

生成式人工智能著作权侵权问题研究

王蕊

贵州大学法学院，贵州 贵阳

收稿日期：2023年12月11日；录用日期：2023年12月21日；发布日期：2024年2月29日

摘要

生成式人工智能技术快速发展，其创作过程中以及最终的生成物均涉及对他人作品的获取和使用，存在的著作权侵权问题包括著作权侵权行为方式认定存在困境、侵权责任主体界定困难、权利受损主体维权艰难。本文先介绍生成式人工智能概念和运行机理，对著作权侵权问题进行梳理，提出分阶段对侵权行为方式进行判断、根据生成物的权利归属和过错责任确定侵权主体、积极完善并拓宽被侵权人的维权渠道，以规制生成式人工智能的著作权侵权问题。

关键词

生成式人工智能，著作权侵权，权属

Research on Copyright Infringement in Generative Artificial Intelligence

Rui Wang

School of Law, Guizhou University, Guiyang Guizhou

Received: Dec. 11th, 2023; accepted: Dec. 21st, 2023; published: Feb. 29th, 2024

Abstract

The rapid development of generative artificial intelligence technology involves the acquisition and use of works by others in both the creative process and the final product. The existing copyright infringement problems include difficulties in determining the method of copyright infringement, difficulties in defining the subject of infringement liability, and difficulties in protecting the rights of those who suffer losses. This article first introduces the concept and operating mechanism of generative artificial intelligence, sorts out the issue of copyright infringement, proposes a phased judgment of infringement behavior, determines the infringing subject based on the ownership and fault liability of the generated products, actively improves and expands the channels for the rights

文章引用：王蕊. 生成式人工智能著作权侵权问题研究[J]. 法学, 2024, 12(2): 1325-1329.

DOI: 10.12677/ojls.2024.122192

protection of the infringed, in order to regulate the copyright infringement problem of generative artificial intelligence.

Keywords

Generative Artificial Intelligence, Copyright Infringement, Ownership

Copyright © 2024 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

随着当今时代互联网科技及其相关产业的飞速发展，人工智能技术也得到大幅提升，其应用从一开始的工业生产领域，逐步扩张到人类生活的文学艺术创作领域，其运行方式也相应地由简单的、预定的代码运算发展到复杂的、主动的学习分析。2022年11月，人工智能公司 Open AI 推出一款名为 ChatGPT 的生成式人工智能，引发全世界广泛关注，咨询、写作、绘画等产品模型应运而生，并进入市场被社会大众使用。诚然，生成式人工智能以其独立的创作能力和效率优势为社会带来了一定的经济效益，但它同时也给以人类为中心的著作权制度带来了许多挑战，其中隐藏的著作权侵权相关问题值得进一步讨论。

2. 生成式人工智能概念及运行机理

《2023年十大新兴技术报告》中指出，生成式人工智能是一种通过学习大规模数据集生成新的原创内容的新型人工智能工作原理。2023年8月15日起施行的《生成式人工智能服务管理暂行办法》中明确生成式人工智能技术的含义是“具有文本、图片、音频、视频等内容生成能力的模型及相关技术”。不难看出，生成式人工智能是一种区别于传统人工智能的新兴技术，具有强大的学习能力，能够利用大量的数据独立创作生成新的内容。

生成式人工智能创作的运行过程可分为三个阶段，即输入阶段、学习阶段和输出阶段，类比人类的智力创作过程，也可将其表述为“机器阅读”、“机器学习”和“机器输出”。首先，机器阅读是生成式人工智能创作的基础和前提，即从外部收集和提取数据并输入人工智能系统中形成数据副本，建立相应的数据内容库，这里的数据内容库涵盖了创作所需的大量的以数据集表现的数字化作品。其次，机器学习是生成式人工智能的核心所在，所谓“学习”即数据的处理和分析，具体而言，人工智能系统在对大量已有数字化作品进行分类和整理的基础上，分析作品中所表达的思想感情、所采用的语言特征、所特有的表达风格等，从中抽取和提炼出这些作品的规则、模式、结构、趋势，再将这些规则和模式应用到具体创作场景之中，其后通过自我观察独立地、持续地改进和优化其分析和处理。最后，机器输出是在经过数据输入和自主学习后，生成式人工智能构建出解决不同目标任务所需要的函数模型算法，只要使用者发出要求和指令，系统便会通过算法进行处理，最终生成输出相应的学习结果。

3. 生成式人工智能著作权侵权问题分析

3.1. 著作权侵权行为方式认定困境

在我国司法实践中，著作权的侵权认定主要采用“接触 + 实质性相似 - 合理使用”规则^[1]，在此项规则中，“接触”较为容易证明或可以从法律上进行推定，而“实质性相似”与“合理使用”的判定

一直以来都是著作权侵权纠纷的重难点，当将该项规则应用于生成式人工智能侵权纠纷时，则使侵权行为为认定之困境愈加凸显。一方面，这两项规则本身在司法适用中存在抽象性和主观性，并且法律规则具有的这种不明确性与生成式人工智能创作所具有的智能化、精确化的技术特点之间产生较大反差，不加区分地直接套用上述规则，很难得出逻辑严密的认定结果。另一方面，这两项规则的认定标准难以与生成式人工智能运行机理相适配，特别是由于生成式人工智能的学习对象呈现出零散和碎片化的特点，现有标准不适应人工智能创作物对每部作品的使用量可能均维持在较低水平的情况，在一定程度上无法有效地应对人工智能生成物在不符合创作规律下进行低质量创作以避免构成侵权的现象[2]。

此外，虽然生成式人工智能具有一定的创造力，但其创作过程仍然高度依赖于收集和输入的数字化作品，侵害这些作品著作权人的著作权利的问题在某种程度上是不可避免的。我们似乎可以据此将其称为一种科技剽窃，但如此简单断定侵权是不可取的，尤其是当前生成式人工智能使用的算法与模型越来越复杂灵活，一般情况下用户无法知晓哪些作品在何时以何种方式被使用，大多数时候开发者也难以对此作出清晰的判断。

3.2. 著作权侵权责任主体界定困难

在著作权法律关系中，主体资格的取得有以下两种情形：一是自然人通过智力创作这一事实行为成为著作权主体，二是法人基于法律的规定而被拟制作者。一般认为，作者概指进行创作行为的自然人，法人只有在特定条件下“视为”作者[3]。然而，人工智能的出现给以人类为中心的著作权制度体系带来挑战，其基于深度学习而生成具有作品外观的思想表达，似乎与人类自主学习、自主思考、自主输出的创作过程如出一辙，有学者据此提出赋予人工智能著作权法律主体的地位的想法。

但是，就生成式人工智能而言，其不能拥有同人类一样的主观意识，不具有行使权力的能力，没有独立的财产，无法成为具有责任承担能力的主体。对于可以创作出作品的生成式人工智能而言，否定了它的法律主体资格后，参与创作并起到了使作品被创作或传播的作用的开发者、使用者、投资人等参与者，都有各自的理由成为著作权主体，在生成过程中引发著作权侵权问题时，如何界定承担侵权责任的主体成为亟待解决的难题。

3.3. 著作权权利受损主体维权艰难

生成式人工智能收集的数字化作品可谓是海量，在创作过程中又对他人作品进行了广泛的学习模仿、使用组合，区别于传统的著作权侵权，人工智能生成物通常不是只对个别或者少数他人原作品构成侵权，而是可能对众多他人作品造成侵害。一方面，按照通说，著作权只保护表达，不保护思想，对于一些“洗稿”“融梗”等行为，有的人认为这是简单的照搬，构成著作权侵权行为，而有的人认为，正如两个作者可能会不谋而合地以相似的思维方式或审美标准进行创作，生成式人工智能通过学习分析得出算法模型具有相似的表达也仅仅是一种巧合，不至于上升到侵权层面。另一方面，对著作权人来说，如果自己的作品被使用的比例较小，是很难发现自己权利受到侵害，证据提取存在较大困难，这也会导致维权诉讼中侵权认定的难度大增，维权成本往往高于维权收益，打击了权利主体的维权积极性。

4. 生成式人工智能著作权侵权规制路径

4.1. 结合运行机理判断侵权行为方式

首先，“机器阅读”阶段需要收集到足够的数据作为后续学习和输出的素材，其中涉及到对受著作权保护的作品的获取和复制，其实现途径主要有三：一是直接在互联网上抓取已经数字化的数据作品；二是将非数字格式的作品转化为机器可读的数字格式；三是将已经数字化但格式上不兼容的数据进行格

式转换[4]。上述途径可根据现行《著作权法》第10条第1款第(五)项之规定,认定是否属于复制权范围,同时应考量该行为的正当性和合理性,以“转换性使用”规则为基础,将大量使用高价值内容、以商业性目的使用和超出法定许可范围的行为判定为侵权。

其次,“机器学习”阶段对各个作者的表达方式进行分析抽取后,形成各种处理模型,这涉及对之前收集的数字化作品的复制、翻译、改编、汇编行为。基于非法方式获取的数据当然构成侵权,但在著作权法意义上,临时复制合法获取的数据的行为不被认为侵犯复制权。机器学习可根据输出结果是否具有独创性,分为表达型机器学习和非表达型机器学习[5],故若属于前者则输出结果具有一定的独创性,属于后者则输出结果不具有独创性,即侵犯翻译权、改编权、汇编权等权利。

最后,“机器输出”的结果具有一定的独创性,但不可避免地会留存数据库中某个作品或者某些作品的基本表达方式,可能会与他人作品构成“实质性相似”。判断“实质性相似”的方法主要有抽象分离法和整体观察法,在司法实践中一般采用前者,然而,如前所述生成式人工智能创作是在庞大的数据库的基础上对分散的作品片段进行处理,针对生成式人工智能著作权侵权行为,抽象的判定规则无法准确认定,对此应当转向以综合性片段的使用量这一更高的认定标准,扩张实质性相似的认定范围。

4.2. 根据权利归属和过错责任确定侵权主体

我国著作权法确认权利归属的一般原则规定著作权属于作者,自然人的作者地位来源于其创作作品的行为,法人或非法人组织则在特殊情况下可以被视为作者。由此看来,生成式人工智能作品的著作权也应当归属于在创作过程中贡献了创造性劳动的主体。根据洛克的劳动理论,成果的价值主要来源于智力劳动、体力劳动及投资的共同作用,分别对应着开发者、使用者和投资者,但他们都直接或间接的参与了生成式人工智能创作,故还应结合“必要安排”¹的实质性贡献原则进行客观考量。当开发者参与程序设计环节,程序的后续运行过程都由使用者所掌控,那么可以认为使用者作出了必要安排;若使用者只是实行简单的程序启动操作,未干预程序运行,则权利应当归属于开发者;当开发者是严格按照投资者的需求设计的运行程序,那么可以将其视为一种雇佣成果,投资者是权利主体。

在生成式人工智能的实际应用中,往往是多方主体共同参与,有必要进一步根据具体的应用场景,结合过错责任对侵权主体加以明确。首先,开发者事前未经著作权人同意,擅自获取使用大量受著作权保护的作品数据用于机器学习,构成侵权,应承担相应的侵权责任;开发者作为著作权侵权预防的第一环,应当负有防抄袭义务,若其无法证明在设计过程中在机器中嵌入了最先进的防抄袭设计,则推定其存在过错[6]。其次,在开发者受到投资人委托,明知却仍然以非法方式获取数据、设计运行程序的情形下,二者均为侵权责任主体。最后,使用者若故意要求并诱导生成式人工智能生成侵权作品,则表明其具有过错;使用者发现生成的作品侵犯著作权,应履行注意义务,暂停使用,是否尽到了注意义务应根据侵权作品的相似性和对侵权作品删除的及时性进行判断,并由使用者证明。

4.3. 拓宽生成式人工智能著作权侵权的维权渠道

面对复杂的生成式人工智能著作权侵权问题,单个著作权人由于诉讼成本高和取证途径难的原因,对维权诉讼失去动力和信心,这不仅会阻碍我国文化事业的发展,打击创作者的创作热情,也会对人工智能技术带去负面影响。目前,被侵权人大多寻求著作权集体管理组织的帮助,或是采用共同诉讼进行维权。著作权集体管理组织借助其行业上的权威性、诉讼上的专业性,可以将分散的被侵权作品及作者汇聚起来,统一搜集相关证据并进行集中诉讼,降低诉讼代理成本和诉讼费用。共同诉讼模式不仅有利于诉讼材料的收集,证据链的完整,提高司法效率,而且可以让各被侵权人更为自由地行使权利,具有

¹Copyright, Design and Patents Act 1988, 9(3): In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken.

更大的意思表示空间。

但是,上述两种方式均存在不足之处。虽然著作权集体管理组织可以作为代理机构,在一定程度上替著作权人降低维权成本,但实际上该组织还尚未在生成式人工智能著作权侵权诉讼领域充分发挥其作用。而若被侵权人采用共同诉讼的形式进行维权,也存在各权利人诉前联络耗时间、产生意见分歧等问题。因此,一方面需加强著作权集体管理组织在侵权诉讼方面的能动性,建立著作权人维权联络平台等方式完善现有的维权模式;另一方面,其他更加便捷、高效的维权渠道有待开发,以维护著作权人权利,规制生成式人工智能著作权侵权现象。

5. 结语

人工智能技术的发展势不可挡,生成式人工智能融合了科技与文艺,在创作领域开辟一条全新的、富有活力的道路,让更多的人参与到创作之中,肯定其创作成果的同时,不应放任其侵害著作权人权益、动摇我国著作权制度。结合生成式人工智能的三个运行阶段判定各种侵犯著作权行为,根据具体情形确定各侵权主体,加强对生成式人工智能著作权侵权的规制,同时,完善并思考可供被侵权人选择的维权路径,以化解该技术发展过程中的法律风险和利益冲突,形成技术革新与权利保护平衡发展的双赢局面。

参考文献

- [1] 官正艳. 论司法实践中洗稿侵犯著作权的认定标准[J]. 电子知识产权, 2018(11): 78-83.
- [2] 刘强, 孙青山. 人工智能创作物著作权侵权问题研究[J]. 湖南大学学报(社会科学版), 2020, 34(3): 140-146.
- [3] 吴汉东. 人工智能生成作品的著作权法之问[J]. 中外法学, 2020, 32(3): 653-673.
- [4] 焦和平. 人工智能创作中数据获取与利用的著作权风险及化解路径[J]. 当代法学, 2022, 36(4): 128-140.
- [5] 刘友华, 魏远山. 机器学习的著作权侵权问题及其解决[J]. 华东政法大学学报, 2019, 22(2): 68-79.
- [6] 赵宏伟, 茹克娅·霍加. 生成式 AI 背景下著作权侵权样态及其风险治理[J]. 网络安全与数据治理, 2023, 42(9): 59-64.