-making logic and setting up result interpretation mechanisms. Further technological development should move toward interpretable AI, making the decision-making process more transparent and fair, thereby enhancing public trust in AI technology.

3.3. Psychological pressure and trust of candidates

The widespread application of AI invigilation systems in the examination has had an undeniable impact on the psychological state of candidates. Under highly intelligent monitoring, candidates may feel nervous and anxious, worrying that every action they make will be misjudged as abnormal behavior by the system. This psychological pressure may not only affect exam performance, but also have a negative impact on candidates’ physical and mental health. Another serious issue involves candidates’ trust in AI systems. If there is a misjudgment or improper operation of the AI system, candidates’ trust in its impartiality will be greatly reduced, thereby questioning the reliability of the entire examination system.

To alleviate the psychological pressure on candidates, educational examination institutions should make the design and application of AI invigilation systems more user-friendly, avoiding excessive reliance on technological means for monitoring. At the same time, it enhances candidates’ understanding and trust in AI systems through education and communication. A transparent and fair error correction mechanism should be established for any misjudgment or abnormal handling of AI systems to ensure the legitimate rights and interests of candidates.

4. Privacy protection challenge of AI in exams

4.1. Data collection and storage

The effective application of AI technology relies on the collection of a large amount of candidate data. However, the scope and legality of collecting the personal data of candidates have attracted widespread attention. During the exam process, the identity information, behavioral data, and even biometric data of candidates may be comprehensively collected. Once these data are abused or leaked, they will pose a serious threat to the personal privacy of candidates. Therefore, the scope of data collection should strictly follow the principle of minimization, that is, only collect necessary data directly related to the exam. At the same time, it is necessary to ensure the legality of the data collection process and fully respect the candidates' right to know and consent.

In terms of data storage, security and access control are the core issues to ensure the privacy of candidates. Due to the sensitivity of candidate data, a data breach will have a widespread negative impact on candidates and society. Therefore, data storage should adopt advanced encryption technology and strictly limit access permissions, only authorizing relevant personnel to access under specific circumstances. In addition, regular security reviews and upgrades of storage systems are necessary to address constantly changing network security threats.

4.2. Data usage and sharing

The purpose and scope of candidate data use is also an important issue in privacy protection. The diversity of AI technology applications makes it possible for candidate data to be used for various purposes, such as exam score evaluation, behavior analysis, educational resource recommendation, etc. However, in the process of using data, it is necessary to clarify the scope of data use, ensure that it is limited to legal and necessary purposes, and avoid infringing on the privacy rights of candidates due to data abuse. In addition, strict auditing and supervision mechanisms should be established to ensure transparency and compliance in the data usage process.

In terms of data sharing, privacy risks and legal responsibilities are particularly prominent. Candidate data may be shared or exchanged by different institutions, such as education departments, research institutions, or third-party technology companies. During this process, the privacy risks of data sharing need to be highly vigilant, especially in the case of insufficient data anonymization and anonymization processing, where the privacy of candidates may still be threatened. At the same time, the issue of legal responsibility in data sharing cannot be ignored. All parties should clarify their responsibilities when sharing data, ensure that data sharing complies with relevant laws and regulations, and be responsible for the privacy of candidates.

4.3. Privacy protection measures of the Education Examination Institute

As an important institution for exam management, the role of the Education Examination Institute in privacy protection is crucial. Currently, the Education Examination Institute has implemented a series of privacy protection policies to a certain extent, including technical means such as data encryption, access control, and data anonymization. However, with the continuous development of AI technology, existing privacy protection measures still have certain loopholes. For example, problems such as insufficient policy implementation, untimely technological updates, and inadequate protection of candidates' right to know and participate in privacy protection still exist.

To address these challenges, the Education Examination Institute needs to further improve its privacy protection policies and take proactive improvement measures. Firstly, the enforcement of privacy protection policies should be strengthened to ensure that all measures are effectively implemented in practice. Secondly, it is necessary to regularly evaluate and update privacy protection technologies to ensure that they can respond to emerging security threats. In addition, the Education Examination Institute should strengthen communication with candidates to increase policy transparency and ensure that candidates have sufficient right to know and participate in their own data processing process, thus achieving a balance between technology application and privacy protection.

5. Strategies for balancing technological progress and ethical privacy protection

5.1. Strengthening the formulation of laws, regulations, and policies

Currently, the application of AI technology in exams has not yet formed a systematic legal norm, and this legal vacuum makes it difficult to effectively address many potential ethical and privacy issues. Therefore, it is urgent to develop and improve a legal framework for the application of AI in exams, clarifying the boundaries and responsibilities of technology use.

Firstly, legislation should be passed to clarify the legality and scope of the application of AI in exams. The law should provide specific requirements for data collection, use, storage, and sharing to ensure the security and privacy of candidates' personal information is not violated. In addition, in response to the bias and injustice that AI technology may bring, the law should clarify the standards for algorithm transparency and fairness and provide corresponding regulatory and error correction mechanisms to prevent the abuse of technology from harming the rights and interests of candidates. In the specific process of formulating the legal framework, it is necessary to widely solicit opinions from various parties, including technical experts, educators, legal scholars, and the public, to ensure the scientificity, rationality, and operability of legislation. In addition, punishment measures for the illegal use of AI technology should be clearly defined in the law to form an effective legal deterrent and prevent potential risks and improper behavior.

As the core institution of examination management, the Education Examination Institute plays an important role and responsibility in policy-making. The Education Examination Institute should first actively participate in the legislative process at the national and local levels, provide professional opinions and technical support, and ensure that the formulation of laws and regulations meets the actual needs and development trends of examination management. On this basis, the Education Examination Institute needs to develop and implement internal policies and operational norms that are compatible with the legal framework to ensure the reasonable use of AI technology in exams. These policies should cover the selection criteria for AI systems, data management standards, invigilation measures, and emergency response plans, ensuring that the rights and interests of candidates can be effectively protected in any situation. In addition, the Education Examination Institute also needs to bear the responsibility of supervision and evaluation in policy formulation. To this end, a comprehensive evaluation mechanism should be established to regularly review the application effectiveness and potential problems of AI technology and adjust policies and measures in a timely manner based on the evaluation results. The Education Examination Institute should also establish communication channels with candidates and various sectors of society, enhance the transparency and credibility of policies, and ensure that the implementation of various measures can truly safeguard the legitimate rights and interests of candidates [4].

5.2. Enhancing technological transparency and fairness

Ensuring transparency and fairness in the integration of artificial intelligence technology into exam management has become a crucial issue. The transparency and fairness of AI algorithms directly affect the credibility and social recognition of exam results. Therefore, developing transparent and fair AI algorithms, as well as introducing multi-party supervision and evaluation mechanisms, are urgent issues that need to be addressed [5].

The transparency of AI algorithms is one of the core issues in the application of technology in exams. Due to the fact that AI systems often rely on complex mathematical models and data analysis, the decision-making process is often difficult for ordinary users to understand. This “black box” effect makes it difficult for candidates, parents, and educators to trust the fairness of AI scoring. Therefore, developing transparent AI algorithms to ensure the interpretability of the decision-making process is the key to solving this problem. Specifically, in the design and development stages of AI systems, the interpretability of algorithms should be fully considered, so that they can not only provide results, but also explain the logic and basis behind scores or judgments. This transparency can be achieved through open algorithm design, open testing standards, and the provision of visualization of algorithm decision-making processes. In addition, the fairness of AI algorithms is an important foundation for ensuring exam fairness. In the process of algorithm development, diverse datasets should be widely collected and used to avoid decision bias caused by data bias. Developers need to establish strict testing and validation procedures, conduct sufficient fairness checks on algorithms, and ensure that they treat candidates of different genders, races, cultural backgrounds, and other groups equally. At the same time, ethical review mechanisms should be introduced in algorithm design to identify and correct potential biases and injustices, thereby ensuring the fairness of the algorithm in practical applications.

Ensuring the transparency and fairness of AI technology in exams through the efforts of technology developers alone is far from enough. It is necessary to introduce a multi-party supervision and evaluation mechanism to form a pattern of social governance. Specifically, this mechanism should involve the participation of various forces such as government regulatory agencies, educational institutions, technical experts, legal experts, and public representatives. Through multi-party supervision, a comprehensive review of the development and application process of AI systems can be conducted to ensure compliance with fair and transparent ethical standards. At the specific implementation level, a multi-party supervision mechanism can be achieved through regular algorithm reviews, independent third-party evaluations, and public hearings. The Education Examination Institute can regularly invite external experts and public representatives to independently evaluate AI systems and make the evaluation results public to enhance transparency.

5.3. Strengthening the protection of candidates’ privacy rights

In the widespread application of AI technology in the examination process, the privacy rights of candidates are facing unprecedented challenges. Ensuring that the privacy rights of candidates are not infringed upon in this emerging technological environment is the key to maintaining educational fairness and social trust. To this end, it is necessary to effectively enhance candidates’ control over personal data and establish a sound data protection and complaint mechanism to safeguard their privacy rights.

In AI-driven exam management systems, candidates’ personal data becomes a critical resource. However, current technological applications often rely on extensive data collection, and candidates have relatively weak control over this data. To safeguard the right to privacy, it is necessary to increase candidates’ control over their personal data, ensuring that candidates have sufficient informed and decision-making power in all aspects of data collection, use, and storage. Firstly, the scope and purpose of collecting candidate data should be clearly defined, and a detailed disclosure procedure should be implemented before data collection to ensure that candidates are aware of how their data will be used and the potential risks they may face. Candidates should have the right to choose whether to agree to the collection and use of data and may withdraw their consent if necessary. In addition, in the process of data storage and processing, candidates should be provided with channels to access their personal data, enabling them to view, modify, or delete inaccurate or unnecessary data.

At the same time, in order to enhance the control of candidates, it is necessary to develop and apply advanced data management tools. These tools should have intuitive user interfaces, facilitate data management for candidates, and further reduce the risk of data leakage through encryption and anonymization technologies. The Education Examination Institute and related technology providers need to work together to promote the popularization and application of these tools, ensuring that every candidate can effectively protect their privacy rights in the AI technology environment.

It is crucial to establish a sound data protection and complaint mechanism to further strengthen the protection of candidates' privacy rights. The data protection mechanism should include multi-level security measures, covering various aspects such as data collection, transmission, storage, and use. This requires not only technical security measures such as data encryption and access control, but also strict regulatory oversight at the institutional level, such as regular security reviews and legal accountability for unauthorized use. In addition, establishing an effective complaint mechanism is a key means of safeguarding the privacy rights of candidates. When candidates believe that their privacy rights have been violated or have doubts about the data processing process, there should be clear and convenient complaint channels. The Education Examination Institute should establish a dedicated complaint-handling department responsible for receiving and handling privacy complaints from candidates, ensuring that the complaint process is open and transparent and the handling results are fair and reasonable. In the process of handling complaints, it is necessary to promptly respond to candidates' doubts, provide detailed investigation reports, and take remedial measures when necessary to maximize the protection of candidates' legitimate rights and interests.

5.4. Management and supervision measures of the Education Examination Institute

In the process of gradually integrating AI technology into exam management, the Education Examination Institute shoulders the responsibility of ensuring the compliance of technology applications and maintaining exam fairness. Therefore, it is necessary to strengthen internal management and supervision measures and strictly regulate the use of AI technology [6]. At the same time, regularly evaluating the application effectiveness and potential risks of AI technology in exams can ensure that it does not harm the rights and interests of candidates while improving efficiency [7].

The Education Examination Institute must establish a strict internal management and supervision system in the application of AI technology to ensure that the application of technology complies with relevant laws, regulations, and ethical norms. Firstly, it is necessary to establish detailed internal policies and operational standards to oversee the entire process of designing, developing, testing, deploying, and using AI systems. These policies should include clear data collection standards, algorithm design requirements, system operation rules, and exception handling procedures to ensure that AI technology is compliant and legal at every step of its application. A multi-level review mechanism should be introduced in the management and supervision process to regularly review and evaluate the application of AI technology. Through internal review, potential violations or technical defects can be identified and corrected in a timely manner. At the same time, the Education Examination Institute should establish a dedicated supervisory agency or position to oversee the daily application of AI technology and ensure that all operations comply with established standards. These supervisory measures not only help prevent technological abuse, but also enhance public trust in the fairness and transparency of AI technology in exams. In addition, the Education Examination Institute needs to strengthen the training and education of internal personnel to enhance their awareness of ethics and privacy protection in the application of AI technology. Through continuous education and training, relevant personnel will have sufficient knowledge and skills to effectively manage and supervise the operation of AI systems, ensuring that every aspect of technology application complies with ethical and legal requirements.

The application effectiveness and potential risks of AI technology are constantly changing, and educational examination institutions must establish regular evaluation mechanisms to comprehensively examine the performance of AI technology in exams. This evaluation mechanism should cover multiple dimensions, such as technical effectiveness, candidate experience, privacy protection, ethical impact, etc., to ensure that AI technology improves exam management efficiency without adversely affecting the rights and interests of candidates and the fairness of the exam. Specifically, regular evaluations should include checks on the accuracy, fairness, and transparency of AI systems to ensure that they accurately reflect the true level of candidates in practical applications and avoid unfairness caused by algorithm bias or data errors. In addition, attention should be paid to the protection of candidates' privacy by AI technology, and the effectiveness of privacy protection measures should be evaluated by analyzing information such as data leakage incidents and complaint records. Risk assessment is equally important. The Education Examination Institute should identify and prevent potential technical risks such as system failures, data loss, or tampering through simulation tests, emergency drills, and other methods. At the same time, it is imperative to establish a rapid response mechanism that can quickly take remedial measures when risks or problems arise, reducing the impact on candidates and the examination system.

6. Conclusion

The application of AI technology in exam management has great potential, but it also brings ethical and privacy risks. The strategies proposed in the article, including strengthening regulatory development, improving technological transparency, safeguarding privacy rights, and improving management supervision, aim to provide support for the Education Examination Institute to promote technological applications while maintaining the fairness of exams and the rights of candidates. In future development, it is crucial to continue paying attention to these issues in order to achieve sustainable application of AI technology in education.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Cheng W, 2023, Difficulties and Implementation Concepts of Paperless Examination Implementation. Zhejiang Examination, (12): 13–16 + 50.

[2] Chen W, Qi Y, 2023, Integration of Virtual Reality and GPT-based Artificial Intelligence: An Innovative Path for Educational Application Development. Computer Knowledge and Technology, 19(34): 129–131.

[3] Liu B, 2023, Digital Transformation of Data-Driven Teaching: Mechanism, Field, and Path. Modern Educational Technology, 33(09): 16–26.

[4] Sun L, 2023, Reflections on Enhancing the Efficiency of National Education Examination Management. Zhejiang Examination, (11): 10–13.

[5] Zhang M, Xue S, 2023, Common Trends and Construction Focus: Global Observation of Digital Transformation in Education. China Distance Education, 43(07): 21–29.

[6] Lu Z, 2023, Research on the Path and Strategy of Digital Transformation of Provincial Examination Institutions.

Enrollment Examination Research, (03): 61–76.

[7] Zhou J, 2023, Implementation Path of Smart Examination Service in National Education Examination. Examination Research, 19(05): 81–86.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.