ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

Research on the Ethical Risk Control Mechanism of AIGC Technology Application in University Education through Ethical Leadership

Yiyuan Sun[1] *and Ting Dong*[2,*]

[1] Human science department, Assumption University of Thailand, Bangkok, 10240, Thailand
[2] School of Marxism, Wuhan Institute of Technology, Wuhan 430205, Hubei, China

SUMMARY: *This article deeply explores the ethical risks caused by the widespread application of AIGC technology in the process of intelligent development of higher education, and is committed to strengthening ethical leadership to achieve effective control of these ethical risks. The first step in building an ethical risk control system is to establish evaluation indicators and a complete system for the ethical risk control mechanism of AIGC technology integrated into university education scenarios. This system is based on ethical principles as its core architecture, with evaluation indicators and specific methods as the supporting foundation. The main content is constructed around the three key steps of risk identification, risk analysis, and risk evaluation. At the same time, attention should be paid to the close connection between the subject and object, and effective linkage should be achieved through prevention and control governance mechanisms to ensure the orderly operation of the entire system. At the level of technical optimization, this article innovatively proposes the use of an improved differential evolution algorithm to optimize RBF neural networks. Specifically, in the process of improving the differential evolution algorithm, a dual archive population optimization strategy was introduced. By creating two different files to record the individual information that was eliminated during the algorithm running process, and applying this information reasonably to the next generation of update iterations. The implementation of this strategy can effectively maintain the novelty and diversity of the population, significantly improve the overall performance of the algorithm, and thereby enhance the accuracy of RBF neural network in predicting ethical risks. Through practical case analysis and verification, an ethical risk assessment index system and grading standards were first constructed for the application of AIGC technology in university education scenarios. Subsequently, the optimized model was used for training and testing, and its evaluation results were compared and analyzed with traditional models. The results showed that the optimized model demonstrated higher reliability and accuracy in evaluating the ethical risks of AIGC technology applied in university education. The series of methods and strategies proposed in this article are of great significance for enhancing the ability to control ethical risks in the application of AIGC technology in university education. They can provide strong theoretical support and practical guidance for the effective prevention and control of ethical risks in the process of intelligent transformation of higher education.*

KEYWORDS: *ethical leadership; AIGC technology; University education; Ethical risk; RBF neural network; Differential Evolution Algorithm*

*gooddt2025@163.com
https://doi.org/10.65102/is20261184
Published: 10 June , 2026

1 Introduction

The intelligent transformation of higher education is a cross era systematic educational innovation process, which refers to the integration of artificial intelligence technology into various aspects of higher education, promoting the innovation and transformation of all elements such as teaching paradigms, organizational structures, teaching processes, and evaluation methods in educational organizations. The intelligent transformation of higher education has gradually become a common trend in the global education industry [1, 2]. Since the proposal of the Digital China strategy, significant achievements have been made in the intelligent transformation of higher education in China. The Ministry of Education is solidly promoting the strategic action of intelligent higher education, continuously improving the top-level design and institutional mechanism of educational informatization, leading the modernization of education with high-level educational informatization, and promoting the high-quality development of education. However, the intelligent transformation of higher education is a continuous process, accompanied by the continuous iteration and innovation of new technologies. ChatGPT, for example, is a specific AIGC (Artificial Intelligence Generated Content) product, which has attracted widespread attention as an example of "text generation" in AIGC's multiple modalities. Its text content generation power and broad application prospects have a huge impact on existing social production relations and productivity [3]. How to seize the transformative opportunities brought by AIGC technology in the field of higher education, continuously explore new models on the path of intelligent transformation, and accelerate the promotion of smart education is a key topic worthy of consideration and exploration. In this context, higher education must balance the relationship between people and technology, technology and education, continuously innovate the ways and ideas of integrating technology and education, update the teaching paradigm of higher education, cultivate intelligent and skilled talents for the future, and build an intelligent moral education system that promotes mutual benefit between humans and machines. Exploring the new driving force of AIGC technology for the intelligent transformation of higher education and how to seize the opportunity to advance the transformation strategy is of great significance for the country's innovative development and enhance international competitiveness [4].

AIGC refers to artificial intelligence generated content, and its application in higher education and entertainment has attracted widespread attention. Some AIGC products have been used to provide interactive experiences, educational assistance, and gaming services for university classrooms, which may have both positive and negative impacts on education [5]. For example, in terms of text generation, ChatGPT is a Chinese dialogue generator based on OpenAI's GPT-3.5/4.0 model, which can generate various texts such as poetry, stories, code, etc. according to user input. ChatGPT can be used in entertainment, education, marketing, and other fields to provide a more natural and user-friendly interactive experience. In terms of audio generation, Waymark is a platform for producing TV commercials and digital video ads. It integrates the ChatGPT model, which can generate video scripts in different styles and languages according to users' needs and preferences, improving the efficiency and quality of video production. In terms of image generation, Midjoumey is an image generator based on OpenAI's DALL-E model [6]. It can generate corresponding images based on user input text descriptions, and can also perform image attribute and partial editing, such as changing color, shape, position, etc. Midjoumey can be used in creative design, artistic creation, educational presentations, and other fields to stimulate users' imagination and creativity. On the one hand, AIGC technology can help university teachers gain a deeper understanding of the needs, interests, and abilities of college students, provide personalized learning plans and resources, and enable students to learn more efficiently. The AI intelligent education tutoring system can provide real-time and accurate learning feedback for college students, helping them identify learning blind spots and improve learning outcomes [7]. Cloud computing and IoT technology enable a large amount of learning resources to be easily accessed and shared, greatly enriching students' learning materials. AIGC technology can assist teachers in teaching evaluation and improvement, enhancing the quality of education. In addition, online education platforms and distance education technology enable high-quality educational resources to no longer be limited by geography, helping to narrow the education gap between universities, industries, and regions. On the other hand, excessive reliance on AIGC technology may also lead to some negative impacts [8]. For example, ChatGPT has certain biases in algorithm and data, which may affect the academic judgment and decision-making of teachers and students. AIGC technology may collect and store a large amount of personal information, including students' learning data, behavioral habits, etc., which may raise privacy and security issues. Zhong Chenglin pointed out that the challenges faced by the current intelligent transformation of higher education include [9, 10]: how to solve the difficulties of integrating artificial intelligence technology with education and teaching, increase classroom teaching interactivity and sense of gain, improve learning effectiveness, and enhance learning efficiency; How to build a faculty team in universities, update educational concepts, transform teacher roles, and reconstruct teacher-student relationships; How to optimize the structure of disciplines and majors in universities, improve discipline construction and scientific research level; How to reform the talent cultivation mode and how to build a modern higher education governance system and governance capacity, and rationally respond to new challenges such as the deviation of artificial intelligence from the laws and values of education [11].

This study mainly focuses on the research of ethical leadership on the ethical risk control mechanism of AIGC technology applied in university education. A method is proposed to use an improved differential evolution algorithm to optimize RBF and enhance ethical leadership, thereby improving the ethical risk control ability of AIGC technology applied in university education. This study introduces a dual archive population optimization strategy in improving the differential evolution algorithm. The establishment of dual archives involves creating archive archive1 for storing discarded historical individuals and archive archive2 for storing discarded experimental individuals. Both files record the individuals eliminated in the current generation and will be used in the next generation. This approach ensures the continuous updating of individuals in the files, thereby maintaining sufficient novelty and diversity, which helps to continuously optimize and improve the performance of the algorithm. The improvement of algorithm performance greatly contributes to the accuracy of RBF ethical risk prediction, and thus helps to enhance the accuracy of ethical risk control in the application of AIGC technology in university education, achieving better ethical control effects.

2 AIGC applied university education ethics risk assessment

2.1 Overall framework for evaluation

Based on the characteristics and evaluation ideas of ethical risk assessment using artificial intelligence, it can provide conceptual guidance for establishing an effective evaluation framework and provide speculative wisdom for conducting evaluation work with a top-level macro perspective. Ethical risk assessment is the intersection of ethical assessment and risk assessment, involving both ethical considerations and risk analysis. Therefore, when building an assessment framework, a comprehensive analysis of the methods of ethical assessment and risk assessment should be conducted to provide theoretical support for a unique path of ethical risk assessment [12, 13]. The commonly used classic ethical assessment framework at present is the "Seven Steps" proposed by Andersen, which is widely used in the field of applied ethics. The "Seven Steps" are: "What are the facts?" "What are the moral issues?" "What are the solutions?" "What are the main stakeholders?" "What are the moral limitations?" "What are the practical limitations?" "What decision should be made in the end. In terms of risk assessment, risk assessment is an important part of risk management, which is inherently included in the risk management system. Common risk management frameworks, such as the NIST Risk Management Framework (RMF) developed by the National Institute of Standards and Technology (NIST) in the United States, include six core steps: Categorize Information Systems, Select Security Controls, Implement Security Controls, Assess Security Controls, Authorize Information Systems, Monitor Security Controls, etc [14].

Through the analysis of classic evaluation frameworks, it is not difficult to find that whether it is ethical evaluation or risk assessment, the basic structure of evaluation is steps and methods, the basis of evaluation is problems, and the core of evaluation is established standards, emphasizing the continuity and fluency of procedures [15]. However, the ethical risk assessment of artificial intelligence is a complex project that is mainly carried out from a third-party perspective. Therefore, it is necessary to include the assessment subject and assessment object, and introduce objective and accurate indicators and methods to improve the assessment of the ecological environment. Based on the classic evaluation framework and combined with the characteristics of artificial intelligence systems, the comprehensive approach of evaluation can be integrated to build the following ethical risk assessment framework. As shown in Figure 1, the ethical risk assessment of artificial intelligence needs to comprehensively consider the subject, object, criteria, indicators and methods, and steps [16]. These five are the backbone of the assessment framework, which mainly focuses on ethical criteria as the core, indicators and methods as the foundation, and risk identification, analysis, and evaluation as the main content. The subject and object are connected by prevention and control governance. Under the guidance of a comprehensive evaluation approach, we aim to construct an ethical risk assessment framework that focuses on effective risk prevention and mitigation, with the core goal of better benefiting humanity through artificial intelligence. At the same time, we must fully consider the multidimensional impact of external social factors such as laws and regulations, policy systems, moral concepts, and public opinion atmosphere on the entire lifecycle of artificial intelligence.

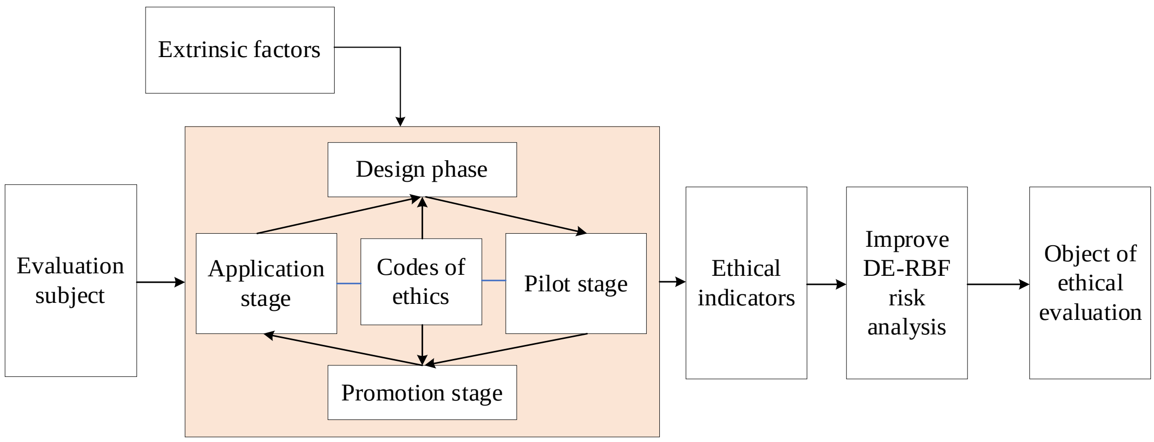


Figure 1: Schematic diagram of the ethical risk assessment framework for artificial intelligence

Ethics, as a research field, is a normative rule with indicative and evaluative research objectives. Evaluation is an inherent connotation of ethics, and the ethical risk assessment of artificial intelligence focuses on evaluation [17]. The primary issue is what standards should be used for evaluation. According to the framework of ethical risk assessment for artificial intelligence, ethical standards are the core content in ethical risk assessment and can be regarded as the criteria for evaluation. The indicator system constructed based on ethical principles and the evaluation methods selected based on the indicator system play a crucial role in the evaluation process and have extremely important significance. They are the foundation and prerequisite for risk identification, risk analysis, and risk assessment work, not only determining the quality and efficiency of the evaluation work, but also directly related to the credibility and effectiveness of the evaluation results [18]. Therefore, it is crucial for the ethical risk assessment of artificial intelligence to establish a comprehensive, scientific, and reasonable ethical code, select objective, diverse, and dynamic evaluation indicators based on the ethical code, and choose effective, applicable, and accurate evaluation methods based on the indicators.

2.2 Object of ethical risk assessment

The objects of ethical risk assessment generally include five categories: data, software, hardware, personnel, and services. As shown in Table 1, the scope of ethical risk assessment for artificial intelligence in the education field is limited to the application security of artificial intelligence technology in the education field. The evaluation criteria mainly include the "Information Security Technology Information Security Risk Assessment Method" (GB/T209842022) and the "Implementation Guidelines for Information Security Technology Information Security Risk Assessment" (GB/T315092015). By defining clear evaluation objects and scope, potential ethical risks related to them can be more accurately identified, and differentiated and targeted management strategies can be developed for ethical issues involved in different evaluation objects.

Table 1: Categories of Objects for Ethical Risk Assessment of Artificial Intelligence in the Education Sector

Category	Specific content	Risk source
Data	Personnel information, academic data, and educational resources.	Data leakage, data abuse, data quality issues, or improper data processing behavior.
Software	Learning management systems and educational applications.	Software vulnerabilities, unauthorized access, malicious software, or unstable systems.
Hardware	Teaching, management, and service equipment	Overreliance on smart devices, algorithmic bias, and discrimination.
Staff	Educators, students, administrators, parents.	Loss of autonomy, excessive reliance on artificial intelligence technology, neglect of technological limitations, surveillance and regulation, privacy breaches, algorithmic discrimination and bias.
Service	Provide technical support and maintenance for artificial intelligence assisted education services.	Low quality teaching resources, limited behavior, bias in artificial intelligence algorithms, and restricted thinking innovation.

Firstly, data. The education system involves a large amount of sensitive data, and risks may arise from data breaches, data abuse, data quality issues, or improper data processing behavior. Inaccurate data may lead to decisions based on incorrect information, affecting teaching management and evaluation. Secondly, software. The security and stability of software are directly related to the smooth progress of the education process. Risks may come from software vulnerabilities, unauthorized access, malicious software, or unstable systems. For example, a learning monitoring software that collects a large amount of student behavior data without their consent may raise privacy protection and ethical controversies. Thirdly, hardware. Hardware includes teaching, management, and service equipment such as intelligent devices, robot teaching assistants, smart cards, smart access control, and intelligent security in the education system [19]. Hardware risks may arise from excessive reliance on devices, algorithmic biases, and discrimination. For example, the convenience of smart devices may lead to an excessive reliance on technology in the educational process, while ignoring other important factors such as the role of teachers, the needs of students, and emotional communication. Fourthly, personnel. Personnel refer to all individuals related to the education system, including educators, students, administrators, parents, etc. On the one hand, risks may come from oneself, such as loss of autonomy, excessive reliance on artificial intelligence technology, and neglect of technological limitations; On the other hand, it may come from external threats such as surveillance and regulation, privacy breaches, algorithmic discrimination and bias. Fifth, service. The service category includes providing technical support and maintenance for AI assisted educational services. The risks may come from poor quality of teaching resources, limited behavior, restricted thinking innovation, and biases in artificial intelligence algorithms. For example, an algorithm used by an artificial intelligence assisted learning platform may exhibit gender bias in student assessment, leading to unfair evaluation of students and affecting their academic development. By comprehensively evaluating these target categories, educational organizations can develop effective risk management strategies, maximize the protection of stakeholders' rights, and ensure the smooth operation of the education system.

2.3 Ethical risk acceptance criteria

The ethical risk acceptance criteria are based on laws, regulations, industry norms, and general social expectations, providing a benchmark for assessing whether risks are acceptable, and serving as the basis for evaluating the importance of risks. In the process of environmental construction, educational organizations need to clearly define their risk acceptance level when facing ethical risks, establish a set of clear standards, and quantify the level of potential ethical risks through risk assessment. Generally speaking, the formula for calculating risk value is: risk value=probability of risk occurrence x degree of risk impact [20]. The range of risk values varies depending on the probability of risk occurrence and the degree of risk impact. This study limits the range of values for risk occurrence rate and risk impact degree to 1 to 5, and calculates the risk acceptance criteria as shown in Table 2.

Table 2: Risk Acceptance Criteria

Risk level		Risk value	Describe
High risk	HRI	Risk value>12	There are serious security threats and potential damages that must be addressed to reduce risks.
Medium risk	MRI	$5 \leq$ risk value ≤ 12	There are certain security threats and potential damages that require further evaluation and decision-making.
Low risk	LRI	Risk value<5	There are minor security threats and potential damages that can be accepted without the need for review.

Risk identification refers to the process of discovering, acknowledging, and describing risk elements. This process aims to comprehensively and systematically identify potential ethical risks that may have adverse effects on the education system, students, and educators. The identification of ethical risks in educational artificial intelligence mainly includes the determination of risk types, analysis of risk triggers, and description of risk factors.

3 Ethical risk analysis based on improved differential evolution RBF network

Risk analysis is the process of considering the likelihood of risk occurrence and the severity of its consequences. Its core goal is to deeply explore the potential threats of various risk types based on risk identification, in order to comprehensively understand their possible impacts. The influencing factors of risk can be divided into two dimensions: probability of occurrence and degree of impact. The ethical risk value of artificial intelligence applications in the education field is also determined by both the probability of risk occurrence and the degree of risk impact. According to the risk value calculation formula mentioned earlier, both are directly proportional to the overall risk size. This article adopts the advocacy of the International Organization for Standardization, which divides the probability of risk occurrence and the degree of impact into five levels. Among them, the probability of risk occurrence is divided into five levels: extremely low, low, medium, high, and extremely high, corresponding to the numbers 1-5; The degree of risk impact is divided into five levels based on the potential serious consequences of the risk, namely extremely small, small, medium, large, and extremely large, corresponding to numerical levels 1-5.

3.1 RBF algorithm

Radial basis function neural network (RBF) is a three-layer feedforward network with a single hidden layer, consisting of an input layer, a hidden layer, and an output layer. The network structure is shown in Figure 2.

When constructing relevant models, introduce risk assessment indicators into the input layer. During the process of inputting data into the input layer, no linear or nonlinear processing operations are performed [21]. Assuming that the input layer of the constructed RBF neural network contains n nodes, each corresponding to n different risk assessment indicators. The data transfer process from the input layer to the hidden layer exhibits non-linear transformation characteristics. Here, φ is used to represent the radial basis functions used in the hidden layer. The number of nodes in the hidden layer is consistent with the dimension of the input layer samples, and the connection between the hidden layer and the output layer is achieved through inertia weight w_i. After the data is transmitted to the hidden layer, the activation function is used to calculate the output data value of the hidden layer. Considering the many advantages of Gaussian function, such as concise representation, excellent radial symmetry, and good smoothness, in this model, the radial basis function of the hidden layer is selected as Gaussian function, and its specific formula is as follows:

$$\varphi_i(x) = \exp\left\{-\frac{x-u_i^{\,2}}{2\sigma_i^2}\right\} \tag{1}$$

where, u_i is set as the center vector corresponding to the hidden layer node, while σ_i^2 represents the width parameter of the hidden layer node. When transferring data from the hidden layer to the output layer, it presents the characteristics of linear transformation. The final output result of the output layer is obtained by weighted summation based on the output values of the hidden layer. In this study, only one node was set up in the output layer to output the risk assessment value.

$F(x)$ is used to refer to the actual output situation, and the output process can be described in detail as follows [22, 23]:

$$F(x) = w_1\varphi_1 + w_2\varphi_2 + w_3\varphi_3 + \cdots + w_i\varphi_i \tag{2}$$

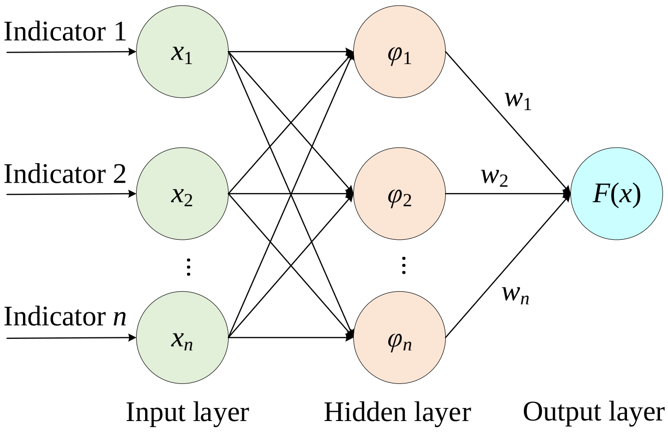


Figure 2: RBF Network Structure

3.2 Dual archive DE algorithm

The establishment of dual archives involves creating an archive archive archive for storing discarded historical individuals and an archive archive archive 2 for storing discarded experimental individuals. Both archives record the individuals eliminated in the current generation and will be used in the next generation. This approach ensures the continuous updating of individuals in the archives, thereby maintaining sufficient novelty and diversity, which helps to continuously optimize and improve the performance of the algorithm [24]. After each iteration of the population, the algorithm in this paper first leaves two archives empty, and adds the original population individuals eliminated by the experimental individuals to archive level. The experimental individuals that cannot replace the original population individuals are added to archive level 2. Through the above approach, it is possible to fully utilize the information generated by discarded potential individuals and experimental individuals in past updates, and use it to maintain diversity in the subsequent evolutionary process. After successfully establishing the archive repository, a resource pool of reusable individuals was created. However, if the archive is used improperly, it may not only hinder the evolution of the population, but may even hinder the population from achieving a balance between exploration and development. In view of this, this article proposes a diversity-based archive usage strategy, which judges the current population status through diversity indicators, aiming to ensure the effective utilization of archives. The definition of population diversity is as follows:

$$DI = \sqrt{\frac{1}{NP}\sum_{i=1}^{NP}\sum_{j=1}^{D}\left(X_{i,j} - \bar{X}_j\right)^2} \tag{3}$$

$$\bar{X}_j = \frac{1}{NP}\sum_{i=1}^{NP} X_{i,j} \tag{4}$$

where, NP is used to characterize the size of the current population, which refers to the number of individuals included in the population; $\bar{X}_j$ specifically refers to the arithmetic mean of the coordinates of each individual in the current population; *DI* represents the diversity index value of contemporary populations, which can reflect the discrete distribution of individuals in the population. Specifically, the larger the value of *DI*, the more dispersed the distribution of individuals within the population, indicating a higher level of diversity; On the contrary, when the value of *DI* is smaller, it means that individuals within the population tend to be more concentrated, and their diversity level is also lower.

In order to accurately measure the trend of population diversity over time or other factors, this study specifically set the relative diversity *RP* value as a measurement indicator, and its specific calculation method is shown in equation (5) [25]:

$$RP = DI / DI_{\text{last}} \tag{5}$$

where, DI_{last} is used to represent the diversity index values presented by the population in the previous iteration process. For the relative diversity *RP* value, there exists a critical threshold α. When the *RP* value is less than α, it indicates that the diversity of the current population is gradually decreasing; When the value of *RP* is greater than α, it means that the diversity of the current population is constantly increasing. In view of this, when a significant decrease in population diversity is detected, this study adopts archive archive and archive 2 data storage structures to enhance population diversity. Through this approach, it is possible to fully explore and utilize individual information generated during the historical iteration process, providing strong support for improving population diversity.

Considering that in the process of population iterative evolution, different stages have different emphasis on optimization objectives. In the early stage of iteration, more emphasis is placed on improving the diversity of the population to explore a wider solution space; In the later stages of iteration, it is usually more inclined to promote the convergence of the population towards the optimal solution. Based on this characteristic, when the number of iterations reaches 0.9 * *maxnfe*, the use of archives to add new individuals to the population will be stopped. This strategic arrangement helps to achieve better convergence of the population and improve optimization efficiency in the later stages of evolution.

The following formula provides a specific calculation method for updating the population size with the iteration process:

$$N_{G+1} = \begin{cases} N_G + 1, RP < \alpha \ \& \ nfe < 0.9 \times maxnfe \\ \left(N^{\min} - N^{\max}\right) \times \dfrac{nfe}{\max nfe} + N^{\max}, \text{otherwise} \end{cases} \tag{6}$$

In the operation rules of population dynamic adjustment, if condition $N_{G+1} > N_G$ is met, an individual is randomly selected from archive archive level and archive 2 and introduced into the current population; When condition $N_{G+1} < N_G$ is met, remove $N_G - N_{G+1}$ individuals with relatively low fitness from the current population; When the condition $N_{G+1} = N_G$ occurs, the size and structure of the population remain unchanged without any adjustments. Through this flexible individual addition and subtraction mechanism, it is possible to fully utilize the information accumulated during the historical iteration process, effectively enhance the diversity level of the population, and further strengthen the exploration and development capabilities of the population in the solution space.

From the perspective of algorithm iteration process, in the early stage of algorithm operation, the population has a large room for improvement, and the number of individuals who successfully achieve updates is relatively high; In the later stages of algorithm operation, the population often approaches the local optimal solution, making it significantly more difficult to further improve, resulting in an increase in the number of individuals who fail to update. Based on this characteristic, when selecting individuals from two archives (archive and archive 2) to join the population, different selection probabilities should be assigned according to the characteristics of different stages [26].

Assuming that in the g-generation iteration process, the number of successfully updated individuals is t1 and the number of failed individuals is t2. The difference in quantity between t1 and t2 can indirectly reflect the amount of valuable information potentially contained in archive 1 (corresponding to archive) and archive 2 (corresponding to archive 2). Generally speaking, archives containing richer information should have a higher probability of individuals being selected to join the population. Based on the above comprehensive considerations, the probability calculation formula for selecting individual X from the archive archive is as follows:

$$\beta = \frac{t1}{t1 + t2} \tag{7}$$

Randomly select individuals from archive archive 2, and in this way, we hope to fully utilize the potential of archive to assist population evolution. It is worth noting that population individuals check every two generations whether the current population needs to add individuals to increase diversity, in order to prevent the population from becoming too large and affecting convergence.

The specific process is shown in the pseudocode below.

Input: The objective function f initializes the population size *NP*, sets the maximum and minimum population sizes to 18 * *D* (where *D* is the function dimension) and 4, sets the maximum iteration count *maxnfe*, initializes empty archive1 and archive2, and sets the initial *DI* value to 1;

Output: The optimal value found in the search.

Line 1: Initialize the position *X* of individuals in the population and initialize control parameters;
Line 2: while (Not meeting the iteration stagnation condition)
Line 3: for i = 1:*NP*
Line 4: Generate a mutation vector based on individual i;
Line 5: Generate experimental vectors based on cross strategy;
Line 6: if $f\left(U_i^g\right) < f\left(X_i^g\right)$
Line 7: Store X_i^g in archive;
Line 8: $X_i^g = U_i^g$;
Line 9: else
Line 10: Store U_i^g in Archive2;
Line 11: end if
Line 12: end for
Line 13: Calculate the *DI* value according to equations (3-4);
Line 14: Calculate *RP* according to equation (5);
Line 15: If *RP*<0.9 and the number of function evaluations is less than maxnfe
Line 16: Select individuals to be added to the population according to equation (7);
Line 17: else
Line 18: Remove a portion of individuals with poor performance from the population;
Line 19: end if
Line 20: Update DE algorithm parameter settings;
Line 21: Empty archive and archive 2;
Line 22: end while

3.3 IDE optimization of RBF neural network implementation

The parameter training method is a key factor affecting the performance of RBF neural networks. Intelligent optimization algorithms are usually used instead of traditional training algorithms to optimize the parameters of neural networks as the optimization objective. This study uses fuzzy clustering algorithm to calculate the center of RBF neural network, optimizes the width and weight parameters of RBF neural network using IDE algorithm, and establishes IDE-RBF neural network model. The key to the convergence speed and ability to find the optimal solution of IDE algorithm lies in the selection of fitness function. In this model, the fitness calculation function for IDE algorithm optimization is RBF neural network mean square error. The formula for the particle fitness function is as follows:

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-y_i\right)^2 \tag{8}$$

where, *n* is the number of training samples; Y_i is the reference output; y_i is the actual output.

The IDE-RBF neural network model is used to assess the risk of sudden water pollution in inland rivers, as shown in Figure 3.

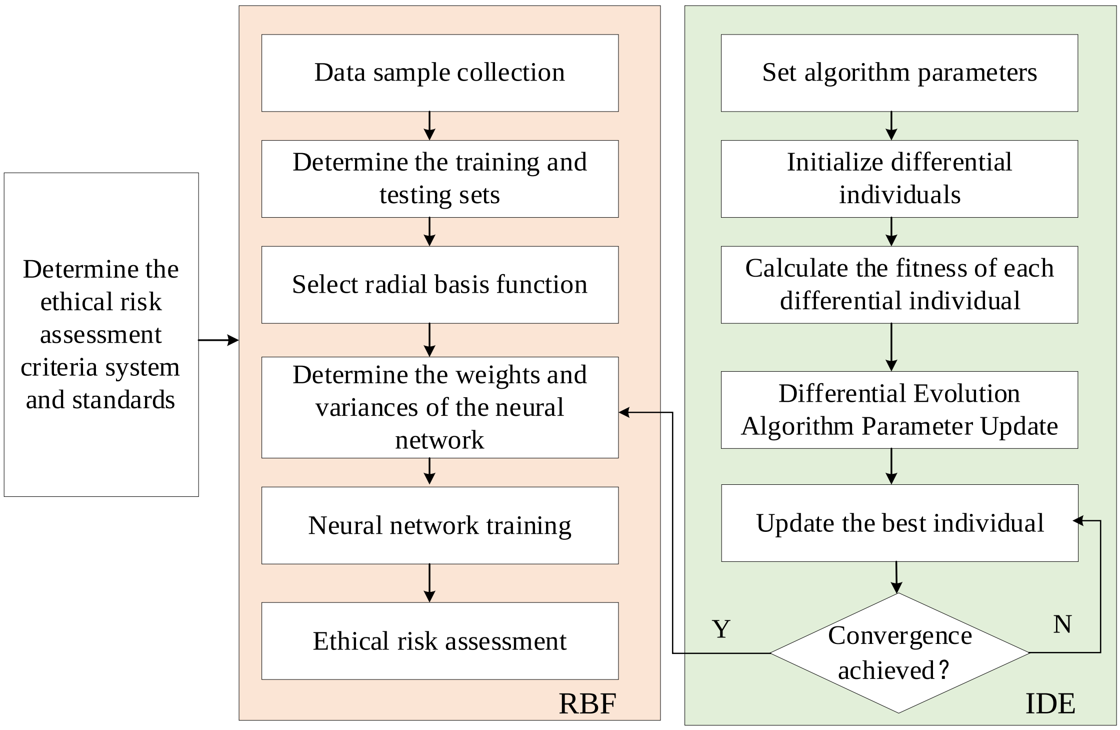


Figure 3: Risk assessment process based on IDE-RBF neural network model

(1) Construct an AIGC technology applied to the ethical risk assessment index system and grading standards of university education, and determine the data samples based on the grading standards, dividing the training set and the testing set.

(2) Train and test the neural network model using the sample set as input vectors, evaluate the prediction performance after training and testing, and conduct error comparison experiments with similar evaluation models to verify the model's effectiveness.

(3) Using the IDE-RBF neural network trained and tested with indicator data from the research area as input, apply AIGC technology to ethical risk assessment in university education, and compare it with traditional evaluation models to verify the reliability of the evaluation

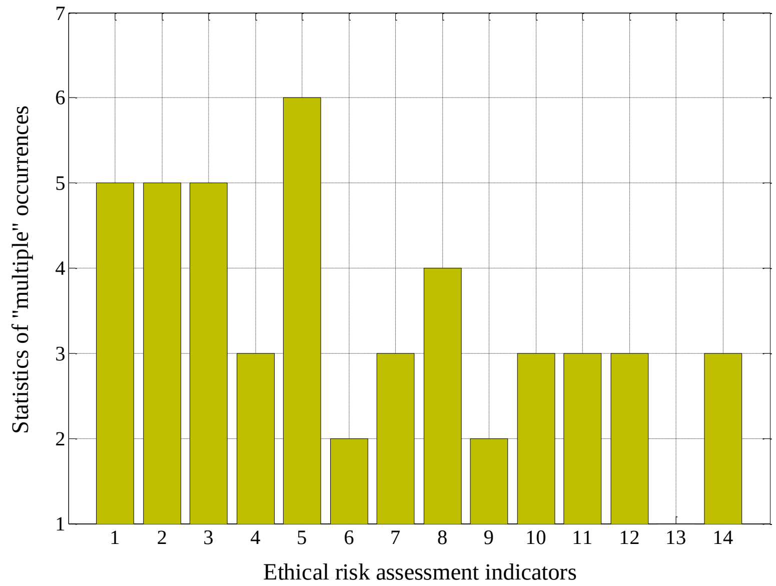
model in this paper.

4 Case studies

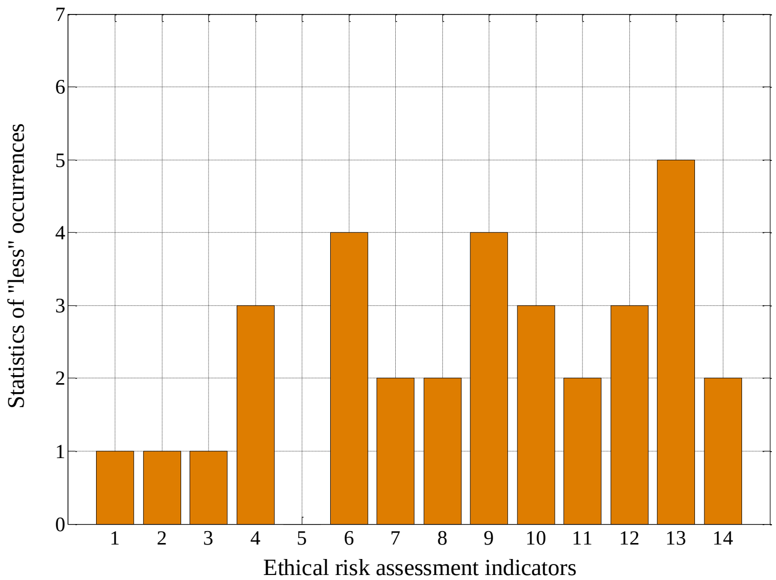
This article provides options for four types of ethical risks in AIGC education applications in each dimension [27]. Among them, Why includes three types: more basis, moderate basis, and less basis, encoded as A1~A3; Who includes seven types: country, school, teacher, student, parent, enterprise, and the general public, coded as B1~B7; What includes three types: point (full staff supervision), line (full process supervision), and surface (comprehensive supervision), coded as C1~C3; Where includes nine types: the promulgation of authoritative mechanisms or laws, the cooperation of school entities, the improvement of ethical literacy of school entities, the implementation of supervision by school managers, the sharp observation and discovery ability of teachers, the awareness of students' courage to raise questions, the implementation of real supervision and reporting by parents, the ethical literacy and self-restraint of enterprises, and the acceleration of the speed of social public assistance in solving problems, coded as D1~D9; When includes six types: timely, timely, real-time, and on-demand, encoded as E1~E4; How to includes six types: controlling ethical risks, transferring ethical risks, reducing ethical risks, eliminating ethical risks, preventing ethical risks, and avoiding ethical risks, coded as F1~F6.

Among the four types of ethical risks in AIGC education applications, this article divides and defines them in detail based on the coding in the previous text, totaling 14 items [28]. Firstly, in the Why theory, all terms are equal to one term. The existing 7-10 governance criteria are divided into A1, the existing 4-6 governance criteria are divided into A2, and the existing 1-3 governance criteria are divided into A3. The maximum number of governance criteria for algorithm bias, discrimination, and algorithm opacity is 10; Secondly, in the Who regulatory body, algorithm bias and discrimination, as well as technology dependence, have the highest number of codes at 7, while algorithm opacity, data loss, and errors have the lowest number of codes at 2; Thirdly, in What regulatory content, algorithm bias and discrimination have the highest number of codes, at 3, while algorithm abuse, data loss and error, technological "fetishism", and loss of autonomy have the lowest number of codes, at 1; Fourthly, among the Where regulated objects, the number of codes without accountability mechanisms is the highest, at 9, while the number of codes for algorithm abuse is the lowest, at 4; Fourthly, in the When regulatory timing, algorithms that are opaque, abused, and leak privacy have the highest number of codes, at 3, while the rest have only two codes; Fifth, in the How to regulate approach, the number of codes that leak privacy is the highest, at 5, while the number of codes for data misuse, data abuse, technology fetishism, technology dependence, and other 8 items is the lowest, at 3.

This article will classify the 14 AIGC education application ethical risks into levels based on the number of codes in each regulatory category. In Why, encoding A1 indicates more, encoding A3 indicates less; In Who, having 4 to 7 codes is considered high, while having 1 to 3 codes is considered low; In What, encoding numbers of 2-3 are considered high, while encoding numbers of 1 are considered low; In Where, codes ranging from 5 to 9 are considered high, while codes ranging from 1 to 4 are considered low; In When, encoding numbers of 3-4 are considered high, while encoding numbers of 2 are considered low; In How to, encoding numbers between 4 and 6 are considered high, while encoding numbers between 1 and 3 are considered low.



(a) Statistics of "multiple" occurrences



(a) Statistics of 'less' occurrences

Figure 4: Horizontal division of 14 ethical risks

From the horizontal division of 14 ethical risks, as shown in Figure 4 on the following page. Finally, the AIGC education application ethical risks will be classified based on the number of times each type of regulation is classified as "more" or "less". The horizontal axis 1-14 indicators in the figure correspond to algorithm bias and discrimination, algorithm opacity, algorithm irresponsibility, algorithm abuse, privacy leakage, data loss and error, data misuse, technology fetishism, technology dependence, AI generated content abuse, reconstruction of teacher-student relationships, loss of autonomy, and lack of accountability mechanisms.

In the grading process, those with 4-6 occurrences of "more" in the six types of supervision are classified as high ethical risk, also known as unacceptable risk; A rating of low ethical risk, also known as limited acceptance risk, with a frequency of 4-6 occurrences of "less"; The remaining ethical risks are classified as moderate ethical risks, also known as limited acceptance risks; Based on this, a classification map of 14 AIGC education application ethical risks is derived, which have different regulatory contents, levels, methods, states, and strategies, as shown in Table 3 on the following page. For example, the five ethical risks of algorithmic discrimination and bias, algorithmic transparency, algorithmic irresponsibility, privacy breaches, and data abuse require high-level regulation (first level warning), strong regulatory measures, mandatory termination and comprehensive inspections, and the use of veto strategies; Seven ethical risks, including algorithm abuse, data misuse, data abuse, and technology dependence, require moderate level regulation (second level warning), strong regulatory methods, high-frequency inspection of regulatory status, and the adoption of warning regulatory strategies; Loss of autonomy, technological "fetishism", data loss and errors require general level supervision (three-level warning), moderate supervision methods, local restrictions and limited inspection supervision status, and the adoption of focused supervision strategies. By formulating graded regulatory measures and implementing risk-based regulatory measures, effective governance of each type of ethical risk in AIGC education applications can ultimately be achieved.

Table 3: Classification of Ethical Risks in AIGC Education Applications

Grade	Regulatory content	Regulatory level	Regulatory approach	Regulatory status	supervise
High ethical risk	Algorithmic discrimination and bias, algorithmic opacity, algorithmic irresponsibility, privacy breaches, data abuse.	High level supervision and first level warning	Strong supervision	Forced termination, comprehensive inspection	Veto strategy
Ethical risk in China	Algorithm abuse, data misuse, technology dependent AI generated content abuse, reconstruction of teacher-student relationships, and lack of accountability mechanisms.	Medium level supervision, Level 2 warning	Strong supervision	High frequency inspection	Warning strategy
Low ethical risk	Loss of autonomy, technological fetishism, data loss and errors.	General level supervision, three-level warning	Moderate regulation	Local restriction, limited inspection	Focus on strategy

5 Exploration of enhancing ethical leadership

In the process of advancing towards intelligence in higher education, ethical leadership plays an irreplaceable and crucial role. The deep integration of AIGC technology has raised a series of ethical issues, such as hidden biases in algorithms, improper leakage of data privacy, and excessive dependence on technology. These situations have significant impacts on the realization of educational equity, the protection of student rights, and the improvement of educational quality that cannot be ignored. The ethical leadership ability not only helps educational organizations accurately identify and evaluate these ethical risks, but also assists them in formulating practical and feasible response strategies to ensure the healthy and sustainable development of AIGC technology in the field of education. The key to ethical leadership ability lies in guiding educational organizations and their relevant personnel to establish correct and clear ethical cognition, strengthen their sense of ethical responsibility, and then consciously follow ethical guidelines in the application of technology. When AIGC is applied in the field of education, ethical leadership can drive educational organizations to establish a comprehensive ethical review system, ensuring that every technology application undergoes strict ethical review. At the same time, it can also encourage educational organizations to strengthen communication and cooperation with stakeholders such as teachers, students, parents, and the general public, work together to address ethical challenges, and create a good ethical environment.

(1) Enhance the level of ethical risk assessment and warning. The optimized differential evolution radial basis function (RBF) network demonstrates excellent accuracy and efficiency in ethical risk assessment, providing strong support for enhancing ethical leadership capabilities. Educational organizations can use this model to conduct a comprehensive and detailed assessment of the ethical risks of AIGC in educational applications. By constructing an ethical risk assessment index system that covers multiple levels such as data, software, hardware, personnel, and services, educational organizations can accurately identify potential ethical risks and conduct quantitative analysis. During the evaluation process, educational organizations should attach importance to the collection and organization of data to ensure the objectivity and accuracy of the evaluation results. Meanwhile, by utilizing the high-precision prediction function of the optimized differential evolution RBF network, educational organizations can establish an ethical risk warning system to promptly detect and respond to potential ethical risks. For example, when algorithm bias or data leakage risk reaches a specific threshold, the system will automatically issue a warning to remind educational organizations to take corresponding measures to intervene.

(2) Improve the ethical decision-making and governance system. Enhancing ethical leadership capabilities also requires educational organizations to improve their ethical decision-making and governance systems. When AIGC is applied in the field of education, ethical decision-making involves numerous stakeholders such as educational organizations, teachers, students, parents, etc. Therefore, educational organizations should establish a diversified ethical decision-making system to ensure that the interests of all parties are fully considered. With the help of optimized differential evolution RBF network, educational organizations can simulate and analyze different ethical decision-making schemes, and evaluate their potential ethical impacts. By comparing the ethical risk values of different schemes, educational organizations can choose the optimal decision-making scheme to reduce ethical risks. At the same time, educational organizations should establish an ethical governance system, clarify the responsibilities and authorities of all parties in ethical decision-making and governance, and ensure the effective implementation of ethical standards.

(3) Strengthen ethical education and training. Improving ethical leadership ability cannot be achieved without conducting ethical education and training for stakeholders such as teachers, students, and management personnel. Educational organizations should incorporate ethics education into their teaching system, and enhance the ethical awareness and literacy of teachers and students through offering ethics courses, hosting ethics lectures, and other forms. At the same time, educational organizations should provide specialized ethics training for management and technical personnel to understand the ethical risks and response strategies of AIGC technology, and to master the methods and skills of ethical decision-making and governance. In the process of ethics education and training, educational organizations can use

optimized differential evolution RBF network models for case analysis, simulate decision-making processes in different ethical scenarios, help teachers, students, and management personnel deeply understand ethical principles, and improve their ethical decision-making abilities. In addition, educational organizations can establish ethical learning communities, encourage teachers, students, and administrators to share ethical experiences, exchange ethical perspectives, and create a good atmosphere for ethical learning.

(4) Create a cultural atmosphere of ethical leadership ability. Improving ethical leadership ability also requires educational organizations to create a cultural atmosphere of ethical leadership ability. The culture of ethical leadership ability is a cultural environment that emphasizes ethical values and advocates ethical behavior. Educational organizations should guide teachers, students, and management personnel to establish correct ethical concepts and consciously abide by ethical standards through the development of ethical norms and the establishment of ethical reward mechanisms. At the same time, educational organizations should strengthen cooperation and communication with external stakeholders to jointly create a cultural atmosphere of ethical leadership ability. For example, educational organizations can collaborate with businesses, research institutions, and other organizations to conduct ethical research projects and jointly explore the ethical applications of AIGC technology in the field of education; We can establish cooperative relationships with government departments, industry associations, etc. to jointly develop ethical norms and standards, and promote the popularization and promotion of ethical leadership culture.

6 Conclusion

This study focuses on the ethical risks arising from the large-scale application of AIGC technology in the intelligent transformation stage of higher education. Aiming to strengthen ethical leadership and achieve proper management of these ethical risks, the following aspects of work have been mainly carried out: 1) Establishment of evaluation system. Clarified the evaluation indicators and complete architecture of ethical risk control mechanisms when introducing AIGC technology in university education scenarios. This system is guided by ethical principles as its core, supported by evaluation indicators and specific methods, and focuses on the three key links of risk identification, analysis, and evaluation. At the same time, we attach great importance to the close connection between the subject and object through the prevention and control governance mechanism, in order to ensure the stable and orderly operation of the system. 2) Algorithm optimization innovation. Innovatively using an improved differential evolution algorithm to optimize the RBF neural network. By introducing a dual archive population optimization strategy, two archives are established to record the information of eliminated individuals and apply this information to the updates of the next generation of individuals. This measure effectively maintains the novelty and diversity of the population, greatly improves the performance of the algorithm, and thereby enhances the accuracy of RBF neural network in predicting ethical risks. 3) Case analysis and research. We have developed an ethical risk assessment index system and grading standards suitable for the application of AIGC technology in university education scenarios. Use the optimized model for training and testing, and compare and analyze it with traditional models. The results indicate that the optimized model has higher reliability and accuracy in evaluating the ethical risks associated with the application of AIGC technology in university education. 4) Enhance strategy discussion. From four dimensions of enhancing ethical risk assessment and warning capabilities, improving ethical decision-making and governance structures, strengthening ethical education and training, and creating a cultural atmosphere centered on ethical leadership capabilities, this article deeply analyzes specific strategies for improving ethical leadership, providing practical guidance for ethical risk prevention and control in the process of intelligent transformation of higher education.

Although this study has achieved phased results, the intelligent transformation of higher education is a dynamic process, and AIGC technology is also continuously iterating and upgrading. In the future, further research can be conducted from the following aspects: 1) Continuous optimization of algorithms. With the continuous expansion of data scale and increasing complexity, it is necessary to continuously improve evaluation models and algorithms, enhance their ability to identify and predict complex ethical risks, and better adapt to the constantly changing educational environment and technological development trends. 2) Expansion of research scope. This study mainly focuses on university education, and in the future, the research scope can be extended to other educational stages, such as primary and secondary education, to deeply explore the ethical risks caused by AIGC technology in different educational scenarios and corresponding prevention and control strategies. 3) Interdisciplinary research reinforcement. Ethical risk prevention and control involves multiple disciplines such as ethics, education, and computer science. In the future, interdisciplinary research should be strengthened, various advantageous resources should be integrated, and a more comprehensive and systematic theoretical and practical system for ethical risk prevention and control should be constructed.

Funding

This work was supported by A Study on the Moral Education Function of General Education Courses in Colleges and UniversitiesX013518

Author's Profile

Yiyuan Sun was borin in Wuhan, China, in 1990. I obtained my bachelor’s degree from Assumption University in Thailand major in tourism management and I also obtained double master degree from Assumption University in Thailand major in business administration and finical management from Wollongong University in Australia. I currently studying the education and leadership as human sciences program in Assumption university of Thailand. My main research direction is the ethical leadership efficacy and AIGC impact in university education

References

[1] Kim M. Abductive Operation and Predictive Privacy Ethics in Artificial Intelligence Generated Content (AIGC): A Case Study of ChatGPT's Ghibli-Style Imagery[J]. Knowledge Management Research, 2025, 26(2): 61-78.

[2] Usher M, Barak M. Unpacking the role of AI ethics online education for science and engineering students[J]. International Journal of STEM Education, 2024, 11(1): 35.

[3] Eden C A, Chisom O N, Adeniyi I S. Integrating AI in education: Opportunities, challenges, and ethical considerations[J]. Magna Scientia Advanced Research and Reviews, 2024, 10(2): 6-13.

[4] Dabis A, Csáki C. AI and ethics: Investigating the first policy responses of higher

education institutions to the challenge of generative AI[J]. Humanities and Social Sciences Communications, 2024, 11(1): 1-13.

[5] Bond M, Khosravi H, De Laat M, et al. A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour[J]. International journal of educational technology in higher education, 2024, 21(1): 4.

[6] Airaj M. Ethical artificial intelligence for teaching-learning in higher education[J]. Education and Information Technologies, 2024, 29(13): 17145-17167.

[7] Mouta A, Pinto-Llorente A M, Torrecilla-Sánchez E M. Uncovering blind spots in education ethics: Insights from a systematic literature review on artificial intelligence in education[J]. International Journal of Artificial Intelligence in Education, 2024, 34(3): 1166-1205.

[8] Khreisat M N, Khilani D, Rusho M A, et al. Ethical implications of AI integration in educational decision making: Systematic review[J]. Educational Administration: Theory and Practice, 2024, 30(5): 8521-8527.

[9] Nguyen K V. The use of generative AI tools in higher education: Ethical and pedagogical principles[J]. Journal of Academic Ethics, 2025: 1-21.

[10] Malik A, Khan M L, Hussain K, et al. AI in higher education: unveiling academicians' perspectives on teaching, research, and ethics in the age of ChatGPT[J]. Interactive Learning Environments, 2025, 33(3): 2390-2406.

[11] Khan W N. Ethical challenges of AI in education: Balancing innovation with data privacy[J]. AI EDIFY Journal, 2024, 1(1): 1-13.

[12] Mumtaz S, Carmichael J, Weiss M, et al. Ethical use of artificial intelligence based tools in higher education: are future business leaders ready?[J]. Education and Information Technologies, 2025, 30(6): 7293-7319.

[13] Aler Tubella A, Mora-Cantallops M, Nieves J C. How to teach responsible AI in Higher Education: challenges and opportunities[J]. Ethics and Information Technology, 2024, 26(1): 3.

[14] Perkins M, Furze L, Roe J, et al. The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment[J]. Journal of University Teaching and Learning Practice, 2024, 21(6): 49-66.

[15] Shahzad M F, Xu S, Zahid H. Exploring the impact of generative AI-based technologies on learning performance through self-efficacy, fairness & ethics, creativity, and trust in higher education[J]. Education and Information Technologies, 2025, 30(3): 3691-3716.

[16] Dieterle E, Dede C, Walker M. The cyclical ethical effects of using artificial intelligence in education[J]. AI & society, 2024, 39(2): 633-643.

[17] Isiaku L, Kwala A F, Sambo K U, et al. Academic evolution in the age of ChatGPT: An in-depth qualitative exploration of its influence on research, learning, and ethics in higher education[J]. Journal of University Teaching and Learning Practice, 2024, 21(6): 67-91.

[18] Acosta-Enriquez B G, Arbulú Ballesteros M A, Arbulu Perez Vargas C G, et al. Knowledge, attitudes, and perceived Ethics regarding the use of ChatGPT among generation Z university students[J]. International Journal for Educational Integrity, 2024, 20(1): 10.

[19] Lérias E, Guerra C, Ferreira P. Literacy in artificial intelligence as a challenge for teaching in higher education: A case study at portalegre polytechnic university[J]. Information, 2024, 15(4): 205.

[20] Nguyen T N T, Van Lai N, Nguyen Q T. Artificial Intelligence (AI) in Education: A Case Study on ChatGPT's Influence on Student Learning Behaviors[J]. Educational Process: International Journal, 2024, 13(2): 105-121.

[21] Agbasi J C, Egbueri J C. Prediction of potentially toxic elements in water resources using MLP-NN, RBF-NN, and ANFIS: a comprehensive review[J]. Environmental Science and Pollution Research, 2024, 31(21): 30370-30398.

[22] Yang X, Zhang Y, Zhao J, et al. A novel active learning method based on matrix-operation RBF model for high-dimensional reliability analysis[J]. Computer Methods in Applied Mechanics and Engineering, 2024, 432: 117434.

[23] Constantinopoulos C, Likas A. An incremental training method for the probabilistic RBF network[J]. IEEE Transactions on Neural Networks, 2024, 17(4): 966-974.

[24] Pesaramelli R S, Sujatha B. Principle correlated feature extraction using differential evolution for improved classification[C]//AIP Conference Proceedings. AIP Publishing LLC, 2024, 2919(1): 070001.

[25] Biswas S, Singh G, Maiti B, et al. Integrating Differential Evolution into Gazelle Optimization for advanced global optimization and engineering applications[J]. Computer Methods in Applied Mechanics and Engineering, 2025, 434: 117588.

[26] de Jesús Rubio J, Orozco E, Cordova D A, et al. Observer-based differential evolution constrained control for safe reference tracking in robots[J]. Neural Networks, 2024, 175: 106273.

[27] Fazil A W, Hakimi M, Shahidzay A K, et al. Exploring the broad impact of AI technologies on student engagement and academic performance in university settings in Afghanistan[J]. RIGGS: Journal of Artificial Intelligence and Digital Business, 2024, 2(2): 56-63.

[28] Allen L K, Kendeou P. ED-AI Lit: An interdisciplinary framework for AI literacy in education[J]. Policy Insights from the Behavioral and Brain Sciences, 2024, 11(1): 3-10.