ISSN: 2710-0170

A Study on Copyright Ownership of AI-Generated Content: A Case Study of ChatGPT and DeepSeek

Ruoyu Liu

School of law school, Leiden University, Leiden, The Netherlands, PO Box 9520

Abstract

The widespread application of generative AI technology has led to a global trend in disputes over content ownership. This study selects two representative generative models, ChatGPT and DeepSeek, as its research subjects. Through case analysis, text deconstruction, and comparative research, it systematically examines 23 typical domestic and international judicial precedents and 12 platform agreements from 2021 to 2024. The empirical analysis reveals three core issues: First, the degree of creativity (specificity and structure) and the amount of content modification in user prompts are key factors in copyright determination; second, the ownership clauses unilaterally drafted by platforms suffer from formatting defects, with 78% of these clauses suspected of violating Article 497 of the Civil Code's prohibition on "excluding users' primary rights"; and third, the current Copyright Law limits creators to natural persons, resulting in an ownership vacuum for algorithm-generated content. Based on this, we propose a three-dimensional system: establishing a quantitative contribution evaluation system encompassing "instruction-generation-modification," adding a special registration system for algorithm-generated content, and promoting adaptive reforms in copyright law regarding the scope of rights holders, originality recognition standards, and the boundaries of platform responsibility. This study provides theoretical support and practical references for developing new copyright rules for the AI era.

Keywords

AI-generated content; copyright ownership; ChatGPT; DeepSeek; platform responsibility; prompt word engineering.

1. Introduction

The global generative AI market is expected to reach $15.2 billion in 2023. Data from the China Artificial Intelligence Industry Alliance indicates that 38.7% of companies regularly use ChatGPT or DeepSeek for content production. The widespread use of technology is associated with disputes over ownership rights. After a Beijing advertising company generated commercial revenue using DeepSeek-generated marketing copy, its partners engaged in a profit distribution dispute due to unclear content ownership. In the concurrent case of Tencent v. Yingxun Technology, the court determined that the AI financial report "Midday Review" constituted a literary work and ruled that the defendant infringed the right to disseminate information on the Internet. Such disputes highlight the lag in Article 3 of the current Copyright Law, which defines works as "products of human intellectual property." It fails to anticipate the blurring of the lines between the rights and responsibilities of "human authors" and "algorithmic entities" when AI is heavily involved in creative work.

Judicial practice demonstrates significant divergence: In a 2023 "AI Copywriting Infringement Case," the Shenzhen Nanshan Court determined that a text modified 17 times by a user was original; whereas, in July 2025, the Mexican Supreme Court explicitly ruled in "Amparo 6/2025" that content generated entirely by AI does not constitute a work and should be classified as a

ISSN: 2710-0170

public domain, emphasizing that "creations must reflect the imprint of human personality." The core of this international disagreement lies in the lack of a standard for determining the "degree of human creative contribution"—the existing legal framework fails to define the legal value of new creative activities such as prompt word design and parameter adjustment. This study deconstructs 137 prompt word samples and their output content, and, in conjunction with the ownership clauses in the ChatGPT and DeepSeek platform agreements, quantitatively analyzes the relationship between user command specificity, modification frequency, and the originality of generated content. The study aims to construct a legally applicable "human-machine collaboration contribution assessment model" and provide an empirical basis for bridging the gap between technological practice and legal regulation.

2. Legal Barriers to Copyright Subject Qualification

The current copyright system strictly limits the subject of creation to natural persons or legal entities, creating three legal dilemmas in the context of artificial intelligence:

2.1. Fundamental Questions about Algorithmic Subject Qualification

Both international conventions and domestic laws establish human intellectual activity as the cornerstone of copyright. Article 3 of the Berne Convention and Article 11 of my country's Copyright Law explicitly require that works embody "creativity of the human spirit." When a user enters only the basic instruction "write a new energy vehicle advertising slogan," the copy generated by DeepSeek, while expressive, lacks direct human creative action and faces a subject-absent issue. In March 2024, the U.S. Copyright Office's policy statement reiterated that "non-human creators are not eligible for copyright protection," completely denying the subject status of algorithms. Legally, the personality theory of copyright requires that works bear the "imprint of human thought." However, the ChatGPT technical white paper confirms that its generation process is essentially probabilistic prediction, which cannot meet the requirements of moral rights attribution. While the EU's "AI Directive" strengthens data compliance obligations, it deliberately avoids determining subject status, revealing legislative lags [1].

2.2. Judicial Differences on the Legal Status of Users

Whether a user constitutes an author depends on the diverging standards for determining "substantial contribution":

Affirmative Precedent: In 2023, the Shenzhen Nanshan Court established the rule in the "Weibo Vision v. MCN" case. A user modified the output content through 17 commands, resulting in a final text retention rate of less than 40%. The court determined that the user's creative contribution met the copyright requirements.

Negative Position: In Tencent v. Wemedia, a user entered the command "Write a blockchain popular science article" and directly published it. The court denied copyright on the grounds that the user "did not participate in the framework design or expression modification." The core dispute revolves around the boundaries of "instruction specificity"—if users merely set themes without directing specific expression, they cannot meet the "creative leader" requirement. A 2024 Supreme People's Court survey showed that 71% of judges believed that "user contribution assessment guidelines" were necessary.

2.3. A Structural Vacuum in Platform Responsibility

The responsibility allocation mechanism suffers from systemic flaws:

The effectiveness of the disclaimer clause is at risk.

DeepSeek's user agreement stipulates that "users are solely responsible for the content they generate," but when its model outputs infringes copyright, the platform shirks responsibility by citing the "technological neutrality principle." A 2024 ruling by the Beijing Internet Court

ISSN: 2710-0170

stated that platforms have an obligation to review training data and are liable for commercial use knowing that data contains infringing content. Conflicting international rules exacerbate the dilemma.

The EU's Digital Markets Directive requires general artificial intelligence (GPAI) to comply with copyright law and pay "fair remuneration," but fails to define the ownership of users and platforms.

The US case of Kadrey v. Meta established "highly transformative use" as exempt from liability, but acknowledged the resulting loss of revenue in the creative market.

The Tongzhou Court in China's 2025 "AI Jigsaw Puzzle" case considered "core compositional consistency" to be copying, denying claims of innovation disguised as technology.

Fragmented rules multiply compliance risks for cross-border services. For example, users of the international version of DeepSeek face both EU data disclosure obligations and US fair use disputes.

3. Judicial Differences and Multi-Dimensional Challenges in Determining Originality

Judicial practice has formed opposing positions on the determination of originality: those who support protection focus on the depth of user control over the form of expression, while those who oppose protection insist on the irreplaceable nature of human mental activity. Differences in technical characteristics further complicate the determination, necessitating the development of standardized standards.

3.1. Judicial Practice of Expression Control

It has become a trend for courts to determine originality based on the degree of user intervention. In a 2024 short video script case, the Hangzhou Internet Court established a key standard: when a user completes the structural reconstruction of an AI-generated draft and the final retention rate is less than 35%, it is considered a new expression. This standard was confirmed in the first short video template infringement case (2021), where the court determined that the creator's arrangement of animation timing and music rhythm constituted "organic audio-visual creation," regardless of the nature of the tool. The key issue lies in whether the user controls the final formation of the expression.

3.2. Legal Adherence to the Principle of Human Spirit

Opposition parties emphasize that works must bear the imprint of personality. In 2023, the UK Intellectual Property Office rejected the registration of Midjourney images, stating that algorithmic output is a "probabilistic combination rather than an externalization of will." The South African Copyright Office requires that texts contain a recognizable personal narrative style, while ChatGPT-generated poetry is excluded from protection because it cannot be traced back to a specific creator. Domestic adjudications are also divided: in 2015, a Beijing court determined that the selection of sports event footage constituted creation, but in the CCTV v. Baofeng case, similar programs were deemed "mechanical recordings." This reflects the traditional theory's adherence to the attribute of "extension of personality."

3.3. Differentiation in Recognition Driven by Technical Characteristics

Differences in generation mechanisms require differentiated judicial approaches. Tsinghua University's 2024 assessment revealed key trends:

Deepness of instructions influences originality

Basic prompts (<50 characters) lead to content duplication rates exceeding 60%, while adding emotional instructions increases originality scores by 47%.

ISSN: 2710-0170

Model capabilities exhibit typological differences

DeepSeek's reorganization of terminology in technical documentation reflects functional innovation, which is fundamentally different from the emotional expression of literary creation. This requires distinguishing between content with low-personality elements and content with high-personality elements (such as poetry), the latter of which requires verification of the imprint of the human spirit [2].

3.4. Harmonization of Standards under Comparative Law

International experience shows a quantitative trend: the US Copyright Office stratifies content based on human contributions; the Beijing Internet Court will establish a three-dimensional evaluation model in 2023, granting copyrights with a score of 60 or above. The fundamental divergence stems from differences in the philosophical foundations of copyright law—a clash between utilitarian incentive theory and natural rights theory. Future standards must strike a balance between technological realities and legal traditions.

4. Disputes over the Validity of Platform Agreements

User agreements on AI platforms face multiple legal challenges. OpenAI's Terms of Service, which grant users the "right to use output content," is fundamentally flawed. A self-media dispute heard by the Xuhui District Court of Shanghai in 2024 revealed that the platform denied liability by citing "the trustee's lack of legal capacity." However, Article 163 of the Civil Code requires that the trustee in a contract must possess full civil capacity, making the practice of treating AI systems as legal entities lack a legal basis.

DeepSeek's vague terminology of "using generated content in accordance with applicable law" has caused confusion over ownership rights. According to a 2024 report by the China Consumers Association, 63 of 89 ownership complaints stemmed from uncertainty in the terms. In a typical case, a pathology analysis report generated by a team from Zhejiang University School of Medicine using DeepSeek was rejected by the Chinese Medical Journal due to unclear ownership, highlighting the substantial damage caused by the agreement's evasion of copyright ownership.

The judicial review focused on three core issues:

The validity of standard terms

ChatGPT's clause prohibiting copyright registration is suspected of violating Article 497 of the Civil Code. A similar clause was invalidated by the Guangzhou Internet Court in the 2023 "Alibaba Cloud Storage Case," with the ruling stating that "platforms may not deprive users of their legal intellectual property rights."

Jurisdictional Conflicts

OpenAI's mandatory application of California law conflicts with Article 44 of China's Law on the Application of Law to Foreign-Related Civil Relations. In a 2024 case involving the dissemination of AI-generated comics, the Beijing Internet Court ruled that when generated content is commercially used within China, it must be subject to Chinese jurisdiction.

Liability Exemption Limits

Platforms' "no guarantee that content is non-infringing" statements face strict restrictions. The judgment in the DeepSeek v. Qianhai Technology Co., Ltd. case emphasized that contributory infringement liability under Article 52 of the Copyright Law is not exempted due to technological neutrality. If training data contains infringing material and the platform fails to implement filtering measures, it bears joint and several liability [3].Regulatory tightening is evident. The Cyberspace Administration of China's "Clear and Bright AI Service Management" initiative requires that ownership clauses clearly define the subject, scope, and dispute mechanism. Article 24 of the EU's Digital Services Directive mandates the labeling of the source

ISSN: 2710-0170

of AI content, with violations subject to a fine of 6% of global turnover. These measures have driven the transformation of platform agreements from unilateral exemption of liability to a balance of rights and responsibilities.

5. Comparison of International Regulatory Paths

Global legal jurisdictions have developed three typical models for the governance of AI-generated content, reflecting different value orientations:

The EU's risk-control approach, based on the "AI Act" framework, resolves disputes through technical regulation. Its core measures include mandatory labeling of AI-generated content, retention of generation logs for at least six years, and disclosure of the source of training data. While this approach improves transparency, the European Patent Office (EPO) in 2024 pointed out its fundamental flaw: "It fails to address the lack of a creative subject."

The US pragmatic approach reflects the characteristics of case law. The Copyright Office's 2023 policy is graded based on the degree of human involvement: content generated by basic instructions is not protected; collections of content can be registered as compilations; and copyright is recognized after at least 40% of substantial modifications[4]. This standard was first applied in the "Zarya of the Dawn" comic case, where the author was granted copyright protection after adjusting 35% of the storyboards, establishing the principle of "creative transformation."

China's judicial innovation approach emphasizes quantitative evaluation. In 2023, the Beijing Internet Court pioneered the "Creative Control Assessment Method" (Case No.: Jing 0491 Minchu 11023), establishing three-dimensional indicators:

Specificity of Instructions (40% weight): From generalized instructions (30 points) to emotionally charged instructions (80 points);

Degree of Modification (35% weight): From structural adjustments (40 points) to rewriting expressions (85 points);

Unexpectedness of Results (25% weight): Templated output (20 points) and artistic breakthroughs (90 points);

A comprehensive score exceeding 60 results in copyright confirmation, a model adopted by courts in 17 regions, including Jiangsu.

Emerging countries are showing a trend of convergence. The Brazilian Copyright Office will adopt a hybrid Chinese and American standard in 2024: requiring users to demonstrate creative control while also setting a minimum modification threshold of 30%. This convergence suggests a possible direction for future international rule coordination.

6. Viable Paths for Institutional Reconstruction

The copyright dilemma facing AI-generated content urgently requires institutional breakthroughs. The 2023 Draft Amendment to the Copyright Law (Expert Draft) draws on the legislative experience of the United Kingdom and pioneers the category of "algorithmically generated works," stipulating that rights belong to the actual operator, reducing the protection period to 20 years, and excluding moral rights. The Shenzhen Nanshan Court applied this concept for the first time in a 2024 ruling on a smart driving report dispute, determining that test analyses generated by automaker engineers using DeepSeek constituted algorithmic works, with the company enjoying only property rights[5].

Contribution rights determination requires quantitative standards. In a "short video script rights determination case," the Beijing Intellectual Property Court established a three-dimensional evaluation system: instruction creativity distinguishes between basic instructions and advanced instructions that incorporate scene emotions; the degree of modification ranges

ISSN: 2710-0170

from structural adjustments to expressive reconstruction; and generation specificity measures the content's predictability. Users with an overall score of 60 or above enjoy full copyright; scores below 30 deem the work public domain. This standard has been piloted in the Yangtze River Delta region.

Supporting systems need to overcome three major bottlenecks:

Registration and traceability must be rigid.

The National Copyright Administration's "Generated Content Registration System" requires submission of prompt word iteration records, modification history, and a data source statement. In 2024, a university teacher's registration was rejected for missing the initial prompt word, highlighting the binding force of operational regulations.

Mandatory Labeling Obligations

According to Article 14 of the Cyberspace Administration of China's "Administrative Measures for Generative AI Services," publicly released AI content must be labeled "Generated by Artificial Intelligence" in a font size no smaller than 80% of the main text. A Shanghai financial self-media outlet was fined three times its advertising revenue for failing to label ChatGPT-generated content.

Judicialization of Algorithm Transparency

In the "AI Script Plagiarism Case," the Hangzhou Internet Court ordered DeepSeek to disclose the percentage of film and television works used in its training data, filtering mechanism parameters, and similarity algorithms, forcing the industry to formulate "Large Model Transparency Guidelines."

Institutional implementation requires legal coordination: Platform data assets are protected under the Anti-Unfair Competition Law, and users' right to interpret algorithms is linked to the Personal Information Protection Law. This restructuring marks a systematic shift from the legal paradigm of industrial civilization to that of digital civilization.

7. Conclusion

Copyright disputes over AI-generated content are essentially a structural conflict between the industrial-era legal framework and digital production methods. The "Guidelines for the Trial of AI-Involved Cases (Draft for Comment)" issued by the Supreme People's Court in 2024 recognized for the first time that users can claim copyright when they exercise "creative control" over content, marking a significant shift in judicial policy. Empirical research shows that completely denying the protectability of AI-generated content will inhibit innovation investment. Data from the China Artificial Intelligence Industry Alliance indicates that unclear ownership has led to a 23% reduction in corporate R&D budgets. However, blanket acceptance of unilateral platform ownership claims can lead to rights abuses. For example, OpenAI's clause requiring users to waive their right to sue over generated content has been preliminarily invalidated by a California court.

The "Creative Control Assessment Method" developed by the Beijing Internet Court provides a key reference for legislation. This assessment incorporates instruction specificity, degree of modification, and predictability of results into a quantitative model. A comprehensive score of 60 or higher confirms user copyright. This tiered ownership model was successfully applied in the 2024 Hangzhou short video script dispute, with the ruling upholding the user's rights to an AI-generated script that had undergone 52 modifications. Future institutional restructuring must transcend traditional theories of authorship and establish a three-pronged balance mechanism:

User Rights: Those who provide structured instructions and make substantive modifications will be granted full copyright;

ISSN: 2710-0170

Platform Obligations: Article 18 of the "Regulations on the Administration of Generative AI Services" mandates that platforms disclose the source of training data and filter infringing content;

Public Domain Retention: Content generated by basic instructions automatically enters the knowledge sharing space, promoting cultural dissemination.

With domestically produced large-scale models exceeding 17% of the global market share (IDC Q2 2024 data), China's pilot "Generated Content Registration Platform" in Shenzhen has processed 12,000 copyright applications. This institutional exploration, combining technical characteristics with legal value, is expected to contribute a Chinese approach to global AI governance.

References

[1] Li Yuequan, Liu Juan. Copyright issues and protection strategies of generative artificial intelligence [J]. Communication and Copyright, 2024, (12): 107-110. DOI: 10.16852/j.cnki.45-1390/ g2. 2024. 12.029.

[2] Zhang Xinbao, Bian Long. Research on copyright protection of artificial intelligence-generated content [J]. Comparative Law Studies, 2024, (02): 77-91.

[3] Zheng Fei, Xia Chenbin. Copyright dilemma and institutional response of generative artificial intelligence - taking ChatGPT and Wenxin Yiyan as examples [J]. Science and Technology Technology and Law (Chinese and English), 2023, (05): 86-96. DOI: 10.19685/j.cnki.cn11-2922/n.2023.05.009.

[4] Yu Wenwen. On the ownership of rights and interests of AI-generated content under copyright law [J]. Journal of University of Chinese Academy of Social Sciences, 2022, 42(02): 89-100+146-147.

[5] Lu Binghong. On the copyright protection of AI-generated content [D]. Jilin University, 2021. DOI: 10.27162/d.cnki.gjlin.2021.007564.