International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 10 Issue 4, Jul-Aug 2026 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Challenges Faced by Generative AI Platforms in Determining Judicial Liability and Corresponding Institutional Measures to Address Them

Chen Han, Gao Zixin

School of Law, Beijing Wuzi University, Beijing, China

ABSTRACT

Generative artificial intelligence has now been widely adopted for commercial use, giving rise to platform-related infringement issues that pose significant legal challenges, including ambiguity regarding the identity of the liable party, difficulty in determining fault, and uncertainty about how responsibilities should be allocated. This article summarizes and organizes typical AI infringement cases observed from 2023 to 2026, and elucidates the key points of contention in courts' assessments of three critical aspects: the nature of the platform, the standard for establishing fault, and the allocation of multi-party liability. By analyzing the European Union's Regulation (EU) 2024/1689 of the European Parliament and the Council-aimed at establishing harmonized rules for artificial intelligence and revising the classification methods for risks set forth in several regulations and directives-we propose a tiered duty-of-care system anchored on two fundamental criteria: the platform's "control" and "benefit." For high-risk content in sectors such as news, healthcare, and finance, prior processing or human review should be mandatory. Moreover, if a platform independently edits or recommends AI-generated content, it will be deemed to bear strict liability. The framework proposed here is both judicially applicable and sufficiently specific, striking a balance between safeguarding users' rights and fostering technological advancement.

KEYWORDS: *Generative AI; Platform Liability; Internet Service Provider; Duty of Care; Judicial Determination.*

How to cite this paper: Chen Han | Gao Zixin "Challenges Faced by Generative AI Platforms in Determining Judicial Liability and Corresponding Institutional Measures to Address Them" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-10 | Issue-4, August 2026, pp.1327-1336, URL: www.ijtsrd.com/papers/ijtsrd142178.pdf

IJTSRD142178

Copyright © 2026 by author (s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (http://creativecommons.org/licenses/by/4.0)

RAISING THE ISSUE

A. Real-world context

Generative AI-technologies such as ChatGPT, DeepSeek, and Doubao-has become a commonplace tool for college students and even new-media professionals. As for generative AI itself, its scope of application is expanding rapidly, and the number of users is growing swiftly. The issue of infringement involving AI-generated content has ceased to be merely a theoretical speculation and has now become a judicial reality. The "first case of AI-generated images," concluded in November 2023 by the Beijing Internet Court, was the first instance in which AI-generated images were included in the scope of analysis under copyright law. In February 2024, the Guangzhou Internet Court handed down the "Ultraman Case," which the industry regards as "the world's first case of generative AI platform infringement." Similarly, in July 2024, the Hangzhou Internet Court concluded the "Hanghu Lian Ultraman Case," appropriately distinguishing the obligations of various types of large-model platforms. Meanwhile, the first nationwide case of "hallucination" infringement involving generative AI, concluded in 2025, addressed the question of liability when generative AI produces false information. Currently, cases involving infringement related to generative AI are on the rise, making it increasingly urgent to clarify the nature of platforms, establish standards for fault determination, and define the responsibilities of all parties involved. In September 2025, the Beijing Internet Court, in releasing statistics on AI-related cases, noted that "the complexity of technological applications makes fact-finding difficult; insufficient adaptation of rules leads to difficulties in legal application; and the diversity of actor roles

complicates liability determination." These represent the three major challenges currently faced in adjudicating cases involving "AI." Ordinary users find themselves caught in a dilemma: they're unclear about whom to sue, unsure how to gather evidence, and uncertain about exactly what responsibilities platforms should bear-a situation that leaves them struggling to protect their rights.

B. Regulatory Background

The "notice-and-takedown" and counter-notice rule system under the Civil Code of the People's Republic of China is premised on the platform being treated as a "network service provider" (Articles 1195–1196). The "Interim Measures for the Administration of Generative Artificial Intelligence Services" were jointly issued by seven departments, including the Office of the National Internet Information Office, on July 7, 2022, and came into effect in August 2022.

Effective from the 15th of the month. Articles 4 through 9 and Article 14 specify the obligations that "providers" must fulfill, and classify AI-based services as falling within the scope of online information content producers. They also introduce systems for security assessments, algorithm archiving, and complaint handling. However, the civil liabilities involved are not explicitly defined; the provisions themselves only indirectly link elements such as algorithm filing and regulations on the management of deep synthesis in internet information services to the "Regulations on the Management of Algorithmic Recommendation Services for Internet Information Services."

The "Method for Identifying AI-Generated Synthetic Content" was jointly issued on March 7, 2022, by the Office of the National Internet Information Office, the Ministry of Industry and Information Technology, the Ministry of Public Security, and the National Radio and Television Administration. The accompanying mandatory national standard was also released simultaneously. The two sets of regulations are scheduled to take effect on September 1, 2022. The management measures outlined in Document 56 provide a relatively comprehensive explanation of the explicit or implicit labeling requirements for content involving artificial intelligence, forming a closed-loop governance system based on the approach of "source identification-distribution review-dissemination verification-user declaration." In line with industry regulatory logic, the National Administration of Financial Regulation issued again-in June 2026, June 2018, and June 2018-the "Guiding Opinions on the Safe Development and Application of Artificial Intelligence in the Banking and Insurance Sectors" to various relevant institutions, proposing full-lifecycle management of artificial intelligence, conducting security reviews of data and algorithms, and ensuring the implementation of ethical responsibilities.

Regarding the filing system, in accordance with the "Interim Measures for the Administration of Generative AI Services," the cyberspace administration authorities and other relevant departments have been conducting filing procedures. As of June 30, 2026, a total of 988 generative AI services had completed their filings. Among these, 72 new filings were registered from March to April 2026; during the same period, local cyberspace administration offices directly filed additional models through API interfaces or other means, resulting in 49 newly filed models. Another 120 new filings were completed from May to June 2026, bringing the total number of newly filed models to 68. On July 15, 2026, the Cyberspace Administration of China publicly disclosed, for the first time, information on 7 AI service filings, including generative AI services available on mobile devices. The services that have completed their filings include Apple Intelligence, Huawei Xiaoyi AI Large Model, OPPO ANDESGPT Large Model, vivo Blue Heart Terminal Large Model, Xiaomi Pengpai AI, Samsung Galaxy AI, and Nubia DouBao Mobile Large Model. Industry insiders believe this marks the official "pass" for generative AI services delivered on the mobile device side.

The regulatory framework mentioned above has rapidly established an institutional backdrop for the operation of generative AI platforms; however, regulatory gaps have yet to be fully eliminated. From a legal perspective, the content "generated" by AI must be fundamentally distinguished from the traditionally understood "stored/transmitted" content-where the former is a product autonomously created by the platform itself, while the latter consists of information merely relayed by the platform. Some scholars argue that, in terms of the duty of care owed by the actor, this distinction constitutes the essential criterion for determining whether an infringement has occurred. Yet, the current Copyright Law and its related higher-level statutes still fall short of providing a comprehensive explanation on this point. As a result, judicial practice often resorts to drawing analogies from rules applicable to online services or interpreting them by analogy.

C. Core Issue Awareness

The central concern of this article is whether the traditional assumptions about platform liability still hold true for today's AI platforms-if, under the guise of AI, these platforms no longer view themselves as mere "channels" following the old internet logic, but instead operate as "generators" or "co-creators." If

these assumptions no longer apply, how should courts respond, and what institutional gaps remain? These are the fundamental questions this article seeks to clarify.

THE THEORETICAL PREMISE FOR THE LIABILITY OF GENERATIVE AI PLATFORMS-FROM "CHANNEL" TO "GENERATOR"

A. Logical Basis of Traditional Platform Liability

The safe harbor rules for "Web Service" are based on technological neutrality, passive transmission, and notice-and-takedown. From a legal perspective, platforms do not have "actual control" over third-party content and thus should bear limited liability. For Web 2.0, the above logic generally holds true-after all, the technologies involved largely consist of storage, retrieval, or linking, while the actual content is mostly information provided by users themselves.

B. The Technical Logic and Legal Qualification Dilemma of Generative AI

Content generation using generative AI operates through three key stages: "input (user prompts) → training (vast amounts of data) → output (AI-generated content)." Regarding the "de facto control over the content produced by these platforms," there remains a critical debate within academia. From Hangzhou Internet Court, according to the court's analysis of the "hallucination" case involving AI, it is inherently difficult for generative AI services themselves to fully anticipate or effectively control the information they generate. On the one hand, content generation can be highly managed through mechanisms such as review, filtering, fine-tuning, and output restrictions; on the other hand, the unpredictability and black-box nature of generated content limit the extent to which platforms can exert control. These challenges pose significant dilemmas in the legal classification of AI platforms. Gao Jingfeng has analyzed the issue of judicial responsibility involved in digital prosecution, arguing that modern technologies such as big data, blockchain, and artificial intelligence are highly beneficial for supervising and handling cases, and their principles of responsibility should be clearly defined. The "bundle-based" approach-integrating both the process of action and its resulting consequences-holds methodological significance in resolving the contradiction between "uncontrollable processes" and "accountability for outcomes" in the context of generative AI platforms. As weak artificial intelligence rapidly evolves toward strong AI, the notion of "technological value neutrality" no longer suffices to address real-world needs. This judgment also applies to scenarios involving the liability of generative AI: the freedom and uncontrollability of AI-generated content pose fundamental challenges to the traditional defense of "technological neutrality."

In his discussion of "Algorithmic Governance," Tang Linyao also argues that algorithmic activities-governed by the logic of "fiduciary duty" or assessed for damages based on the principle of contractual relativity-are currently unable to fully enumerate all the obligations that intelligent agents should fulfill, leading to a situation in which algorithmic manipulation, information rent-seeking, and regulatory arbitrage are simultaneously spiraling out of control. The attribute of "public nuisance," highlighted by Tang Linyao, precisely underscores the point that generative AI platforms should not rely solely on contractual relativity as the sole standard for liability-after all, the content generated by these platforms has broad dissemination potential and exerts an impact on non-designated third parties that far exceeds the scope of responsibility typically borne by ordinary online service providers. As for acts of "public nuisance" inflicted on the public after breaking through the technological black box, the algorithm itself constitutes a public phenomenon; therefore, the law must design corresponding regulatory requirements tailored to each responsible entity. The conclusions drawn from this analysis can serve as a theoretical foundation for the diversification of liability applicable to generative AI platforms.

(EU) With regard to the relevant parties, Regulation (EU) No 2024/1689 of the European Parliament and of the Council establishes harmonized rules on artificial intelligence and amends certain provisions and directives contained in the AI Act (Regulation (EU) 2024/1689-AIACT). The regulation adopts a risk-based approach, assigning AI systems to one of four risk levels: unacceptable, high, limited, and minimal. The underlying principle is "risk-based," with classification and handling of AI systems according to the actual risks they pose.

C. The Binary Tension in Platform Identity Determination

The identification of the status between "technical service providers" and "content producers/co-producers" serves as the logical starting point for all subsequent liability analyses. This tension has already been fully evident in judicial rulings-different courts and different cases have reached inconsistent conclusions regarding the nature of platforms.

TYPOLOGICAL ANALYSIS OF JUDICIAL PRACTICE BASED ON TYPICAL CASES

A. Source and Selection Criteria for Cases

The typical cases cited in this article focus on incidents occurring between January 2023 and June

2026. The primary data sources include the Judgment Documents Network, Peking University LawInfo, white papers related to the three major Internet courts, and typical case examples from their official websites. Cases involving the use of generative AI platforms and questions regarding whether such platforms bear responsibility are the key focus of our selection. Given the limited number of publicly available judgment documents, this article primarily relies on in-depth analyses of representative cases, supplemented by theoretical discussions.

B. Case Typification Presentation

a. Copyright Infringement Type

For relevant cases, please refer to those cited in the "First Case Involving AI-Generated Images," as published by Beijing Wangyuan. The plaintiff generated images using the Stable Diffusion model and then published them without obtaining authorization from the defendant, who also removed the watermark. This case primarily addresses whether AI-generated content can be considered a work of authorship and whether platforms should be deemed "participants in the creative process." According to the court's ruling, the plaintiff created the relevant images in a distinctive manner, meeting the criteria for works under copyright law.

The Guangzhou Online University's "Ultraman Case" has sparked heated debate over platform liability. Images strikingly similar to the "Ultraman" character were generated by the defendant's AI platform. Based on this, the court ruled that the defendant had committed an infringement-specifically, improperly exercising the plaintiff's rights of reproduction and adaptation.

The 2024 "Hangzhou Internet Court 'Hanghu Ultraman Case'" provides more specific details. The court held that the generative AI services provided by the defendant company should be regarded as a potential threat to others' rights of online information dissemination, and the defendant company failed to take necessary precautions against such use. Failure to exercise reasonable care constitutes contributory infringement. After examining various large-model platforms separately, the article provides a relatively comprehensive summary of the duty of reasonable care applicable to application-oriented AI services.

b. Infringement of Personality Rights and Infringement by False Information

The nation's first case involving infringement due to AI-generated "illusions" (handled by the Hangzhou Internet Court in 2025). In July 2025, Liang used a certain generative AI tool to inquire about admission information for universities. The AI provided inaccurate information and promised to compensate 100,000 yuan if the content turned out to be incorrect. At this point, the case primarily revolved around whether the AI's alleged "promise" constituted an expression of intent by the platform and whether the platform had been at fault. The court ruled that AI is not a civil subject and its "promise" should not be regarded as an expression of intent by the service provider. As for the generative AI involved in the case, it should be treated as a service and analyzed according to the principle of fault liability. The defendant had clearly disclosed the relevant risks and employed retrieval-augmented generation (RAG) technology as a technical supplement; thus, the defendant had properly fulfilled its duty of care and did not commit infringement. Some commentators noted that this case provides a relatively comprehensive discussion on the attribution of liability for infringement caused by false information generated by generative AI, with its liability framework both inheriting traditional principles and adapting to new circumstances. It should be noted that the requirement mentioned in the judgment-that risks must be conspicuously indicated-can be systematically compared with the explicit labeling provisions set forth in the "Measures for Identifying AI-Generated Synthetic Content," which will take effect starting September 2025. According to these measures, service providers shall use conspicuous labels at appropriate locations when generating various types of synthetic content (whether textual or multimedia).

c. Platform Handling Mechanism-Controversial Type

In the case "First Case on Platform Determination of AI-Generated User Content" (2025), the court explicitly stated that, in disputes involving the identification of AI-generated synthetic content, both users and platforms bear their respective burdens of proof. The case analyzed the institutional challenges encountered by platforms when using algorithmic tools to determine whether content was AI-generated. Meanwhile, several precedents from courts in Shanghai have shown that parties may use AI-based methods to draft appeal briefs and cite hypothetical cases.

C. Consensus and Disagreements in Judicial Views

Whether a platform should be exempt from liability depends on an analysis of the duty of care required under the circumstances. In the case of generative AI, it should be treated as a service provider and subject to the principle of fault-based liability.

The main points of disagreement lie in the following three areas: First, there is currently no unified view on

the nature of the platform. While the Guangzhou Internet Court regards the platform as a provider of network technology services, some academics argue that generative AI services are more akin to content service providers. Second, the standards outlined in the "Should Have Known" provision cannot be applied uniformly; specifically, the extent to which platforms should audit the legality of training data remains unresolved. Third, there is no uniformity in the forms of liability-issues such as whether liability should be based on direct infringement, aiding infringement, or apportioned liability remain unsettled.

THE THREE MAJOR CHALLENGES IN DETERMINING JUDICIAL ACCOUNTABILITY

A. Difficulty No. 1: The determination of the platform's legal nature and the applicability prerequisites of the "notice-and-takedown" rule remain uncertain.

The traditional safe harbor rules were designed on the premise that platforms "do not actively participate in content creation." However, today's generative AI platforms exert a relatively deep level of intervention in content formation by means of model training, parameter tuning, and sophisticated output filtering. From the perspective of judicial practice, there are two main approaches: The conservative approach holds that platforms remain mere providers of technological services; whereas the more progressive view regards both the platform and the generated content as exercising "substantial control."

The implementation of the filing system provides a new regulatory reference for determining the nature of platforms. For generative AI services that have been filed, it is required to indicate the model name, filing number, or online launch number, and to publicly disclose the status of such filed or registered services in a prominent location or on the product details page. The very process of filing implies that the platform will be brought under the regulatory scope of "service providers." However, in civil tort law, whether a "service provider" should be regarded as a "network service provider" or a "content provider" still needs to be determined on a case-by-case basis.

The view expressed in the article is that it is inappropriate to treat ISPs and ICPs as mutually exclusive categories. As illustrated by the "Hangzhou Ultraman Case," the court has differentiated among various types of large-model platforms and assigned different responsibilities accordingly. In analyzing judicial liability, Gao Jingfeng once proposed a methodological approach: judicial liability-and its corresponding accountability-should be understood as a combination of behavioral responsibility and outcome-based responsibility, rather than being treated as discrete, "block-like" entities. Instead, such liability should be categorized as "bundle-like." In short, the degree of control platforms exert over each stage of content generation is not always fixed; accordingly, the forms of liability should also be dynamically arranged along these lines, and one should avoid forcibly fitting them into pre-established, rigid patterns. This approach is particularly relevant to generative AI.

The same principle of liability applies-platform liability cannot be simply categorized as that of a "technical service provider" or a "content producer." Instead, a dynamic standard must be established based on the degree of control and the extent of benefit received, thereby enabling a nuanced allocation of liability in the form of a "bundle of responsibilities."

B. Difficulty No. 2: The standard for determining fault is vague-specifically, the boundary between "knowing or ought to know" is unclear.

In the context of Scenario A, the provision in Article 119 of the People's Republic of China's Code-"knowing or ought to know"-is relatively difficult to specify concretely. There are currently three key points that need to be discussed: first, whether infringement of training data should be treated separately from infringement of output content; second, after a platform has properly implemented the input of infringement warnings, whether it "ought to know" that the content it sends might still infringe upon rights; and third, whether current technology is sufficient to effectively carry out infringement-filtering tasks. In the AI "hallucination" case, the Hangzhou Internet Court established an important benchmark: the duty of care owed by service providers is subject to a dynamic adjustment process. If a service provider has fulfilled its legal duty of care, its duty of care in social interactions, and its contractual duty of care, yet the generated information remains inaccurate, it generally cannot be deemed to have been at fault.

The "Measures for Identifying AI-Generated Synthetic Content" provide a new normative interpretation regarding the determination of fault. The document incorporates both explicit and implicit identification methods into its regulatory framework, stipulating that service providers must "mark" content at the source and that distribution services must verify the information involved based on these identifiers. It follows from this that whether or not the

identification obligation has been fulfilled is likely to be a key factual element upon which courts rely when determining whether a platform has “exercised reasonable care.” Tang Linyao’s view is that governing public harm caused by algorithms should encourage greater transparency and comprehensibility of the algorithms themselves; however, tracing the responsible parties and allowing audiences to actively opt out represent even more fundamental safeguards. The relevant identification system focuses on the traceability of the chain of responsible parties, thereby providing the technical prerequisites for establishing fault.

The "Measures for Identifying AI-Generated Synthetic Content" provide a new normative interpretation regarding the determination of fault. The document incorporates both explicit and implicit identification methods into its regulatory framework, stipulating that service providers must "mark" content at the source and that distribution services must verify the information involved based on these identifiers. It follows from this that whether or not the identification obligation has been fulfilled is likely to be a key factual element upon which courts rely when determining whether a platform has “exercised reasonable care.” Tang Linyao’s view is that governing public harm caused by algorithms should encourage greater transparency and comprehensibility of the algorithms themselves; however, tracing the responsible parties and allowing audiences to actively opt out represent even more fundamental safeguards. The relevant identification system focuses on the traceability of the chain of responsible parties, thereby providing the technical prerequisites for establishing fault. With regard to generative AI in the output stage, given its high degree of autonomy or lack of controllability, the principle of strict liability should be applied to hold service providers accountable. However, strict liability itself carries the risk of stifling innovation-especially considering that the so-called generative AI currently exhibits inherent technological limitations such as the "black-box" phenomenon (a limitation not intentionally tolerated by the platform). From a practical operational perspective, it would be more feasible to adopt different standards of liability for the input and output stages, respectively. Moreover, some scholars argue that it is inappropriate to classify generative AI under the scope of product liability; instead, tortious acts involving generative AI should be addressed by taking into account both violations of safety obligations and considerations of fault and risk simultaneously. The rationale behind this approach stems from an understanding of

The affirmation of the reality that AI systems "pose risks" does not negate the flexibility of the fault element; rather, it allows courts to flexibly determine the degree of care required based on the maturity of the technology. Some scholars further argue that if an AI system itself inherently generates risks due to its intrinsic characteristics-such as the phenomenon of "hallucinations"-then the service provider deploying such a system actually has the capacity to establish control mechanisms and should, accordingly, fulfill its duty of care by striking a balance between benefits and risks. This principle of balancing is a practically feasible approach that provides courts with a relatively objective benchmark for determining whether a platform has “failed to take feasible technical measures.” It prevents the duty of care from becoming an aimless moral requirement devoid of practical relevance. The discussion treats the "reasonable technical preventive obligation" as a supplementary criterion for determining fault-aiming neither to impose excessive liability on platforms nor to allow them to evade their obligations by invoking technological neutrality. The duty of care owed by platforms should be assessed in light of their actual control capabilities and must not be unduly restricted. The guiding principles outlined in the "Guidelines on the Safe Development and Application of Artificial Intelligence in the Banking and Insurance Sectors"-namely, “adhering to the principle that whoever uses AI is responsible” and "improving business management processes for human-machine collaboration and scientifically defining the functional boundaries of AI"-can both serve as useful frameworks for considering liability issues generally associated with AI (i.e., platform liability).

C. Difficulty No. 3: In the dilemma of allocating responsibilities among multiple parties, the chain of "who is responsible for AI" has broken down.

There is a lack of a clear mechanism for allocating responsibilities among the three parties: foundational models, application platforms, and end users. When open-source models are subject to secondary development by third parties, there is a risk of producing infringing derivatives. How responsibility should be allocated under the API interface invocation model represents a typical scenario.

The Hangzhou Internet Court once issued a statement regarding the "Hanghu Ultraman Case": When users upload training materials, the platform itself cannot be considered directly infringing. However, if the platform fails to exercise reasonable care, it may be deemed to be aiding and abetting infringement. Yet, as of now, there is still no definitive conclusion on the

responsibilities that model developers and platform operators should bear. Some scholars have pointed out that generative AI involves the joint participation of multiple actors in the value-creation process. Therefore, applying traditional approaches to determining online infringement liability-under which service providers are held fully responsible-may not necessarily be fair. Instead, we should consider infringement issues within the broader value chain of generative AI, focusing particularly on the three key components: the foundational model, the system, and the service itself. The proposed approach takes responsibility allocation as its central focus, redefining "types of actors" as "links in the value chain." This way, the parties involved in the foundational model, system design, and actual application would each assume responsibility according to their respective levels of control, rather than treating the end-user platform as the sole party accountable.

The "API Interface Registration" under the filing system provides a preliminary institutional solution to the issues involved. As for the filing announcement itself, all generative AI applications-whether directly using the APIs of already-filed models as interfaces or employing other methods-should be filed and registered with the local cyberspace administration offices. The "filing plus registration" arrangement here is designed to clarify the responsibilities of both the foundational models and their downstream applications in a two-tiered approach: the filer and the registrar each bear their own responsibilities. However, how these responsibilities are specifically allocated in civil law still requires further clarification through judicial interpretation based on actual cases.

Some scholars have argued that providers of generative AI services-not merely as mere providers of online technical services-possess both the technical capability and legal obligation to ensure the authenticity and accuracy of the generated content. Consequently, they should not simply rely on the notice-and-takedown or knowledge-based rules applicable to online infringement to evade responsibility; rather, they must assume liability for their own actions according to the principle of fault-based responsibility. The view that platforms cannot evade responsibility by invoking "technological neutrality" is indeed well-founded. However, the notion of "technical control" employed in this argument is overly broad: Does the platform's control over content primarily involve macro-level adjustments to the model during training, or does it entail micro-level filtering at the output stage? These two approaches differ significantly in terms of their effectiveness and the degree of predictability they offer. If we were to unthinkingly dismiss the applicability of notice-and-takedown rules solely on the grounds that a platform "exerts control," it would compel platforms to excessively censor content merely to avoid liability, thereby stifling the vitality of generative AI. Therefore, a more reasonable stance would be to affirm that platforms do have a duty of care, yet not to completely abolish the notice-and-takedown rules. Instead, the requirements for responding to notices should be determined based on the actual strength of the platform's control.

By drawing on the EU's AI Act's two-part framework of "provider-deployer," and taking "control" and "benefit" as the fundamental criteria for allocating responsibilities, this analysis examines China's current situation. Tang Linyao argues that, at a stage when intelligent robots have not yet acquired independent legal personality, long-arm jurisdiction rules, fiduciary duties, and general tort law can adequately meet the various legal requirements arising from algorithmic direct harm to parties under contract law. This argument is highly explanatory for cases involving "direct harm"; however, its focus remains largely on contractual relativity, leaving it insufficiently equipped to fully address "indirect harm" or "public nuisance" caused by algorithms and directed toward unspecified individuals. A typical example of infringement in generative AI is the widespread output that simultaneously affects a large number of users-information fabricated by an AI may be viewed by tens of thousands of users, yet these users have no contractual relationship with the platform. Therefore, the two criteria adopted in this paper-"control" and "benefit"-can effectively complement and rectify the shortcomings in Tang Linyao's approach: If a platform indeed exercises control over content and derives benefits from it, even in the absence of a contractual relationship, it should still bear responsibility for any harm inflicted upon unspecified third parties. This represents a newly established "public liability" mechanism beyond traditional contract logic.

Zhang Linghan and Yu Lin propose analyzing tort liability based on the value chain of generative AI, identifying three fundamental stages-foundation model development, system integration, and service application-and prescribing appropriate duties of care for each of the involved parties. They suggest addressing issues according to a fault-based approach, while allowing equitable liability as a supplementary measure in cases involving model "hallucinations." This proposal provides a relatively clear explanation of how liability should be allocated, shifting the focus

of discussion from "who the platform is" to “what the platform has done.” The proposed value-chain approach is indeed valuable. However, the suggestion to supplement the treatment of model "hallucinations" with equitable liability still requires careful consideration: equitable liability can only be invoked when neither party is at fault. Since "hallucinations" themselves are inherent technical flaws, it is inappropriate to automatically attribute fault to the platform simply on the basis of its design, material selection, or control over outcomes. Relying solely on equitable liability to substitute for fault determination would undermine the court’s need to conduct a thorough examination of the platform’s technical negligence, rendering the liability analysis overly formalistic. A more reasonable approach would be to generally apply fault-based liability, reserving equitable liability as a supplementary measure only when the platform has exhausted all feasible technical measures yet remains unable to prevent “hallucinations.” Moreover, the application of equitable liability should be strictly limited.

THE PATH TO INSTITUTIONAL IMPROVEMENT: FROM "UNIFIED RULES" TO "HIERARCHICAL AND CATEGORIZED" SYSTEMS

A. Application and Limitations of the Current Rules

Article 14 of the "Interim Measures for the Administration of Generative Artificial Intelligence Services" stipulates that providers, upon discovering illegal content, shall promptly take measures such as ceasing generation, halting transmission, and removing the content. A limitation of these measures is that they only prescribe administrative obligations without clearly defining civil liabilities; moreover, they fail to distinguish between the obligations of foundation model providers and those of vertical application providers. Some scholars point out that the public law obligations prescribed by these measures, together with the contractual obligations set forth in user agreements, can serve as a basis for generative AI service providers to fulfill their duty of care in ensuring safety and security.

The "Measures for the Identification of AI-Generated Synthetic Content" have established a collaborative governance loop encompassing “source identification-distribution review-dissemination verification-user declaration.” However, these measures primarily regulate identification obligations under administrative law, and civil liability provisions remain absent. The "Guiding Opinions on the Safe Development and Application of Artificial Intelligence in the Banking and Insurance Sectors" put forward 32 guiding recommendations covering aspects such as governance architecture, development and application, data governance, computing power infrastructure, risk management, capability enhancement, and safeguards and supervision; yet their scope of application is limited to the financial industry. Zhang Linghan, Yu Lin. Allocation of Infringement Liability in the Value Chain of Generative Artificial Intelligence [J]. Journal of Beijing Administrative College, 2025(5).

B. Drawing on and Adapting from Overseas Experiences

In terms of overseas experience, Regulation (EU) 2024/1689-the Artificial Intelligence Act-adopted by the European Parliament and the Council of the European Union, which establishes harmonized rules for artificial intelligence and amends several regulations and directives, entered into force on August 1, 2024, as scheduled. It is currently the world’s only comprehensive piece of legislation that provides a holistic regulatory framework for artificial intelligence. The Act adopts a "risk-based" approach to categorize and regulate AI systems, placing them into four risk categories: unacceptable risk, high risk, limited risk, and minimal risk. AI practices involving unacceptable risks have been prohibited since February 2, 2025; social scoring and real-time biometric identification are among the high-risk scenarios explicitly listed.16 Obligations for general-purpose AI models will take effect starting from August 2, 2025.17 As for high-risk AI systems, key compliance requirements were originally scheduled to be fulfilled by August 2, 2026, but several of these requirements have now been postponed until December 2027. The Act also provides a more detailed classification of responsible parties: “providers of general-purpose AI models,” “system providers,” and "system deployers" are distinct roles. Moreover, in cases where a "role conversion" occurs-meaning that a deployer is treated as a provider-such entities assume corresponding responsibilities. With regard to generative AI, Article 50 of the Act further clarifies transparency requirements, stipulating that providers must embed technically identifiable, machine-readable markers in the outputs of their AI systems. The very design of this risk-based classification system, coupled with the layered approach to defining responsible parties, offers valuable insights for China’s consideration of tiered duty-of-care obligations. From the perspective of liability attribution, some scholars have proposed that generative AI services should be subject to product liability. using the concept of product defects to bridge the gap between tort liability and regulatory oversight.

Traditionally, the United States has relied on Section 230 of the Communications Decency Act of 1996 as a basis for granting online platforms immunity from liability. The underlying logic is that platforms "are not considered publishers or speakers merely because of information provided by third parties." However, with the emergence of generative AI, this logic is now beginning to face fundamental challenges. Chatbots based on transformers engage in creative activities that are neither replicative nor referential-rather, they generate entirely new, organized content. As a result, these platforms "appear far from being neutral intermediaries; instead, they increasingly resemble authors producing original content." In terms of legislation, in June 2023, Senators Josh Hawley and Richard Blumenthal introduced a bipartisan bill-the "No Section 230 Immunity for AI Act"-which would exclude Section 230's immunity provisions from claims involving generative AI. Although the proposal did not pass the Senate,

Indeed, this policy stance has already been clearly articulated. From a judicial perspective, in the 2025 case of Garcia v. Character Technologies, Inc., the Middle District of Florida Federal District Court agreed to proceed with the trial along product liability lines, treating large language models as "products" rather than "services." As a result, it became easier to apply the strict product liability doctrine to AI19. Even if the exemption under Section 230 is not directly mentioned here, this determination that AI constitutes a "product" could potentially hinder future attempts to invoke Section 230. Thus, it is evident that current U.S. thinking is evolving from a position of "broad immunity" toward a "generative AI exception"-a scenario where complete immunity would fail to meet user needs, while overly stringent regulation would stifle innovation. Striking the right balance between these two extremes is precisely the central challenge that both legislative and judicial authorities must address.

C. Constructing a "Tiered Duty of Care" Framework

All generative AI platforms shall establish user complaint mechanisms, label AI-generated content, and retain records of generation activities. They shall also comply with the general duty of care and the "notice-and-takedown" rule. Providers of platform services shall use "industry-standard technical measures" as the benchmark for determining their duty of care, thereby ensuring that service functionalities meet the market's average standards. The explicit and implicit labeling obligations stipulated in the labeling guidelines shall serve as the fundamental duty of care applicable to all platforms. High-risk scenarios-such as news, healthcare, and finance-should bear heightened duties of care, including pre-emptive content filtering, human review, and verification of the legality of training data, and should be subject to a presumption of fault. The "Guiding Opinions on the Safe Development and Application of Artificial Intelligence in the Banking and Insurance Sectors," which call for "establishing and improving a full-lifecycle management system for artificial intelligence" and "implementing risk-based classification and tiered management of AI applications," can serve as a reference for setting standards of care in high-risk scenarios. When a platform proactively recommends or edits AI-generated content, it shall exercise substantial control over the output and bear strict liability or joint and several liability.

The hierarchical structure of liability attribution principles can be conceptualized by distinguishing between an input stage and an output stage: In the input stage, fault-based liability serves as the fundamental principle; in the output stage, a certain degree of recognition is given to the autonomy and uncontrollability inherent in AI, leading to a slight relaxation of liability standards. Currently, the filing system is progressing according to plan (as of June 2026, a total of 988 services have been filed). This represents the institutional foundation for the proposed tiered approach. The usage scenarios, model types, and service contents recorded in the filings can serve as clear reference points for courts when analyzing platform-level classifications and determining the appropriate duty of care.

D. Procedural Recommendations for Ordinary Users to Protect Their Rights

The fixation of evidence can be documented using screenshots as well as the AI-generated content, prompts, interfaces, and timestamps. Once the relevant identifiers have been adopted, users shall determine whether preliminary evidence is established based on the identifiers-whether explicitly or implicitly-contained in the generated content.

The methods used to protect rights include platform complaints, administrative reports (submitted to the Cyberspace Administration), and civil litigation.

The key points of evidence include objective circumstances demonstrating that the platform "should have known" or "knew" about the infringement, such as repeated complaints remaining unresolved or the platform actively recommending infringing content. The relevant registration information of platforms that have already been filed can be publicly accessed on the website of the Cyberspace Administration, and can serve as

foundational materials for determining the competent court and identifying the defendant's identity.

CONCLUSION

A. Summary of Key Points

Generative AI platforms cannot be hastily categorized under the established classifications of "internet service providers" or "content producers." The judicial concept of a "one-size-fits-all safe harbor" is gradually shifting toward a more nuanced, "classification-based approach," yet regulatory resources in this area remain insufficient. The analysis conducted by the Hangzhou Internet Court in the AI "hallucination" case-defining obligations through fault-based liability and dynamic due diligence-is a highly valuable reference for judicial methodology. Gao Jingfeng's approach of limiting judicial liability assessments to a "stringent, layered" rather than a "block-like" framework provides an excellent starting point for thinking about the liability issues surrounding generative AI: the responsibilities of each stage and each actor must be considered separately and implemented item by item; it is inappropriate to lump all responsibility together under a single, uniform model. The scaled-up implementation of the filing system-by June 2026, 988 services had already completed their filings-is an important institutional preparation upon which judicial liability determinations rely. The explicit and implicit identifiers listed in the relevant labeling guidelines can serve as a technical starting point or foundation for assessing fault. It would be advisable to use the "Interim Measures for the Administration of Generative AI Services" to regulate the issue of "tiered duty of care" in judicial proceedings, bringing the foundational models, system integration, and the services themselves all within the scope of liability analysis and defining or attributing responsibility accordingly at each respective stage. References

REFERENCES

[1] Beijing Internet Court (2023) Jing 0491 Min Chu 11279 No.

[2] Guangzhou Internet Court (2024) Yue 0192 Min Chu 113 No.

[3] Hangzhou Internet Court's "Hanghu Ultraman Case" (2024)

[4] Hangzhou Internet Court AI "Hallucination" Case (2025)

[5] The National Administration of Financial Regulation's "Guiding Opinions on the Safe Development and Application of Artificial Intelligence in the Banking and Insurance Sectors"

[6] National Internet Information Office, "Announcement on the Publication of Filing Information for Generative Artificial Intelligence Services (March to April 2026)"

[7] National Internet Information Office, "Announcement on the Release of Filing Information for Generative Artificial Intelligence Services (May to June 2026)"

[8] The Cyberspace Administration of China's "Announcement on the Publication of Filing Information for 7 Mobile-End Generative AI Services"

[9] Gao Jingfeng. Practical Approaches for Deepening the Identification and Accountability of Judicial Responsibility within the Prosecutorial Organs [J]. Journal of the National Academy of Prosecutors, 2025(1).

[10] Tang Linyao. Algorithmic Governance and the Governance of Algorithms [M]. Beijing: Social Sciences Literature Press, 2024.

[11] Zhang Linghan, Yu Lin. Allocation of Infringement Liability in the Value Chain of Generative Artificial Intelligence [J]. Journal of Beijing Administrative College, 2025(5).

[12] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) [EU AI Act].

[13] European Commission. Commission Guidelines on Prohibited Artificial Intelligence Practices [EB/OL]. (2025-02-04). https://digital-strategy.ec.europa.eu.

[14] European Commission. Guidelines on the scope of the obligations for providers of general-purpose AI models established by Regulation (EU) 2024/1689 (AI) Act)[EB/ O L]. (2025-07-24). https://digital-strategy.ec.europa.eu.

[15] No Section 230 Immunity for AI Act, S. 1993, 118th Cong.(2023).

[16] Garcia v. Character Technologies, Inc., No. 6: 24-cv-01903 (M.D. Fla. 2025).