The Ethical Risks and Standardized Path of AI Empowering Ideological Course Teaching in Universities

Liang Han[1]

[1]*Chongqing College of Humanities, Science and Technology, Chongqing 401524, China*
[1]*Email: 375851229@qq.com*

ABSTRACT
With the rapid development of Artificial Intelligence Generated Content (AIGC), its application in teaching ideological courses in universities brings both innovative opportunities and ethical risks. This article systematically analyzes the ethical risks of data leakage, algorithm bias, value deviation, and subject displacement in the process of AI empowered ideological education teaching, and points out that the inherent tension between technological logic and the value attributes of ideological education is the root cause of the risks. By constructing a hierarchical and progressive governance framework, it is proposed to establish an algorithm value calibration system and intelligent audit mechanism at the technical level, formulate application norms and regulatory mechanisms at the institutional level, and strengthen the cultivation of ethical and critical thinking abilities of teachers and students at the literacy level. This framework aims to achieve compliant and scientific application of AI technology, ensuring the implementation of the fundamental task of cultivating morality and talent.

Keywords: *AI empowerment, Ideological courses, Ethical risk, Standardized path, Governance Framework.*

1. INTRODUCTION

At present, a new round of technological revolution represented by Artificial Intelligence Generated Content (AIGC) is profoundly reshaping the global higher education ecology [1]. Its core ability to understand, generate, and create multimodal content not only revolutionizes the way knowledge is produced and disseminated, but also presents systematic challenges and historic opportunities for university ideological education with value shaping and ideological guidance as its core tasks. As a key link in implementing the fundamental task of "cultivating morality and talent", how ideological education can actively adapt to this change, effectively prevent and control potential risks while grasping the technological dividends, has become an important issue that urgently needs to be addressed.

AIGC is based on massive data and complex algorithms, capable of simulating human cognitive patterns for logical reasoning, content generation, and contextual construction. This essential characteristic and the inherent laws and development needs of ideological education demonstrate profound strategic compatibility in multiple dimensions show the combination of content generation ability and educational resource innovation. AIGC can generate case materials with the characteristics of the times based on real-time social dynamics, transform abstract theory into concrete narrative, and solve the problem that theoretical teaching lags behind social practice [2]. The alignment between situational simulation ability and practical teaching needs. This technology can construct highly simulated historical scenes and value decision-making scenarios, allowing students to deepen their theoretical understanding in virtual practice and make up for the shortcomings of traditional practical teaching in terms of time and space limitations. Through the integration of personalized interaction ability and the principle of teaching students according to their aptitude, and by

analyzing the learning trajectory and cognitive characteristics, AIGC can provide differentiated learning paths and coaching solutions, providing unprecedented technological possibilities for the transformation of the education model from "flood irrigation" to "precision drip irrigation".

However, technology is a sharp weapon for development, and it may also become the source of risk [3]. When generative AI is deeply integrated into ideological education, there are inherent tensions and contradictions between its inherent characteristics of technical logic, algorithm black box and data dependence, and the distinctive ideological attributes, value-oriented goals and humanistic care needs of ideological education, and then a series of unique and complex ethical risks are derived [4]. These risks have gone beyond the instrumental application of ordinary disciplines and penetrated into the whole process of educating people, forming multi-dimensional and deep-seated educational risks.

Therefore, this study aims to systematically identify the unique ethical risks that arise in the process of AI empowering ideological teaching in universities, and reveal their generation mechanisms and transmission logic. Build a hierarchical, progressive, and operable standardized path and governance framework aimed at managing technology, preventing risks, ensuring that AI technology always serves the fundamental task of cultivating morality and promoting the safe, compliant, and beneficial application of AI technology in ideological education.

2. ETHICAL RISK ANALYSIS OF AI EMPOWERING IDEOLOGICAL COURSE TEACHING IN UNIVERSITIES

2.1 Risk Dimension Identification and Typical Performance

The process of AI empowering the teaching of ideological courses in universities is essentially an interactive process between technical logic and ideological education value logic [5]. Due to the special attributes of ideological courses, the uncertainty of AI application may lead to unique ethical risks that are different from general educational scenes, which can be analyzed from four dimensions: data layer, algorithm layer, value layer and subject layer. Different levels of risk have different generation logic and external manifestations, as follows:

2.1.1 Ethical Risk of Data Layer

The dynamic collection of students' classroom behavior, speech content, learning trajectories, etc. may have the problem of excessive collection and exceeding necessary boundaries. Sensitive information input by students using AI tools may be leaked or abused. There are annotation errors and ambiguous expressions in the training corpus of historical figures, events, and red resources.

2.1.2 Ethical Risks at the Algorithm Level

The ideological content generated by AI has implicit biases such as historical nihilism. Algorithms excessively cater to students' preferences and weaken their guidance on depth and higher-order ideological values. The black box algorithm makes it difficult to trace the reasons for the generation of erroneous content.

2.1.3 Value Based Ethical Risks

Excessive reliance on AI generated content has led to a weakening of teacher subjectivity and the emergence of the problem of "technology replacing value guidance". The homogenization of AI content undermines the personalization and infectiousness of ideological course teaching. The application of technology squeezes the value empathy space for face-to-face interaction between teachers and students.

2.1.4 Ethical Risks at the Subject Level

Teachers lack the ability to use AI tools or rely on technology, making it difficult to identify and correct erroneous content. Students overly rely on AI to complete assignments and even assessments, weakening their ability to think independently. In the scenario of human-machine collaboration, the boundary of responsibility is blurred, making it difficult to trace responsibility when problems arise.

2.2 Risk Transmission Logic

Various risks do not exist in isolation, but gradually propagate from the technical level to the value level, ultimately affecting the implementation of the fundamental task of cultivating morality and talent. The transmission process is shown in Figure 1. The core characteristic of risk lies in the natural

tension between the tool attributes of AI technology and the value attributes of ideological courses. Without effective norms, the application of technology may deviate from the educational goals. Therefore, it is necessary to explore systematic standardized paths from the levels of rules, mechanisms, and abilities to ensure the positive empowerment of AI.

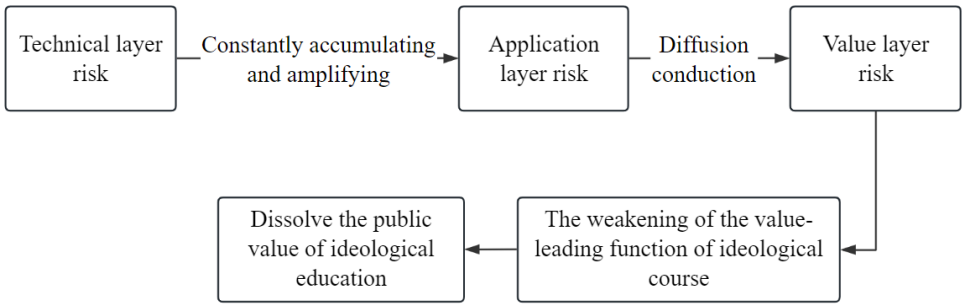


Figure 1 Risk transmission process.

3. CONSTRUCTION OF STANDARDIZED PATH AND GOVERNANCE FRAMEWORK

In response to various ethical risks such as algorithm bias, value deviation, data security, subject displacement, and ethical misconduct in the teaching process of AI empowered ideological courses in universities, based on the core essence of moral education and talent cultivation in university ideological education, and combined with the application characteristics of AIGC technology, a hierarchical collaborative and closed-loop standardized path and multidimensional governance framework are constructed from the three dimensions of technology, system, and literacy. The three dimensions are mutually supportive, interconnected, and complementary. By building a solid safety bottom line through technological foundations, establishing a constraint system through institutional regulations, and activating subject capabilities through literacy empowerment, we aim to achieve the compliant, scientific, and beneficial application of AI technology in ideological education, resolve the conflict between technological empowerment and value guidance, and promote the deep, healthy, and high-quality integration of AI and university ideological education. The overall governance logic is presented through a closed-loop governance process diagram (as shown in Figure 2).

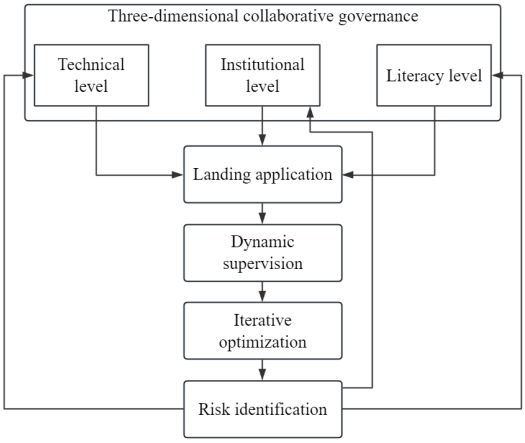


Figure 2 AI empowers the closed-loop governance framework for ethical risks in ideological course teaching in universities.

3.1 Intelligent Empowerment, Building a Solid Ethical and Safety Technology Defense Line

The technical level is the basic support for preventing and controlling the ethical risks of AI ideological teaching. The core is to avoid the risks of algorithm bias, content anomie, data leakage and value deviation from the source through technical optimization, technical adaptation and technical control, so that AI technology can adapt to the ideological attributes and value education attributes of ideological courses, and build a full-chain technical governance system of prevention and control in advance, monitoring in the process and tracing back afterwards [6].

Firstly, establish an ideological exclusive AI algorithm value calibration system. The general

AIGC algorithm has problems such as lack of neutrality, value diversification, and ideological ambiguity, which can easily lead to ideological teaching content deviating from mainstream values. In this regard, it is necessary to rely on the Marxist College of the university, the team of ideological education experts, and joint technical research and development personnel to build an algorithm value constraint model in the ideological field, embedding core value criteria such as socialist core values, mainstream ideology, national education policies, and ideological curriculum teaching outlines into the underlying logic of the algorithm. Through fine-tuning algorithm, value weight assignment, sensitive content filtering training and other technical means, the value of courseware, cases, Q&A content and classroom materials generated by AI is screened, so as to prevent historical nihilism, western erroneous thoughts, one-sided interpretation and other problems, and ensure that the output content of AI always conforms to the educational orientation of ideological courses [7]. Establish a dynamic iterative mechanism of the algorithm, and regularly update the algorithm value database in combination with current affairs policies, social hotspots and ideological education reform requirements, so as to improve the accuracy of the algorithm's adaptation to ideological education.

Second, build an intelligent audit and risk control system for the whole process content. Aiming at the risk of fragmentation, emotion and distortion of AI-generated content, a dual content review mechanism of AI preliminary review and expert review was established [8]. Relying on natural language processing and multimodal recognition technology, all teaching contents generated by AI, such as teaching copy, short video material, classroom question and answer, exercise analysis, etc., are intelligently screened to accurately identify politically sensitive content, false value orientation, vulgar expression, historical deviation and other issues, and automatically intercept illegal content and mark the risk level. For high-risk content, it is directly sealed and banned, and medium and low-risk content is pushed to ideological backbone teachers and subject experts for manual review. Only after correcting the content deviation can it be put into teaching use. Establish a content traceability ledger to fully record the generation time, review process, and usage scenarios of all AI teaching materials, achieving traceability, rectification, and accountability of problematic content.

Thirdly, strengthen the security technology protection of ideological teaching data. Upgrade the data security architecture of the university's smart teaching platform to address risks such as the leakage and abuse of students' personal information, classroom learning data, and teaching case materials in AI teaching applications. By using technologies such as data encryption, desensitization, access permission grading, and cloud security backup, we provide comprehensive protection for the student's ideological dynamic data, classroom interaction data, and learning feedback data collected during the ideological course teaching process. Strictly distinguish between public teaching data and private academic data, prohibit AI technology from privately capturing, sharing, and using core ideological teaching data from commercial universities, block technical vulnerabilities for data leakage, abuse, and tampering, and ensure the security of teacher and student information and the standardized use of ideological teaching data.

3.2 Establishing Regulations And Systems, Building A Systematic Governance Constraint System

The system level is the core guarantee to standardize the application of AI ideology teaching. In view of the current system blank of non-standard, non-standard and non-supervision in the application of AI ideology education in universities, based on the requirements of university teaching management, ideology management and AI application management, an institutional governance framework of four in one, namely, standard and standardization, supervision mechanism, responsibility system and evaluation, is constructed [9], and the teaching of AI empowerment ideology course is rule-based, evidence-based and accountable.

3.2.1 Developing Standards for the Application of AI Enabled Ideological Education in Teaching

Based on the educational attributes of ideological education in universities and the application characteristics of AIGC technology, a special "AI Application Management Specification for Ideological Courses in Universities" has been issued, clarifying the application boundaries of AI technology in various scenarios such as ideological lesson preparation, classroom teaching, after-school tutoring, learning situation analysis, and teaching

evaluation. It is strictly forbidden for AI technology to replace teachers' core teaching, value guidance, ideological guidance and other core educational work, and it is clear that AI is only used as a teaching aid [10]. Simultaneously standardize the selection standards, content review standards, data usage standards, and technical operation and maintenance standards for AI teaching materials, unify the application guidelines for AI ideological teaching in universities, and eliminate problems such as technology abuse, excessive dependence, and disorderly application.

3.2.2 Establishing a Multi-party Collaborative Regulatory Governance Mechanism

Break down the departmental barriers between the university academic affairs office, Marxist college, network information center, and propaganda department, establish a cross departmental AI ideological teaching supervision working group, and clarify the supervisory responsibilities of each department. The Marxist College is responsible for controlling the value orientation of teaching content and reviewing AI teaching application plans; The Academic Affairs Office is responsible for standardizing the teaching process and coordinating the implementation of teaching applications; The Network Information Center is responsible for technical operation and maintenance, as well as data security control; The Propaganda Department is responsible for ideological risk investigation and public opinion monitoring. Establish a normalized inspection mechanism, conduct regular spot checks and evaluations on the application of AI in ideological courses throughout the school, promptly identify value deviations, teaching misconduct, safety hazards, and other issues, and form a regulatory pattern of each performing their own duties and coordinating efforts.

3.2.3 Establishing a Sound Responsibility Traceability System with Clear Rights and Responsibilities

Establish a four level responsibility system consisting of school coordination, college supervisors, teacher responsibility, and technical operation and maintenance support, and clarify the responsibility boundaries of each subject in the application of AI ideological teaching. As the primary responsible person for classroom teaching, the instructor is responsible for the applicability, value, and accuracy of AI teaching materials; The college is responsible for coordinating the AI ideological teaching reform and risk prevention and control work of the college; The technical operation and maintenance team is responsible for technical security, system stability, and data protection; The functional departments of the school are responsible for overall supervision and institutional implementation. Establish a hierarchical accountability mechanism for teaching accidents, ideological issues, and data security issues caused by AI applications, and implement measures such as rectification, accountability, and rectification notification based on the severity of the problems, forcing all parties to standardize the application of AI technology.

3.2.4 Improving the AI Ideology Teaching Assessment and Evaluation Mechanism

Incorporate the application of AI technical norms and the effectiveness of ethical risk prevention and control into the teaching quality evaluation, teacher performance evaluation and curriculum construction evaluation system of ideological courses in universities [11]. Abandoning the evaluation orientation solely based on technology utilization rate, the focus is on assessing the educational adaptability, value leadership effectiveness, and risk prevention effectiveness of AI applications. Teachers are encouraged to use AI technology reasonably, moderately, and scientifically to optimize teaching, eliminate formal, homogeneous, and utilitarian technology applications, and guide AI technology to return to the essence of education.

3.3 Empowering the Subject and Cultivating the Core Ability of Cultivating Good People

Teachers and students are the core subjects of AI empowerment ideology teaching, and their AI ethical literacy, technical application ability and value judgment ability are the key cores to resolve ethical risks [12]. The literacy level focuses on both teachers and students. Through systematic training, normalized cultivation and normalized guidance, the risks of teachers' technical abuse, weakened value, students' lack of critical thinking and technical dependence are solved, and the core goal of technology serving people and educating people to lead technology is realized.

Firstly, strengthen the cultivation of ideological teachers' AI ethics and teaching application literacy. Establish a normalized and systematic teacher training system, incorporating AI ethical norms, AI ideological teaching application skills, risk identification and prevention capabilities into the core content of ideological teacher continuing education, pre service training, and research and development training. On the one hand, conducting technical application training to help teachers master practical abilities such as AI material screening, content verification, algorithm identification, and data security protection, and enhance their level of intelligent teaching; On the other hand, emphasis should be placed on conducting ethical literacy training, strengthening teachers' ideological bottom line thinking and value oriented subject consciousness, guiding teachers to recognize the tool attributes of AI technology, eliminating excessive reliance on AI for lesson preparation, AI Q&A, and AI teaching, and adhering to the core position of teachers in ideological guidance, value shaping, and personality education. Build an AI teaching and research exchange platform for ideological teachers, and enhance their practical ability to deal with ethical risks in AI teaching through case studies, public course presentations, risk reviews, and other forms.

Secondly, cultivate college students' AI critical thinking and ethical self-discipline. To address issues such as students' excessive reliance on AI to complete assignments, blind adherence to AI perspectives, and weakened value judgments, we will integrate the cultivation of AI ethics, digital literacy, and critical thinking abilities into the entire process of ideological classroom teaching. Add expanded content such as AI ethics, digital civilization, and network ideology to ideological courses, guiding students to have a correct understanding of the tool value and potential risks of AI, and cultivating students' ability to distinguish, critically think, and judge the value of AI generated viewpoints, materials, and conclusions. Through themed class meetings, special lectures, social practices, and other forms, guide students to establish correct digital ethics, consciously resist AI false information and erroneous value orientation, eliminate plagiarism of AI homework, blind obedience to AI viewpoints, and cultivate students' core competencies of independent thinking, independent judgment, and adherence to mainstream values.

Finally, establish a two-way collaborative ethical self-discipline system between teachers and students. Establish a normalized interactive feedback mechanism between teachers and students, encourage students to provide feedback and supervision on AI application issues, incorrect guidance, and unreasonable content in ideological classrooms, and form a two-way supervision and common standardization of education between teachers and students. Cultivate a digital culture of good in ideological education in universities, establish the AI application concept of technology for good, education oriented, and ethics first through campus propaganda, teaching and research guidance, and atmosphere creation, and make the standardized use of technology, rational use of technology, and safe use of technology a consensus among teachers and students, avoiding the ethical risks of AI ideological teaching from the ideological root.

4. CONCLUSION

AI technology and ideological education have strategic compatibility in terms of content generation, situational simulation, personalized interaction, and other dimensions. However, the technical characteristics of algorithm black box, data dependence, and value diversity have inherent tension with the distinct ideological attributes of ideological education, which gives rise to ethical risks in four dimensions: data layer, algorithm layer, value layer, and subject layer. Through the transmission chain of technology, value, and education, it affects the implementation of the fundamental task of cultivating morality and talents. To this end, this study has constructed a standardized path that integrates technology foundation, institutional regulation, and literacy empowerment. Through value calibration, intelligent auditing, and data protection, the technical defense line is strengthened. The governance system is improved through standard specifications, collaborative supervision, responsibility tracing, and assessment evaluation. The internal driving force is activated through the cultivation of ethical literacy between teachers and students. The collaborative linkage and closed-loop operation of the three can effectively manage technological risks, ensure that AI technology always serves the purpose of education, and promote the high-quality integration of AI with university ideological education in terms of safety, compliance, and goodness.

REFERENCES

[1] Wang J .Research on the Reconstruction of the Practical Teaching Model of Digital Intelligence Empowerment-driven Applied Undergraduate Ideological and Political Courses[J].Scientific Journal of Economics and Management Research,2026,8(8):98-105.DOI:10.54691/F685TD77.

[2] Li D ,Wang K .The Application of Generative AI in English Learning: Opportunities, Ethical Risks and Standardization Paths[J].Journal of Educational Research and Policies,2026,8(4):29-34.DOI:10.53469/JERP.2026.08(04).07.

[3] Ding X .Artificial Intelligence Application and Ethical Risk Governance in Public Decision Making[J].Scientific and Social Research,2026,8(6):119-125.DOI:10.26689/SSR.V8I6.15412.

[4] Ejoh I C ,Akeke U C ,Oladoyinbo B O , et al.Ethical and Secure Deployment of Generative AI: Balancing Innovation, Data Privacy, and Enterprise Risk Governance[J].Journal of Engineering Research and Reports,2026,28(7):48-68.DOI:10.9734/JERR/2026/V28I71943.

[5] Sheppard M S .Artificial Intelligence in the Behavioral Health Professions: Ethical and Risk Management Issues. Frederic G. Reamer[J].Health & Social Work,2026,51(2):170-172.DOI:10.1093/HSW/HLAG006.

[6] Sun W .Big Data Technology Empowers the Improvement of Teaching Ability of Ideological and Political Course Teachers in Colleges and Universities[J].Contemporary Education Frontiers,2025,3(9):69-74.DOI:10.18063/CEF.V3I9.1077.

[7] Paglioni P V ,Mortenson T .A Systems-Level Approach to Address Risks and Ethics in Artificial Intelligence Systems[J].Systems Engineering,2026,29(5):793-806.DOI:10.1002/SYS.70062.

[8] Jin J ,Wang L ,Wu S .Ethical Risks and Prevention of Generative AI Embedded in Ideological and Political Education in Colleges and Universities[J].Journal of Education and Educational Research,2026,18(2):69-75.DOI:10.54097/HBWX6A90.

[9] Wilkins E .Ethical integration and artificial intelligence in speech-language pathology: balancing risks and opportunities to deliver safe, person-centred care[J].Journal of Clinical Practice in Speech-Language Pathology,2026,28(1):55-61.DOI:10.1080/22000259.2026.2665861.

[10] Mao Y .Ethical Risks of Artificial Intelligence in Financial Management: Identification and Governance Based on Stakeholder Theory[J].Proceedings of Business and Economic Studies,2025,8(8):115-121.DOI:10.26689/PBES.V8I8.13356.

[11] Bellaby R .Managing the ethical risks of AI in intelligence: a multi-layered framework[J].AI & SOCIETY,2026,41(7):1-18.DOI:10.1007/S00146-026-03045-2.

[12] Su N .Research on the Innovation Mechanism of Ideological and Political Course Teaching in Universities Empowered by Artificial Intelligence[J].Journal of Education and Educational Research,2026,18(2):9-14.DOI:10.54097/W8VY8R51.