

浅析AIGC的可版权性与权利归属

俞祉彤

南京理工大学知识产权学院, 江苏 南京

收稿日期: 2024年7月3日; 录用日期: 2024年7月30日; 发布日期: 2024年8月8日

摘要

网络内容生产迈入人工智能生成内容阶段, 以ChatGPT、Stable Diffusion为代表的生成式人工智能在模型训练与内容生成方面均展现出显著的人机交互、人机协同的特征。根据生成式人工智能的特征, 重新构思中文屋测试, 从过程论的视角可以认定AIGC可能具备主观论下的独创性, 具有可版权性。意志能否直接决定表达不应当成为判定创作的核心要素, 创作更应聚焦于独创性的贡献之上, 人工智能在创作中不再是工具的角色定位, 在事实上与人类合作创作作品。作品的著作权应当归属于人工智能的使用者。

关键词

生成式人工智能, Stable Diffusion, ChatGPT, 独创性, 中文屋测试

A Brief Analysis of the Copyrightability and Ownership of AIGC

Zhitong Yu

School of Intellectual Property, Nanjing University of Science and Technology, Nanjing Jiangsu

Received: Jul. 3rd, 2024; accepted: Jul. 30th, 2024; published: Aug. 8th, 2024

Abstract

The production of online content has entered the stage of AI-generated content. Generative artificial intelligence, represented by ChatGPT and Stable Diffusion, exhibits significant characteristics of human-computer interaction and collaboration in both model training and content generation. Based on the features of generative artificial intelligence, rethinking the Chinese Room Test from a process-oriented perspective suggests that AIGC may possess originality under the subjective theory and thus be copyrightable. Whether or not the will directly determines expression should not be the core criterion for judging creation; instead, creation should focus more on the contribution of originality. In the process of creation, artificial intelligence is no longer positioned merely as a tool, but collabo-

文章引用: 俞祉彤. 浅析 AIGC 的可版权性与权利归属[J]. 争议解决, 2024, 10(8): 37-43.

DOI: 10.12677/ds.2024.108350

rates with humans to create works. The copyright of the work should belong to the user of the artificial intelligence.

Keywords

Generative Artificial Intelligence, Stable Diffusion, ChatGPT, Originality, Chinese Room Test

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

人们主观的感觉、认知、思想、创造和表达，以及人文科学、艺术和自然科学都是以实质性的内容作为基础和前提。网络内容生产方式则经历了门户网站整合内容(1994年至2003年)、用户生成内容(2004年至2015年)、专业用户生成内容(2016年至2021年)、人工智能生成内容(2022年至今)四个阶段的变化[1]。2022年以来，Disco Diffusion、Midjourney等人工智能类创作平台和工具在互联网上风靡一时，人工智能生成内容(Artificial Intelligence Generated Content，简称AIGC)已成为数智环境下网络内容生产方式发展变迁的重要趋势。代表AIGC最新进展的是美国OpenAI公司开发的ChatGPT，完成了机器学习算法发展中自然语言处理领域的历史性跨越，即通过大规模的预训练模型，形成人工智能技术理解自然语言和文本生成的能力，可以生成文字、语音、代码、图像、视频，且能完成脚本编写、文案撰写、翻译等任务[2]。这消除了职业技术领域所带来的门槛，使任何人都能够以自然语言和人工智能实现交流，并且生产出多样化的内容产品。在此背景下，以“人类创作”为基石建构的著作权法[3]无疑面临巨大冲击。

2023年11月，北京互联网法院审理了一起被诉侵犯著作权案件，并公布了一审裁判结果。法院认为，涉案人工智能生成图片受著作权保护，但作者或权利人并非人工智能(AI)，也不是AI模型的设计者，而是向AI模型输入提示词并设定相关参数的使用者。判决明确指出人工智能生成图片是否构成作品需要“个案判断，不能一概而论”。原告通过参数、提示词调整修正图片的过程体现了其审美选择和个性判断，即图片能够体现使用者的个性化表达。但法院将“独创性”理解为个性化表达是否恰当？人工智能在生成图片过程中的功能定位如何？图片的个性化表达是来源于使用者抑或人工智能？来源于人工智能的个性化表达还能否被认为具有“独创性”？上述问题均值得思考与探究。

2. 生成式人工智能特征——显著的交互性与协同性

2017年，人工智能微软“小冰”便已能够“创作”诗歌，并出版了诗集《阳光失了玻璃窗》。相较于“小冰”，以ChatGPT、Midjourney为代表的AIGC阶段的生成式人工智能，其模型训练与生成内容的过程均具有显著的交互性与协同性。

模型训练方面，谷歌AI团队采用了人类反馈(Human Feedback)精调Stable Diffusion来提升生成效果，其采用了人机交互的学习机制，优化方式是从人类反馈中进行强化学习[4]。第一步是收集人类反馈数据，由人类对预训练好的Stable Diffusion v1.5基于预定义文本生成的图像进行打分；第二步是学习奖励函数(Reward Function)：给定生成的图像和输入的文本，预测生成图像的评分。人工智能以此获得的泛化与扩展，显然就不再是算力和数据、人与人类经验、先验和社会历史现实的简单叠加，而是交互的深度融合[5]。微软“小冰”创作诗歌的核心则是深度学习，即通过模拟大脑神经网络的多层结构，使计算机对于输入

的数据进行分层处理。在预先设定程序代码后，微软“小冰”对数千首诗歌进行的上万次深度学习过程中，程序反复运行，实现诗歌表达规律性特征的发现与重组，并不断累积、更新。这一学习训练的过程，人类并未直接干预。

人机交互、人机协同在当前内容生成的过程中体现尤为显著，AIGC 是人机协同参与的数字内容生成方式。ChatGPT 根据人类输入的提示词(Prompt)生成所需内容，同时具有强大的记忆、多轮对话与连接上下文的能力，用户可以通过衍生提问，在初始生成内容的基础上引导 ChatGPT 进行修改完善。微软“小冰”则对用户输入的图片进行图像特征的识别和抓取，特征经处理后由映射函数转换为相应词汇，再通过对词汇的排列组合，最终输出成型的诗歌语句。这一过程采用了符号到符号的生成机制[6]，实为机器学习完毕后的算法接收输入值后输出的过程，生成内容并无人类意志作为中介。

生成式人工智能的发展打破了人、机器与信息资源之间的边界，重塑了模型训练与数字内容生成的范式，实现了由人与信息、人与人的互联，到人机交互、人机协同创作的延伸。拥有一定程度的认知和交互能力，是 AIGC 发展的重要趋势[7]。

3. 独创性的判断

《著作权法》第 3 条对于“作品”的定义为：“文学、艺术和科学领域内具有独创性并能以一定形式表现的智力成果”。唯有具备“独创性”的外在表达，方能构成著作权法意义上的作品，“独创性”是作品区别于其他人类劳动成果的关键[8]。但是独创性的含义阐释与标准界定却并未得到明确，独创性的判断是仅仅评价客观的表达，还是在此基础上并考量主观的思想及其背后蕴含的人格与个性，学者们对此持有相异的观点。持客观论的学者强调独创性可以不源自于人，内容本身具有最低限度的创造性，在形式上产生了与人类作品相互混淆的结果，能够为公众提供与人类作品相同的利益，便可认定该内容具备了独创性[9] [10]。持主观论的学者以“人格价值观”作为著作权之哲学基础，认为作品表达的是人类的思想情感，人工智能生成内容并未体现人类在创作领域的意志能力[11]。人工智能产生相关内容的过程是算法、规则和模板的应用[12]，而主体在创作过程中将自己的精神与意识融入作品之中，作品是作者主观构想的再现[13]。上述客观论与主观论之分歧，可追溯至两者采取了不同的判断视角——主观论在客观论的结果视角之上考察了内容的生成过程。相较于人类创作，人工智能生成的内容与人类的思想之间存在更远的间隔，算法取代思想的运用、演绎而将思想转化为表达，人工智能的这一转化不需要再以理解文字符号的含义为基础，人类的思想是否融入到人工智能生成的内容中存在疑问。

著作权的人权意义不可隐没，知识产权的保护对象是体现人类价值和尊严的智力成果[14]，创造与表达本身就隐含了主体的意向，故笔者更为认同独创性认定以人格主义要素为基点的主观论。但笔者认为即使采取过程视角，在人机交互与人机协同不断深化的 AIGC 时代，也不宜否认当前以 Stable Diffusion、ChatGPT 为代表的人工智能生成内容的内在人类意志。

主观论的论证路径可溯源至人工智能哲学领域的中文屋测试，塞尔对结果主义的图灵测试提出质疑，认为人工智能不具有意向性，机器的形式转化能力并不是理解力[15]。塞尔假设一个对中文一窍不通的人被锁在房间内，给予他批量的中文文本，这些文本对于他而言仅仅是毫无意义的形式符号；同时，给予他一套操作中文字符的规则(类比程序)，该套规则用他可以理解的英语书写。此时，屋外有人以中文提出问题(输入)，房间内的人能够依据规则在中文文本中组合出词句(输出)回答输入的问题，以致从外部看来其回答与掌握中文的人毫无区别。通过该思想实验，塞尔反驳了“如果机器能以无法区别于人类的方式输出，则可判定机器具有了像人一样的智能”[16]，认为计算机的程序编码系统是纯语法、纯符号、纯形式化的过程，人类的大脑是智能的——拥有意向性和语义，两者完全不同，故计算机不具有智能[17]。

在“中文屋测试”条件设置中，融入当前 AIGC 的具体情境，恰能为人工智能生成内容具有主观意

义上的独创性提供可能的解释路径。上文已经论证了以 ChatGPT 为代表的生成式人工智能,在模型训练和生产内容的过程中都具有显著的人机交互与人机协同的特征。在此逻辑前提下,笔者认为 AIGC 背景下的房间内不再仅有一位毫不理解中文的屋内人 A,而是多了两位掌握中文(人类意识)的主体 B、C。主体 B 对 A 运用规则操作中文字符的行为进行指导(人类反馈强化模型),使 A 能够熟练地掌握运用这一技能;主体 C 对 A 组织形成的词句提供修改完善建议(内容生成过程中用户不断给予指令引导人工智能),在三者的合作下房间作为整体向外界输出回答(人工智能生成内容)。人工智能生产内容的过程始终体现人类的选择与安排,人的意志与个性已经内嵌入人工智能生成的内容中,将包含有多个主体、中文文本、中文操作规则的房间视为系统,该系统整体具备意向性,系统生成的内容也就具备主观独创性。

诚然,人工智能生成内容中人格要素的渗透较之传统作品似乎有所淡化。但随着技术的发展,作品独创性判断偏向弱化作者个性这一趋向,在国际立法与司法实践中均已有所体现[18]。计算机软件、数据库等作为功能性的实用技术成果,作者的个性、人格特征难以彰显,亦被纳入著作权保护的范畴。“追气球的熊孩子”案件中,技术因素导致人类无法参与创作过程,创作内容中作者个性弱化,法官通过人类介入、参与的路径论证作品的独创性,对于前述中文屋测试的论证亦具有一定的借鉴意义。在自由飞行的气球上固定有自动拍摄的相机,人类无法进行操控,但在照片拍摄、形成的过程中人为因素的参与使人以独创性的方式在拍摄过程中发挥作用,可以满足构成作品的独创性要求。美国 Time Incorporated v. Bernard Geis Associates 一案法官的说理也有相似之处。

综上,在人类训练人工智能、人工智能与人类互动生成内容的过程中,人类的引导、条件设定、修改润色等人类参与行为,使人机交互的创作整体具备意向性与理解力,人类的主观意志融入人工智能生成内容之中。人工智能生成内容因而可能具备主观论下的独创性,成为著作权法保护的作品。

4. 人工智能作品的权利归属

4.1. 生成式人工智能的角色定位

网络内容生产迈入 AIGC 时代,借助生成式人工智能,人类在内容生产中的直接参与度、贡献度显著下降,且生成内容几近人类创作作品。由此,生成式人工智能在内容生产中的角色定位发生模糊,人工智能究竟是直接创作内容的主体抑或是辅助人类创作的工具?明晰人工智能的角色定位,是探讨人工智能生成内容权利归属的逻辑前提。众多著作权法学者已就此进行了充分的讨论,其观点可凝练为人工智能“创作工具”之辩。肯定者认为:人工智能生成内容更类似于人类利用技术设备创作作品,与人工智能直接产生的成果相比仅是在技术实现步骤方面具有多层性、梯次性和间接性[19]。创造的概念可以容纳技术的变革,就如同摄影技术诞生后,尽管图像的精确复制由机器完成,人类依然保有取景的自由意志[6]。否定者认为:人工智能“输入”与“产出”的过程类似于人类的创作,人类思考创造的机理尚未明确,不能断然否定人工智能存在创作行为[20]。能够生产自主性内容的人工智能,具备深度学习能力,在创作上超越了机械延伸的范畴,已经摆脱了内容创作辅助工具这一地位[21]。在否定论的基础上也存在分歧,部分学者认为人工智能对 AIGC 的独创性表达做出了主要贡献,成为 AIGC 的单独创作主体[3],另有学者认为,AIGC 由“机器作者”与人类作者合作创作,是“智能版权”时代合作作品的特殊类型[22]。人工智能图片案中,法官认为 AIGC “本质上仍是人利用工具进行创作”的同时,强调 AI 模型画图与“人们之前使用画笔、绘图软件有很大的不同”,但并未具体指出不同的本质及其对人工智能“创作工具”地位的影响。

《著作权法实施条例》第 3 条第 1 款规定:“著作权法所称创作,是指直接产生文学、艺术和科学作品的智力活动。笔者认为对该条文创作的理解中,片面强调“自由意志直接决定内容的表达性要素”

等意志(思想)对于表达控制力的观点似乎偏离了创作之本义[23]。思想与表达之间本就存在距离,最贴近思想的语言和文字表达也偶有言(词)不达意之感,创作工具更是增加了两者之间的层次性和间接性,意志没有直接决定表达不能成为否认创作的有力论据。例如,一摄影师设置相机的参数意欲拍摄出明亮温暖风格的图片,但其对于环境光线的判断出现偏差,最终呈现的效果是阴暗忧郁风格的图片,其表达偏离了意志,但不可否认这一拍摄过程仍是创作。

创作仍应追溯至独创性,从功能主义的视角探究人工智能对于生成内容的独创性的贡献,这是明确人工智能角色定位的基础和关键。

人工智能对于独创性表达的贡献无可否认,设计者设置的模板、参数与使用者输入的提示词仅仅只是一个粗略的框架或概念,最终呈现的符号表达的选择、取舍、编排几乎都是由人工智能完成的。但若认为生成内容中人类的贡献程度显著下降,人工智能已经单独对于独创性表达做出主要贡献的观点亦是不妥。首先,人工智能尚不能模仿人类智慧中最具创造力的部分。我国著名人工智能专家钟义信认为,人类智慧最为充分的表现就是:为了实现自己的目标而不断地发现问题、定义问题和解决问题的能力。其中,发现问题与定义问题的能力是人类最具创造力的能力,这种能力依托于人的想象、知识、审美、灵感等抽象的能力[24]。在人工智能生成内容的过程中,迸发灵感欲解决问题并将问题具象定义的步骤是由人类完成的,人工智能无法自发地发现问题。其次,从人工智能的运行机理出发,如 ChatGPT 运用名为转化器的神经网络架构进行语言的概率性分析,选择最可能的搭配内容作为答案高效地提供给人类提问者[25]。倘若没有或欠缺关键词选择、引导为生成的内容注入人类创造性的灵感与思想,人工智能依据最可能的概率分析生成的通常只会是惯常性表达,而非著作权法保护的独创性表达。因此笔者认为,对于 AIGC 时代背景下的人工智能生成内容,事实层面上是人工智能与人类“合作创作”的结果,应当将人工智能进行深度学习,自行选择、识别、分析数据并输出结果,人工智能的使用者对结果输出给予特定指令或目标,不断引导作为一个相互交融的创作过程的整体。两者不可割裂,缺乏任一都无法产生前述著作权法意义上的“独创性表达”。此外,有学者将使用者对生成内容的贡献类比为导师指导学生论文——导师在学生论文研究方向的确定、篇章结构的安排、论证思路的完善等方面都起到了重要的指引作用,但导师贡献的是不受著作权法保护的思想和方法,认为使用者同样只是为独创性表达贡献了思想,人工智能如同学生一样对独创性表达做出了主要贡献。学生具有人工智能所缺乏的主观能动性,即使没有老师加以指导,其可能不足以写出结构严谨论证翔实的学位论文,但其论文能够成为著作权法保护的作品应当并无疑问。但当前的人工智能欲产生独创性表达,则必须依附于某一人类创作主体,这一类比可能并不具有足够的解释力。综上,在事实层面,人工智能介入人类创作,人机合一实施了合作生成作品的行为,两者都对独创性表达做出了实质贡献,缺一不可。

4.2. 权利归属

人工智能与人类在事实上以密不可分的协作关系生成了具备主观独创性的作品,但人工智能的著作权主体资格不应被认定。其一,在康德“主客体统一认识论”和“人是目的”哲学观点下,无论人工智能发展到何种阶段,都只能作为人利用的客体和工具处理,而不能将其拟制为与人享有平等地位的法律主体[26]。其二,人工智能日益先进的智能背后也蕴藏着巨大的法律和社会风险,必须作为客体保证其处于人类的支配和控制之下。其三,人工智能没有独立的意识能力和特定的社会角色,不能作为著作权人去理性并真实地享有权利、履行义务和承担责任。

生成式人工智能的研发依赖大量的资本投入,出于鼓励和保护投资的需要,部分学者支持将人工智能生成作品的著作权归属于投资者[27]。但向投资者赋予人工智能生成作品的著作权,并不必然带来收益的利好[28],反而可能因为收费阻碍信息在流转中的升值。投资者会自发地进行利益权衡,选取合适的收

益模式。开发 ChatGPT 的 OpenAI 就自发地放弃了生成内容的版权而采取了订阅制，同时 ChatGPT 也被集成到微软的必应搜索引擎和 Office 办公软件中开辟新的收益渠道。

使用者拥有人工智能生成作品的著作权，契合了以劳动财产学说与激励理论为基石的著作权制度[20]。洛克“财产权劳动学说”将劳动作为财产产生的根据，著作权是基于作者的创造性劳动作为法律保护的正当性基础，并脱离自然状态而成为一项私有财产权利。人工智能的使用者与人工智能生成作品的联系最为紧密，并且在内容生成过程中进行一定的选择、编排与加工。激励作品创作与传播是著作权立法的核心目的。人工智能生成内容与使用者的需求最为密切相关[29]，因此对于使用者的激励也最为有效，使用者最易因此产生强烈的创作欲望。人机协作的创造范式得到激励与推广后，无疑能够以更为高效的方式促进社会知识产品总量的增加，实现著作权法增进社会福祉的立法目标。

5. 结论

新的技术革命无疑会带来创作形态的变化，不应固守非人工智能时代基础上的著作权法的概念和逻辑，应当以一种开放的心态拥抱 AIGC 时代的到来，并在此基础上对著作权法中的概念和规则进行全新的解释。当前的生成式人工智能具有显著的交互性与协同性，客观上人工智能生成内容本身能够具备最低限度的创造性，人类的主观意志在人机互动的过程中融入生成的内容使之具备主观层面的“独创性”。人工智能生成内容的独创性来源于人类，但人工智能也对其做出了实质性贡献，故可视为两者合作创作，权利归属于使用者所有。

参考文献

- [1] 谢新洲, 韩天棋. 基于数字技术的网络内容生产方式变迁研究[J]. 新闻爱好者, 2023(8): 14-20.
- [2] 杜雨, 张孜铭. AIGC: 智能创作时代[M]. 北京: 中译出版社, 2023: 4-5.
- [3] 丁文杰. 通用人工智能视野下著作权法的逻辑回归——从“工具论”到“贡献论”[J]. 东方法学, 2023(5): 94-105.
- [4] Lee, K. (2023) Aligning Text-to-Image Models Using Human Feedback. <https://arxiv.org/pdf/2302.12192v1.pdf>
- [5] 涂良川. “生成式人工智能”逼近通用智能的哲学叙事——ChatGPT 追问智能本质的哲学分析[J]. 东北师大学报(哲学社会科学版), 2023(4): 40-47+93.
- [6] 李琛. 论人工智能的法学分析方法——以著作权为例[J]. 知识产权, 2019(7): 14-22.
- [7] 李白杨, 白云, 詹希旎, 等. 人工智能生成内容(AIGC)的技术特征与形态演进[J]. 图书情报知识, 2023, 40(1): 66-74.
- [8] 王迁. 知识产权法教程[M]. 第7版. 北京: 中国人民大学出版社, 2021: 58-59.
- [9] 谢琳, 陈薇. 拟制作者规则下人工智能生成物的著作权困境解决[J]. 法律适用, 2019(9): 38-47.
- [10] 郑远民, 贺翔溪. 结果视角下人工智能生成物的保护路径检讨[J]. 科技与法律, 2020(3): 14-21.
- [11] 李俊. 论人工智能生成内容的著作权法保护[J]. 甘肃政法学院学报, 2019(4): 77-85.
- [12] 王迁. 论人工智能生成的内容在著作权法中的定性[J]. 法律科学(西北政法大学学报), 2017, 35(5): 148-155.
- [13] 杨述兴. 作品独创性判断之主观主义标准[J]. 电子知识产权, 2007(7): 64-65.
- [14] 吴汉东. 知识产权法总论[M]. 北京: 中国人民大学出版社, 2020: 59.
- [15] J·R·塞尔. 心灵、大脑与程序[M]. 上海: 上海译文出版社, 2006: 74-76.
- [16] A·M·图灵. 计算机与智能[M]. 上海: 上海译文出版社, 2006: 56-57.
- [17] 孙会. 意向性与人工智能: 基于对“中文屋论证”的批判性考察[J]. 科技管理研究, 2019, 39(10): 200-205.
- [18] 周澎. 算法生成物的著作权保护研究[M]. 长春: 吉林大学出版社, 2022: 96-98.
- [19] 冯刚. 人工智能生成内容的法律保护路径初探[J]. 中国出版, 2019(1): 5-10.
- [20] 卢海君. 著作权法意义上的“作品”——以人工智能生成物为切入点[J]. 求索, 2019(6): 74-81.
- [21] 易继明. 人工智能创作物是作品吗? [J]. 法律科学(西北政法学报), 2017, 35(5): 137-147.

-
- [22] 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.
- [23] 王迁. 再论人工智能生成的内容在著作权法中的定性[J]. 政法论坛, 2023, 41(4): 16-33.
- [24] 钟义信. 人工智能:“热闹”背后的“门道”[J]. 科技导报, 2016, 34(7): 14-19.
- [25] 邓建国. 概率与反馈: ChatGPT 的智能原理与人机内容共创[J]. 南京社会科学, 2023(3): 86-94+142.
- [26] 李扬, 李晓宇. 康德哲学视点下人工智能生成物的著作权问题探讨[J]. 法学杂志, 2018, 39(9): 43-54.
- [27] 陈晨. 论人工智能生成内容的可版权性问题——以 ChatGPT 生成内容为例[J]. 科技与出版, 2023(6): 98-106.
- [28] 刘鑫. 人工智能对知识产权制度的挑战与破解——洛克“财产权劳动学说”视角下的路径选择[J]. 云南社会科学, 2020(6): 138-145+185.
- [29] 李杨. 人工智能生成物著作权问题的一点胡思乱想——作品创作和利用的世俗化、自由和平等[EB/OL]. <https://mp.weixin.qq.com/s/Xalpru4uSTNy23GOiSmOQ>, 2023-12-01.