

生成式人工智能生成内容的侵权风险与 规制路径

赵 鑫

青岛大学法学院, 山东 青岛

收稿日期: 2025年12月30日; 录用日期: 2026年1月13日; 发布日期: 2026年1月23日

摘 要

当今生成式人工智能的发展迅速, 其生成内容的质量受到了广泛的关注, 该技术的飞速发展与法律的滞后性相矛盾, 因此生成内容的侵权风险难以适用传统的法律体系加以规制, 给治理带来了困境。为厘清生成式人工智能生成内容侵权风险的规制路径, 本文以动态体系论为基础, 聚焦生成内容的侵权风险问题, 结合目前存在的治理困境展开分析。研究发现需构建能动分工的分层治理机制, 同时明确生成内容著作权的归属标准, 并且需要完善生成侵权内容的过错归责原则。为生成内容的侵权风险规制路径提供具体参考, 也为生成式人工智能的后续研究拓展了动态治理的研究方向。

关键词

生成式人工智能生成内容, 侵权风险, 动态体系

Infringement Risks and Regulatory Paths of Content Generated by Generative Artificial Intelligence

Xin Zhao

School of Law, Qingdao University, Qingdao Shandong

Received: December 30, 2025; accepted: January 13, 2026; published: January 23, 2026

Abstract

Nowadays, generative artificial intelligence is developing rapidly, and the quality of the content it generates has attracted widespread attention. The rapid advancement of this technology conflicts

with the inherent lag of legal systems. Consequently, the infringement risks associated with generative content cannot be effectively regulated by traditional legal frameworks, creating predicaments for governance. To clarify the regulatory approaches to the infringement risks of generative AI content, this paper, based on the theory of dynamic systems, focuses on the issue of infringement risks of generative content and conducts an analysis in combination with the current governance dilemmas. The study finds that it is necessary to establish a hierarchical governance mechanism featuring active division of labor, clarify the ownership criteria for the copyright of generative content, and improve the fault liability principle for the creation of infringing content. This research provides specific references for the regulatory paths of generative content infringement risks, and also expands the research direction of dynamic governance for subsequent studies on generative artificial intelligence.

Keywords

Content Generated by Generative Artificial Intelligence, Infringement Risks, Dynamic Systems

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. 问题的提出

《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》提出：“加快人工智能等数智技术创新，突破基础理论和核心技术，强化算力、算法、数据等高效供给。全面实施‘人工智能+’行动，以人工智能引领科研范式变革，加强人工智能同产业发展、文化建设、民生保障、社会治理相结合，抢占人工智能产业应用制高点，全方位赋能千行百业。”^[1]新一代人工智能作为前沿引领技术，对经济的高质量发展，国家的高效治理有着重要作用。目前，新一代人工智能的典型代表——生成式人工智能在全球掀起发展热潮，并广泛应用于越来越多的领域。然而，技术的突飞猛进与法律的滞后性造成了生成式人工智能生成内容的侵权风险日益加剧，亟待创新的法律规范与治理理论加以规制。

学界围绕生成式人工智能侵权问题已构建多维度研究框架。归责原则方面，过错责任说获“广州奥特曼案”等司法实践认可，主张以注意义务客观化平衡权益保护与产业发展；无过错责任说则强调将生成内容视为“产品”，以风险控制能力为核心分配责任。责任主体界定上，研究已从单一服务提供者扩展至价值链全环节，明确基础模型开发者、系统集成者、使用者的共同责任，提出按控制力与收益比例划分责任权重。风险类型聚焦知识产权侵权、人格权侵害、虚假信息传播等，成因指向训练数据瑕疵、算法“幻觉”与标识缺失。规制路径上，形成法律适用优化(类推“通知-删除”规则)、技术治理(双重标识体系)、主体义务设置(数据审核、内容过滤)的多元方案。

现有研究的核心争议集中于三方面：治理层面上，在现有技术确定责任主体困难；在生成内容著作权上，权力归属尚未形成明确共识；归责原则上，过错责任与无过错责任的取舍本质是创新激励与受害人保护的平衡难题。研究不足则表现为：治理层面过于单一；著作权体系亟待更新完善；多主体责任缺乏量化分配框架，实操性不足。因此，面对生成式人工智能生成内容的侵权风险，需要遵循动态体系论理论，在分层治理机制、生成内容著作权归属标准以及过错归责原则方面对侵权风险加以规制。

2. 生成式人工智能生成内容的侵权风险检视

随着生成式人工智能技术的蓬勃发展，人们对其生成内容的依赖也在不断提高，而生成内容的质量

已成为亟需关注的重点。我国《生成式人工智能服务安全基本要求》第5条规定：“应采取技术措施提高生成内容响应使用者输入意图的能力，提高生成内容中数据及表述与科学常识及主流认知的符合程度，减少其中的错误内容；应在用户服务协议中，向使用者告知使用生成内容时的知识产权相关风险，并与使用者约定关于知识产权问题识别的责任与义务。”^[2]从生成内容的风险类型来看，其包括信息失范风险、著作权虚置风险与救济不明风险。

2.1. 信息失范风险

随着生成式人工智能的广泛应用，数字内容生成的门槛显著降低，伴随着大量虚假诽谤信息的滋生，信息失范的风险日益严重。在实践领域中，除了信息不对称等传统不正当竞争行为外，经营者利用人工智能生成合成信息开展竞争，生成合成混淆性信息、虚假性信息和暴力性信息^[3]。这种行为凭借自动性算法、信息快速生成、定向传播的优势，摆脱了人工操作效率低下的限制。违法成本显著降低，容易对市场生态造成不可估量的破坏。同时，生成式人工智能的算力在不断提升，容易导致虚假误导信息的涌现。不法行为人利用该技术，生成诽谤性内容并在网络空间大肆传播，进而侵犯了相关主体诸如人身权等权力。生成式人工智能技术的普及给予了不法侵害者们可乘之机，容易对社会的公序良俗造成广泛影响。再者，由于生成式人工智能技术实现了从静态模仿向动态模仿的转变。生成式人工智能通过模仿人类的语言与用户沟通，这一形式给用户带来了沉浸体验，用户对信息内容较容易产生信任^[4]。使得公众难以辨别一段信息中是否有生成式人工智能的参与，从而侵害了公众的知情权，进而可能造成社会的信用危机。

2.2. 著作权虚置风险

生成式人工智能通过多模态建模等技术实现了内容的智能生成，其生成内容的著作权呈现出责任主体难界定、权益来源多元化等特征^[5]。生成式人工智能生成内容的底层逻辑是包含两个过程，首先通过Transformer模型¹进行自然语言处理，主要是将信息转化为数字，再利用Diffusion模型²进行算法生成特定的新型内容。上述技术所生成的内容已经难以适配现行的著作权体系。一方面，“人类作者观”表明，著作权的权力主体应围绕人类来划定，明确只有自然人才具备著作权法层面“作者”的合法主体资格。尽管我国著作权体系拟制法人或其它组织可以成为著作权的主体，但是也没有将生成式人工智能纳入到著作权主体中。另一方面，我国著作权体系表明了作品需具有独创性。如果无法厘清并衡量人工智能生成内容过程中所涉及的自然人贡献者(如程序员、用户等)对生成内容的独创性联结强度，似乎很难肯定人工智能生成内容为著作权保护客体^[6]。实践中用户和生成式人工智能配合输出生成内容，用户输入提示语，引导生成式人工智能通过模型输出符合用户意愿的内容，生成内容的输出过程难以与传统的独创性表达相适应。

2.3. 救济不明风险

当面临自身权力遭受侵害的情形时，如何快速制止不法侵害，令自身法益恢复到圆满状态是法律价值的体现。生成式人工智能价值链上各主体身份易发生重叠与转换、上下游相互依赖且具有极强的风险传导性，生成式人工智能侵权行为主体复杂、人类行为与侵权损害结果之间的因果关系高度分散、主观过错认定困难^[7]。这将导致当侵害行为发生时相关主体难以及时有效地维护自身合法权益。首先，生

¹Transformer模型是一种基于自注意力机制(Self-Attention)的深度学习模型架构，由Google团队在2017年的论文Attention Is All You Need中提出，最初用于自然语言处理(NLP)的机器翻译任务，如今已成为生成式AI、计算机视觉(CV)等领域的核心基础架构，是大语言模型(LLM)如GPT、BERT的技术基石。

²Diffusion模型(扩散模型)是一种基于概率生成的深度学习模型，核心原理是模拟“逐步加噪-逐步去噪”的扩散过程，从而生成高质量、多样化的输出内容，是目前文生图(如Stable Diffusion)、图生图、音频生成等多模态生成任务的主流技术之一。

成式人工智能生成内容侵权的主体特殊。在生成内容的整个过程中涉及系统开发者、数据管理者、云服务提供者、服务提供者以及用户等多方主体参与，他们都对侵权内容的产生有不同程度的贡献，谁应向损害结果负责难以确认。其次，人机行为对生成侵权内容的因果关系复杂。用户虽然通过输入提示词主导侵权内容的产生，但直接实施侵权内容生成的行为的主体是生成式人工智能，因果关系具有分散性。最后，生成侵权内容的主观过错认定模糊，由于“算法黑箱”³的存在，被侵权的主体为了证实系统开发者等相关主体有过错，但是难以搜集证据。

3. 生成式人工智能生成内容侵权风险的治理困境

生成式人工智能生成内容的质量越高，将提升人们对该项技术的预期，从而对其不断进行完善，形成良性循环，但监管的分配、著作权的归属以及归责原则的选择对生成内容风险的治理带来了挑战。具体而言，多元主体的审查的复杂性给监管的分配提出了更高的要求；著作权体系中以人为中心的创作理念与人机共同创作的模式相冲突；哪种归责原则可以更好地让被侵权人高效地维护自身的合法权益是一个亟待解决的问题。

3.1. 多元主体审查复杂性的监管困境

传统网络内容的产出一般依赖于单一主体，但生成式人工智能改变了这一产出模式。整个服务提供与内容生成过程需要多主体参与和配合，也就意味着生成内容真实与否，可能与各主体的参与行为都存在关联^[8]。参与主体的多元性就给监管的分配带来了挑战。一方面，生成式人工智能生成内容依赖每一个主体对单独的模块负责，并且由于模块的分割，任一主体难以对其它主体负责的模块进行控制。这就给生成内容的监管提出了更高的要求，即需要在整体上进行把控。另一方面，每个主体对生成内容的真实与否的影响都有侧重，意味着每个主体对生成内容的贡献程度可能存在差别，因此需要据此分配不同的注意义务，更需要对每一个主体做出具体可行的监管规定来平衡各方利益，引导生成式人工智能健康发展。

3.2. 人机共创模式的著作权归属困境

明确生成式人工智能生成内容的权力归属，是规范相关主体权责统一的重要前提。生成式人工智能生成内容可能具有相当于自然人创作作品的质量^[9]。在生成式人工智能深度参与的应用场景中，人机协同的创作模式深刻挑战了著作权制度关于创作主体、作品标准、独创性认定等方面的预设，其技术自主性、创作过程非线性和成果生成概率性特征，都使得智能生成物的独创性判定、权利主体界定等面临系统性的规范重构挑战^[10]。从生成式人工智能生成内容的机制来讲，其不仅依赖用户输入的创造性指令，而且取决于模型对数据的个性化编排。因此难以仅将用户作为生成内容著作权归属的唯一主体，其更类似于人机智慧的共同成果。深层次讲，诸如数据供给方、模型设计者以及用户等都对生成内容有不同程度的贡献，其都有可能对生成内容主张著作权的归属。在这种情况下，目前著作权的规定难以明确人机共创模式下各主体的权力交集的解决方案。

3.3. 生成内容侵权的归责原则选择困境

应当看到，生成式人工智能不同于传统的网络服务，其在运行原理与责任认定方面存在一定的特殊性^[11]。生成式人工智能生成侵权内容，主要需要明确服务提供者应当承担何种责任，为此首先应当界定生成式人工智能服务提供者的性质。有些观点认为应当对生成式人工智能服务提供者适用产品责任，虽

³算法黑箱(Algorithm Black Box)是指内部运行逻辑、决策依据和计算过程不透明、不可解释的算法系统——尤其是以深度学习为代表的复杂机器学习模型，外界只能观察到输入数据和输出结果，却无法理解或追溯从输入到输出的中间推理链路。

然能更好地保护弱势群体的合法权益,但这种做法也可能会对生成式人工智能的发展带来不良的影响。此外,如果适用过错责任原则,由于多个主体共同参与了生成内容的产出,那么就需要明确每个主体的过错程度。同时,区别于传统的网络服务提供者,生成式人工智能属于人机共创输出内容,并不属于独自提供网络信息,这就导致其可能并不适用针对传统的网络服务提供者的归责原则。

4. 生成式人工智能生成内容侵权风险的规制路径

新一代人工智能技术已成为大国博弈的焦点,生成式人工智能在降本增效、提高生产力方面有着深远影响。其深刻改变了数字网络内容的生产方式,系统开发者、服务提供者与终端使用者等多方主体均共同参与了生成内容的产出。因此应贯彻动态体系论的治理理论,对生成式人工智能生成内容的侵权风险进行有效的规制。要构建分层治理的机制,促进多元主体共同减少虚假信息的产生与传播;要动态回应“独创性表达”要求,找到生成内容著作权归属的解决方案;要弹性设置侵权领域内的归责原则,改变“全有或全无”的判断标准,结合多种要素的评价来选择归责原则,实现各方主体利益的动态平衡。

4.1. 构建能动分工的分层治理机制

目前的监管方式已经难以适配生成式人工智能中多方主体共同参与生成内容的应用场景。在传统框架下,责任主要集中在初始行为人身上,而在数字环境中,由于信息传播路径的多样化和参与者角色的多重性,确定具体责任主体变得较为困难^[12]。为此构建符合生成式人工智能工作逻辑的分层监管方式有利于治理生成内容的信息失范风险。在数据层面,完善数据输入阶段的责任分配。基础模型开发者应建立训练数据白名单与三重审查机制(来源合法性、内容安全性、完整性存证)。虚假诽谤信息的生成根源在于数据的损害与篡改,应利用区块链技术控制生成内容的准确性和可预测性。同时生成式人工智能服务提供者应完善信息披露体系,确保监管的可追溯性。在内容层面,完善标识监管方式,采用隐形或显性方式对生成内容进行标注,建立分级内容审查机制(低风险时放行、中风险交叉验证、高风险 24 小时人工复核)。通过明晰的标签来回应公众的知情权,让公众明确该网络内容来自于生成式人工智能。同时完善内容审查义务,采取事前审查加事后处置的监管策略来尽可能阻断虚假信息的传播。在服务层面,完善信用评价治理机制,促进生成式人工智能服务提供者自觉完善内容生成技术,惩罚与激励方式并用,引导生成式人工智能向着生成高质量内容方向发展。在应用层面,职业用户提交生成内容合规证明,提供生成指令日志留存至少 6 个月。在全环节,监管机构应建立标识合规性抽检机制,搭建跨平台信用评价体系,对高风险主体实施穿透式监管。

4.2. 规范生成内容著作权归属标准

在实践中,对于生成式人工智能生成内容的著作权归属尚且存在争议。在“漫画《黎明的扎里亚》”⁴一案中,美国版权局并未批准其版权注册,而在我国司法实践中,法院肯定了“春风图案”⁵属于作品。如果依照传统的法律进路,人工智能生成物不具有可版权性的原因在于客体与主体之间存在适格的难题^[13]。因此需要对著作权中客体的定义和主体的识别进行明确。首先在客体的定义方面,由于生成式人工智能生成内容的外在形式仍是人机共创的产物,而且生成式人工智能目前仍不属于通用型人工智能,其不具备完全自主创作的能力,因此其产出本质上属于人类主导,机器辅助创作的内容,相当于照片作品属于人类拿着摄像机拍出来的成果,那么生成式人工智能就不能法律拟制为著作权的归属对象。其次在主体的识别层面,需要设立动态的权力配置模式。我们不能否定系统开发者、服务提供者在生成内容中

⁴《黎明的扎里亚》(Zaravia)是一部由 Kris Kashtanova 创作的漫画,它被认为是首部由 AI 创作的漫画。

⁵正式名称为李某诉刘某侵害作品署名权、信息网络传播权纠纷案,它是我国首例“AI 文生图”著作权纠纷案。该案首次以司法裁判形式确认 AI 生成内容可构成作品,明确使用者享有著作权,对 AI 创作相关的知识产权界定具有里程碑意义。

贡献。因此在具体情况下，他们可以通过合同的方式与使用者划定权益的分配。同时使用者对生成式人工智能的干预程度决定了其对生成内容的贡献大小。对于“强人类干预型”的生成内容，表明了使用者在深度干预下的独创性表达，为此可以考虑将使用者作为著作权归属的对象；对于“弱人类干预型”的生成内容，由于使用者的贡献程度较低，因此可以将著作权归属于系统开发者等主体。

具体而言，可以构建一个多维度的“人类贡献度”评估指引：创作意图主导权重 30%，考察提示词具体性、创作目标明确度；提示词质量权重 25%，关注其创新性、细节丰富度和逻辑完整性；后续编辑程度权重 20%，明确其修改篇幅占比、创造性修改占比；技术依赖程度权重 15%，关注其对模型的依赖；适用场景权重 10%，考察其是否为商业使用以及是否获利。对于总分大于或等于 70 分的情况，规定其属于“强人类干预型”，小于 70 分的情况，规定其属于“弱人类干预型”。

4.3. 完善生成侵权内容的过错归责原则

在确定了生成式人工智能服务提供者不宜适用产品责任的前提下，服务提供者是否尽到注意义务是认定过错的关键，但仍需要明确过错的判断标准。一是其有没有对训练数据尽到了合理审核的义务，但由于训练数据过于庞大，因此不应对其设置过重的审核义务；二是其针对明显侵权的生成内容是否尽到了采取了必要的措施。首先应当建立信息过滤机制，识别明显的侮辱、暴力词句并禁止该内容的生成，其次在生成内容已造成侵权的情况下，服务提供者应当采取相应的措施防止损害结果的进一步扩大；三是其应当遵守有关监管规定，根据《暂行办法》以及《基本要求》的规定，服务提供者在数据收集、数据处理、内容标识、算法训练、内容生成、信息披露等全过程中，应当履行一系列注意义务^[14]。如果服务提供者违反了相关的监管规定，那自然可以认定其对生成的侵权内容具有过错。但总体来讲，生成式人工智能技术本身仍在不断发展完善，所以应当随着其技术的发展动态地调整过错的判断标准。在司法实践中，法官应该凭借现有的技术标准来判断服务提供者是否具有过错，在现有技术难以避免侵权内容生成的情况下，自然不能过分追究服务提供者的责任，但需要服务提供者对系统性风险采取更高层次的风险防范措施^[15]。另一方面也需要考虑应对侵害的成本，包容审慎的态度是促进生成式人工智能技术快速安全发展的关键，如果服务提供者采取应对侵害发生的措施成本过重，以至于超出了生成的侵权内容所造成的损害，那么也不应认定其具有过错，否则不利于生成式人工智能技术的良性成长。

建立具体注意义务清单，明确服务提供者在数据、算法和内容审核方面的注意义务：在数据层面，训练数据来源需核查授权证明，数据安全保障要求实施匿名化处理；在算法层面，披露算法运行原理、局限性，建立偏见纠正机制；在内容审核方面，建立禁止生成内容清单，对医疗、金融等专业内容添加风险提示，接到侵权通知后 24 小时内采取删除/屏蔽措施，留存处置记录至少一年。

5. 结语

生成式人工智能作为新一代人工智能的典型代表，已成为衡量一个国家科技水平的重要指标之一。但是由于生成式人工智能技术的特殊性，传统的法律体系已难以精准规范该项技术，难免会有信息失范、著作权虚置、救济不明的风险的发生。多元主体审查复杂性的监管困境、人机共创模式的著作权归属困境、生成内容侵权的归责原则选择困境要求我们需要贯彻动态体系论的治理理论，构建能动分工的分层治理机制，规范生成内容著作权归属标准，完善生成侵权内容的过错归责原则，来引导生成式人工智能技术的健康发展。

参考文献

- [1] 中共中央关于制定国民经济和社会发展第十五个五年规划的建议[J]. 前进, 2025(11): 9-27.
- [2] 全国网络安全标准化技术委员会(SAC/TC260). GB/T45654-2025 网络安全技术生成式人工智能服务安全基本要

- 求[S]. 北京: 中国标准出版社, 2025.
- [3] 丁国峰, 寿晓明. 经营者滥用人工智能生成合成信息不正当竞争的技术治理路径[J]. 福建论坛(人文社会科学版), 2025(5): 47-67.
 - [4] 林雨佳. 生成式人工智能对信息治理的挑战与应对[J]. 苏州大学学报(哲学社会科学版), 2025, 46(2): 104-115.
 - [5] 姚志伟, 李卓霖. 生成式人工智能内容风险的法律规制[J]. 西安交通大学学报(社会科学版), 2023, 43(5): 147-160.
 - [6] 刘杰勇. 论人工智能生成内容的著作权保护——基于比较法的视角[J]. 比较法研究, 2024(4): 176-193.
 - [7] 张凌寒, 于琳. 生成式人工智能价值链上的侵权责任划分[J]. 北京行政学院学报, 2025(5): 72-83.
 - [8] 邓宏光, 王雪璠. 生成式人工智能虚假信息治理的风险与应对[J]. 理论月刊, 2024(9): 115-129.
 - [9] Elgammal, A., Liu, B.C., Elhoseiny, M. and Mazzone, M. (2017) Can: Creative Adversarial Networks, Generating “Art” by Learning about Styles and Deviating from Style Norms. arXiv:1706.07068.
 - [10] 位涛涛. 生成式人工智能训练数据的侵权风险与规制路径[J]. 青海民族大学学报(社会科学版), 2025, 51(4): 179-189.
 - [11] 王利明. 生成式人工智能侵权的归责原则与过错认定[J]. 中国法律评论, 2025(4): 15-30.
 - [12] 黄尹旭. 《人工智能生成合成内容标识办法》的规范功能与法律适用[J]. 法律适用, 2025(9): 136-147.
 - [13] 杨瑞琦. 人工智能生成物可版权性的法律因应: 适格难点、版权获得与公共对策[J/OL]. 北京交通大学学报(社会科学版): 1-10. <https://doi.org/10.16797/j.cnki.11-5224/c.20251204.008>, 2025-12-17.
 - [14] 张华韬. 动态体系论下生成式人工智能侵权的归责与构成[J]. 法学杂志, 2025, 46(6): 82-101.
 - [15] Helberger, N. and Diakopoulos, N. (2023) ChatGPT and the AI Act. *Internet Policy Review*, 12, 1-6. <https://doi.org/10.14763/2023.1.1682>