

人工智能生成内容的著作权刑事保护边界： 从“深度伪造”侵权到刑法规制路径

韩阳

泰国格乐大学，泰国 曼谷 10220

DOI:10.61369/SE.2025060003

摘 要： 本文聚焦人工智能生成内容（AIGC）著作权的刑事保护边界，以“深度伪造”这一 AIGC 技术滥用的典型样态为切入点展开研究。文章分析了“深度伪造”侵权行为的类型及其带来的多元危害，并指出当前刑法在评价该行为时面临的困境。文章提出构建 AIGC 著作权刑事保护边界与完善规制路径的方案，应遵循谦抑性、平衡保护、技术中立原则；立法上修改相关罪名、增设特定罪名、完善司法解释；司法上明确裁判规则与证据标准；治理上构建政府监管、行业自律、社会监督相结合的多元协同规制生态，以实现对著作权的有效保护与 AIGC 技术的健康发展。

关 键 词： 人工智能生成内容；著作权；刑事保护；深度伪造

Criminal Protection Boundaries of Copyright for AI-Generated Content: from "Deepfakes" Infringement to Criminal Law Regulatory Pathways

Han Yang

Kirk University, Bangkok, Thailand 10220

Abstract： This article focuses on the criminal protection boundaries of copyright for Artificial Intelligence Generated Content (AIGC), using "deepfakes," a typical form of AIGC technology abuse, as a starting point for research. The article analyzes the types of "deepfakes" infringements and their multifaceted harms, and points out the difficulties faced by current criminal law in evaluating such behavior. It proposes a scheme to establish criminal protection boundaries for AIGC copyright and improve regulatory pathways, following the principles of modesty, balanced protection, and technological neutrality. Legislative changes include modifying relevant charges, adding specific charges, and improving judicial interpretations. Judicially, clear verdict rules and evidentiary standards should be established. In terms of governance, a diversified and collaborative regulatory ecosystem combining government supervision, industry self-discipline, and social supervision should be constructed to achieve effective copyright protection and healthy development of AIGC technology.

Keywords： artificial intelligence generated content; copyright; criminal protection; deepfakes

引言

AIGC 技术在降低创作门槛的同时引发了复杂的著作权争议，“深度伪造”等滥用行为挑战传统保护体系。现行著作权刑事保护体系针对传统侵权构建，面对 AIGC 催生的新型侵权行为，存在罪名适配不足、责任主体模糊、危害量化困难等问题。如何划定鼓励技术创新与强化权利保护的边界，以刑法规制遏制 AIGC 侵权恶性发展，是亟待回应的问题。本文以“深度伪造”为切入点，分析 AIGC 著作权刑事保护现实困境，探索规制路径与边界，为构建兼顾权利保障与技术发展的刑事法律体系提供参考。

一、“深度伪造”：AIGC 侵权的典型样态与刑事可 责性分析

（一）“深度伪造”侵权行为的类型化与危害性评估

人工智能生成大模型的成熟大大降低了创作的门槛，创作者通过人工智能生成小说、漫画、视频的案例屡见不鲜，这些创作有一定的市场价值甚至受众群体广泛^[1]。“深度伪造”作为 AIGC

技术滥用的典型表现，其侵权行为可根据侵害对象与表现形式进行多维度类型化划分。在侵害个人权益方面，肖像权侵权最为常见，即通过技术将他人肖像植入色情、暴力场景或制作虚假代言视频；名誉权侵权亦频发，表现为生成模仿他人言行的虚假音视频传播不实信息，损害其社会评价。而在侵害社会公共利益层面，虚假信息传播是主要类型；知识产权侵权同样值得警惕，像模仿知名创作者风格生成作品并冒充原作售卖，就侵犯了著作

权^[2]。“深度伪造”的危害深远且多元，对个人而言，既会造成精神创伤，也可能引发财产损失；对社会来说，虚假信息泛滥会破坏信任体系、降低公众辨别力、影响社会稳定^[3]。在政治领域，伪造政治人物不当言论视频可能引发外交冲突或国内动荡；在经济领域，企业虚假负面信息视频或导致股价暴跌，扰乱市场秩序。

（二）“深度伪造”侵权行为的刑法评价困境

现有刑法罪名在应对“深度伪造”行为时存在滞后性，虽然可适用诈骗罪，但当“深度伪造”仅用于传播虚假信息，未直接骗取财物时，难以找到完全匹配的罪名^[4]。对于侵犯肖像权、名誉权的“深度伪造”行为，现行刑法中相关的侮辱罪、诽谤罪要求情节严重，且多针对自然人之间的行为，对于利用AI技术实施的大规模、匿名化侵权行为，在认定和追责上存在困难。“深度伪造”技术的普及使得普通人也能轻易制作虚假内容，部分行为人可能出于娱乐或好奇心理实施了侵权行为，其主观上是否具有犯罪故意难以界定。一些“深度伪造”内容经过多次传播后，源头难以追溯，导致无法确定具体行为人的主观心态，增加了刑法评价的难度^[5]。“深度伪造”行为引发的危害结果往往具有间接性和扩散性，其因果链条复杂，难以准确量化行为人的责任。同时对于精神损害等非物质性危害结果，刑法在评价时缺乏明确的标准，导致处罚力度难以把握。AIGC技术本身具有中立性，其开发者和提供者是否应对“深度伪造”侵权行为承担刑事责任存在争议。

三、我国AIGC著作权刑事保护的现状与适用障碍

（一）现行著作权刑事保护体系概览

我国著作权刑事保护体系以《中华人民共和国刑法》为核心，辅以相关司法解释和行政法规，形成了多层次、多维度的保护网络^[6]。在刑法层面，主要通过三个罪名对著作权侵权行为进行规制，分别是侵犯著作权罪、销售侵权复制品罪以及非法经营罪。侵犯著作权罪规定在《刑法》第二百一十七条，明确了以营利为目的，未经著作权人许可，复制发行其文字作品、音乐、美术、视听作品、计算机软件及法律、行政法规规定的其他作品等情形，违法所得数额较大或者有其他严重情节的，将面临刑事处罚^[7]。这一罪名直接针对著作权侵权的核心行为，为打击盗版、抄袭等行为提供了刑事威慑。销售侵权复制品罪则聚焦于侵权产品的流通环节，《刑法》第二百一十八条规定，以营利为目的，销售明知是本法第二百一十七条规定的侵权复制品，违法所得数额巨大或者有其他严重情节的，构成该罪。此罪名与侵犯著作权罪形成衔接，从生产到销售的全链条打击著作权侵权行为。此外，在司法实践中，对于一些严重扰乱市场秩序的大规模盗版发行、销售等行为，若不符合上述两个罪名的构成要件，但符合非法经营罪的相关规定，可能会以非法经营罪定罪处罚，进一步扩大了著作权刑事保护的覆盖范围。

（二）现行刑法在规制AIGC侵权时的适用障碍

我国现行著作权刑事保护体系虽较完善，但在AIGC侵权领域存在诸多适用障碍。犯罪构成要件适配难，现行罪名构成要件基于传统侵权设定，AIGC的“模仿式侵权”是否属“复制发行”

存在争议，其生成过程复杂，难直接套用现有规定^[8]。“以营利为目的”认定困局，部分AIGC侵权出于非营利目的，如免费传播却危害大，因缺营利目的难追刑责，形成规制漏洞。侵权主体确定难，涉及生成者、传播者、平台及AI模型开发者等多主体，各角色和责任难界定，尤其AI模型开发者责任及主观过错认定缺乏明确规定^[9]。违法所得数额计算难，AIGC侵权盈利模式多样隐蔽，现行数额标准针对传统模式，不适用于新型盈利模式，影响处罚精准度。证据收集与认定挑战大，生成过程技术隐蔽，电子证据易篡改删除，证据固定分析需专业知识，且实质性相似及侵权因果关系的证明缺乏明确指引。

三、AIGC著作权刑事保护边界的构建与规制路径完善

（一）确立AIGC著作权刑事保护的基本原则

AIGC著作权刑事保护需遵循谦抑性原则，即刑事手段应作为最后的保障，只有当AIGC侵权行为的社会危害性达到相当严重的程度，且民事、行政救济无法有效遏制时，才启动刑事追责。这是因为AIGC技术仍处于快速发展阶段，过度的刑事干预可能抑制技术创新与产业发展。平衡保护原则也至关重要，要在保护著作权人合法权益与促进AIGC技术创新、保障社会公众信息获取自由之间找到平衡点^[10]。一方面需通过刑事手段严厉打击恶意利用AIGC技术大规模侵犯他人著作权、扰乱市场秩序的行为，另一方面也要避免对正常的AIGC研发、实验性创作等行为设置过多刑事障碍，为技术发展预留合理空间。比如对于科研机构为改进AIGC模型而进行的合理使用他人作品的行为，应予以包容，不轻易纳入刑事规制范围。此外，技术中立原则不可或缺，即不能仅因AIGC技术本身可能被用于侵权而对技术开发者或提供者施加刑事责任，刑事责任的追究应聚焦于具体的侵权行为及其主观过错。只要技术开发者在技术设计、提供过程中已履行必要的注意义务，采取了合理的防侵权措施，就不应为他人利用该技术实施的侵权行为承担刑事责任。。

（二）立法层面：完善刑事法律规范体系

针对AIGC侵权的特殊性，应对现行刑法中的相关罪名进行适应性修改。在侵犯著作权罪中，明确将AIGC“模仿式侵权”等新型侵权行为纳入“复制发行”的范畴，通过立法解释或修正案的形式，界定“模仿式侵权”的构成要件，如规定当AIGC生成内容与原作品在表达上存在实质性相似，且这种相似性源于对原作品的不当利用时，即构成侵权。同时适当放宽“以营利为目的”的限制，对于那些虽无直接营利目的，但通过免费传播AIGC侵权内容造成著作权人重大损失或严重扰乱市场秩序的行为，也应纳入刑事规制范围。增设AIGC特定罪名，对于一些利用AIGC技术实施的、传统罪名难以涵盖的严重侵权行为，如大规模生成并传播伪造的享有著作权的作品，情节特别严重的，可单独设立罪名，明确其构成要件和法定刑，增强刑事打击的针对性。完善相关司法解释与配套规定，细化AIGC著作权刑事案件中“违法所得数额”“严重情节”等定罪量刑标准，考虑AIGC

侵权的传播范围、对原作品市场价值的影响等新型因素，制定符合 AIGC 特点的量化标准，为司法实践提供明确指引。

（三）司法层面：明确裁判规则与证据标准

在裁判规则方面，应明确 AIGC 侵权行为的认定标准，对于 AIGC 生成内容与原作品的“实质性相似”判断，需结合 AIGC 技术特点，不仅考察两者在表达形式上的相似性，还应分析 AIGC 模型是否不当利用了原作品的核心创意、独特风格等实质性要素。同时清晰界定各侵权主体的责任，如 AIGC 内容生成者承担直接侵权责任，平台服务提供者在明知或应知侵权行为存在而未采取必要措施时承担相应责任，AI 模型开发者一般不承担责任，除非其与侵权行为人存在共同故意。在证据标准方面，建立适应 AIGC 特点的证据收集与认定规则。明确电子证据的固定、提取方法，如要求 AIGC 服务提供者留存生成内容的相关数据记录，包括提示词、生成过程日志等，以便在侵权案件中作为证据使用。对于 AIGC 生成内容与原作品的实质性相似性鉴定，可引入专业的技术鉴定机构，制定专门的鉴定规范和流程，确保鉴定结果的科学性和公正性。同时明确举证责任分配，在一些复杂的 AIGC 侵权案件中，可适当减轻著作权人的举证负担，如允许其通过初步证据证明侵权可能性后，由侵权行为人举证证明其行为的合法性。

（四）治理层面：构建多元协同的规制生态

构建政府监管、行业自律、社会监督相结合的多元协同规制生态。政府相关部门如版权局、工信部、公安部等应加强协调配合，明确各自在 AIGC 著作权保护中的职责，形成监管合力。例如版权局负责对 AIGC 著作权侵权行为的日常监测与行政处理，

公安部负责对涉嫌刑事犯罪的 AIGC 侵权行为进行侦查打击。推动行业协会制定 AIGC 行业自律规范，引导 AIGC 服务提供者建立健全内容审核机制，对生成内容进行著作权合规审查，主动过滤侵权内容。鼓励行业内建立侵权投诉与处理机制，及时响应著作权人的维权请求。同时，行业协会可组织开展 AIGC 著作权保护培训，提高从业者的法律意识和合规水平。加强社会监督与公众参与，建立便捷的侵权举报渠道，鼓励公众对 AIGC 著作权侵权行为进行举报。利用大数据、人工智能等技术手段构建 AIGC 侵权监测平台，实现对网络上 AIGC 侵权内容的实时监测与预警。此外，通过宣传教育提升社会公众的 AIGC 著作权保护意识，引导公众自觉抵制侵权内容，形成全社会共同保护 AIGC 著作权的良好氛围。通过以上多个层面的协同发力，构建起科学合理、切实可行的 AIGC 著作权刑事保护边界与规制路径，既有效保护著作权人的合法权益，又促进 AIGC 技术的健康发展。

四、结束语

本文通过对“深度伪造”侵权样态的剖析、现行刑事保护体系的检视，以及规制路径的探索，揭示出 AIGC 著作权刑事保护的核心在于平衡——既要通过必要的刑事手段遏制恶性侵权行为，为著作权人筑起法律屏障，又要避免过度干预技术创新，为 AIGC 产业预留健康发展的空间。未来的研究与实践需进一步关注技术发展与法律规制的协同演进，不断优化刑事保护的精准性与适应性，最终实现“保护创新”与“促进创新”的双重价值目标，为数字时代的著作权保护与技术进步提供坚实的法治保障。

参考文献

- [1] 刘慧. 人工智能生成创作版权权属认定的困境与突破路径——从个案认定到创作者公示[J]. 出版广角, 2024, (12): 38-43. DOI: 10.16491/j.cnki.cn45-1216/g2.2024.12.007.
- [2] 杨利华, 王诗童. 人工智能生成内容的著作权署名规则探究[J]. 延边大学学报(社会科学版), 2024, 57(03): 100-113+143. DOI: 10.16154/j.cnki.cn22-1025/c.2024.03.009.
- [3] 冯晓青, 李可. 人工智能生成内容在著作权法中的定位[J]. 中国版权, 2024, (03): 20-30.
- [4] 钟莉. 人工智能生成内容的权属及侵权认定研究[J]. 中国市场监管研究, 2024, (06): 85-88.
- [5] 苏紫阳. 人工智能生成物著作权保护研究[D]. 内蒙古科技大学, 2024. DOI: 10.27724/d.cnki.gnmngk.2024.000205.
- [6] 王小雨. 人工智能创作内容(AIGC)的著作权合理使用探析[J]. 法制博览, 2024, (16): 36-38.
- [7] 刘云开. 人工智能生成内容的著作权侵权风险与侵权责任分配[J]. 西安交通大学学报(社会科学版), 2024, 44(06): 166-177. DOI: 10.15896/j.xjtuskb.202406015.
- [8] 孙文嘉. 人工智能生成内容著作权问题研究[D]. 吉林财经大学, 2024.
- [9] 牛诗妍. 人工智能生成内容(AIGC)著作权保护研究[D]. 山西财经大学, 2024. DOI: 10.27283/d.cnki.gsxc.2024.000560.
- [10] 赵天月. 人工智能生成内容著作权侵权治理研究[D]. 大连海事大学, 2024. DOI: 10.26989/d.cnki.gdlhu.2024.000892.