This paper analyzes the copyright attributes of artificial intelligence generated content in education and teaching

Xiaoli Wang*

Yibin University, Yibin, 644007, China
**Corresponding author: 18780691198@163.com*

Abstract: *With the extensive involvement of artificial intelligence-generated content in education and teaching activities, the determination of its copyright attributes faces theoretical disputes and application difficulties. This paper analyzes the issue from three dimensions: the criteria for determining originality, the logic of ownership attribution, and the boundaries of fair use. At the level of originality, this paper proposes shifting the reference framework from general works to the set of possible expressions under the same type of teaching task. At the level of ownership attribution, this paper advocates allocating rights between users and developers based on the standard of "substantive creative control" and constructs a joint copyright structure under the collaborative generation model. At the level of fair use, this paper clarifies the exemption conditions for reproduction acts, citation norms, and the interest balancing mechanism for non-commercial dissemination. The research aims to provide an analytical framework for the copyright determination of artificial intelligence-generated content in education and teaching scenarios.*

Keywords: *artificial intelligence-generated content; education and teaching; copyright attributes; originality criteria; ownership attribution; fair use*

Introduction

Artificial intelligence-generated content has become increasingly common in the field of education and teaching. Its functions, ranging from exercise generation and lesson plan design to the automatic production of classroom presentation materials, have been deeply embedded into the teaching process. However, the current copyright system is primarily constructed around the creative acts of natural persons. It lacks clear institutional responses to questions such as whether algorithm-driven content output possesses the attributes of a work, whether the rights should belong to the developer or the user, and to what extent the reproduction and dissemination in teaching activities can be exempted from liability. This theoretical gap not only affects the legitimate use of generated content by educational subjects but may also inhibit the positive functions of artificial intelligence technology in teaching innovation. Therefore, it is necessary to start from the basic principles of copyright law and systematically analyze the originality determination, ownership allocation, and fair use boundaries of artificial intelligence-generated content in education and teaching scenarios, so as to provide theoretical support for resolving relevant disputes.

1. Criteria for Determining the Originality of Artificial Intelligence-Generated Content

1.1 Definition of the Creative Act for Generated Content in Education and Teaching Scenarios

Content generation in education and teaching activities often serves the goals of knowledge transmission and cognitive construction. When artificial intelligence intervenes in this process, whether the text, exercises, or teaching plans it outputs constitute a "creation" in the sense of copyright law depends on the existence of intellectually creative decision-making nodes in the generation process. Under the typical teaching instruction input mode, the user provides subject keywords, structural requirements, or difficulty parameters, and the artificial intelligence completes the content arrangement based on its training data. If such arrangement merely reflects the algorithm's reorganization of existing patterns, it is difficult to satisfy the minimum requirement of individual expression for a creative act. Conversely, when the user filters, revises, or reorders the generated content through multiple rounds of

interaction and injects selective judgments at the level of teaching logic, such intervention can be regarded as an extension of the creative act.

The core of defining the creative act lies in distinguishing between "triggering generation" and "dominating expression." In education and teaching scenarios, teachers or learners often issue binding instructions based on specific teaching objectives. These instructions define the direction of the generated content but do not directly determine the language style, selection of examples, or interpretive pathways of the content. If the artificial intelligence demonstrates a diversity of expression within its parameter space that exceeds the expectations of the instructions, and such diversity cannot be attributed to the user's specific choices, then the expressive form of the generated content should be attributed to the algorithm's autonomous output. In this case, the user's role is closer to a "tool user" rather than an "author." Only when the user imposes modifications or reorganizations that involve original judgments on the intermediate generation results can the final output meet the definitional criteria of a creative act[1].

1.2 Dimensions for Distinguishing Between Algorithm-Driven Generation and Human Intellectual Input

The basic characteristic of algorithm-driven generation lies in the repeatability and statistical correlation of its output results. The same input instructions, when executed multiple times under the same model parameters, tend to produce highly similar or even identical sequences of content. This characteristic forms an essential difference from the non-replicability, intuitive leaps, and experiential biases that accompany human intellectual input. From the technical logic of copyright attribution, the key dimension for distinguishing between the two lies in the "source of expressive choices." If every replaceable word, structure, or example in the content has multiple equally possible alternatives under the internal probability distribution of the algorithm, and if the algorithm does not perform a non-deterministic selection based on external aesthetic or logical criteria, then this selection process lacks the intentionality unique to human intellectual activity.

Furthermore, human intellectual input typically manifests as the active shaping of expressive forms, including the selection of rhetorical strategies, the arrangement of argumentative sequences, and the construction of conceptual hierarchies. In the generation of education and teaching content, a teacher may ask an artificial intelligence system to generate three different explanations of a certain knowledge point, then select the version that best matches the class's cognitive level and adjust its word order and examples. In this process, the algorithm completes the parallel generation of multiple draft versions, which constitutes a computational act at the level of data processing; the user's version selection and word order adjustment, on the other hand, constitute expressive decisions at the level of intellectual input. A clear dimensional boundary exists between the two: the algorithm contributes generative potential and permutations, while the human contributes value judgments and the finalization of expression[2].

1.3 Applicable Difficulties of the Objective Originality Standard and Their Adjustment

The objective originality standard in copyright law requires that a work manifests individualized expression that deviates from the routine, and that such expression cannot be reduced to rules or templates. However, artificial intelligence-generated content in education and teaching scenarios often serves the purpose of standardized knowledge transmission, and its output naturally tends toward expressive forms with high reusability and low variability. For example, when generating a sample problem for solving a quadratic equation, most models follow the narrative structure of "standard form - discriminant calculation - quadratic formula substitution." Although this structure satisfies the functional requirements of teaching, it is difficult to recognize it as possessing objective originality due to its high similarity to the arrangement in similar textbooks. Applying the traditional originality standard may result in a large amount of generated content that has practical teaching value but routine expression being excluded from copyright protection.

To resolve this difficulty, it is necessary to adjust the reference framework of the originality standard. In the field of education and teaching, originality should not be understood as an absolute departure from routine expression, but rather as the production of "non-unique reasonable expression" within the constraints of specific functions. If, for the same teaching objective and the same input conditions, different users can obtain generation results with significantly different expressive structures through differentiated instructions or interactive strategies, then this generation process

contains room for originality. In this case, the object of evaluation should not only be the degree of difference between the final content and existing templates, but also the expressive guidance imposed by the user during the generation process that goes beyond the minimum necessary range. By shifting the reference framework from "general works" to "the set of possible expressions under the same type of teaching task," the originality standard can maintain operability and moderate inclusiveness in the teaching application of artificial intelligence-generated content.

2. The Logic of Ownership Attribution for Artificial Intelligence-Generated Content in Education and Teaching

2.1 Rules for the Initial Allocation of Rights Between Users and Developers

Regarding the issue of copyright ownership of artificial intelligence-generated content, the core dispute over the initial allocation of rights lies in the "source of creative contribution." The developer, by designing the algorithm architecture, constructing the training dataset, and optimizing model parameters, lays the technical foundation for content generation. However, the developer's contribution stops at constructing a system capable of producing expression, rather than directly generating a specific work. In education and teaching scenarios, the user inputs teaching task instructions and proposes functional or stylistic requirements for the output content. The user's behavior is closer to that of a commissioning party in a commissioned creation. If the copyright is attributed solely to the developer, the user's actual control over the expressive form of the generation result is ignored. If the copyright is attributed solely to the user, the creative potential embedded in the algorithm's internal expressive choices is underestimated[3].

The existing copyright system generally takes the "natural person who creates the work" as the basic rule for the initial ownership of rights. For artificial intelligence-generated content, the standard of "substantive creative control" may be introduced for consideration. When the user sets elements in the instructions that are sufficient to influence the specific form of expression, such as requesting the generation of a teaching explanation containing a specific analogy, personified phrasing, or non-structured argumentation logic, and the artificial intelligence output faithfully reflects such settings, the user should be regarded as the actual maker of the expressive choices. Conversely, if the user only makes a broad thematic request, and the artificial intelligence exhibits autonomous variation at the expressive level that is unrelated to the instructions, the creativity in such variation should be attributed to the algorithmic mechanism established by the developer. This allocation rule emphasizes the assignment of rights to the party who can anticipate and guide the details of expression, rather than solely based on the physical operator.

2.2 The Impact of Teaching Instruction Input on the Copyright Ownership of Artificial Intelligence-Generated Content

The technical characteristic of teaching instruction input lies in its inclusion of both functional constraints and expressive guidance. From the perspective of copyright ownership, the degree of refinement of the instructions directly determines the user's level of control over the generated content. When the instruction is merely "generate an example question about Newton's first law of motion," the artificial intelligence system must independently determine the numerical values, object types, narrative order, and language style. In this case, most of the choices at the expressive level are determined by the algorithm's probability distribution, and the user can hardly claim a creative contribution to the specific combination of words and sentences in the example question. Conversely, if the instruction explicitly requires "using the scenario of a soccer ball continuing to roll after being kicked, with an initial velocity of 10 meters per second, a friction coefficient of 0.2, and narrated in the first person," the user has in effect prefabricated the core expressive elements of the content, and the artificial intelligence only performs the executive function of linguistic filling and logical coherence.

Furthermore, incremental instructions in multi-round interactions have a cumulative effect on ownership attribution. The user may first obtain a draft and then modify the argumentative structure, replace examples, or adjust the tone through feedback instructions. Each modification instruction that has an expressive orienting function injects new intellectual choices into the generated content. When the expressive part contributed by the original algorithm and the part modified by the user cannot be clearly separated, the entire generated content may be considered a collaborative result dominated by the user. It should be noted that if the teaching instruction input itself possesses originality, for example,

forming a complete instructional design framework or a unique sequence of questions, such instruction may be protected as an independent literary work. However, the ownership attribution between the instruction and the output content should still be determined separately[4].

2.3 The Construction of Joint Copyright Under the Collaborative Generation Mode

In education and teaching, there are situations where multiple subjects collaborate in using artificial intelligence to generate content. Examples include a teaching and research group jointly designing a sequence of instructions, a teacher and a student interactively adjusting the generation results, or different users successively iteratively modifying the same draft. Under such a collaborative generation mode, if the intellectual inputs of each participant make inseparable contributions to the original expression of the final content, a joint copyright relationship may be formed. The difference between collaborative generation and simple relay generation lies in the fact that the instruction inputs of each subject in the collaborative process are not linearly superimposed but are fused within the same unit of expression (such as a paragraph or a set of teaching steps). For example, Teacher A sets the analogy framework for concept explanation, Teacher B supplements three specific analogy instances within that framework, and the artificial intelligence generates a complete paragraph accordingly.

For the construction of such joint copyright, two elements need to be clarified: "the intention of joint creation" and "the identifiability of contributions." In education and teaching scenarios, the intention of joint creation can be proven through operation records on a collaboration platform or a clear division of labor agreement. The identifiability of contributions faces challenges: the intermediary nature of artificial intelligence generation often embeds the expressive elements of each participant into the final text in an inseparable manner. In response to this, the "joint ownership" model may be adopted, which means that each co-owner holds an indivisible right to the entire generated content. However, when exercising the rights externally, the co-owners need to reach a consensus, and the internal distribution of benefits shall be determined based on the relative proportion of each subject's contribution to the originality of expression. To prevent the rigidity of use caused by joint ownership, a presumptive license mechanism may be established within the scope of non-commercial teaching activities, allowing any co-owner to freely use the content in teaching activities. Only when the content is to be published externally or used commercially, the consent of all co-owners shall be obtained.

3. Fair Use Boundaries of Artificial Intelligence-Generated Content in Education and Teaching

3.1 Exemption Conditions for the Reproduction of Generated Content for Teaching Purposes

The reproduction of artificial intelligence-generated content in teaching activities includes behaviors such as classroom display, distribution of learning materials, and uploading to teaching platforms. If such reproduction behaviors are not licensed by the copyright holder, the possibility of exemption depends on whether they fall within the scope of fair use exemption. The core elements for determination lie in the "transformative purpose" of the reproduction behavior and its substitution effect on the original work market. When the reproduction behavior serves non-display functions such as in-class teaching explanation, knowledge consolidation, or student homework reference, its purpose deviates from the original functional setting of the artificial intelligence-generated content. The generated content originally exists as an independent output, while teaching reproduction transforms it into a teaching medium. Such transformation of purpose can reduce the infringement risk of the reproduction behavior.

Furthermore, the establishment of the exemption conditions also requires that the scope of reproduction be commensurate with the teaching purpose. In education and teaching scenarios, reproduction should not exceed the length or quantity necessary to achieve a specific teaching objective. For example, reproducing a paragraph of an artificial intelligence-generated concept definition for classroom explanation, as opposed to reproducing an entire long text containing multiple independent sections for the same explanation, makes the former more likely to satisfy the exemption conditions. The frequency and scope of dissemination also serve as influencing factors. Internal reproduction for a single class, compared to public reproduction on a platform accessible to the entire school, has a smaller impact on the potential market of the copyright holder, and therefore is more likely to meet the exemption standard. It should be emphasized that the establishment of an exemption does not require that the reproduction behavior completely preclude any loss of revenue to the copyright holder; rather, it requires that such loss be within a tolerable limit when weighed against the public interest in

teaching.

3.2 Citation Norms and Limits for Using Generated Content as Teaching Materials

Citing artificial intelligence-generated content as teaching material involves the selection and embedding of original generated text, images, or structured data. The normative requirements for citation are reflected in three aspects: the teaching relevance of the citation purpose, the necessity of the citation length, and the obligation to identify the citation source. Teaching relevance requires that the cited generated content have a direct demonstrative or exemplary relationship with the subject of the lecture, rather than merely filling space. For example, citing a typical output paragraph generated by an artificial intelligence system when explaining the logic of algorithm generation constitutes teaching relevance; whereas citing an artificial intelligence-generated poem unrelated to the subject in a literary appreciation course exceeds the relevant scope[5].

The criterion for determining the citation limit lies in the distinction between "core expression and non-core expression." If the citation only involves factual statements, routine examples, or data lists in the generated content, such elements lack uniqueness in expressive form, and even if the citation length is large, its impact on the original copyright holder's market exclusivity is relatively limited. Conversely, if the citation involves the most personalized rhetorical structures, metaphorical systems, or argumentative paths in the generated content, even a small length may constitute the appropriation of a substantial part of the original work. In education and teaching scenarios, the citation limit may refer to the teaching substitution test: whether the cited material is sufficient to make students no longer need to access the original generated content. If the citation provides the most original expressive core of the original content, causing a significant decrease in the necessity for students to obtain the complete original content, then such citation exceeds the reasonable limit.

3.3 Non-Commercial Teaching Dissemination and the Interest Balancing Mechanism with the Copyright Holder

Non-commercial teaching dissemination refers to the diffusion of artificial intelligence-generated content through teaching network platforms, school resource libraries, or teaching and research exchange channels without direct economic return. The essential difference between such dissemination behavior and commercial use lies in the fact that the disseminator's motivation does not involve obtaining monetary benefits from reproduction or distribution, and the scope of dissemination is usually limited to within the educational community. However, non-commercial nature does not automatically equate to fair use. When the large-scale implementation of dissemination behavior causes the generated content to substitute for the licensing opportunities of the original copyright holder in the teaching authorization market, even if the disseminator does not directly profit, the copyright holder's potential loss of revenue may still reach a level requiring legal intervention.

The core approach to constructing an interest balancing mechanism is to introduce a dual framework of "cumulative effect control" and "reserved authorization areas." Cumulative effect control requires an assessment of the total number of disseminations and the population coverage of the same generated content in similar teaching scenarios. If a certain content is independently disseminated by hundreds of teachers in different schools, its combined effect may substantially erode the original work market. In this case, even if each individual dissemination meets the non-commercial characteristic, the overall dissemination behavior should still trigger the licensing obligation. The reserved authorization areas delineate the forms of use over which the copyright holder retains absolute control, including the incorporation of generated content into fee-based teaching materials, its use for secondary training of generative artificial intelligence, and its removal from the teaching context into the realm of popular cultural consumption. For non-commercial teaching dissemination outside such areas, a statutory licensing or collective management mechanism may be established, allowing users to obtain dissemination authorization at a low cost, thereby achieving a balance between teaching convenience and the copyright holder's revenue.

Conclusion

The copyright attributes of artificial intelligence-generated content in education and teaching cannot be simply determined by applying traditional rules for the identification of works. Instead, adaptive adjustments should be made to the originality standard, the logic of ownership attribution, and the

boundaries of fair use according to the specific characteristics of teaching scenarios. At the level of originality, the requirement of absolute novelty should be replaced by the concept of "non-unique reasonable expression," and the user's expressive guidance should be incorporated into the scope of evaluation. At the level of ownership attribution, the contributions of users and developers should be distinguished based on the degree of refinement of the instructions, and a joint copyright structure along with an internal presumptive license mechanism should be introduced for multi-party collaborative generation. At the level of fair use, the exemption for reproduction acts should satisfy the requirements of transformative purpose and scope limitation; citation should be limited by the teaching substitution test; and non-commercial dissemination should achieve interest balancing through cumulative effect control and reserved authorization areas. Future research directions may include the differentiated needs of various types of educational institutions (such as basic education and higher education) regarding the copyright attributes of generated content, as well as the impact of generative artificial intelligence technology iteration on the criteria for determining the "source of expressive choices."

References

[1] Zhao, Z. Q., and Yu, Y. F. "Generative Artificial Intelligence Empowering Vocational Education and Teaching: Theories and Cases." Chinese Vocational and Technical Education, no. 02 (2026): 86-93.
[2] Jia, N., and Zhang, T. M. "Research on the Risk Evaluation Model of 'Private Education-Generative Artificial Intelligence Content Application'." Big Data Age, no. 12 (2025): 30-40.
[3] Chen, X. Research on Legal Governance of Risks in Educational Application of Generative Artificial Intelligence. 2025. Southwest University, MA thesis.
[4] Tong, F., et al. "Generative Artificial Intelligence Driving the Paradigm Transformation of Adult Education and Teaching: Mechanisms, Practical Paths, and Risk Strategies." Chinese Adult Education, no. 22 (2025): 40-48.
[5] Bai, J. "Mechanisms, Risks, and Responses of Generative Artificial Intelligence Empowering the Transformation of Foreign Language Education and Teaching." Chinese University Teaching, no. 11 (2025): 83-89.