

生成式人工智能侵权的责任主体认定研究

程语童

华东政法大学, 上海, 201620;

摘要: 生成式人工智能的快速发展在推动内容生产变革的同时也带来诸多侵权风险。当所生成的内容侵害他人权益之时, 因人工智能缺乏法律主体资格而无法明确产品背后的担责人成为司法适用之难点。文章首先梳理了生成式人工智能的侵权类型与技术特征, 指出其侵权行为的独特性源于算法黑箱、内容不可预测性以及人机交互的复杂性。研究认为, 服务提供者作为技术平台的管理者应承担主要关注义务并应对侵权行为负责; 而算法设计者的责任程度则与其所犯过失的程度及范围相关; 对于使用者来说, 在进一步传播或使用该生成物的时候有可能会侵犯他人的权利。最后, 本文从归责原则、注意义务标准以及举证责任分配等方面出发, 提出了科技发展的权利认定路径, 以实现保护受害人权益和技术进步之间的平衡。

关键词: 生成式人工智能; 侵权责任; 责任主体; 服务提供者; 算法设计者

DOI: 10.64216/3104-9672.25.04.043

随着深度学习的发展, 生成式人工智能既为使用者提供便利地进行创作并获取建议的同时也带来一系列侵害名誉权、著作权等。由于人工智能本身不具备法律主体资格, 如何确定侵权发生时的责任主体成为司法与理论研究的紧迫问题, 传统侵权法以人的行为为核心, 然而人工智能介入之后, 侵权路径更加多元。由于人工智能本身不具备法律主体资格, 如何确定侵权发生时的责任主体成为司法与理论研究的紧迫问题。因此, 有必要对生成式人工智能侵权的责任主体进行探析, 这是理论的要求, 也是实践的需要。本文将结合生成式人工智能侵权技术机理及类型, 分别对服务提供者、算法设计者以及使用者三种身份进行论述, 在此基础上探讨其在侵权过程中承担的责任认定基础与归责原则, 旨在为立法完善与司法裁判提供有益参考, 助力构建公平合理

1 生成式人工智能侵权的技术基础与法律挑战

1.1 生成式人工智能的技术机理与侵权成因

就技术层面而言, 生成式人工智能的核心在于其深度学习算法与大规模参数模型, 通过对海量数据中蕴含的规律性、关联性的信息进行归纳总结而获得生产新内容的能力。其具有典型的“黑箱”特征, 即开发者无法准确预测模型对特定输入条件下的产出结果是什么。侵权可能发生在模型学习阶段吸收了包含侵权内容的数据或者生成过程中不正确地组合某些特征之时。正是由于这一不确定性, 传统归责原则遭遇适用困境, 因为损

害结果无法被预知。而且生成式人工智能的功能很大程度上依赖于使用者指令输入, 由使用者以关键词指引生成的内容路径, 决定了侵权事态往往是人机合谋所致, 并由此拉长了侵权环节, 涉及角色包括算法训练师、服务提供者乃至发布指令的使用者, 其行为与最终损害结果之间的因果关系需要仔细分析方可确认。

1.2 现行法律框架应对新型侵权面临的困境

在法律层面, 生成式人工智能引发的首要难题是主体资格的缺失。目前的法律体系下, 能够成为主体并承担法律责任的前提是有意识有能力的存在者, 人工智能作为技术产品, 无论其表现多么智能, 都不能成为承担法律责任的主体。因此, 只能在其开发者、提供者以及应用者之间进行责任分配。不过, 这些合理的分割没有明确具体的规定作为指引, 民法典关于产品责任的规定主要针对有体物, 而算法作为无形产品是否适用存在争议; 而通知删除制度本意在于保障网络平台不受知识产权侵害, 然而, 在实践中, 由于人工智能自动生成的文本并非使用者所上传的内容, 这就造成了与此原则间的不兼容性; 其次, 人工智能技术日新月异的发展态势以及在基本构架上的多元化及各领域的具体化, 各阶段参与者管理能力和义务承担的差异性, 也让认定责任变得异常复杂。因此在对具体的责任主体进行深入探讨前我们首先要了解的是相关技术及法律知识。

2 生成式人工智能侵权责任主体的类型化分析

2.1 服务提供者的法律地位与责任基础

服务提供者是生成式人工智能面向市场的直接出口,也是使用者接触和使用技术的第一界面。无论是通过应用程序编程接口调用大模型提供服务的平台,还是开发并运营聊天机器人、绘图工具的企业,都属于服务提供者的范畴,在法律上通常会缔结服务协议:进而保证其提供的服务合法安全。如果人工智能生成的文本侵害他人权益,则作为管理者及提供方的服务提供者具有对算法、输入输出进行控制以及对结果负责的能力,在此前提下其是否在合理的限度内尽到注意义务,例如是否存在必要的审查机制来保障训练素材本身合法等等,都将决定其是否应当承担过错责任;因服务提供者可通过服务运营获得经济效益,遵循风险收益相当原则,让其承担主要的监管义务以及相应的侵权损害赔偿也是合理的。实践中,要求服务提供者在每一次生成内容之前都进行实质性的审查既不可能,也不利于促进技术创新,因此其注意义务应限定在合理范围内,关键在于其是否已尽到与其技术能力和所从事的服务相对应的必要风险防范措施。

2.2 算法设计者的过错形态与责任认定

生成式人工智能模型的设计者是算法设计者,他们通常是研究所或者科技公司的成员,负责设计、选择并加工用于训练的数据集以及改进算法等等。虽然这些工作是在提供服务之前进行的,但这些决策对最终的服务效果有非常重要的影响。如果在选取训练素材的过程中犯了大忌,例如选用大量存在侵权性质的内容,或者没有在算法研发环节采取相应技术手段规避侵权风险,比如未设置安全过滤程序模块,则可能为后续的侵权纠纷埋下隐患。此时,应当借鉴产品责任或者专家责任原则来确定算法设计师所应承担的责任,当因设计失误或数据瑕疵而造成侵权行为出现时,算法设计者就须承担相应的责任。但在追究责任的同时,必须谨慎规定标准,因为其作为一种新锐前沿科技,人工智能产品可能有不可预测的行为结果,设计者无法保证此模型不会在任何场景下产出侵权性内容。只有当设计者明知违反公认的设计原则或者业已成规,例如明知存在重大侵权危险仍不得不用,或虽具可避免之技术手段而不采取之时,方能被认定为有过错。

2.3 使用者的行为边界与侵权构成

使用者作为人工智能生成内容的发起者和接收者,

其行为同样可能构成侵权。他们可以输入关键字引导生产的指向性,如果该关键字包含对他人意志的侵害,则可能构成侵权,那么作为侵权故意的来源者,其构成对他人权益的主动侵害,在这种情况下应将其认定为直接侵权人。尽管生成的信息并非来源于本人的创作,但因其启动生成过程中发挥了作用,利用人工智能实现了侵权目的,这就需要承担相应的法律后果。还有一种情况是,当使用者得到生成的信息后会将其公开并且商业性使用或者以著作者身份发表,该步骤可能会单独侵犯他人名誉权、著作权等权利。就使用者而言,其责任在于预判自己指令及其后续行为可能产生的侵权后果以及避免利用技术手段从事违法犯罪行为。由于普通使用者的非专业性,故在判定其法律责任时应更多关注其主观目的和客观行为而非对算法本身的探究。

3 生成式人工智能侵权责任认定的规则构建

3.1 归责原则的适用是责任认定的逻辑起点

对于生成式人工智能侵权,应当坚持以过错责任为基本原则,在特殊情况下考虑采用过错责任要求受害人证明侵权主体存在过错。如果服务提供者已采取了能力范围内必要的技术和管理措施,例如进行了必要的安全检查、建立了内容审核机制及时处理侵权投诉,就不能轻易认定其有过错;同理,如开发者遵循行业的一般性研究规范以及道德准则,他的创造性行为也不应被归咎。然而,由于人工智能系统的复杂性,受到伤害的人往往并不知道或者无法理解算法如何运作,因此对于受害者来说,证明被告方有过错确实是困难的。因此在特定情形下我们可运用错误的假设原则,即首先假设相关主体存在过错,再由其自证其已尽到必要的注意义务。例如,当有明显违法的信息存在,并且在服务提供者的技术能力及控制范围内可以采取有效防范措施而未采取,则可认定其有过错。但推定错误的规则又未滑向无过错责任,从而加大了对技术提供方的压力,也能够适当缓解被害方举证难的问题,平衡了各方的利益。

3.2 注意义务的具体化是责任认定的核心内容

对于服务提供者,其注意义务应当贯穿于服务的事前、事中、事后全过程。事前阶段,对其训练素材的内容来源进行审查,防止侵权;对人工智能算法安全进行测试,用来判定其是否存在生成侵权信息的可能。事中阶段,在接收客户指令及提供服务的过程中应当存在对

生成内容进行识别的系统机制,可有效预判是否构成侵权并及时制止或提示。事后阶段,在侵权发生后还应具备便捷有效的反馈渠道,例如接到举报后及时处理,包括但不限于停止传播相关信息等。同时应当注意,上述标准应随着技术进步而不断调整,并不苛责平台服务提供者进行超出当前技术水平所能实现的严格审查。而对于算法工程师来说,则是在合理研究的基础上以公开透明的方式做出相应的设计。这包括明确披露训练数据的信息以及任何潜在的偏差和危害;遵循算法伦理原则,在算法的设计过程中每一步都考虑防范生成的内容的危险性。对使用者而言其注意义务较为明晰,即不得使用该生成服务从事恶意侵权,在后续传播中也应进行初步的内容合法审查,对存在明显侵权可能的生成物不应对外公布。

3.3 因果关系的认定与举证责任的分配是责任认定的关键环节

在生成式人工智能的侵权行为中,导致损害的原因可能是多因素作用的结果:客户指令、人工智能运算、互联网平台传播……在此过程中,要厘清不同主体的行为对损害结果发生的作用力有多大。如果损害主要是由于人工智能算法缺陷引起的,如果其本身没有恶意使用也能生成侵权作品,则应由算法的设计者和服务提供者承担主要责任;但如果损害来源于使用者恶意使用,而算法只是被动地响应,则应当由使用者作为直接责任人,而服务提供商则应在明显恶意使用方面尽到必要防范义务,亦应承担次要责任。关于何者提出证据的问题,则首先由受害人证明其权利被侵害的事实,并提出初步证据证明其侵权内容系由某一具体生成式人工智能服务所生成;其次可以适当将进一步举证的责任转移给服务提供者,让其通过提供技术日志、合法记录等文件证明自身并未出错。基于证据距离及举证能力考量,在证据责任上作出这样的调整有利于查明案件事实进而做出公平裁判。

3.4 责任形式的确定与责任分担的机制需要与现有法律体系相衔接

无论是对服务提供者、算法设计者、使用者而言,都可能因自己的行为承担多种多样的责任类型:停止侵害、消除影响、赔礼道歉或赔偿损失等。当有多个主体都应应为损害担责的时候,应当如何区分是分则抑或连

带?如果服务提供者与使用者达成合谋行为,例如平台鼓励使用者利用其技术进行侵权以获得更多的流量,则应承担连带责任;但如果各主体的行为没有交流关联,但是共同造成了同一损害结果,就要根据过错程度以及因果关系大小确定按份责任。关于最终的责任承担,也可以借鉴比较过失原则来综合考虑各方行为对损害后果的发生及可归责程度,并公平地确定各人的责任份额。同时也要鼓励服务提供者采取保险契约的方式分散风险,形成社会化事故分担机制,如此一来既保证了受灾人得到及时救助,又减轻单个公司赔付的压力,维持行业的活力。

4 结语

生成式人工智能作为新一轮科技革命的重要标志,对作品创作过程以及传播路径产生深刻的影响,其引发的知识产权问题是法律规则面临的一个现实困境,同时也提供了一个法律改革契机。从本文的研究结果来看,在界定生成式人工智能技术的作品侵权主体方面,需要明确其特点,并运用基础侵权责任理论将其责任进行合理地分配,将主要的注意义务及赔偿责任赋予平台,因其具备管理和获益的能力,而对算法设计者则应当限于开发过程中的疏漏责任。而对于使用者来说,其需要承担其发布指令的目的以及后续的扩散行为。未来在该技术发展成熟并被广泛使用之后,相应的法律法规也需要不断地细化和完善,立法机关可以考虑出台专门的人工智能责任指南,司法机关可以通过典型案例积累裁判经验,行业组织可以制定技术标准与伦理规范。唯有法律与技术协同共进,才能为人工智能的创新发展营造稳定、公平、可预期的法治环境。

参考文献

- [1] 郑世贤. 人工智能生成内容著作权侵权责任主体认定问题研究[J]. 中国价格监管与反垄断, 2025(7): 125-127.
- [2] 吕姝洁. 生成式人工智能侵权责任认定规则: 伦理困境、成因分析、伦理规则及优化路径[J]. 北京邮电大学学报(社会科学版), 2025, 27(2): 87-96.
- [3] 王安琪. 生成式人工智能侵权责任主体研究[J]. 高科技与产业化, 2025, 31(9): 30-33.
- [4] 涂藤. 生成式人工智能服务提供者的著作权侵权责任认定[J]. 电子知识产权, 2025(5): 82-92.