

On the Tort Liability Identification of Generative AI Service Providers

Ziyu Li

School of Law, Tianjin University of Finance and Economics, Tianjin, China
liziyu20051121@stu.tjufe.edu.cn

Abstract. Generative AI service providers should be positioned as indirect infringement subjects, not directly subject to product liability or traditional "safe harbor" rules. Liability attribution should be based on fault liability, with fault presumption introduced in scenarios such as the "algorithm black box" where decision-making processes are opaque. The duty of care must be dynamic, requiring a "type + hierarchy" system that distinguishes between basic obligations applicable to all providers and strengthened obligations for high-risk scenarios. The traditional notification-removal rule should be reshaped into a "notification-shielding" mechanism aimed at "preventing re-generation" of infringing content rather than mere post-hoc removal. Additionally, protective norms in public law should be identified to establish a linkage mechanism between public and private law. This approach balances technological innovation with rights protection, ensuring legal responses remain adaptive to generative AI's unique challenges while maintaining accountability and fairness. A reasonable duty of care applies across different service provider types and application scenarios, promoting responsible innovation without overburdening providers. The proposed framework avoids imposing strict liability while ensuring effective protection for infringed parties through fault-based attribution and dynamic obligations tailored to specific risk profiles.

Keywords: Generative Artificial Intelligence, Service Providers, Tort Liability, Duty of Care, Notification-Masking

1. Introduction

The rapid development of generative artificial intelligence has profoundly changed the content production mode, but also brought a series of unprecedented legal challenges. As the large language model is widely used in various scenarios, incidents such as "AI illusion" generating false information and AI output content infringing copyright occur frequently, causing widespread concern in society. These practical problems force the legal research to face a core proposition: under what conditions and to what extent should the generative AI service provider bear the tort liability for the infringement of content generated by its system. Although this issue has been widely discussed, there are still obvious deficiencies in the specific structure of the duty of care and the connection between public law and private law norms.

Academic circles have reached a certain consensus on the legal positioning of service providers, which are generally defined as "new network service providers" that are different from traditional content providers and pure technology intermediaries. However, the consensus basically stops here. On the choice of imputation principle, the differences are still obvious: some scholars advocate adhering to the principle of fault liability, and worry that excessive liability will stifle the vitality of technological innovation; Some scholars also focus on the difficulty of proof caused by "algorithm black box", and advocate the introduction of fault presumption to balance the position of both parties. In the field of duty of care, although some scholars have carried out normative analysis from the perspective of risk control theory and the "rational person" standard, the specific content of duty of care in different risk scenarios is still vague, and the judgment standard of breach of duty has not yet formed a systematic construction. At the same time, the "notice delete" rule in the traditional safe haven system is limited by the characteristics of real-time dynamic output of generative AI - the key is no longer to delete a specific file, but how to prevent the re-generation of infringing content. At a deeper level, what is the relationship between the public law obligations set for service providers by administrative regulations, such as the Interim Measures for the administration of generative artificial intelligence services and the identification of civil tort fault? The convergence of public and private law norms also needs to be clarified.

The identification of tort liability of generative AI service providers needs a systematic integration analysis framework. In terms of methods, the comprehensive use of legal doctrinal analysis, case studies (focusing on typical judgments such as the "Ai illusion" case of Hangzhou Internet court and the "Altman" copyright case of Guangzhou Internet court), and comparative law for reference. Through these system designs, a reasonable balance is sought between encouraging technological innovation and protecting legitimate rights and interests.

2. Concept definition and problem discussion

2.1. Legal characterization of generative AI service providers

To determine the tort liability, the subject of liability must first be clarified. This subject is the "generative AI service provider". According to the definition of Article 22 of the Interim Measures for the administration of generative AI services, such subjects refer to "organizations and individuals that use generative AI technology to provide generative AI services (including providing generative AI services by providing programmable interfaces and other means)" [1] This definition includes both technology companies that develop large language models and applications that integrate AI capabilities into their own service platforms through API interfaces. In short, as long as the AI content generation service is provided to the public, it is within the scope of this definition.

As for the legal identity of service providers, scholars believe that service providers are neither traditional content providers nor pure technical service providers, but a "new type of network service provider". This judgment has two meanings.

First, the service provider is not a content provider. Traditional content providers bear direct responsibility for the content they publish, because the content is completely under their control. However, the content generated by AI is "unpredictable", and the service provider does not know what the model will output in advance. If service providers are allowed to bear the same strict responsibilities as traditional content providers, technological innovation will be stifled.

Second, service providers are not pure technical service providers. AI service providers are deeply involved in the content generation process, with subjective involvement in data training,

algorithm design and output optimization. Therefore, it cannot simply claim exemption by "technology neutrality".

Therefore, service providers are positioned as "indirect infringement subjects". This means that it does not bear direct liability for the infringement of content, but indirect liability for its failure to fulfill the reasonable duty of care. This positioning is the logical starting point for the subsequent discussion.

2.2. Application dilemma of the current legal framework

The question of the applicability of the current legal framework is whether the existing legal rules can be directly applied. Let's start with two core systems: product liability and the "safe haven" rule.

Consider first product liability. The judicial practice has given a negative answer to the question of whether AI can be regarded as a "product" to enable service providers to bear strict liability without fault. In the first domestic "Ai illusion" infringement case concluded by Hangzhou Internet court at the end of 2025, generative AI was clearly defined as "service" rather than "product". The court held that the generative AI system lacked standardized and verifiable quality indicators, and defining it as a "product" would hinder technological innovation [2].

Consider next the "safe harbor" rule. The "notification and deletion" rule stipulated in Article 1195 of the Civil Code is designed for traditional network service providers. This rule faces fundamental obstacles under the AI scenario. Some scholars pointed out that the "notice" rule can be applied by analogy in the infringement of generative AI works, but it needs to be modified in combination with the characteristics of AI technology [3]. First, AI infringement content is not stored statically on the platform, but generated in real time. When the obligee discovers the infringing content, the content may have been disseminated. Second, as for the object of "deletion", the infringement content on the traditional platform is a specific file, but the problem of AI infringement lies in the model itself. Even if the content generated this time is deleted, the infringement content may reappear next time the user asks a different question.

3. Problem analysis-legal perspective and practical dilemma

3.1. Choice of imputation principle

Fault liability or fault presumption is the focus of controversy at present.

Some scholars advocate the principle of fault liability and believe that AI is still in the early stages of development, and excessive responsibility will hinder the initiative of innovation. The judgment of the Hangzhou "Ai illusion" case also adopted this position.

Some scholars hold different views. Wutaixuan and Dengchaohui believe that in AI infringement, the obligee faces the difficulty of proof caused by the "algorithm black box". Training data, algorithm design, and other information are under the control of the service provider, and the obligee simply cannot get them. Therefore, fault presumption should be introduced [4].

These two views are not completely opposite. Presumption of fault can be used as a specific way of application under the principle of fault liability. In the general scenario, fault liability is applied, and the obligee provides evidence. In the high-risk scenario, where the obligee is unable to obtain internal information, the presumption of fault is applied, and the service provider proves that it has fulfilled its duty of care.

3.2. The vague duty of care

The duty of care is the core of fault identification. Some scholars pointed out that the legitimacy of the duty of care lies in the risk control theory and the interest balance theory. The judgment should be carried out from both subjective and objective aspects: the objective level should consider predictability, and the subjective level should reconstruct the "rational person" standard. From the perspective of comparative law, some foreign scholars have proposed that the fundamental challenge of AI to tort law is that the traditional "rational person" standard is difficult to apply to the black box decision-making system, and the theoretical basis of fault identification needs to be reconsidered [5]. The case of "AI hallucination" in Hangzhou provides a specific reference framework. The court analyzed the duty of care from three dimensions: first, the "result obligation" for illegal and harmful content; Second, the "obligation to inform" of functional limitations; Third, the "means obligation" for content accuracy does not require ensuring correctness, but requires the adoption of industry-wide technical measures [2].

This framework has not yet formed universal rules. The "rationality" determination of technical measures, whether the standards are fixed or dynamic, and the application of different standards to different service types, need to be addressed. From the perspective of interest balance, some scholars proposed that the standard of duty of care should be dynamically adjusted according to the control ability and technical possibility of service providers [6].

3.3. Acclimatization of notification rules

The notification rules are not acclimatized. The court required the defendant to take measures such as "keyword filtering" to prevent the subsequent generation of similar content [7]. The Altman case is an important precedent in this regard. The Guangzhou Internet court held that the AIGC service provider was liable for copyright infringement because it could generate pictures substantially similar to Altman's works.

This shows that it is necessary to rethink the connotation of "necessary measures". In the context of AI, the goal of necessary measures is not to "delete" the generated content, but to "prevent the re-generation of infringing content". Some foreign scholars have studied this issue from the perspective of comparative law, pointing out that different jurisdictions have taken different paths in dealing with generative AI tort liability, but they are trying to balance technological innovation and right protection [8].

3.4. Difficulties in the convergence of public and private law norms

Finally, there is a deeper problem: the relationship between public law obligations and private law responsibilities. The Interim Measures for the administration of generative artificial intelligence services set public law obligations for service providers, such as content identification, a complaint reporting mechanism, etc. There will be administrative penalties for violating these obligations, but whether they can be recognized as civil tort liability needs to be further classified. Some scholars pointed out that it is necessary to identify which public law norms are "protective norms", that is, their legislative purpose includes protecting the rights and interests of others. Only when these norms are violated can they be used as a reference for civil fault.

4. Identification path and suggestions

Based on the previous analysis of the three controversial focuses of the imputation principle, the duty of care, and the notification rule, the identification of the tort liability of generative AI service providers needs to be constructed at three levels. The first level is the entity standard, that is, what kind of duty of care the service provider should perform to be recognized as no fault; The second level is the procedural mechanism, that is, how the obligee effectively notifies and how the service provider reasonably responds; The third level is the source of norms, that is, whether and how obligations in public law can enter the fault judgment in private law. These three parts answer the three progressive questions of "how to identify the fault", "how to remedy the right," and "how to connect the norms", and jointly serve the general proposition of the identification of tort liability. These are described below.

4.1. Typed and hierarchical duty of care system

As for typology, it should be horizontal scene differentiation. This means that the duty of care is divided into different types according to the different scenarios and risk levels applied to the service, and corresponding to the obligation requirements of different intensities. It also contains two contents. One is the basic obligation, which is a universal obligation applicable to all service scenarios (such as content identification, complaint mechanism, and usage tips). This is the "prescribed action", and everyone must do it. The second is the enhanced obligation, which is a special obligation applicable to high-risk scenarios (such as medical, legal professional fields or involving minors). This is an "upgrade action", requiring service providers to assume a higher level of attention, such as algorithmic filtering, sensitive content intervention and other measures. As for hierarchy, it is a vertical standard floating. This means that the examination standard of the duty of care is not fixed, but dynamically adjusted with the level of technological development. It introduces a time variable to judge whether the service provider has fulfilled its obligations. The reference frame is "the industry technical level at the time of the behavior", rather than an unchanging absolute standard. Yesterday's "reasonable measures" may not be enough to exempt from responsibility today. This provides flexibility for fault identification and avoids the legal standard lagging behind technological progress.

The first layer is the basic obligation, which applies to all service providers. It mainly includes: content identification obligations; a complaint reporting mechanism; Use the reminder obligation to inform the user that AI may be wrong. Some scholars have systematically studied the problem of fault identification of service providers and proposed that judgment should be made from the two dimensions of foreseeing possibility and avoiding possibility [9].

The second layer is the enhanced obligation, which is applicable to high-risk scenarios (such as medical, legal and other professional fields or minor users). In high-risk scenarios, service providers need to assume a higher duty of care, including algorithm filtering, sensitive content intervention, etc.

The so-called hierarchy means that the standard of the duty of care should be dynamically adjusted according to the technical level. To judge whether the duty of care is fulfilled, it should be based on the "technical level of the industry at that time", rather than a fixed standard. Some foreign scholars analyzed the input materials of generative AI from the perspective of training data, providing theoretical resources for the hierarchical construction of the duty of care [10].

4.2. Construction of "Notification shielding" rule

It advocates changing from "Notification deletion" to "Notification shielding". The core is to redefine the goal of "necessary measures": from "removing the infringing content" to "preventing the regeneration of infringing content". A qualified "notice" should provide keywords that can "locate the infringing content", so that the service provider can identify and block the generation of similar infringing content. After receiving the qualified notice, the service provider should take the following necessary measures: blocking the output of infringing content in the current conversation; Add the prompt word to the filter library to prevent subsequent generation; When necessary, the model is optimized and trained.

At the same time, "consequentialism" should be used to determine whether responsibility exists: it is not required to ensure that the infringement will never happen again, but rather to examine whether "a reasonable and feasible scheme under the technical conditions at that time" has been adopted. As long as effective measures recognized by the industry are adopted, even accidental infringement should not be investigated for responsibility.

4.3. Connection mechanism between public law obligations and private law responsibilities

The aforementioned duty of care system and the "notice shield" rule solve the internal problems of "how to judge the fault" and "how to relieve the obligee" in the determination of tort liability. However, there is a more basic premise to clarify: the fact that service providers comply with or violate the norms of public law, and whether and how to enter the fault judgment in private law. This involves the normative connection between public law and private law, and is an unavoidable systematic issue of responsibility determination.

Some scholars analyze from the perspective of "Referral", and establish the connection mechanism by identifying whether public law norms belong to "protective norms". On this basis, a more specific differentiation scheme is proposed.

The first category is the norms for the purpose of protecting private interests. For example, the direct purpose of the content marking obligation in Article 12 of the interim measures is to protect users' right to know. Breach of this obligation can be used as the basis of a civil fault presumption.

The second category is the norms for the purpose of administrative management. For example, the algorithm filing obligation is designed to facilitate the regulatory authorities in grasping the situation. Violation of such obligations should not directly presume civil fault, but can be used as a reference factor for determining "due knowledge" or "failure to pay reasonable attention". This distinction not only respects the diversity of public law norms but also provides an operational path for private law relief.

5. Conclusion

Focusing on the core proposition of tort liability determination of generative AI service providers, this paper systematically analyzes the application dilemma and system outlet under the current legal framework. Generative AI service providers should be defined as indirect infringement subjects, and their liability determination must break through the traditional framework, and build an integrated system with the core of "layered fault Attribution-dynamic duty of care-notice shielding rules-public-private law convergence", so as to balance technological innovation and rights protection. The research is mainly based on the interpretation theory and system fine-tuning of the current legal framework, and the construction of the legislation theory of the tort liability of generative AI service

providers has not been fully carried out. Future research can further explore the differential configuration of different types of generative AI services (such as basic model development, API access, and terminal applications) in tort liability. The problem of causality identification caused by generative AI, especially when the damage is caused by multiple subjects (user tips, model training data, algorithm design), how to allocate the responsibility. The in-depth discussion of these issues will be helpful to the improvement of the generative AI tort liability system.

References

- [1] Yuan, X., Wang, H. (2025) Copyright law regulation of infringement content generated by artificial intelligence-- Also on the applicable basis and rule construction of "license infringement" in China. *Electronic Intellectual Property*, (03), 4-16.
- [2] King & Wood Mallesons (2026) Chinese AI Service Provider Found Not Liable for Generating AI "Hallucinations". Available at: <https://www.kwm.com/cn/en/insights/latest-thinking/chinese-ai-service-provider-found-not-liable-for-generating-ai-hallucinations.html>. (Web document, adjusted to target format style)
- [3] Wang, L., Bao, D. (2025) On the analogy application of the "notice" rule in the infringement of generative AI works. *Comparative Law Research*, (04), 1-13.
- [4] Wu, T., Deng, C. (2025) Comparison, selection and use of generative AI infringement imputation principle. *Electronic Intellectual Property*, (08), 4-18.
- [5] Abbott, R. (2024) *The Reasonable Robot: Artificial Intelligence and the Law*. Cambridge University Press.
- [6] Wu, C., Huang, J. (2025) Identification of copyright infringement of AIGC service providers from the perspective of interest balance. *Intellectual Property*, (03), 88-104.
- [7] Zhu, K. (2024) Research on copyright infringement liability of AIGC service providers. *Journal of Shanghai Normal University (Philosophy and Social Sciences Edition)*, 53(06), 39-49.
- [8] Richardson, M. (2024) Tortious Liability for Generative AI Outputs: A Comparative Perspective. *European Journal of Law and Technology*, 35, 1.
- [9] Xu, W. (2024) Identification of infringement fault of generative AI service providers. *Law*, (07), 110-124.
- [10] Sobel, B. L. W. (2023) A Taxonomy of Training Data: Disentangling the Legal Status of Generative AI Inputs. *Cardozo Law Review*, 44, 1549.