Governing Logic and Path Reconstruction of AIGC Abuse in Educational Contexts: A Three-Dimensional Analysis Based on Technology-Ethics-Institution

Qinqin Li*

College of Education, Guangxi Normal University,Guilin Guangxi, China541000

Abstract: The widespread application of generative artificial intelligence in education has brought about abuse risks such as academic fraud, AI-assisted ghostwriting, data privacy leakage, and teaching ethics violations.With the rapid iteration and extensive application of generative artificial intelligence (AIGC) technology, the risks of its abuse have evolved from potential threats into multiple practical problems, including technological out-of-control, ethical disorder, and institutional deficiency. Grounded in collaborative governance theory, risk society theory, and technological ethics framework, this paper adopts case analysis and comparative research methods to dissect the essential risks of AIGC abuse from three dimensions: technically, it manifests as the unbounded expansion of instrumental rationality and the imbalance between prevention and control capabilities; ethically, it reflects the dissolution of social value consensus by algorithmic logic; institutionally, it highlights the inadequate adaptation between regulatory systems and technological application scenarios. Furthermore, this paper analyzes the core dilemmas in AIGC governance, such as the dynamic imbalance between technological innovation and risk prevention, conflicts of interest among multiple stakeholders, and difficulties in constructing cross-cultural ethical consensus. It proposes a technology-ethics-institution collaborative governance system and ultimately constructs a multi-stakeholder co-governance network characterized by government leadership, enterprise primary responsibility, and social participation. This study provides theoretical references and practical paths for balancing the innovative development and risk prevention of AIGC.This study is of great practical significance for educational digital governance, campus AI ethics, and academic integrity construction.

Keywords: AIGC; Technological Prevention and Control; Ethical Consensus; Institutional Regulation; Collaborative Governance

DOI:10.12417/3029-2328.26.03.007

1.Introduction

With the iterative evolution of artificial intelligence technology, generative artificial intelligence (AIGC) has expanded from a single dimension of text generation to the field of multi-modal content creation.In educational scenarios, AIGC has been increasingly applied in teaching, learning, assessment, and academic research, yet its abuse has also triggered prominent problems such as ghostwriting, academic plagiarism, privacy breaches of students and teachers, and erosion of educational ethics. while technology empowers social development, the risks of AIGC abuse have become increasingly prominent. In 2024, incidents of AIGC abuse witnessed explosive growth. The Stanford University Human-Centered AI (HAI) Report explicitly states that the number of global AI-related harmful incidents surged to 233 in 2024, a sharp increase of 56.4% compared with 2023, including numerous malicious cases such as deepfake private images and chatbots inducing adolescent suicide.

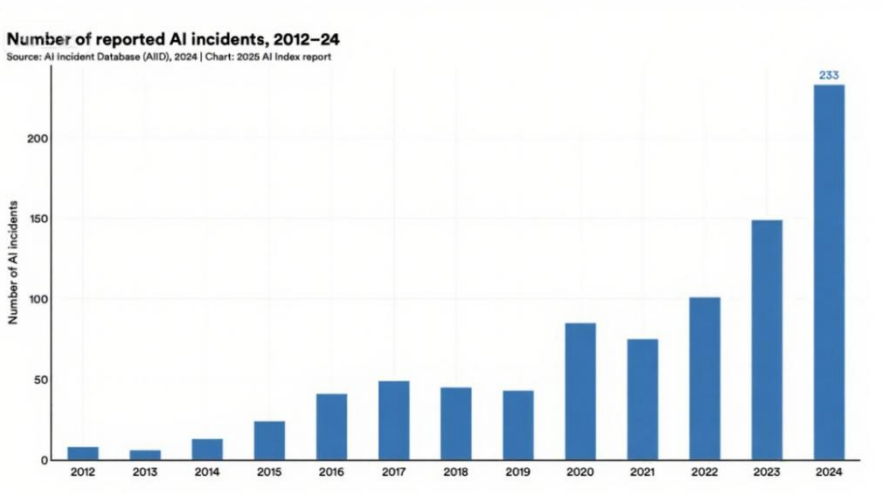

Figure 1. Global AI-Related Harmful Incidents

（Data source: Stanford University HAI Report）

In the early 2025 Shigatse earthquake in Tibet, an image depicting a child wearing a hat in earthquake ruins triggered widespread public concern and numerous comments from netizens worried about the child's safety. It was not until the topic "The Picture of the Child in the Earthquake Ruins is AI-Generated" trended on Weibo that the public realized they had been misled by false information. After the earthquake, some self-media deliberately removed the AI-generated label when reposting the image, added relevant earthquake topic tags and guiding copy, leading the image to be misinterpreted as real disaster footage. This not only violated the ethics of information dissemination but also dissipated the public's goodwill and concern for disaster relief. In 2025, the Cyberspace Administration of China launched the special campaign "Clean Network · Rectifying AI Technology Abuse", directly targeting the damage of AI-generated false content to the online ecosystem, which confirms the urgency of addressing such issues. In July 2025, the World Artificial Intelligence Conference (WAIC 2025) was held in Shanghai. Premier Li Qiang of the State Council, in his keynote speech at the opening ceremony, listed "AIGC false content and deepfakes" as the primary task of global AI governance for the first time.

Figure 2. The Picture of the Child in the Earthquake Ruins is AI-Generated

（Source: CCTV News Client）

https://news.cnr.cn/native/gd/20250110/t20250110_527036359.shtml

2.Analysis of the Three-Dimensional Risk Nature of AIGC Abuse

2.1 *Technological Out-of-Control: Expansion of Instrumental Rationality and Imbalance of Prevention and Control Capabilities*

The out-of-control risk of AIGC technology stems from a structural imbalance between the speed of technological evolution and the capacity for risk identification, with its core manifestation being the unbounded expansion of instrumental rationality. On the one hand, the open-source and lightweight nature of generative tools has significantly lowered the threshold for technology adoption: GitHub's annual report shows that 65,000 new generative AI projects were added in 2023, representing a year-on-year increase of 248%, and the total number of such projects in 2023 was more than double that of 2022. On the other hand, breakthroughs in the parameter scale of large models—from tens of millions of parameters in 2017 to trillions of parameters in 2023—have enhanced the performance of generative models such as Generative Adversarial Networks (GANs), significantly improving the fidelity of generated content. [1]For instance, in March 2025, a "government document generator" leaked from an overseas open-source community was exploited by criminals to forge a notice suspending public transportation in a city. Due to the high consistency of the document's format and official seals with authentic government documents, it was reposted over 100,000 times within two hours of spreading in grassroots communities, triggering panic buying among citizens.[2]Difficult identifiability also makes AIGC abuse in education, such as automatic paper-writing and homework-generating tools, more concealed and difficult to identify.

2.2 *Ethical Disorder: Dissolution of Social Value Consensus by Algorithmic Logic*

The data-driven generation logic of AIGC reconstructs existing social ethical norms, with its core contradiction lying in the conflict between algorithmic instrumental rationality and human value rationality, which is specifically manifested in three ethical dilemmas. First, the dissolution of authenticity ethics. False information spreads faster and wider on social media platforms and is more likely to be shared than factual content. [3] Second, the transgression of rights ethics. For example, using AI to extract celebrities' portraits and voice data to create virtual idols may constitute unauthorized infringement of portrait rights and privacy rights. [4] In education, this is reflected in the unauthorized collection and use of students' behavioral data, learning records, and facial information, which seriously violates privacy and educational ethics.Third, the bias in value ethics. For instance, gender and racial biases embedded in word embedding models can be amplified and distort social cognition. [5] From the perspective of ethical theory, the root cause of AIGC ethical disorder is the virtualization of responsible subjects.In educational practice, it further leads to the blurring of responsibility for academic integrity and the alienation of teacher-student relationships.

2.3 *Institutional Deficiency: Inadequate Adaptation Between Regulatory Systems and Technological Scenarios*

The existing AIGC governance institutions have failed to keep pace with the expansion of technological application scenarios, resulting in a passive situation where "rules lag behind practice", which is primarily reflected in three adaptability contradictions. First, the contradiction of ambiguous responsibility definition. Current laws lack clear standards for delineating rights and responsibilities among various stakeholders in the AIGC industrial chain. AIGC-generated content lacks originality; the technical embedding of AIGC is influenced by designers' and users' 'cognition and personal preferences, and the opacity and inexplicability of algorithm operations further exacerbate ethical risks [6]. In the education sector, the absence of special regulatory norms for AIGC applications makes campus governance and educational supervision face institutional gaps.Second, the contradiction of cross-border governance limitations. The borderless dissemination of AIGC content conflicts with the principle of territorial jurisdiction, and the complexity of cross-border regulatory cooperation may lead to time delays. [7] Third, the contradiction of inefficient rights relief. Victims in "AI face-swapping fraud cases" face the dilemma of "difficult proof and high cost" in safeguarding their rights. According to reports, victims need to commission professional institutions for technical identification, which costs over 50,000 yuan and takes six months to complete the evidence

collection process. From the perspective of institutional change theory, this institutional deficiency stems from the path dependence effect.

3.Governance Dilemmas and Core Tensions of AIGC Abuse

3.1 *Dynamic Imbalance Between Technological Innovation and Risk Prevention*

AIGC governance first confronts the dilemma of balancing innovative incentives and risk constraints. On the one hand, technological innovation delivers significant social value: in government affairs, AI applications in emergency response improve real-time decision-making and communication efficiency;[8]in healthcare, AI-generated personalized diagnosis and treatment recommendations reduce misdiagnosis rates in primary medical institutions.In education, AIGC can optimize teaching design, personalized learning, and intelligent evaluation, but excessive supervision will restrict its positive role in educational digital transformation. Excessive regulations, such as strict restrictions on training data usage and cumbersome approval procedures, may stifle the positive application of technology in public services. [9] On the other hand, lax regulation risks systemic harm. AIGC still faces challenges in regulation, technological maturity, and ethics, and may trigger risks including false information dissemination, corporate losses, and market disruption. [10] This tension between innovation and prevention represents a trade-off between short-term technological dividends and long-term social costs, and existing governance tools struggle to achieve a dynamic equilibrium.

3.2 *Conflicts of Interest Among Multiple Stakeholders*

AIGC governance involves multiple stakeholders—government, enterprises, the public, and academia—whose divergent interest demands form a complex governance game. As technology R&D entities, tech enterprises pursue commercial profits and technological monopolies, resisting governance measures that increase operational costs. For example, in 2025, the EU mandated full-process data traceability for AIGC enterprises, and companies including Google and Meta jointly opposed the requirement, claiming it would increase R&D costs by 30%. [11] As representatives of public interests, governments prioritize maintaining social order and national security, favoring strengthened regulatory measures. As technology users and risk bearers, the public desires both the convenience of AIGC and the protection of personal rights and information security. However, they lack effective voice in governance: only 40% of students follow government AI regulation news, and 32% express interest in AI policy as a career path, indicating limited public participation and engagement in AI governance.[12] In educational scenarios, schools, teachers, students, and EdTech companies also have divergent demands, making it difficult to form a unified governance consensus.The fragmentation of stakeholder interests hinders consensus formation on governance policies and weakens implementation effectiveness.

3.3 *Difficulties in Cross-Cultural Construction of Ethical Consensus*

The formulation of AIGC ethical guidelines faces the fundamental contradiction between universality and cultural diversity. Globally, cultural differences shape distinct perceptions of AI ethics: Western countries emphasize individual privacy and freedom of speech, adopting a cautious stance toward government content regulation; East Asian nations prioritize social order and collective interests, advocating strict norms for AIGC content dissemination. International divergence on AI ethics, rooted in cultural disparities and stakeholder conflicts, impedes the establishment of unified global standards, requiring multilateral cooperation and respect for cultural diversity.[13] Domestically, social groups hold divergent ethical expectations: the cultural and artistic community worries about copyright infringement by AI-generated content, advocating strict restrictions on AI applications in creative fields; the tech industry promotes ethical inclusiveness to avoid stifling innovation; elderly groups, with limited ability to identify AI false information, call for enhanced content review; young people focus on algorithmic bias and data privacy protection.In the field of education, the differentiation of ethical cognition is more prominent: academic integrity, teacher ethics, student privacy protection, and fair education have become key points of ethical game.The fragmentation of ethical cognition poses substantial challenges to the implementation of unified ethical norms.

4.Path Reconstruction of AIGC Collaborative Governance System

4.1 *Technological Governance: Source Prevention Centered on Controllability*

Technological governance constitutes the foundation of AIGC risk prevention, requiring the construction of a full-process technical supervision system. First, establish a mandatory traceability mechanism. All AIGC content must embed tamper-proof digital watermarks containing core information such as the generating entity, timestamp, model version, and data source. The watermark technology adopts a dual-mode design: explicit labels inform the public of AI generation, while implicit coding enables regulatory authorities to trace the source. This can be applied to educational assignments, papers, and teaching materials to realize traceable identification of AI-generated content.Second, implement risk classification management. Differentiated regulatory standards are formulated based on application scenario risk levels: low-risk scenarios (e.g., entertainment text generation) adopt a filing system; medium-risk scenarios (e.g., commercial advertising) require third-party security assessments; high-risk scenarios (e.g., government announcements) implement an approval system. Third, standardize open-source tool management. Sensitive content generation tools (e.g., face/voice synthesis) are subject to "filing-classification-accountability" oversight: developers must register tool functions and usage scope; high-risk tools are restricted to qualified institutions; developers bear joint liability for abuses. A dynamic monitoring platform is established to intercept and trace illegal generative tools in real time.

4.2 *Ethical Reconstruction: Consensus Cultivation Guided by Value Embedding*

Ethical reconstruction must transcend the myth of "technological neutrality" and integrate value rationality into the entire AIGC lifecycle, establishing a multi-stakeholder ethical governance mechanism. First, establish a cross-domain ethics committee. A national AIGC ethics committee, composed of government regulators, university ethics researchers, enterprise technical representatives, and public representatives in a 2:3:3:2 ratio, formulates the AIGC Ethics Guidelines, defining prohibitions and advocacy principles. It can absorb educational experts and frontline teachers to strengthen the guidance of educational ethics and academic integrity.Second,implement algorithmic ethics pre-review. High-risk AIGC applications must submit ethical impact assessments before launch, focusing on training data fairness, algorithm transparency, and content value orientation. An expert review pool is established, with panels selected randomly by field; review outcomes are mandatory for launch approval. Third, build a public participation platform. Ethical hearings, online opinion solicitation, and AI ethics education in schools/communities enhance public awareness and participation.

4.3 *Institutional Improvement: Rule System Aimed at Clear Rights and Responsibilities*

Institutional reform must break path dependence and construct a legal regulatory network covering the entire "R&D-application-dissemination" chain. First, advance special legislation. Accelerate the formulation of the Generative Artificial Intelligence Governance Law, clarifying tripartite responsibilities: developers must conduct technical safety reviews, correct algorithmic biases, and embed traceability, bearing joint liability for abuses; platforms establish AIGC content review centers, implementing dual human-technical audits for high-risk content, with penalties of 5–10 times illegal gains for non-compliance; users are liable for generated content legality, with severe penalties for criminal misuse. At the same time, supporting policies for educational digitization should be improved, and special norms for AIGC in campus scenarios should be formulated to strengthen educational supervision and campus governance.Second, establish cross-border collaborative governance. Participate in global AIGC rule-making, join the International AI Content Supervision Alliance, and establish cross-border information sharing, evidence mutual recognition, and joint enforcement mechanisms. Third, innovate rights relief channels. Establish an AI Infringement Rapid Rights Protection Center integrating technical identification, legal consultation, mediation, and litigation support for one-stop relief.

4.4 *Multi-Party Collaboration: Building a Government-Led Governance Network*

AIGC governance must transcend single-subject limitations and form a collaborative pattern of government leadership, enterprise primary responsibility, and social participation. At the government level, clarify inter-departmental regulatory division and coordination: the Cyberspace Administration oversees overall governance and cross-departmental coordination; the Ministry of Science and Technology regulates R&D, focusing on model safety and data compliance; market regulators oversee commercial applications, cracking down on false advertising and infringement; judicial authorities refine trial rules for AIGC infringement cases to unify judicial standards.Educational administrative departments can participate as special governance subjects to promote the coordinated governance of AIGC in education.The National AIGC Governance Center is established to integrate data resources and build an integrated platform for risk early warning, real-time monitoring, and emergency response. At the enterprise level, fulfill primary responsibilities: tech enterprises embed ethics and compliance into R&D, establishing internal ethics review departments; platforms optimize review algorithms and train auditors to improve false content detection; data enterprises establish compliant training data systems to prevent privacy breaches. At the social level, leverage multi-party supervision: media expose abuse cases and interpret policies; industry associations develop self-regulatory norms; academia conducts governance research; NGOs promote digital literacy to enhance public critical thinking about AIGC content.

5.Future Outlook

In the future, the integration of AIGC with emerging technologies such as the metaverse and brain-computer interfaces will spawn novel applications, including AI-generated virtual social rules and brain-computer interface-driven AIGC content. These scenarios may trigger new risks, such as blurred virtual-real boundaries and disputes over consciousness control, requiring proactive risk warning and governance research to explore preventive governance models. Additionally, strengthening quantitative evaluation of governance tools (e.g., watermark detection rates) and institutional measures (e.g., burden-of-proof reversal effects) will optimize governance frameworks through cost-benefit analysis. Against the background of educational digitization, the application of AIGC in teaching, research, and management will be deeper, and its governance needs to be closely integrated with educational scenarios to better protect academic integrity, student rights, and educational fairness. With the increasing cross-border dissemination of AIGC, deepening international cooperation and comparative research will facilitate the formation of globally recognized governance rules, avoiding technological fragmentation and regulatory arbitrage, and realizing global collaborative governance of AIGC technology.

References:

[1] Lin,R.(2024).Exploration and Thinking on Intelligent Computing Technology Innovation by China Telecom. Communications World,(3),36–38.

[2] Strategic Study of CAE.(2024).Construction of a Technological Security Risk Monitoring and Early Warning System.Chinese Academy of Engineering,16(2),112–120.

[3] Vosoughi,S.,Roy,D.,&Aral,S.(2018).The spread of true and false news online.Science,359(6380),1146–1151.

[4] Zhou,X.(2024).Research on Copyright Infringement Risks in Generative AI Data Use.Open Journal of Legal Science,12(2),89–102.

[5] Nadeem,M.,Bethke,A.,&Liu,Y.(2020).StereoSet:Measuring stereotypical bias in pretrained language models. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics,5426–5433.

[6] Liu,Z.,&Fan,M.(2024).Opportunities and challenges:AIGC empowers the protection and inheritance of traditional culture.Frontiers in Humanities and Social Sciences,8(1),145–158.

[7] Xu,Y.L.(2022).Research on Customs Supervision Issues and Countermeasures for Cross-Border E-Commerce in

China.Master's Thesis,Soochow University,Jiangsu.

[8] Li,Z.S.(2024).Leveraging AI automated emergency response with natural language processing:Enhancing real-time decision making and communication.Applied and Computational Engineering,5(3),210–225.

[9] Wang,P.J.(2023).Challenges and Transformations Brought by AIGC to the Media Industry.Audio-Visual Circle, (3),34–39.

[10] Wang,T.Y.,Li,L.,&Zhang,H.(2024).A Study on the Risks and Countermeasures of False Information Caused by AIGC.Journal of Electrical Systems,20(2),345–358.

[11] Heilongjiang Provincial Library.(2024).Escaping the"Colling ridge Dilemma"of AIGC:Full-Process Dynamic Data Compliance Governance.Library and Information Knowledge,(2),56–63.

[12] Weichert,J.,Kim,D.,&Chen,L.(2025).Assessing Computer Science Student Attitudes Towards AI Ethics and Policy.Journal of Educational Technology&Society,28(2),123–138.

[13] Goffi,E.R.,&Momčilović,A.(2022).Respecting cultural diversity in ethics applied to AI:A new approach for multicultural governance.Misión Jurídica,14(3),78–92.