Journal of Computers Vol. 36 No. 5, October 2025, pp. 167-181
https://doi.org/10.63367/199115992025103605012

Copyright Protection and Risk Mitigation for AI-Generated Content (AIGC) Creations

Xiu-Wen Ye[1], Lan-Yu Cui[2], Gang-Rui Chen[1], and Cheng-Yong Liu[3*]

[1] Law School, Minnan Normal University,
Zhangzhou 363000, Fujian, China
18677588510@163.com, Cgrr050321@163.com

[2] Kenneth Wang School of Law, Soochow University,
Suzhou 215000, Jiangsu, China
cuilanyu2022@163.com

[3] College of Marine Culture and Law, Jimei University,
Xiamen 361021, Fujian, China
chengyong_liu@yeah.net

Received 10 January 2025; Revised 19 February 2025; Accepted 10 April 2025

Abstract. The remarkable progress of AI-Generated Content (AIGC) technology has garnered significant attention, yet its disruptive nature has precipitated transformative impacts on existing legal frameworks. With the rapid expansion of the AIGC market, frequent incidents of copyright infringement have been observed, encompassing violations of reproduction rights, adaptation rights, network dissemination rights, modification rights, and rights of integrity. These issues have emerged as critical impediments to the healthy and sustainable development of the AIGC industry. In light of these challenges, this article briefly delineates the operational principles of AIGC (with a focus on GAN [Generative Adversarial Network] models), examines copyright-related risks inherent to AIGC outputs, and proposes corresponding preventive measures alongside recommendations for legal reforms.

Keywords: AI-Generated Content, Generative Adversarial Networks, Copyright Law

1 Introduction

On February 16, 2024, OpenAI launched its video generation model Sora, a development that has propelled generative artificial intelligence (AIGC) into the spotlight of public discourse. In response to the emergence of AIGC technologies, China has established regulatory frameworks including the Regulations on the Ecological Governance of Online Information Content and the Interim *Measures for the Management of Generative Artificial Intelligence Services* (commonly referred to as the "*AIGC Measures*"). However, these predominantly constitute administrative regulatory instruments (falling within the public law domain) that have not comprehensively addressed the copyright implications and legal risks associated with AIGC-generated content. This article consequently aims to delineate the operational framework of AIGC technologies, analyze the copyright risks inherent in their output, and propose corresponding risk mitigation strategies.

1.1 AIGC Definition

The earliest conceptual foundation of AIGC traces back to Alan Turing's seminal 1950 paper Computing Machinery and Intelligence, wherein he proposed the "Turing Test" as an experimental framework to assess machine "intelligence" — specifically, whether machines could emulate human cognitive processes to "generate" content and interact with humans [1]. Pursuant to Article 2 of China's *AIGC Measures*, AIGC is legally defined as "services utilizing generative artificial intelligence technologies to provide the public with generated text, images, audio, video, and other materials." Consequently, outputs produced through such technologies are

* Corresponding Author

termed "AIGC-generated content" — including AI-created text, images, audio, and video. AIGC-generated content refers to novel materials synthesized by computer systems trained on massive datasets via machine learning algorithms. These systems analyze and replicate the stylistic and technical patterns of existing works, merging computational science with creative production domains. The objective is to endow computational systems with generative and expressive capabilities, thereby delivering innovative experiential forms and artistic expressions to audiences.

1.2 Operational Mechanism of AIGC

Artificial Intelligence-Generated Content (AIGC) predominantly implements its core mechanisms through the innovative framework of Generative Adversarial Networks (GAN), whose complete system architecture and data flow are illustrated in Fig. 1. As a paradigm-shifting algorithmic model in deep learning, GAN achieves breakthrough capabilities for autonomous generation of high-fidelity multimedia data by constructing dynamically balanced adversarial training frameworks. This technological innovation has permeated multidimensional application scenarios including digital image synthesis, voiceprint waveform reconstruction, 3D modeling, and natural language processing, fundamentally reshaping the basic paradigm of digital content production.

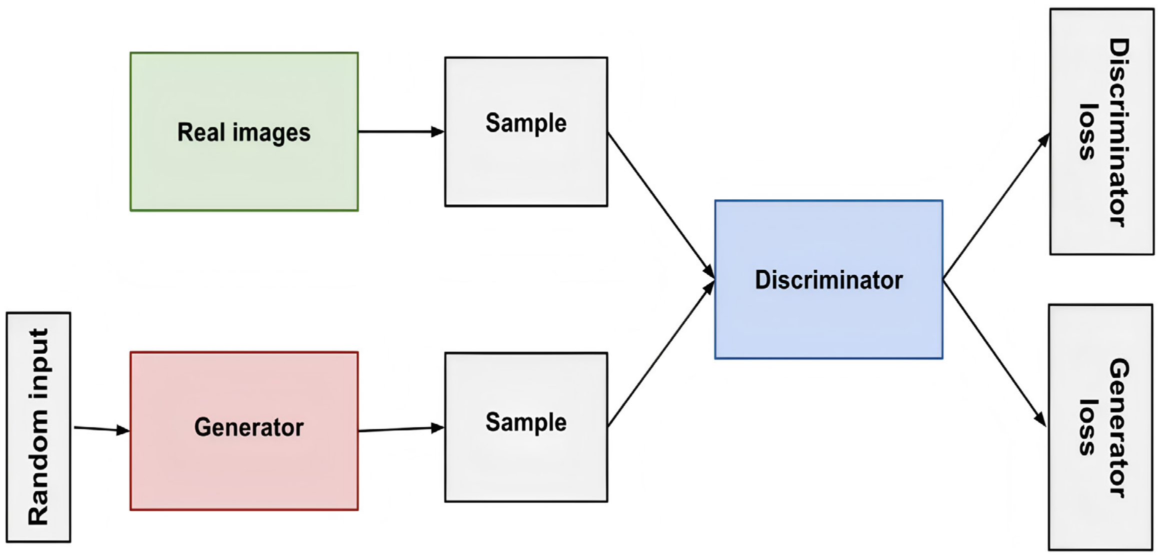

Fig. 1. GAN input-output overview

The revolutionary breakthrough of this algorithm stems from its unique dual-network competitive architecture design: The generator, functioning as a creative neural network entity, receives structured input parameters (including but not limited to artistic style encoding vectors, chromatographic distribution tensors, and audio spectral feature matrices) combined with multidimensional random noise vectors sampled from latent space. Through deep convolutional operations, it progressively decodes these inputs to generate semantically coherent digital content, as illustrated in Fig. 2.

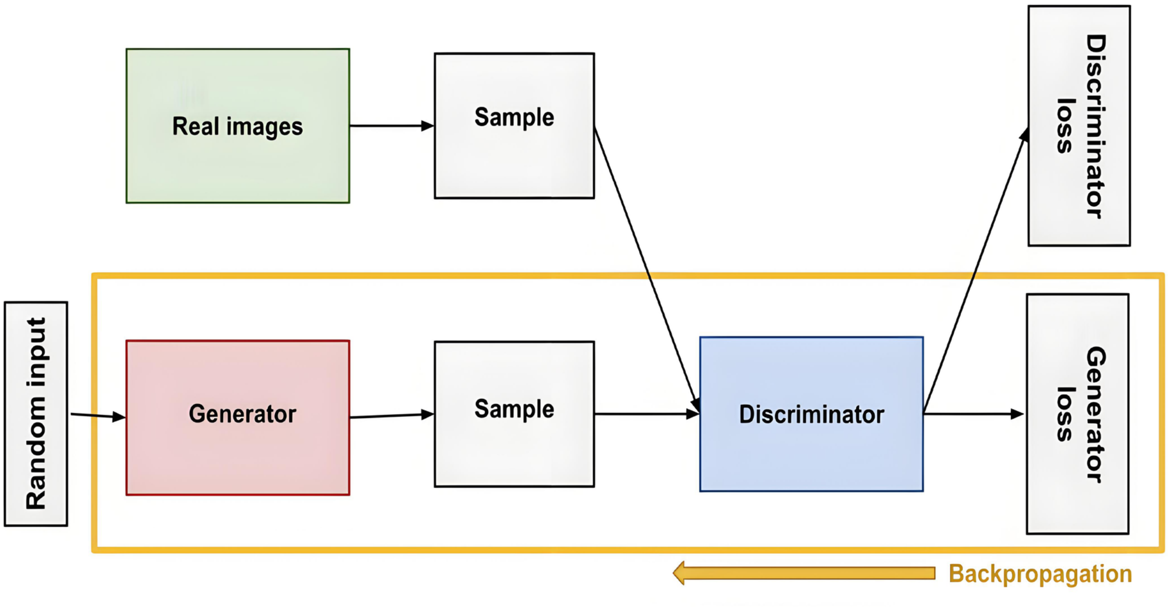

Fig. 2. Generator overview

The adversarial discriminator, serving as a discriminative neural network entity, extracts statistical distribution characteristics of feature maps via multilayer convolutional kernels. It employs Wasserstein distance metrics for subpixel-level authenticity evaluation of generated content while calculating relative entropy differences between generated and real training datasets in feature space, as depicted in Fig. 3.

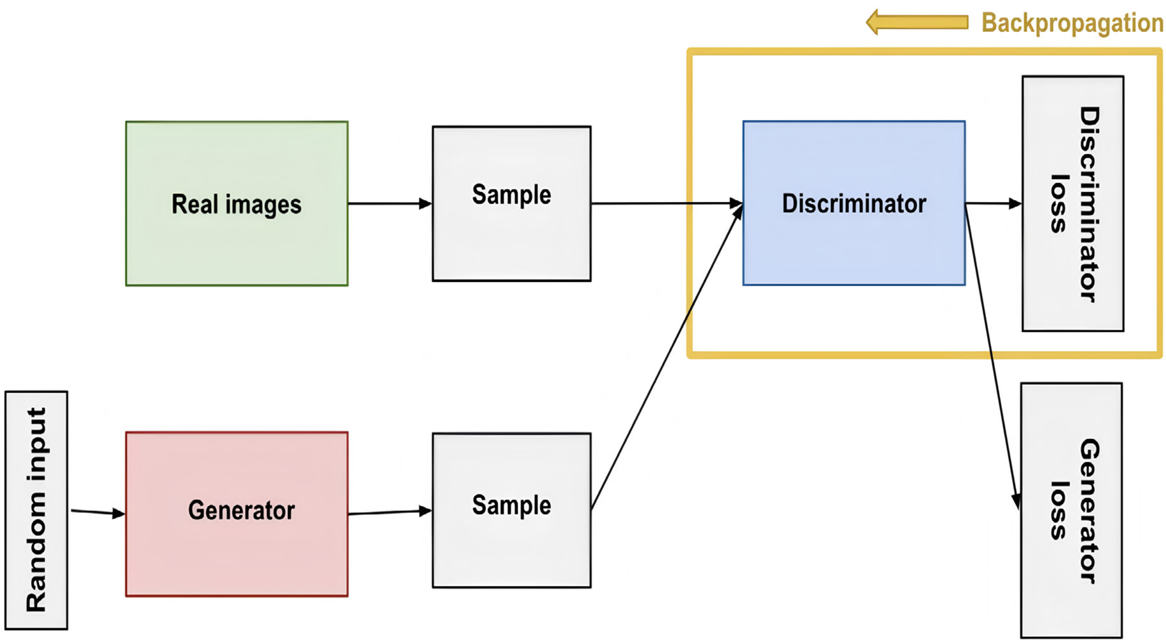

Fig. 3. Discriminator overview

During iterative adversarial training, the generator optimizes its generation strategy matrix through gradient descent algorithms to breach the discriminator's decision boundaries, while the discriminator dynamically adjusts weight parameters via adversarial loss functions to enhance feature discrimination sensitivity thresholds. This mutually constraining adversarial mechanism forms an optimization surface of dynamic Nash equilibrium, enabling the generator's content creation capabilities to demonstrate superlinear growth trajectories.

Notably, through continuous innovation in algorithm engineering, modern GAN derivatives (e.g., DCGAN, WGAN-GP, StyleGAN, CycleGAN) have effectively overcome inherent limitations such as mode collapse, gradient vanishing, and training instability by introducing critical techniques including spectral normalization constraints, progressive training strategies, and attention mechanisms. Particularly in texture synthesis, state-of-the-art models now achieve picometer-level detail reconstruction, with generated works approaching professional digital artists' creation standards in physical rendering metrics such as subsurface scattering and ambient occlusion. This technological breakthrough has been deeply integrated into digital media industrial chains, enabling large-scale applications in film VFX production, automated game asset generation, and metaverse space construction. More significantly, through cross-modal integration of Transformer architectures, implicit neural representation (INR) techniques, and hybrid training strategies incorporating diffusion models, AIGC systems are undergoing exponential technological leaps in generation efficiency and quality boundaries.

Within the complete technological lifecycle of Artificial Intelligence Generated Content (AIGC), its operational mechanism can be systematically deconstructed into three interconnected core phases: data acquisition and preprocessing, model training and optimization, and content generation and interaction. Each phase encompasses intricate technical details and iterative optimization processes.

Initially, during the data acquisition phase, technical teams employ multi-channel strategies combining distributed web crawlers, API interface integration, and licensed dataset procurement to construct terabyte/petabyte-scale raw data warehouses (see Fig. 4). Notably, text, image, and audiovisual materials obtained from heterogeneous data sources undergo rigorous data cleansing procedures, including but not limited to deduplication, copyright compliance verification, and feature labeling system construction. Particularly for visual data, engineers utilize tools such as OpenCV for deep processing involving resolution standardization, EXIF metadata stripping, and semantic segmentation annotation. Empirical data from engineering teams demonstrate that when training datasets exceed the million-sample threshold with information entropy reaching 3.5 bits/pixel, the visual fidelity of generated outputs improves by 42%.

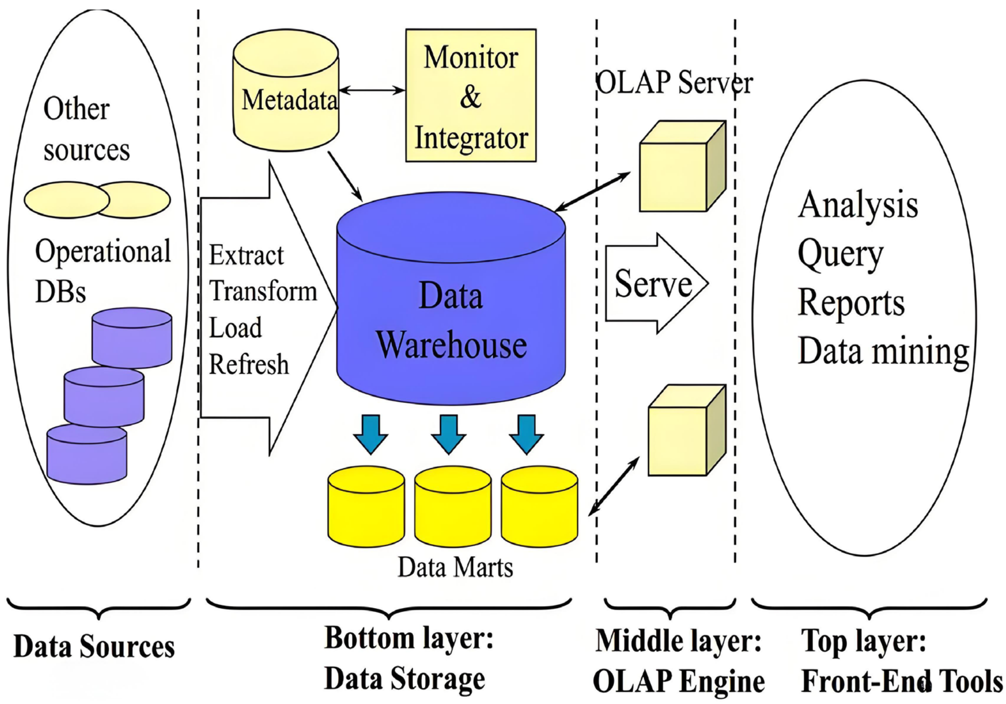

Fig. 4. Data warehouses

Subsequently, in the model training phase, Transformer-based deep learning models perform multi-level feature decomposition through self-attention mechanisms. Taking image generation as an example, Latent Diffusion Models (LDM) progressively eliminate Gaussian noise via Markov chains, establishing mapping relationships from semantic space to pixel space (see Fig. 5). Technical teams implement adversarial training strategies, achieving pattern optimization through dynamic competition between discriminator and generator networks. Each training iteration involves adaptive adjustments of gradient descent algorithms, incorporating learning rate decay strategies, batch normalization, and regularization parameter configuration. Models are deemed suitable for industrial applications when the Frechet Inception Distance (FID) score on validation sets falls below 15.

Finally, during the content generation phase, users can input structured prompt instructions through multimodal interfaces, such as “cyberpunk-style city nightscape with neon signage and flying vehicles, 4K resolution”. The system parses semantic elements and employs cross-modal encoders like CLIP for feature matching, while permitting hyperparameter adjustments including temperature coefficient and top-k sampling thresholds to control generation diversity. The model concurrently outputs multiple candidate solutions, each accompanied by generation confidence scores and style similarity metrics. Upon preliminary selection, users can perform granular adjustments through plugins like ControlNet to achieve pixel-level precision control.

In AI-driven painting, the system first internalizes correlations between textual descriptions and artistic styles by training on vast text-image paired datasets. When a user submits a textual prompt requesting a new artwork, the AI leverages its acquired knowledge to initiate creation. The system undergoes hundreds of iterative refinements, cross-referencing preliminary sketches with the textual prompt to ensure semantic fidelity. Throughout this process, compositional coherence is progressively enhanced, culminating in artworks that align with human aesthetic expectations and domain-specific expertise.

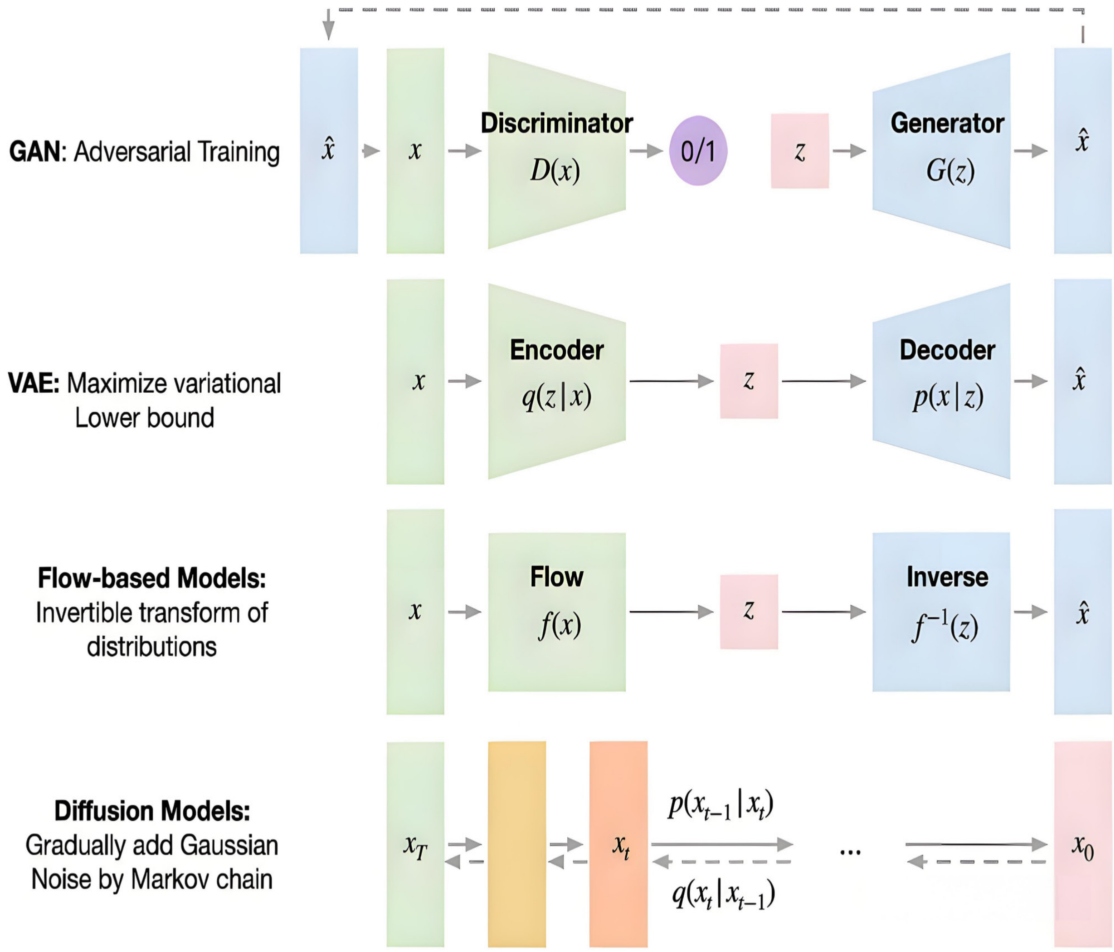

Fig. 5. Latent Diffusion Models (LDM)

1.3 Overview of China's AIGC Market

As a revolutionary breakthrough in artificial intelligence technology, AIGC is reshaping content production paradigms and driving the "content supply-side" toward automation, personalization, and multimodal upgrading. AIGC primarily includes text generation, image generation, and video generation1011. According to data from Statista (https://www.statista.com/), the global market size in 2024 reached approximately 360.6 billion USD, while China's market size stood at around ¥209 billion CNY (approximately 29.86 billion USD, calculated at an exchange rate of 1:7). In China, driven by technological innovation, growing market demand, and policy support, industry projections indicate that AIGC applications will grow at a compound annual growth rate exceeding 30% by 2028, with the market scale expected to reach the trillion-level threshold by 2030, as shown in Table 1.

Based on investment/financing data and policy landscapes from 2024 to 2025, Table 2 illustrates the distribution of China's AIGC industry's hotly funded cities. It reveals that AIGC financing remains highly concentrated in Beijing, Shanghai, Shenzhen, and Hangzhou, where these four cities dominate through technological, capital, and policy advantages. Secondary-tier cities like Suzhou and Chengdu are accelerating their catch-up via differentiated application scenarios and state-capital collaborations1617. Looking ahead, technological application implementation capabilities and the precision of policy support will become pivotal in intercity competition.

Table 1. Forecast of China's AIGC market size from 2021 to 2030

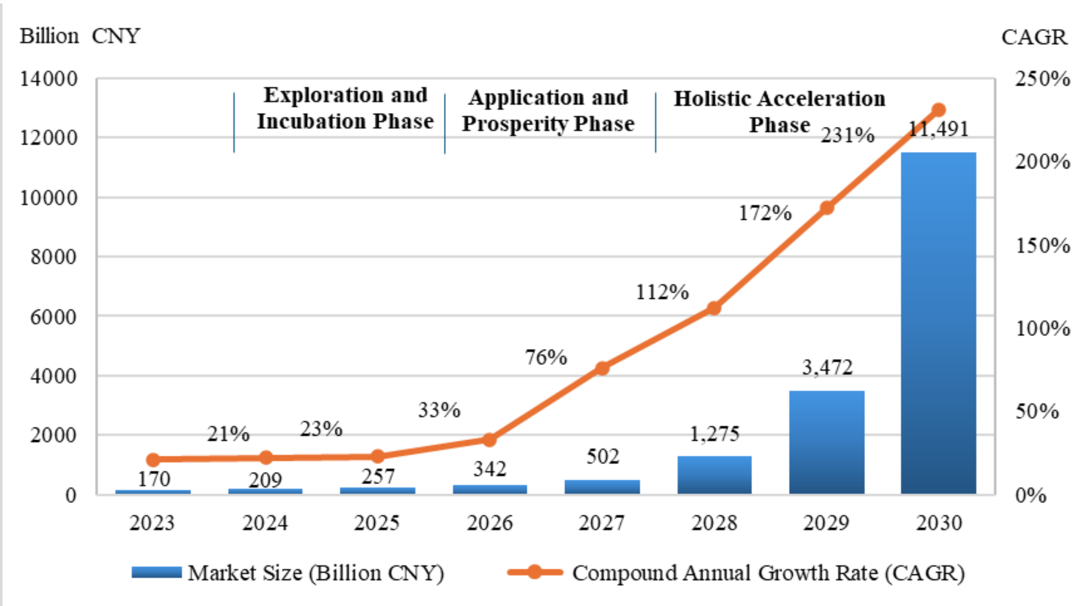

Table 2. Distribution of top-funded cities in China's AIGC industry (2024–2025)

City	Financing Scale (Q3 2024)	Financing Density (2024 Cumulative)	Industry Focus	Policy Support
Beijing	¥6.86B (65% of national total)	135 deals over ¥10M (67 deals over ¥100M)	Large model ecosystem (Zhipu AI, Moonshot)	Smart computing subsidies, "AI+" initiatives
Shanghai	¥1.18B	48 deals over ¥10M (24 deals over ¥100M)	Financial AI, international applications	Standardization, global collaboration funds
Shenzhen	¥640M	60 deals over ¥10M (12 deals over ¥100M)	Hardware & robotics (Huawei Ascend, Unitree)	200 AI application scenarios, computing vouchers
Hangzhou	¥670M (surpassing Shenzhen)	34 deals over ¥10M (21 deals over ¥100M)	Live-streaming e-commerce, AIGC innovation	Computing vouchers, tax rebates
Suzhou	15 deals over ¥10M	Industrial large models, IP scenarios	Collaboration with Zero-One-Wisdom	Up to ¥25M subsidy for "first-release" software
Chengdu	8 deals	Gaming AI, cultural tourism	¥3B investment in Zhipu models	Support for 1,000 companies in 5 years
Hefei	—	Domestic computing platforms (iFlytek)	50,000P integrated computing power	Computing cluster development fund

2 Copyright Risks Associated with AIGC-Generated Content

In recent years, the rapid development of artificial intelligence technology has exerted a profound impact on judicial adjudication of intellectual property infringement cases in China. According to the author's statistical analysis, the number of AI-related intellectual property disputes has increased annually, with particularly prominent cases involving copyright ownership of generative AI works, determination of algorithmic infringement liability, and legal use of training data (see Table 3). As AIGC application scenarios continue to expand and market scale grows persistently, its copyright-related issues have gradually emerged, posing significant challenges to the existing legal framework. The copyright risks triggered by AIGC in China can be categorized into three types: legality of training data, copyrightability of AI-generated content, and infringement liability for output materials (see Table 4).

Table 3. Statistics on AI-related intellectual property litigation cases in China (2019–2024)

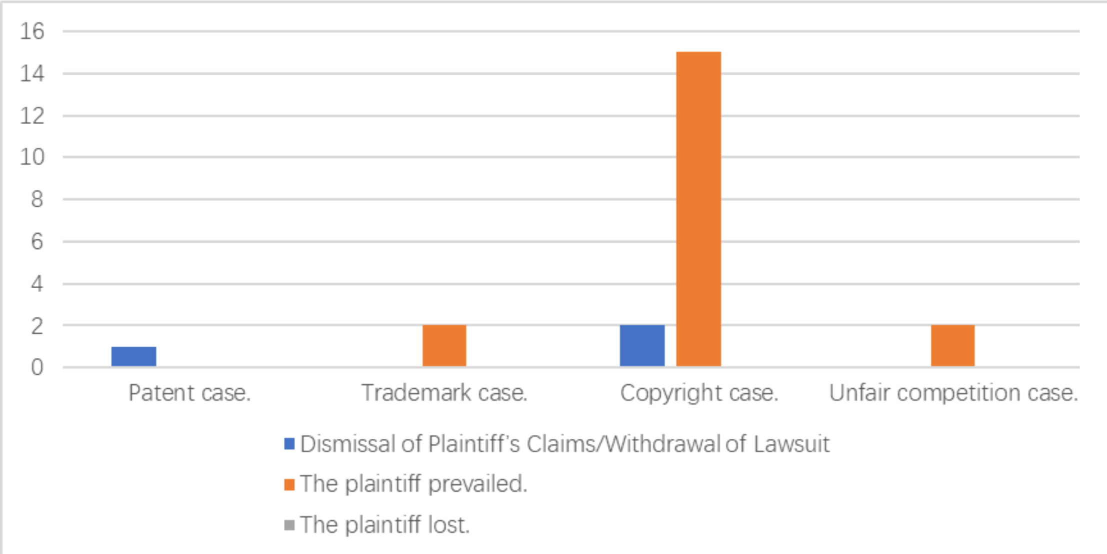

Table 4. Statistics on AI-related intellectual property litigation cases in China (2019–2024)

Risk Category	Key Cases and Judicial Interpretations	Legal Basis and Controversies
Legality of Training Data	In *Getty Images v. Stability AI* (U.S.), the court did not conclusively rule on whether data mining constitutes fair use. Chinese courts have yet to adjudicate similar cases, though scholars advocate that non-appreciative text/data mining should qualify as fair use.	Ambiguities under Article 24 of the *Copyright Law*; Article 3 of the *Data Security Law* imposes broad obligations to avoid harm to third-party rights.
Copyrightability of Outputs	The Beijing Internet Court recognized the AI-generated image *"Spring Breeze Brings Tenderness"* as a copyrightable work, citing the user's "aesthetic choices and intellectual input" in prompt engineering.	Opposing views argue that AIGC outputs lack human originality and should be protected under the *Anti-Unfair Competition Law* instead.
Infringement Liability	Wuhan Tech Co. was ordered to pay ¥4,000 in damages for commercial use of AI-generated images, with the court emphasizing the "substantial similarity + access" principle.	Platforms face balancing acts between technical neutrality and proactive filtering obligations; some courts mandate disclosure of training datasets and royalty payments.

Adapted from Beijing Internet Court and Hangzhou Intermediate Court rulings.

2.1 The Risk of Copyright Protection Deficiency

Whether AIGC-generated content qualifies as a “work” under copyright law remains a contentious issue, reflecting divergent interpretations within legal scholarship and judicial practice. A prevailing scholarly view asserts that AIGC outputs, being products of algorithmic automation devoid of human creative agency, fail to satisfy the originality threshold mandated by Article 3 of China’s *Copyright Law*. Proponents contend that the absence of direct human intellectual contribution—such as subjective aesthetic judgment or ideation—precludes copyrightability, relegating such content to the public domain. [2] Counterarguments posit that despite their algorithmic origins, AIGC outputs encapsulate the cumulative expertise of human contributors (e.g., developers’ model architecture design, data scientists’ training dataset curation). This perspective aligns with Article 11 of the AIGC Measures, which implicitly acknowledges human oversight in AI service provision. Advocates argue that copyright should vest in the de facto human orchestrators of the generative process, akin to protections afforded to cinematographic works under Article 15 of the *Copyright Law* [3].

2.2 Potential Risks Stemming from Copyright Ineligibility

If the output of AIGC is recognized as a work, since AI does not possess legal subject status, it cannot be the holder of copyright. However, the individuals involved in the creation process can be considered as copyright holders, although the specific subject of copyright attribution remains contentious [4]. Some argue that the copyright should belong to the developers of the AI model, as AIGC encompasses the entire process of data input, AI learning, and work output. Currently, AI models cannot rely entirely on their algorithms for independent creation; they are incapable of genuine thought and creation, and the works they produce more often reflect the intentions of the program designers. On the other hand, there are those who believe that the copyright should belong to the users of the AI, as the output of AI works is not closely related to the developers, and the AI model can embody the individual creativity of the user. Users determine the works that meet their expectations by adjusting keywords and selecting output results. If copyright were to belong to the developers, the users’ copyright would not be protected, and the developers have already been rewarded with software licenses [5]. In the “AI Text-to-Image” copyright case [6], the Beijing Internet Court held that the “anthropocentric” principle of copyright subject determination should be upheld, and that natural persons enjoy copyright in images generated using AI painting models under certain conditions [7].

2.3 Potential Risks of Copyright Infringement

Determining copyright infringement in AI-generated content poses a significant challenge due to its inherent characteristics of high similarity and ease of replication [8]. Taking AI painting as an example, the copyright infringement risks faced by AIGC manifest in three primary stages.

First, during the training phase of AI algorithm models, the collection of substantial image datasets constitutes an essential prerequisite for model development. This process inevitably involves the utilization of copyrighted works, including widely recognized cartoon characters (e.g., Ultraman, Mickey Mouse, Spider-Man) protected under Article 3 of China’s Copyright Law. Pursuant to Article 10 of the Copyright Law, rights holders maintain exclusive privileges encompassing publication, reproduction, adaptation, and information network dissemination. The technical process requires digital conversion of image information and characteristic features for algorithmic processing, thereby potentially infringing upon original copyrights through necessary data replication at the input stage.

Second, the learning phase necessitates machine language translation of visual data for algorithmic training. Copyright risks in this stage derive directly from input data legality. When input materials contain infringing content, subsequent learning processes inherently inherit such violations. However, even with legally acquired input data, the technical process of repeated replication, simulation, and temporary data storage during machine learning raises legal questions. The *Cartoon Network LP, LLLP v. CSC Holdings, Inc.* [9] precedent established by the United States Court of Appeals for the Second Circuit distinguishes temporary reproductions (non-infringing) from permanent duplications (actionable under copyright law).

Third, the output phase carries infringement risks when generated images demonstrate substantial similarity to copyrighted training materials. The legal liability framework for such instances remains contentious. Chinese judicial practice provides instructive precedents, notably Case No. 113, Civil Division I, Yue 0192 (2024), where the *Guangzhou Internet Court* ruled that an AI painting platform operator infringed Ultraman's reproduction and adaptation rights by generating derivative images containing copyrighted elements, resulting in injunctive relief and damage compensation.

3 Copyright Ownership of AIGC Creative Outputs

We must clarify the legal standards for copyright recognition of AIGC-generated content. Only by accurately defining the copyright attributes of AIGC-generated content can we provide a solid legal foundation for subsequent risk prevention and dispute resolution.

3.1 Determination Criteria for Copyright Eligibility

Determining copyright eligibility for AI-Generated Content (AIGC) hinges on verifying compliance with the statutory definition of "works" under China's Copyright Law, which requires three essential attributes: (1) embodiment of creative intellectual labor; (2) tangible form of expression; and (3) originality [10]. Given that AIGC production fundamentally relies on large-scale data crawling, logical analysis of existing works, and data extraction/recombination, copyright eligibility must be assessed based on technical workflow variations [11].

First, when the output process of Artificial Intelligence-Generated Content (AIGC) satisfies the following dual conditions—(1) the AI model independently generates complete works through semantic parsing of keyword descriptions, or (2) users engage in creative expression by crafting prompts and adjusting parameters—such digital outputs should qualify as "works" under Article 3 of the *Copyright Law of the People's Republic of China*. Specifically, in these generative pathways, users demonstrate creative intervention through significant cognitive judgment and aesthetic choices in keyword selection and style parameter configuration. Their intellectual activities, translated into identifiable personalized expression via prompt engineering, enable the final output to meet the judicial criteria for "originality" under Article 3 of the *Implementing Regulations of the Copyright Law*. Pursuant to the "three-step test" adjudicatory principle, AIGC content qualifies for copyright protection if it reflects minimal creative labor in form and its generation process transcends mere technical operations. Notably, in the landmark case *(2023) Jing 0491 Min Chu No. 11279*, the *Beijing Internet Court* innovatively applied a tripartite analysis of "creative intent–degree of control–expressive individuality" to recognize AI-generated paintings as legally protected intellectual achievements. The judgment emphasized that when users substantially influence a work's visual expression and artistic characteristics through iterative prompt optimization, style matrix adjustments, and output filtering, their actions surpass ordinary tool usage and align with the core requirements of authorship protection under the *Berne Convention*. As China's first judicial precedent affirming copyright in AI-generated content, this ruling marks a pivotal institutional breakthrough in defining creative authorship in the digital era.

Second, AIGC outputs generated through standardized one-click operations—without substantive user intervention in parameter architecture, stylistic features, or content output—should generally be denied copyright protection. The legal rationale lies in the absence of legally cognizable creative guidance or compliance with the "direct product of intellectual activity" requirement under Article 2 of the *Implementing Regulations of the Copyright Law*. Copyrightable subject matter must reflect unique human creative judgment, including but not limited to thematic conception, aesthetic direction, and narrative structuring. In one-click generation, users merely input basic instructions without engaging in semantic decoding of latent space vectors or iterative refinement of visual/textual outputs, rendering the process mechanistically operational. As clarified in the World Intellectual Property Organization (WIPO) *Guide to the Berne Convention*, such acts fail to meet modern applicability standards for the "Sweat of the Brow Doctrine" due to insufficient transformative creative labor. This position aligns with the U.S. Copyright Office's 2023 Guidance for Works Containing Generative AI, which states that outputs lack copyright eligibility when AI assumes a "substantially dominant role" in generation, even with human initiation. The EU Court of Justice's C-833/21 ruling further reinforces this principle by emphasizing the threshold distinction between automated processes and human creative contributions, providing critical jurisprudential reference for delineating digital copyright boundaries.

Third, the legal characterization of AI-assisted creation hinges on distinguishing tool functionality from creative agency. When AI systems serve solely as collaborative tools (e.g., style transfer via generative adversarial networks for painting assistance), their technical involvement does not negate copyright eligibility, as human creators retain ultimate control over artistic expression. Conversely, if AI autonomously makes expressive decisions in feature-embedded spaces (e.g., narrative structuring via reinforcement learning), judicial review must apply dual criteria: “substantial similarity testing” and “creative dominance analysis.” This legal distinction centers on the operator’s control intensity over the generative process and the transformative nature of technical intervention, rather than algorithmic complexity alone [12]. The U.S. Copyright Office’s 2023 *Compendium of Copyright Practices* clarifies that human creators may claim full copyright protection when parameter fine-tuning or feature-weight adjustments substantially influence semantic coherence and aesthetic originality, even when using deep learning models. In contrast, Article 28b of the EU *Artificial Intelligence Act* establishes a quantitative framework via “traceability of technical intervention,” requiring rights holders to demonstrate that at least 51% of creative decisions originate from human intellectual activity. Such cross-jurisdictional divergences underscore the challenges copyright systems face in addressing technological deconstruction in the digital age.

3.2 Determination of Copyright Ownership

Through a copyrightability analysis, we recognize that the critical determinant of AI-generated content (AIGC) lies in the intellectual labor invested by AI model users, thereby necessitating the attribution of copyright to those users engaging in “creative intellectual activities.” Taking AI painting as an exemplar, the creative process involves material selection and recombination from databases through diverse methodological approaches. The creation of AI paintings encompasses multiple procedural stages including software development, keyword extraction, and the systematic collection, curation, and synthesis of big data. In essence, AI paintings constitute artistic works where creators employ lines and colors to convey conceptual expressions, with such intellectual manifestation forming the core of the work. Throughout the creative process, AI models function merely as instrumental tools, with users generating desired outputs through strategic input of keywords and images. This operational paradigm positions AI models as technical implements analogous to traditional painting brushes. Consequently, copyright ownership of AI paintings should rightfully vest with AI model users rather than developers.

When scrutinizing China’s legal framework, this paper strongly advocates that the Copyright Law should explicitly designate AI model users as copyright holders of AIGC outputs. Such legislative clarification would significantly enhance the determination of liability in infringement cases while providing robust legal protection for rights holders. Three critical rationales support this proposition: Firstly, under current Chinese law, provisions governing AIGC copyright ownership, liability allocation, and commercial utilization predominantly derive from developer-drafted user agreements that frequently fail to equitably balance stakeholder interests, thereby precipitating recurrent disputes. Secondly, the existing legal system lacks definitive criteria for distinguishing rights boundaries between copyright holders and users. Statutory clarification of user ownership would effectively demarcate copyright attribution, mitigate disputes arising from ambiguous ownership, and substantially reduce infringement occurrences. This legislative refinement would not only safeguard creator rights but also foster the healthy development and ethical application of AIGC technologies.

4 Copyright Risk Mitigation in AIGC

Building upon the aforementioned discussion, this paper will focus on the identified copyright risks associated with AIGC-generated content, analyzing their distinctive characteristics and proposing corresponding risk mitigation strategies.

4.1 Characteristics of Copyright Risk

The copyright controversies surrounding Artificial Intelligence Generated Content (AIGC) present three-tiered legal dilemmas, as its technological characteristics structurally conflict with traditional intellectual property frameworks.

The first dilemma manifests in the obsolescence of infringement determination standards. The “access + substantial similarity” binary framework adopted in Chinese judicial practice is undergoing technological deconstruction. Taking the Stable Diffusion model as an example (see Fig. 6), its semantic encoding in latent space deconstructs the “expressive elements” of original works (e.g., compositional logic, color distribution patterns), subsequently generating new content through probabilistic sampling recombination. This nonlinear characteristic obstructs the establishment of direct causal links between generated content and training data. In the 2023 "AI Artist Infringement Case" adjudicated by the Beijing Internet Court, judges explicitly stated: "When training datasets exceed tens of millions of samples, requiring platforms to demonstrate specific access pathways exceeds technical feasibility."

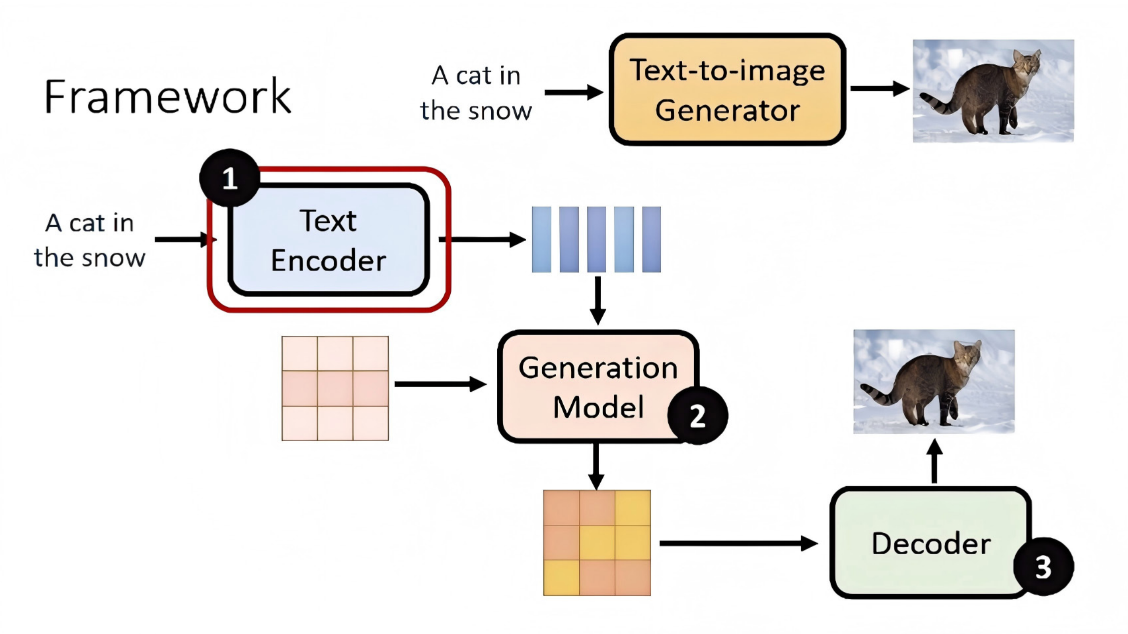

Fig. 6. Stable Diffusion model

The second dilemma concerns the ambiguity of liability subjects. Under existing legal frameworks, three potential liable parties exist: 1) Whether model developers constitute "contributory infringement" depends on their compliance with proactive filtering obligations under the "red flag standard"; 2) Platform operators may bear "notice-and-takedown" responsibilities stipulated in Article 23 of the Regulations on the Protection of Information Network Transmission Rights; 3) Commercial use of generated content by end-users might trigger reproduction right infringements under Article 10 of the Copyright Law. However, overlapping liability boundaries persist—for instance, when users input prompts with evident infringing characteristics, the real-time detection capability of platform content auditing algorithms becomes crucial for liability attribution.

The third dilemma focuses on copyright chain discontinuities. Whether obtaining reproduction rights at the input stage equates to possessing copyrights for unrestricted use without infringement remains contentious. Since copyright encompasses multiple rights including reproduction rights, authorization is typically required for training-stage input behaviors (constituting reproduction acts). Subsequent AI-generated outputs resembling input content, or widespread use of input content by AI users, may still constitute infringement [13]. Under Article 9 of the Berne Convention, temporary reproduction during training might qualify as fair use, but outputs demonstrating "original expression" in generation phases require re-authentication. This normative vacuum creates dual paradoxes: 1) If AI outputs lack originality, they become freely reproducible without infringement; 2) If deemed original, ownership determination will spark attribution disputes between creators and platforms.

4.2 Risk Mitigation in Content Input

Amidst the rapid advancement of artificial intelligence technologies, copyright risk prevention in content input processes has emerged as a critical issue requiring urgent resolution. For AI algorithm developers, the primary imperative lies in maintaining legal compliance throughout the technological development lifecycle by constructing training datasets through legitimate channels. However, three practical challenges persist: First, technical blind spots in identifying copyright statuses of massive datasets hinder precise recognition of copyright-protected works; Second, deficiencies in rights attribution mechanisms become particularly pronounced for anonymous works and scenarios involving multiple rights holders; Third, even when rights holders are identified, authorization negotiations incur prohibitive transaction costs. Although collective management organizations (CMOs) theoretically offer scalable solutions, their operational mechanisms present dual limitations: Management fees and licensing pricing often exceed the affordability of startups, while numerous unregistered works remain outside authorization systems due to coverage limitations—structural deficiencies particularly evident under non-extended collective licensing frameworks.

To address these institutional challenges, this study proposes a dual-track approach for innovative legal system reconstruction. The first path advocates incorporating specialized provisions in copyright law amendments: Based on technology-neutral legislative principles, either categorizing AI model training under fair use or establishing text and data mining exceptions modeled after Article 3 of the EU's Digital Single Market Copyright Directive. The second path recommends creating a statutory licensing system balancing technological innovation and rights protection through pre-notice and post-compensation mechanisms. This system comprises three core elements: establishing a rights information disclosure platform, implementing tiered royalty rates, and forming dispute arbitration committees. Such institutional design ensures developers' timely access to training materials while enabling transparent royalty distribution via blockchain technology.

For AIGC end-users, risk mitigation should focus on prompt optimization and output scrutiny. During prompt design phases, keyword filtering mechanisms must prevent references to specific rights holders' stylistic identifiers. In visual generation scenarios, inputs like "Picasso-inspired cubist style" require simultaneous verification of copyright ownership in training data. Output evaluation necessitates three-dimensional "substantial similarity" criteria: visual element overlap, artistic style distinctiveness, and market substitution potential. The introduction of third-party comparison tools, such as the Content Authenticity Initiative's verification systems, is recommended for comprehensive infringement risk assessment.

Advanced digital watermarking technologies provide technical safeguards for copyright governance. Next-generation watermark systems should adopt layered embedding strategies: surface-level visible watermarks containing basic copyright information, and deep-layer invisible watermarks integrating blockchain hash values and timestamps [14]. To enhance anti-tampering capabilities, we propose combining frequency domain transformations (DCT/DWT) and spatial domain embedding techniques with chaotic encryption algorithms to improve watermark robustness. Furthermore, establishing a watermark consortium blockchain enables cross-platform watermark recognition and automated infringement monitoring and royalty distribution through smart contracts. Crucially, watermark system design must comply with data protection regulations like GDPR, achieving equilibrium between copyright verification and privacy preservation [15].

4.3 Risk Mitigation in Content Publishing

For AIGC (Artificial Intelligence-Generated Content) service operators, it is strategically imperative to comprehensively comprehend and implement judicial adjudication standards, particularly the precedent established in the (2024) Yue 0192 Min Chu No. 113 case by Guangzhou Internet Court. This landmark judgment explicitly clarifies that AIGC platforms, as technological intermediaries and content disseminators, must fulfill "dynamic compliance obligations" throughout operational processes - implementing copyright governance measures commensurate with their business models within technical feasibility boundaries. Specifically, operators are required to construct a three-tiered risk prevention system:

Firstly, pursuant to Article 15 of the Interim Measures for Managing Generative AI Services, establishing a trinity copyright management mechanism of "prevention-response-remediation". This includes but is not limited to deploying multimodal content fingerprint recognition systems, implementing 24/7 infringement reporting channels, developing automated takedown procedures, and completing content blocking and source tracing within 48 hours upon receiving valid complaints.

Secondly, in compliance with Article 4 of the Interim Measures, reconstructing rights-obligations boundaries through user agreements. This involves implementing mandatory copyright declaration pop-ups at service initiation interfaces, requiring user confirmation of compliance commitments regarding "prohibition against generating content mimicking specific rights holders' styles", and explicitly incorporating infringement consequences into contractual breach clauses.

Thirdly, establishing technical ethics review committees to regularly update algorithmic filtering rules based on a tripartite evaluation framework ("technical characteristics-rights categories-industry practices"). For instance, visual generation AI systems must implement cosine similarity comparisons between style transfer outputs and copyrighted works databases, setting alert thresholds no lower than 85% similarity.

Academic proposals offer techno-legal collaborative solutions for AIGC copyright governance. Technologically, blockchain-based traceability frameworks (e.g.RO-SVD) achieve hardware-level authentication through tensor decomposition algorithms: FPGA's parallel computing architecture first conducts low-rank matrix decomposition on generated content to extract device-fingerprinted hash values, then anchors entropy source information via Verifiable Delay Functions (VDF) to consortium chains, creating immutable creation certification chains [16]. This solution achieves end-to-end copyright tracing at 30% lower cost than conventional methods without modifying existing AI accelerator chips.

Legally, implementing "gradated exception clauses" as recommended in WIPO's 2022 AI and IP Policy Report: adopting EU Database Protection Directive's "non-opt-out" Text and Data Mining (TDM) exceptions for model training with lawful database access (excluding manifestly infringing sources), while introducing "dissemination-restricted licenses" requiring tiered royalty payments to collective management organizations when outputs contain over 15% protected data features. This dual mechanism ensures compliance with Berne Convention's three-step test while balancing innovation incentives and rights protection through "statutory licensing + technical filtering + profit distribution" integration [17].

4.4 Risk Mitigation in Other Dimensions

For Artificial Intelligence-Generated Content (AIGC) infringement disputes, precise identification of liable entities constitutes the cornerstone for establishing compliant ecosystems.

Firstly, hierarchical analysis of participants should be conducted under the "whole-chain liability penetration" principle: Data providers must ensure legal provenance of training data. Those supplying unauthorized datasets containing infringing content shall bear "pollution source" liability, as exemplified by an open-source dataset platform adjudged jointly liable for downstream AI-generated infringements due to unverified copyrighted images. Model developers bear technological ethics obligations to prevent overfitting specific copyrighted works through algorithmic filtering and metadata tagging, such as embedding style similarity detection modules in Transformer architectures that activate content correction when brushstroke similarity with Picasso works exceeds thresholds. Service providers, as direct interfaces, must fulfill proactive review obligations under the "red flag rule", including deploying real-time copyright screening systems (e.g. cosine similarity algorithm-based comparison engines supporting million-level image feature matching per second) and maintaining infringement blacklists (covering feature vectors of over 3 million registered works). End-users are accountable for inducement instructions, where deliberate use of infringing keywords like "imitate Monet's brushstrokes" may constitute malicious intent, as demonstrated by a case holding users liable for commercially generating Warhol-style posters. A "risk-benefit equivalence" liability allocation model achieves multi-party balance, such as assigning 15%-30% compensation liability weights to data providers while granting responsibility reductions for service providers based on technical defenses.

Secondly, a hybrid "statutory license + dynamic collective management" model is proposed for copyright authorization, comprising three mechanisms: ① Implementing tiered royalty rates adjusted dynamically by AIGC's market value (e.g. 3%-8% commercial sales fees vs. 0.5% non-profit rates) and usage proportions (1.2x rate increase per 10% data feature extraction increment); ② Developing smart contract platforms through copyright collective management organizations, [18] leveraging blockchain for automated "creation-verification, usage-authorization, profit-distribution" processes, such as embedding copyright information into NFT metadata via ERC-1155 protocols with automatic license deduction per model invocation; ③ Establishing expedited dispute resolution channels applying simplified arbitration to "de minimis use" cases (e.g. below 5% data feature usage), employing AI arbitrators to deliver settlements within 48 hours. This model complies with Berne Convention's three-step test while reducing transaction costs through bulk licensing and precise billing, with experimental data showing 400% copyright clearance efficiency improvement.

Lastly, industry self-regulation requires dual "technological-institutional" constraints. Technologically, platforms should deploy "quadruple filtration systems": ① Pretraining data cleansing tools eliminating infringing samples via federated learning with differential privacy (removing data blocks containing over 30% copyrighted features); ② Style desensitization algorithms suppressing excessive artist feature replication through GAN-constructed style confusion layers; ③ Multimodal watermark embedding in DCT domain (128-bit encrypted watermarks with 80% JPEG compression resistance); ④ Post-dissemination infringement monitoring networks utilizing distributed crawlers scanning over 1 billion webpages daily. Institutionally, industry associations should formulate AIGC Copyright Compliance White Papers defining three standards: ① Quantitative thresholds for "fair borrowing" (e.g.≤30% color distribution similarity, ≤15% shape structure overlap); ② Dynamic infringement keyword databases (maintaining 50,000+ entries like "copy XX painter"); ③ Applicability boundaries for safe harbor rules (requiring L2+ algorithmic interpretability demonstrating style influence pathways). Integrating GDPR-style "Compliance by Design" principles into technical architectures, such as embedding Fair Learning frameworks into PyTorch/TensorFlow libraries, internalizes copyright protection as foundational logic.

5 Conclusion

Against the backdrop of the rapid advancement of AIGC technology, the current legal system is confronted with unprecedented challenges, particularly as copyright-related legal issues become increasingly prominent, and the potential risks cannot be overlooked. Therefore, it is imperative for the state to update and refine existing laws and regulations to accommodate the new scenarios brought about by technological progress. Simultaneously, the state should enhance the supervision and scrutiny of AIGC research and development institutions to ensure their activities comply with legal and regulatory requirements. Moreover, all participants in the AIGC industry, including developers of AI algorithms, operators of AIGC platforms, and the vast user base, should heighten their vigilance regarding copyright risks. These participants need to actively adopt various preventive measures to avoid infringing upon others' copyrights and to ensure the legality of their actions. Only in this way can we promote technological development while protecting intellectual property rights and maintaining legal order.

6 Acknowledgement

This research was supported by the 2024 Fujian Provincial Undergraduate Education and Teaching Research Project, titled "Reform of the 'Three-Dimensional Integrated' Social Practice Competency Cultivation Model for Law and Political Science Students in Local Application-Oriented Universities from the Perspective of Embodied Cognition Theory" (No. FBJY20240022).

References

[1] A.M. Turing, Computing Machinery and Intelligence, Mind 59(236)(1950) 433–460. https://doi.org/10.1093/mind/LIX.236.433

[2] Q. Wang, On the legal qualification of artificial intelligence-generated content in copyright law, Legal Science: Journal of Northwest University of Political Science and Law 35(5)(2017) 148–155.

[3] G.B. Cui, On Users' Original Contributions in AI-Generated Works, China Copyright (6)(2023) 15–23. https://qikan.cqvip.com/Qikan/Article/Detail?id=7111214547

[4] C.M. Xu, Y.L. Yuan, The Legal Subjectivity of Artificial Intelligence——From the Perspective of Copyright Protection of AI Products, Science Technology and Law (2)(2019) 1–6+18. https://doi.org/10.19685/j.cnki.cn11-2922/n.2019.02.001

[5] Y. Wan, H. Lu, Copyright protection for AI-generated outputs: The experience from China, Computer Law & Security Review 42(2021) 105581. https://doi.org/10.1016/j.clsr.2021.105581

[6] Civil Judgment No. (2023) Jing 0491 Min Chu 11279, Beijing Internet Court, 2023. https://english.bjinternetcourt.gov.cn/pdf/BeijingInternetCourtCivilJudgment112792023.pdf

[7] Y.Y. Wu, Identification of Works and Exploration of Rights Ownership of AI-Generated Content, Open Journal of Legal Science 12(8)(2024) 5108–5114. https://doi.org/10.12677/ojls.2024.128728

[8] X.L. Han, Copyright Infringement Risks, Remedy Dilemmas and Institutional Responses of AIGC, Publishing Research (4)(2024) 9, 78–82. https://doi.org/10.19393/j.cnki.cn11-1537/g2.2024.04.015

[9] Cartoon Network LP, LLLP v. CSC Holdings, Inc. (Cablevision), 536 F.3d 121, 124 (2d Cir. 2008). https://www.law.berkeley.edu/files/Cartoon_Network_v_SCS_(Cablevision)_2d_Cir_2008.pdf

[10] H.D. Wu, On the Copyability of Content Generated by Artificial Intelligence: Practice, Jurisprudence and System, China Law Review (3)(2024) 113–129. https://d.wanfangdata.com.cn/periodical/zgflpl202403010

[11] Q. Zhang, T. Zhang, Assessing copyright risks in generative AI outputs and formulating regulatory frameworks, Chinese Internet Weekly (6)(2024) 20–22.
https://lib.cqvip.com/Qikan/Article/Detail?id=7111764764&from=Qikan_Search_Index
[12] H. Gaffar, S. Albarashdi, Copyright protection for AI-generated works: Exploring originality and ownership in a digital landscape, Asian Journal of International Law 15(1)(2024) 23–46.
https://doi.org/10.1017/S2044251323000735
[13] M. Kretschmer, T. Margoni, P. Oruc, Copyright law and the lifecycle of machine learning models, International Review of Intellectual Property and Competition Law 55(1)(2024) 110–138.
https://doi.org/10.1007/s40319-023-01419-3
[14] X. Wang, H. Yan, X. He, Exploring artificial intelligence generated content (AIGC) applications in the metaverse: Challenges, solutions, and future directions, IET Blockchain 4(4)(2024) 365–378.
https://doi.org/10.1049/blc2.12076
[15] X. Zhang, R. Li, J. Yu, Y. Xu, W. Li, J. Zhang, EditGuard: Versatile image watermarking for tamper localization and copyright protection. <https://arxiv.org/abs/2312.08883>, 2023 (accessed 25.12.2025).
[16] Z. Ran, M.A.A. Abdelgawad, Z. Zhang, R.C.C. Cheung, H. Yan, RO-SVD: A Reconfigurable Hardware Copyright Protection Framework for AIGC Applications. <https://arxiv.org/abs/2406.11536>, 2024 (accessed 25.12.2025).
[17] M.K. Odeh, Copyright in the age of artificial intelligence: Navigating access to algorithmic training materials and the three-step test for text and data mining in Nigeria, The Journal of World Intellectual Property 28(2)(2025) 428–470.
https://doi.org/10.1111/jwip.12342
[18] H. Zhang, H. Wang, Copyright or Technology First: Trends and Implications of France's Response to AIGC's Copyright Risks, Editorial Friend (5)(2024) 103–112.
https://doi.org/10.13786/j.cnki.cn14-1066/g2.2024.5.014