

生成式人工智能侵权责任的分层配置研究

——基于责任主体与归责原则的协同视角

胡 淼

青岛科技大学法学院, 山东 青岛

收稿日期: 2026年7月1日; 录用日期: 2026年8月3日; 发布日期: 2026年8月11日

摘 要

生成式人工智能技术的快速发展推动了内容生产方式的深刻变革,同时也引发了著作权侵权、人格权侵害、个人信息泄露及虚假内容传播等多元侵权风险。传统侵权责任制度以行为人实施侵权行为为前提构建责任体系,而生成式人工智能具有自主生成、技术链条复杂、主体协同参与以及风险来源多元等特点,使开发者、模型提供者、服务提供者、平台运营者及用户共同参与内容生成过程,传统以单一主体为中心的侵权责任体系面临适用困境。当前学界围绕生成式人工智能侵权责任形成了统一归责、无过错责任、类型化归责以及分层责任等不同理论路径,但现有研究多聚焦于归责原则、证明责任或特定主体责任配置,缺乏责任主体划分、归责原则选择与证明责任配置之间的体系化协调。基于此,本文从技术运行流程与风险形成机制出发,以责任主体分层为基础,以归责原则协同适用为核心,探讨兼顾技术创新与权益保护的生成式人工智能侵权责任分层配置体系,以期为我国生成式人工智能侵权责任制度的完善提供理论参考。

关键词

生成式人工智能, 侵权责任, 责任主体, 归责原则

A Study on the Layered Allocation of Tort Liability for Generative Artificial Intelligence

—From the Synergistic Perspective of Liability Subjects and Imputation Principles

Miao Hu

Law School, Qingdao University of Science and Technology, Qingdao Shandong

Received: July 1, 2026; accepted: August 3, 2026; published: August 11, 2026

文章引用: 胡淼. 生成式人工智能侵权责任的分层配置研究[J]. 争议解决, 2026, 12(8): 70-76.
DOI: 10.12677/ds.2026.128238

Abstract

The rapid advancement of generative artificial intelligence technology has profoundly transformed modes of content production, while simultaneously giving rise to diverse tort risks such as copyright infringement, infringement of personality rights, disclosure of personal information, and dissemination of false content. Traditional tort liability systems construct liability frameworks on the premise that an actor commits a tortious act. However, generative AI is characterized by autonomous generation, complex technological chains, collaborative participation of multiple subjects, and diverse risk sources, involving developers, model providers, service providers, platform operators, and users in the content generation process. Consequently, the traditional tort liability regime centered on a single subject faces difficulties in application. The current academic community has developed various theoretical approaches to tort liability for generative AI, including unified imputation, strict liability (liability without fault), typified imputation, and layered liability. Nevertheless, existing research predominantly focuses on imputation principles, the burden of proof, or the allocation of liability to specific subjects, lacking systematic coordination among the division of liability subjects, the choice of imputation principles, and the allocation of the burden of proof. On this basis, proceeding from the technical operation process and risk formation mechanism, this paper takes the stratification of liability subjects as the foundation and the synergistic application of imputation principles as the core, and explores a layered allocation system of tort liability for generative AI that balances technological innovation and the protection of rights and interests, with a view to providing theoretical reference for the improvement of China's tort liability regime for generative artificial intelligence.

Keywords

Generative Artificial Intelligence, Tort Liability, Liability Subjects, Imputation Principles

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来，以 ChatGPT、DeepSeek 等为代表的生成式人工智能进入大规模商业应用阶段，生成式人工智能逐渐从辅助工具演变为内容生产的重要参与者。与传统互联网平台相比，生成式人工智能不仅能够自主生成文本、图像、音视频等内容，而且能够通过持续学习不断优化生成能力，推动人工智能产业迅速发展。在生成式人工智能的赋能下，相关的行为和决策从人类转向智能系统，损害也开始从由人类和人类行为造成转向由智能系统及其行为造成，致使以过错为中心的责任追溯与分配机制失灵，传统侵权责任责任的归结认定范式面临结构性挑战。为应对这一挑战，迫切要求对生成式人工智能服务提供者侵权责任归责原则予以细致的分层化研究。

2. 生成式人工智能侵权责任分层配置的理论基础

传统侵权责任制度建立于“侵权行为 - 损害结果”的二元结构之上，侵权责任通常由直接实施侵权行为的主体承担。然而，生成式人工智能的运行过程具有显著的链条化特征，其内容生成涉及训练数据收集、模型训练、算法优化、模型部署、平台运营以及用户交互等多个技术环节，不同主体分别控制不同阶段的风险来源，共同决定最终生成内容的合法性。侵权风险的形成并非单一主体行为所致，而是多

个主体共同作用的结果。

从技术运行逻辑来看,模型开发者负责模型架构设计、训练算法及安全机制构建,是人工智能底层风险的重要控制者;服务提供者承担模型部署、运营管理及内容审核职责,对模型输出具有持续管理能力;平台运营者负责传播渠道管理,能够通过技术措施降低侵权内容传播风险;用户则通过提示词输入直接影响生成内容,其行为具有直接原因力。由于各主体控制能力、获利程度以及风险预防能力存在显著差异,如果继续采用统一责任主体模式,既可能扩大部分主体责任,也可能导致真正具有风险控制能力的主体逃避责任,不利于侵权风险的有效治理。随着开源模型、本地部署及模型微调技术的发展,开发者与服务提供者逐渐分离已成为产业发展的重要趋势。责任主体进一步呈现多元化、协同化特征,传统将服务提供者作为唯一责任主体的制度安排已难以适应人工智能产业的发展现实。因此,以风险控制能力、技术控制能力及利益获取能力为基础,对责任主体进行分层配置,实现责任承担与风险控制能力相一致,已成为完善生成式人工智能侵权责任制度的重要前提[1]。

归责原则是侵权责任制度的核心内容,也是生成式人工智能侵权责任配置中争议最为集中的问题。由于生成式人工智能兼具自主生成、算法黑箱及持续学习等技术特征,其侵权风险既来源于开发阶段的数据训练,也可能产生于模型部署、平台运营以及用户使用过程中,不同侵权行为对应的风险来源和控制能力存在明显差异[2]。自罗马法以来,民事责任的分配就遵循着“各负其责”的基本精神——每个人应当对自己能够控制的行为及其后果负责,而不对超出其控制范围的事项承担法律义务。从制度效率视角看,法律责任在原则上应配置给最有能力以最低成本预防、限制或消除损害的一方。在风险治理的归责选择上,更具解释力的路径是遵循“最小成本避免者”思路[3],其判断核心在于某一主体是否具备对相关风险的认知以及对该风险的控制程度,且此种认知与控制足以预防由此产生的损害。简言之,由具备信息优势与治理工具、能够以更低成本识别并降低风险的一方承担更高阶的注意义务,才可能实现“成本-效益”的最优匹配。责任配置是与预防成本、损害规模、信息条件相绑定的,在直接侵权人分散、权利识别困难、执行成本高的情形下,若中间主体能以较低成本改变侵权发生概率或损害扩散路径,则使其承担一定责任的规则,更可能实现更优的预防性投入效果。这一路径的逻辑是谁能以更低成本减少更多风险,而非谁更靠近内容就当然应审查内容。因此,统一适用一般过错责任、过错推定责任或者无过错责任,均难以兼顾公平与效率,责任分层配置更具现实可行性。

3. 生成式人工智能侵权责任配置的现实困境

3.1. 责任主体多元化导致侵权责任认定困难

生成式人工智能的技术运行突破了传统侵权法以单一行为人为中心的责任模式,其内容生成过程涉及数据集建构、模型训练、算法优化、模型部署、平台运营以及用户交互等多个环节,不同主体在技术流程中承担不同功能,共同决定最终输出内容是否合法。侵权风险的形成不再来源于某一主体的单独行为,而是多主体共同参与、共同作用的结果,责任主体呈现明显的链条化、协同化特征。

从技术运行机制来看,模型开发者负责算法架构设计、训练数据处理、安全机制构建及模型优化,对模型底层风险具有决定性的控制能力;服务提供者承担模型部署、运行维护及内容审核等职责,是生成内容输出环节的重要管理主体;平台运营者负责内容传播及用户管理,能够通过过滤、删除等技术措施降低侵权内容扩散风险;用户则通过输入提示词直接影响生成内容,其行为可能成为侵权结果发生的直接诱因。随着开源模型、本地部署及模型微调技术的发展,开发者、服务提供者及平台运营者逐渐相互分离,传统由服务提供者集中承担责任的制度模式已难以适应人工智能产业发展的现实需要[4]。

然而,我国现有法律规范尚未针对生成式人工智能建立完整的主体责任划分体系。《中华人民共和国民法典》侵权责任编仍以传统产品责任、网络侵权责任及一般侵权责任为基础展开制度设计,《生成

式人工智能服务管理暂行办法》¹虽然规定了服务提供者的内容安全管理义务，但更多体现的是行政监管要求，并未系统解决开发者、模型提供者、平台运营者及用户之间的责任边界问题。司法实践中，法院通常基于服务提供者直接提供生成服务这一事实，将责任集中于服务提供者承担，而对开发者等前端主体的责任认定相对不足。这种责任配置模式虽然有利于受害人寻求救济，却容易忽视不同主体风险控制能力的差异，既可能导致服务提供者承担超出其控制能力范围的责任，也难以充分发挥开发者等主体事前预防侵权风险的制度功能。

从侵权法风险分配理论来看，责任承担应与风险控制能力、利益获取程度以及注意义务范围保持一致。生成式人工智能侵权责任主体已经由传统的单一主体逐渐演变为多元主体协同参与的新型结构，若继续适用统一责任主体模式，不仅难以实现责任公平分配，也可能削弱不同主体主动履行安全保障义务的积极性。因此，如何依据不同主体的技术控制能力及风险来源合理划分责任边界，已成为完善生成式人工智能侵权责任制度亟需解决的问题。

3.2. 归责原则适用标准尚未形成统一规则

归责原则决定侵权责任承担的基本依据，也是生成式人工智能侵权责任制度中争议最为集中的问题。由于生成式人工智能兼具自主生成、算法黑箱及持续学习等技术特征，其侵权风险既可能源于模型训练阶段的数据使用，也可能产生于模型部署、平台运营乃至用户使用过程中，不同侵权行为之间存在明显差异。目前，理论界围绕生成式人工智能侵权责任形成了几种代表性观点：一是坚持一般过错责任，认为生成式人工智能并未突破传统侵权责任制度的基本框架，应继续依据注意义务判断服务提供者是否存在过错^[5]；二是主张适用无过错责任，认为人工智能技术具有高度风险性，应通过风险社会理论强化受害人保护^[6]；三是提出类型化、分层化归责思路，根据不同侵权类型、不同责任主体及风险来源分别适用不同归责原则，以实现风险控制能力与责任承担之间的动态平衡^[7]。

此外，随着人工智能技术持续迭代，不同应用场景的风险程度亦存在显著区别。例如，普通文本生成与医疗、金融、自动驾驶等高风险领域人工智能系统在风险来源、社会危害及控制能力方面均存在较大差异。如果一律适用统一归责原则，不仅可能增加人工智能产业创新成本，也可能降低受害人的司法救济效率。因此，建立能够兼顾风险防控、产业发展与权益保护的差异化归责体系，已经成为完善生成式人工智能侵权责任制度的重要方向。

3.3. 证明责任配置失衡影响侵权救济效果

除责任主体及归责原则之外，证明责任配置同样是影响生成式人工智能侵权责任制度运行效果的重要因素。根据传统侵权法的一般规则，受害人应当对侵权行为、损害结果、因果关系及行为人过错承担举证责任。然而，生成式人工智能具有算法黑箱、技术复杂及持续学习等特点，普通用户通常难以接触模型训练数据、算法参数及系统运行日志等核心信息，导致受害人在侵权诉讼中面临较高的举证成本^[8]。

一方面，生成式人工智能系统的技术复杂性导致侵权行为具有较强隐蔽性。受害人虽然能够证明生成内容存在侵权事实，却难以进一步证明侵权内容究竟源于训练数据、模型设计、算法缺陷还是用户提示词输入，更难证明开发者或者服务提供者违反了何种具体注意义务。另一方面，人工智能模型训练数据、算法结构以及安全机制通常涉及企业商业秘密，开发者和提供者出于保护商业利益的考虑，不愿主动披露相关信息，也进一步加剧了双方之间的信息不对称。信息优势集中于开发者及服务提供者，而举证责任却主要由受害人承担，容易造成责任认定失衡。

从制度功能来看，证明责任不仅关系到案件事实认定，更直接影响侵权责任制度的威慑效果。如果

¹https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm

举证责任配置过重，受害人可能因举证不能而无法获得有效救济，从而削弱侵权责任制度的风险预防功能；反之，如果过度降低受害人的举证负担，则可能导致开发者、服务提供者承担过重责任，影响人工智能技术创新。因此，在生成式人工智能侵权责任制度中，证明责任配置应当与责任主体划分、归责原则选择保持协调，实现风险控制能力与举证能力相统一，为侵权责任制度的有效运行提供程序保障。

4. 生成式人工智能侵权责任的分层配置路径

4.1. 以风险控制能力为基础构建责任主体分层体系

生成式人工智能侵权责任制度的完善，应首先解决“谁承担责任”的问题。传统侵权责任法遵循“行为人责任”模式，即由直接实施侵权行为的主体承担侵权责任。然而，生成式人工智能的运行机制决定了内容生成并非单一主体行为，而是数据收集、模型训练、模型部署、平台运营以及用户交互等多个环节共同作用的结果。因此，责任主体的认定不能再局限于直接输出内容的主体，而应结合技术控制能力、风险来源以及利益获取情况建立分层责任体系。

首先，应明确模型开发者作为基础层责任主体。开发者负责训练数据选择、模型架构设计、安全机制构建以及算法优化，是人工智能系统运行规则的制定者，也是侵权风险最初的形成者。对于因训练数据违法使用、模型设计存在重大缺陷或者未建立必要安全机制导致侵权结果发生的，应认定开发者承担相应侵权责任。开发者虽然并不直接参与内容输出，但其对模型具有最强的技术控制能力，因此应承担与其控制能力相适应的基础责任。

其次，应将服务提供者定位为运行层责任主体。与开发者相比，服务提供者并不决定模型底层架构，但其承担模型部署、持续运营、内容审核以及安全维护职责，能够通过建立提示词过滤、风险识别、内容审核及侵权投诉机制降低侵权风险。因此，当服务提供者未履行合理审核义务或者未及时采取删除、屏蔽等必要措施时，应依法承担相应侵权责任。需要注意的是，服务提供者责任并非来源于模型开发，而是来源于持续运营过程中的管理义务，因此其责任范围应区别于开发者^[2]。明确平台运营者与模型服务提供者之间的责任区分。随着大模型开放接口、本地部署及第三方应用不断发展，平台运营者更多承担传播渠道管理职责，其风险控制能力主要体现在侵权内容传播阶段。对于平台已经建立有效投诉机制并及时采取必要措施的，不宜因用户实施侵权而承担当然责任；反之，对于明知或者应知侵权而放任传播的，应依据网络侵权规则承担相应责任。

最后，应合理界定用户责任。生成式人工智能本质上仍属于工具，用户通过提示词输入直接决定生成目的和使用方式。当用户故意利用生成式人工智能实施著作权侵权、人格权侵害或者传播违法信息时，应作为直接侵权人承担主要责任。尤其是在利用提示词诱导生成特定作品、名人形象或者商业秘密等情形下，用户行为已经超出合理使用范围，应依法承担直接侵权责任。只有将开发者、服务提供者、平台运营者及用户纳入统一责任体系，并依据各主体风险控制能力配置责任，才能实现侵权责任公平分配。

4.2. 构建差异化归责原则协同适用机制

明确责任主体之后，还需进一步解决不同主体应适用何种归责原则的问题。当前理论争议主要集中于一般过错责任、过错推定责任与无过错责任之间的选择。然而，不同责任主体承担的技术职责、风险控制能力及注意义务均存在明显差异，如果统一适用某一种归责原则，不仅难以体现公平原则，也可能影响人工智能产业健康发展。因此，生成式人工智能侵权责任制度应突破单一归责模式，根据责任主体及风险来源建立差异化归责原则协同适用机制。

对于模型开发者，应坚持一般过错责任作为基本归责原则。开发者虽然具有较强技术控制能力，但其并不直接参与具体内容生成，亦无法预见所有用户行为。若适用过错推定，将举证不能的风险移转于

开发者,因其难以预见全部生成内容而极易陷入无法自证无过错的困境,在效果上与无过错责任趋同,同样会过度抑制创新。如果一律适用无过错责任,将显著提高模型研发成本,抑制人工智能技术创新。因此,应围绕开发者是否履行合理注意义务进行判断,包括是否依法获取训练数据、是否建立风险过滤机制、是否持续更新安全模型等。只有开发者违反合理注意义务并与损害结果之间具有因果关系时,方承担侵权责任。

对于服务提供者,则宜采用一般过错责任与过错推定相结合的归责模式。服务提供者掌握模型运行数据、内容审核机制及投诉处理流程,具有明显的信息优势。单纯的一般过错将使受害人因无法穿透技术黑箱而几乎不可能完成举证,全面无过错责任则置服务提供者于无论尽到何等注意义务均须负责的境地,有违比例原则;两者结合的过错推定模式,恰能在矫正信息不对称的同时保留服务提供者的抗辩空间,构成救济受害人与维系产业创新之间的最佳平衡。当受害人能够初步证明侵权事实存在,而相关技术资料完全掌握于服务提供者时,应适当减轻受害人举证负担,由服务提供者证明其已经建立合理审核机制并履行必要注意义务。此种配置既能够缓解人工智能“黑箱”导致的举证困难,又不会完全突破传统侵权法过错责任原则^[9]。对于平台运营者,则应继续坚持“通知-必要措施”规则。平台责任以收到通知后未能及时采取必要措施为过错认定核心,已为权利人提供了相对明确的举证节点和救济路径,若引入过错推定或无过错责任,势必迫使平台对用户内容进行普遍性、预防性审查,与“通知-必要措施”的制度功能相悖。平台责任本质上属于网络服务提供者责任,其注意义务主要发生于侵权内容传播阶段。平台在收到权利人通知后未及时采取删除、屏蔽或者断开链接等必要措施的,应承担相应侵权责任;已经履行合理注意义务的,则不宜承担过重责任,以维持网络平台正常运营秩序。

此外,对于医疗诊疗、自动驾驶、金融投资等高风险人工智能应用,应根据法律特别规定探索适用更严格的归责规则。由于上述领域一旦发生侵权将直接危及生命、健康或者重大财产安全,仅依赖一般过错责任可能不足以实现风险防范功能。固守一般过错责任将使受害人因高度技术复杂性而难以证明系统缺陷,而直接引入一般性无过错责任又可能过度遏制这些社会必需的高风险技术创新;依托特别法引入危险责任等强化归责,能在不对整体制度造成冲击的前提下,更有针对性地保护重大法益。因此,可以依据危险责任理论,在特定高风险领域适当强化经营主体责任,以实现科技创新与公共安全之间的平衡。

4.3. 建立证明责任动态配置机制,实现实体责任与程序规则协调统一

责任主体划分与归责原则配置最终仍需通过诉讼程序加以实现,而证明责任正是连接实体法与程序法的重要纽带。生成式人工智能侵权具有算法黑箱、信息高度集中以及技术专业性强等特点,如果继续机械适用“谁主张、谁举证”的一般规则,将导致受害人因无法取得核心证据而难以获得司法救济。因此,证明责任配置应服务于责任主体分层与归责原则协同,实现实体规则与程序规则的统一。

应坚持受害人承担侵权事实的初步证明责任。受害人仍应证明侵权内容的存在、损害事实以及生成内容与人工智能系统之间具有基本关联,以防止滥诉现象发生,维护正常创新秩序。但当案件涉及模型训练数据、算法结构、安全审核机制等由开发者或者服务提供者控制的核心证据时,应适当适用举证责任转移。若受害人已经完成初步证明,而开发者或服务提供者拒绝提供相关技术资料,法院可以结合案件事实推定其未尽合理注意义务。这种动态举证机制符合证据控制理论,也能够有效缓解双方信息不对称问题。

法院需完善技术调查机制。生成式人工智能侵权案件通常具有较强技术专业性的特点,仅依赖双方举证难以查明事实。未来可借鉴知识产权案件技术调查官制度,引入第三方技术鉴定、专家辅助人及技术调查官等机制,提高司法机关认定算法缺陷、训练数据来源及风险控制措施的能力,为侵权责任认定提供专

业支持。

5. 结语

生成式人工智能作为新一轮科技革命的重要成果，在推动内容生产方式变革、促进数字经济发展的同时，也深刻改变了传统侵权行为的发生机制。与传统侵权不同，生成式人工智能侵权具有技术链条长、责任主体多元、风险来源复杂以及算法运行高度专业化等特征，使传统侵权责任制度建立于“行为人-受害人”二元结构之上的责任配置模式面临新的挑战。特别是在开发者、服务提供者、平台运营者及用户共同参与内容生成的背景下，侵权责任已经从单一主体责任逐渐演变为多主体协同责任，传统统一归责模式难以兼顾风险治理、技术创新与权利保护之间的利益平衡。

生成式人工智能侵权责任制度的完善，不应局限于归责原则或者单一主体责任的调整，而应从人工智能技术运行全过程出发，建立责任主体、归责原则与证明责任相互衔接的分层配置体系。具体而言，应首先依据不同主体在人工智能运行过程中的技术控制能力、风险预防能力及利益获取程度，对开发者、服务提供者、平台运营者及用户进行责任分层，实现责任承担与风险控制能力相匹配；其次，应根据不同主体承担职责的差异，构建一般过错责任、过错推定责任及特殊领域严格责任相协调的归责原则体系，避免统一归责带来的制度失衡；最后，应针对人工智能“算法黑箱”导致的信息不对称问题，建立与责任主体及归责原则相适应的动态证明责任配置机制，通过合理分配举证责任，提高侵权救济效率，实现实体法与程序法之间的协调统一。

参考文献

- [1] 章文丽. 人工智能应用的类型化侵权责任研究[J]. 上海师范大学学报(哲学社会科学版), 2026(2): 58-69.
- [2] 张文睿, 刘颖泽. 生成式人工智能服务提供者归责原则的分层建构[J]. 河北法学, 2026(7): 178-194.
- [3] 吴伟光. 网络企业对网络用户版权侵权的替代责任[J]. 知识产权, 2025(1): 3-24.
- [4] 刘友华. 生成式人工智能版权侵权责任分层配置研究[J]. 法学杂志, 2026(3): 67-87.
- [5] 熊琦. 生成式人工智能平台的著作权侵权责任[J]. 环球法律评论, 2025(2): 5-22.
- [6] 戴昕. 无过错责任与人工智能发展——基于法律经济分析的一个观点[J]. 华东政法大学学报, 2024(5): 38-55.
- [7] 胡巧莉. 人工智能服务提供者侵权责任要件的类型构造——以风险区分为视角[J]. 比较法研究, 2024(6): 120-135.
- [8] 何泽昊. 从信息成本到信息生产: 人工智能侵权过错责任的理论证成[J]. 法商研究, 2026(3): 65-79.
- [9] 朱晓峰. 论生成式人工智能服务提供者侵权的过错证明责任[J]. 比较法研究, 2026(3): 19-32.