

Research on Copyright Definition and Ethical Norms of Creative Content in Generative Artificial Intelligence

Aisheng Zhang*

Guangzhou University of Software, Guangzhou, 510990, China

*Corresponding author: zas219@126.com.

Abstract: *With the rapid advancement of generative artificial intelligence technology, the extensive application of its generated content in literary, artistic, and scientific fields has posed profound challenges to existing copyright systems and ethical norms. This paper conducts a systematic study on the copyright definition and ethical norms of generative AI-created content. It begins by theoretically analyzing whether AI-generated content meets the originality requirements under copyright law, examines the ownership of rights under different subject models, and thereby reveals the impact on traditional author-centric theories. Furthermore, this paper analyzes the divergences in the judicial recognition of AI-generated content, the legality of training data sources and associated infringement risks, as well as the ambiguity of rights attribution under current legal frameworks. It proposes comprehensive solutions combining legal interpretation expansion, institutional innovation, and technological governance. On the ethical front, this paper examines risks related to the lack of transparency, algorithmic bias, and the erosion of cultural diversity in generative AI, and constructs a collaborative governance framework encompassing technical standards, industry self-regulation, and social supervision, aiming to provide theoretical support for the improvement of relevant laws and ethical development.*

Keywords: *generative artificial intelligence; copyright definition; ethical norms; copyrightability; algorithmic bias; governance framework*

Introduction

The rise of generative artificial intelligence is reshaping the fundamental models of content creation and dissemination. While it enhances production efficiency and creativity, it also poses systematic challenges to the copyright system and ethical order established on human subjectivity. Current laws face multiple difficulties in defining the copyright attributes of AI-generated content, including ambiguous originality standards, dilemmas in rights attribution, and complexities in infringement identification. Moreover, ethical risks such as technological black boxes, data bias, and accountability gaps further highlight the urgency of normative construction. Against this backdrop, systematically studying the copyright definition pathways and ethical governance mechanisms for generative AI-created content not only holds academic significance in advancing intellectual property theory to adapt to technological developments but also addresses the institutional need to build a trustworthy, fair, and sustainable AI innovation ecosystem. Through theoretical analysis, controversy clarification, and normative exploration, this paper aims to provide academic insights for constructing a legal and ethical framework that balances technological incentives with rights protection, and innovation freedom with social responsibility.

1. Theoretical Analysis of the Copyrightability of Generative AI-Created Content

1.1 The Theory and Connotation of Originality in Copyright Law

The cornerstone of copyright protection lies in the originality of a work. The standard of originality generally comprises two core dimensions: "independent creation" and "a minimal degree of creativity." Independent creation requires that the work originates from the author itself, without plagiarizing others; a minimal degree of creativity demands that the work reflects certain unique intellectual choices, judgments, and personal expression. This standard is often aligned with the "imprint of personality"

theory in civil law systems, while in common law systems, it manifests as the modern evolution of the "sweat of the brow" doctrine.

In the context of traditional human creation, the originality standard is closely linked to the author's mental activities, thoughts, and emotions, reflecting the fundamental characteristic of copyright law in protecting human intellectual creation. However, the operational mechanism of generative artificial intelligence poses a fundamental challenge to this theoretical framework. The content it generates is the result of the combined effects of human-preset algorithms, model architectures, and massive data training, wherein the extent of direct and specific human intellectual input is often difficult to define, and the creative process exhibits a high degree of automation and randomness. This situation has sparked profound theoretical debates: when the creative process is driven by algorithms, can the output still meet the traditional originality requirements of copyright law, which are premised on a direct connection to human intellect? Determining the copyrightability of AI-generated content necessitates not only examining whether its external form of expression meets the standards of originality but also delving into an analysis of whether its generative mechanism contains sufficient human intellectual activity to constitute creativity under copyright law^[1].

1.2 Theories on Rights Attribution of AI-Generated Content under Different Subject Models

Assuming that AI-generated content meets the standards of originality, the issue of rights attribution subsequently arises. Currently, the theoretical community has proposed various rights attribution approaches for different subject models. The first is the "user-oriented approach," which advocates attributing rights to the user who provides specific instructions, adjusts parameters, and ultimately selects the generated content for the AI model, arguing that the user's operational actions constitute a decisive intellectual contribution.

The second is the "developer-oriented approach," which posits that the true creative labor lies in the design of algorithmic models, the selection of training data, and the optimization of model parameters; therefore, copyright should belong to the developer or trainer of the artificial intelligence. Additionally, there exists the "tool theory" perspective, which regards AI as an advanced creative tool, with its generated content akin to photographs taken with a camera, where rights naturally belong to the user operating the tool. The debate among these different approaches essentially revolves around which intellectual contribution in the complex "human-machine collaboration" chain is legally regarded as the decisive "creative act," the conclusion of which will directly impact investment incentives and the distribution of benefits^[2].

1.3 The Challenge of AI-Generated Content to Traditional Author-Centrism

The emergence of generative artificial intelligence has profoundly deconstructed the traditional "author-centrism" copyright theory established on human subjectivity. Classical copyright theory regards the author as the creative core with autonomous consciousness and free will, while the work is considered an extension of the author's personality and thoughts. The creative mechanism of generative AI undermines this foundation. Its content generation is an automated process based on probability statistics and pattern matching, devoid of the emotions, intentions, or personal imprint characteristic of human authors. This creates interpretive difficulties when applying the current legal system, which centers on protecting "authorial rights."

When the creative subject shifts from a conscious individual to a complex computational system, the philosophical foundation of copyright law — namely, incentivizing individual creation to enrich societal culture — requires reexamination. AI-generated content compels legal theory to consider a potential paradigm shift from the "human author" to "functional output," exploring a new intellectual property framework capable of addressing technological realities and transcending traditional notions of authorship, thereby adapting to a future creative ecosystem where human-machine collaboration becomes the norm.

2. Real-world Disputes and Legal Responses in Copyright Definition for Generative Artificial Intelligence

2.1 Divergence in Originality Determination of AI-Generated Content in Judicial Practice

Different jurisdictions exhibit significant jurisprudential divergence regarding the determination of

originality in AI-generated content, reflecting the interpretative challenges faced by traditional copyright theory when confronting emerging technologies. Some judicial rulings adhere to a strict subjective interpretative approach, emphasizing that a work's originality must be rooted in the spiritual creation and expressive intent of a human author. These rulings posit that AI serves merely as a tool for executing human will, and its autonomously generated content lacks copyrightability due to the absence of direct human intellectual investment^[3]. This stance strictly confines copyright protection to the realm of human intellectual activity, maintaining internal consistency within the legal system, yet potentially fails to adequately address new forms of creation brought about by technological development.

In contrast, judicial perspectives adopting a functional assessment demonstrate greater flexibility. This approach shifts the evaluative focus from the creative subject to the objective expressive form of the work itself, examining whether the generated content demonstrates unique arrangement and expression distinguishable from existing works. From this viewpoint, the creative contributions of users during model training, prompt design, and output selection are considered crucial factors in satisfying originality requirements. This judgment criterion prioritizes the external characteristics of creative output over the internal generation mechanism, creating possibilities for protecting AI-generated content, while simultaneously introducing new challenges in accurately defining the extent of human involvement.

Such judicial discrepancies not only create uncertainty in legal application but also pose substantive challenges to intellectual property management in a globalized context. Content creators and users face the risk of conflicting rulings across different legal jurisdictions, increasing legal risks in cross-border transactions and dissemination. The persistence of these divergences underscores the urgent need to reexamine and clarify originality standards in the digital age.

2.2 Legitimacy of Training Data Sources and Copyright Infringement Risks

The data acquisition and utilization during the training process of generative artificial intelligence raise multi-layered copyright law issues. The core controversy centers on the legal characterization of unauthorized copying and use of copyrighted materials for model training. From a technical perspective, the training process necessarily involves temporary reproduction and processing of original data, and whether such reproduction constitutes an infringement of the right holder's exclusive reproduction right has become a primary legal dispute. The applicable boundaries of the fair use doctrine in this field warrant particularly in-depth examination, with determinations requiring comprehensive consideration of factors including the purpose and character of the use, the nature of the copyrighted work, the amount and substantiality of the portion used, and the effect of the use upon the potential market for or value of the work^[4].

The commercial nature of training data and its potential substitution effect on the original work's market further complicate fair use assessments. When generated content demonstrates substantial similarity at the expression level to specific works within the training data, direct infringement disputes may arise. In such circumstances, it becomes essential to carefully analyze whether the generated content merely draws inspiration from the style and ideas of the original work or actually replicates its distinctive expressive elements. Furthermore, the potential inclusion of personal information and privacy rights within training datasets adds another layer of complexity to these issues, necessitating the establishment of more robust data governance and rights clearance mechanisms.

2.3 Ambiguities in Copyright Attribution under Existing Legal Frameworks

Traditional copyright law is premised on the assumption that natural or legal persons are the creative subjects, a theoretical foundation that faces interpretive challenges in the context of AI-generated content. When multiple parties participate in the creative process, the issue of rights attribution presents unprecedented complexity. Model developers claim rights based on their fundamental role in designing algorithms, selecting training data, and optimizing model parameters, arguing that their investments constitute the foundational conditions for creation. End-users emphasize their creative contributions in prompt design, parameter adjustment, and output selection, asserting that these intellectual inputs should be regarded as decisive factors.

In certain complex application scenarios, successive improvements or collective collaborations among multiple users further obscure rights attribution, making traditional principles of joint works or works made for hire difficult to apply directly. This legal uncertainty directly affects the legal

expectations of relevant parties regarding copyright transactions, licensing agreements, and infringement accountability, potentially inhibiting industrial innovation and investment willingness. Establishing clear and predictable rules for rights attribution has become an institutional requirement for promoting the healthy development of the AI-generated content field.

2.4 Exploration of Legal Response Paths for Copyright Disputes in Generative Artificial Intelligence

Confronted with the copyright challenges posed by generative artificial intelligence, the legal community is exploring systematic solutions across multiple dimensions. At the level of legal interpretation, there are efforts to redefine the concept of originality in accordance with technological development, establishing clear standards for determining the critical nature of human contributions within the AI creative process to provide a basis for rights attribution. This approach maintains the stability of the legal system but requires judicial bodies to possess corresponding technical comprehension capabilities.

At the institutional innovation level, considerable attention is being given to the potential introduction of new rights protection models specifically designed for AI-generated content. Approaches such as neighboring rights protection schemes could provide appropriate legal protection for generated content that does not qualify as traditional works, acknowledging its economic value while distinguishing its protection level from traditional copyright. This path requires legislators to carefully balance the interests of all parties and design reasonable protection terms and rights limitations [5].

Technical solutions and contractual arrangements serve as important supplements to legal regulation. The development and application of technological measures, including content provenance tracking and rights marking, can enhance transparency in training data usage and authorization processes. Simultaneously, industries are encouraged to adopt standard contractual terms to clarify the distribution of rights and obligations among developers, platforms, and users, thereby establishing effective private ordering governance. The coordinated development of these diverse approaches will help construct a legal environment that both promotes technological innovation and ensures balanced rights protection.

3. Ethical Risks and Normative Construction in Generative AI Creative Activities

3.1 Lack of Creative Transparency and Challenges in Traceability Mechanisms

The opacity of the generative artificial intelligence creative process constitutes a fundamental ethical challenge. The complexity of its model parameters and generative logic often results in output lacking explainable decision pathways, creating what is termed a technological black box state. This transparency deficit directly leads to difficulties in traceability mechanisms, making it challenging within existing technical frameworks to accurately trace the training data sources or identify specific generative algorithm versions for particular outputs.

When content triggers disputes, this lack of traceability impedes effective verification of content authenticity, originality, and potential biases. The integrity of the information ecosystem consequently faces risks, potentially eroding public trust in the credibility of digital content. Establishing transparency standards and traceability mechanisms compatible with technical characteristics represents a necessary precondition for addressing this dilemma. Current technical explorations include digital watermarking, content fingerprinting, and research on explainability in model training processes. However, these solutions still face significant challenges in practicality, universal applicability, and anti-interference capabilities. From an information ethics perspective, the transparency deficit not only affects individual judgment of information but also undermines the social foundation of knowledge production and dissemination.

3.2 Potential Impact of Algorithmic Bias on Creative Fairness and Cultural Diversity

Inherent social biases within training data can be learned and amplified by generative artificial intelligence models, systematically influencing their output content. These algorithmic biases may cause generated content to reproduce or even reinforce existing stereotypes and unequal structures across dimensions such as gender, ethnicity, and culture. From the perspective of creative fairness, this biased content generation mechanism may marginalize certain groups or perspectives, depriving them of equal representation in digital cultural expression.

At the level of cultural diversity, models trained on large-scale internet data tend to generate content aligned with mainstream narratives or majority-group aesthetic paradigms, potentially suppressing the expression and dissemination of non-mainstream cultures and local knowledge. This poses long-term risks to the diversity and richness of the global cultural ecosystem. The ethical dilemma of algorithmic bias lies in its concealed and systematic operation, creating dual challenges in technical remediation and social awareness. Addressing this issue requires incorporating multicultural perspectives and inclusive design principles from the initial stages of dataset construction, while establishing specialized evaluation metrics during model assessment to monitor fairness in cultural representation.

3.3 Social Accountability Challenges Arising from Ambiguity in Liable Entities

The circulation and application of AI-generated content may produce a range of social impacts, and when these impacts prove negative, identifying liable entities becomes exceptionally complex. Traditional chains of accountability become blurred among multiple parties including developers, operators, and end-users. Developers may argue they merely provide general-purpose tools without foreseeing specific application outcomes, while users may contend they cannot comprehend or control the model's complex internal mechanisms. This uncertainty in responsibility allocation creates a de facto accountability vacuum, making it difficult to initiate effective recourse and remedy mechanisms when content involves misinformation, personality rights infringement, or other societal harms [6].

The ambiguity in determining liable entities not only challenges existing legal attribution principles but also raises urgent ethical questions regarding social governance and protection of individual rights. The complexity of this issue lies in the autonomous and unpredictable nature of generative AI systems, which creates interpretive difficulties for traditional causation models. Potential solutions include establishing risk-based layered responsibility frameworks and exploring collaborative governance models where both development and application stakeholders share social accountability.

3.4 Constructing a Multi-stakeholder Collaborative Ethical Governance Framework for Generative Artificial Intelligence

Addressing the complex ethical risks of generative artificial intelligence requires establishing a collaborative governance framework that transcends single entities or isolated measures. The construction of this framework relies on the co-evolution of technology, industry standards, and social mechanisms. At the technological level, research and development priorities should include explainable AI methodologies, bias detection and mitigation tools, and embedded content provenance identification technologies. At the industry standardization level, it is crucial to promote the establishment of ethical guidelines and self-regulatory standards covering the entire lifecycle from data collection and model design to content generation and deployment.

Broad societal discourse and ethics education can enhance public critical awareness and form social oversight of technological development. This multi-stakeholder collaborative approach aims to create a dynamically adjusting governance ecosystem, ensuring that generative AI creative activities develop healthily along a path that respects human values and safeguards public interests, while providing a solid ethical foundation for relevant copyright definitions and liability allocations. An effective ethical governance framework should possess adaptive characteristics, enabling continuous optimization as technological capabilities and social understanding evolve. International dialogue and cooperation are equally essential to address the global nature of generative AI's ethical challenges and promote culturally sensitive transnational governance consensus.

Conclusion

Generative artificial intelligence poses fundamental challenges to existing institutional systems at both copyright and ethical levels. In the copyright domain, it is necessary to enhance the adaptability and interpretive power of the legal system by redefining originality, exploring new rights protection models, and improving rights distribution rules. In the ethical domain, efforts should focus on constructing a multi-dimensional governance system integrating technological transparency, algorithmic fairness, clear accountability, and cultural inclusiveness to prevent risks of technological misuse and social division. Future research should further address the legal coordination of generative AI in global governance, pathways to achieving technological transparency and explainability, and establishing dynamic equilibrium mechanisms that balance innovation incentives with rights protection,

thereby providing continuous theoretical support and practical guidance for humanity's institutional transition into the intelligent era.

References

- [1] Hou Kun, and Sun Yubin. "Ethical Reflection and Risk Mitigation of Generative Artificial Intelligence from an Ecological Perspective." *Journal of China University of Geosciences (Social Sciences Edition)* 1-9.
- [2] Wan Ruihong. "Academic Ethics Issues and Preventive Measures in Universities under the Background of Generative Artificial Intelligence." *Education Exploration* .10(2025):7-13.
- [3] Han Xiaoyu, and Meng Saifu. "Research on Copyright Issues of AI-Generated Content (AIGC)." *Hebei Vocational Education of Law* 3.09(2025):117-122.
- [4] Cui Ting, and Liu Zhiyi. "The Impact and Challenges of AI-Generated Content (AIGC) Technology on the Intellectual Property System." *Proceedings of the 18th Library Management and Service Innovation Forum 2025*. Ed. Beijing Union University;, 2025, 9-10.
- [5] Lin Jinhui. "On the Copyrightability of Strong AI-Generated Content under Creative Alienation." *Journal of Dalian University of Technology (Social Sciences)* 1-8.
- [6] Zhu Surong. "From Prompt to Work: Exploring the Origin of AIGC Works' Copyright and Legal Reconstruction." *Science and Society* 15.02(2025):116-133.