

人工智能大模型训练数据知识产权问题研究

刘 慧

天津外国语大学 天津 300204

摘要: 人工智能大模型正在重塑我们的生产生活方式, 引起巨大社会变革, 与此同时, 其训练数据涉及版权问题引起诸多争议。论文从行动者网络理论视角审视大模型训练数据知识产权问题, 不仅带来深刻哲学启示, 同时为促进该技术发展提供新的对策思路。

关键词: 人工智能大模型; 行动者网络理论; 知识产权

引言

生成式人工智能训练依赖海量数据, 数据的数量和质量深刻影响模型效能, 开发者在获取数据的角逐中大显神通, 由此引发诸多纠纷。其中, 训练数据的知识产权侵权纠纷尤为突出, 正在成为知识产权保护的新战场。

人工智能大模型强调交互性, 模糊了人与技术之间的主客体关系。技术的能动性增强, 使得传统主体-客体的二分法难以充分解释实体间的联系。鉴于此, 当前需引入新的理论框架对多元主体进行系统分析, 进而探索出一条合作共赢的解决路径。行动者网络理论以其对多实体互动分析的优势, 为解构此类博弈提供了独特视角。

一、行动者网络理论: 理论框架及其适用性分析

(一) 广义对称性: 异质行动者的网络构成

ANT由米歇尔·卡龙等学者提出的一种社会学理论框架, 认为社会现象是以行动者为节点构成的网络, 人类行动者与非人类行动者在行动能力上具有对等地位, 被平等地视为潜在网络的构成者、参与者与改造者^[1]。ANT超越传统的人类中心主义分析范式, 强调非人类行动者的能动性。具体而言, 在人工智能大模型的行动者网络中, 行动者包括模型开发者、投资人、训练数据提供者和AI产品的消费者等为人熟知的主体, 也包括训练数据、算法系统、硬盘和版权法等非人类行动者。其中,

作为实现大模型的主导, 开发者是网络的核心行动者。其行动面临着来自技术、法律以及伦理等多方面来自其他行动者的压力。在行动者世界中, 行动者基于不同的社会行动目标, 与其他行动者形成流动性的连接关系, 共同塑造技术、法律、伦理的冲突景观。

(二) 转译: 行动者网络的动态特征与复杂互动

在行动者网络中, 开发者通过技术创新与生态协作, 持续推动人工智能大模型从实验室研究走向规模化应用。为实现其目标, 开发者对网络内其他实体进行识别和解释, 赋予其在行动者网络中角色与关系, 这一过程被称为转译。具体而言, 他们转译AI产品使用者为消费者、转移创作作品为训练数据、转译创作者为训练数据提供者等。然而转译可能会受阻: 如果其它行动者并不认同开发者赋予的角色, 以不同的角色定义自己的未来, 转译将会失败, 行动者网络也将分崩离析。以创作者为例, 开发者倾向于自由使用将其作品来提升训练模型, 创作者随着被赋予训练数据提供者的角色。但创作者往往视为其作品为私人资产, 而开发者被视为无偿劳动侵占者, 创作家们断然会拒绝开发者为他们准备的角色, 并通过发起诉讼、抵制平台等方式来进行反击, 开发者也不会坐以待毙。为应对网络中的冲突与不稳定性, 开发者实施新的转译策略, 卡龙将转译过程分为四个关键环节: (1)“问题化”(如何成为不可或缺的部分); (2)“收益共享”(如何留住同盟者); (3)“招募成员”(如何定义和协调各种角色); 以及(4)“动员”(代言人是否具有代表性)^[2]。

二、转译路径: 从对抗到协作的实践策略

(一) 问题化: 构建“数据获取合法性”的必经之点
“问题化”是指将不同的行动者的关注对象问题化,

基金项目: 本文系2024年天津外国语大学研究生科研创新项目“大语言模型时代中国知识产权保护体系的构建研究”(2024YJSS009)研究成果。

作者简介: 刘慧(1999—), 女, 汉族, 山东人, 天津外国语大学硕士在读, 研究方向: 美国社会与文化。

通过一系列转化使实现某个目标成为所有行动者要达到的目的必经之点。人工智能大模型的行动者网络中，开发者面临的核心问题是：“数据获取的合法性”。为了使数据为训练模型所用成为各行动者实现目标的必经点，开发者事实了一系列策略，包括通过利用算法不可逆性和分布式存储，使数据一旦被嵌入到大模型中，内容生产者便无法追溯或移除他们的数据，构建起人工智能大模型的数据依赖型技术生态。例如Raw Story对openAI的“移除相关作品副本”诉求，因技术上不可行，从而被迫依赖开发者提供的平台和解决方案，这加强了开发者的控制权。其二，开发者通过话语重构将“数据贡献”转化为“参与人类文明进程的机遇”，通过描述大模型是科技领域的一项重大突破，创作者若试图融入人工智能的叙事，则必然经过开发者设定的强制通过点。

基于当前技术生态与价值框架重置，创作者无论以自主参与或被迫卷入的方式进入人工智能大模型的行动者网络，均被纳入开发者的问题化框架，从而在实质上承认其必经之点的地位。需要注意的是，问题化阶段隐藏的权力不对称性为网络动态稳定性埋下了张力，此阶段的局限性需要转译的其他阶段协同作用来弥合。

（二）收益共享：开发者双重行动策略

为强加和稳定问题化阶段中所定义的其他行动者的身份，行动者可以通过物理或者话语等一系列行动，防止其他行动者被其他竞争性利益吸引，从而确保他们与开发者的目标保持一致。具体而言，开发者采取两种策略来吸引创作者：其一，技术重构。开发者依托数据清洗、去标识化与合成数据生成等技术，将原始数据转化为法律认可的“非侵权材料”，声称已剥离版权信息，以此剥离艺术家、作家的版权持有者身份，他们被转译为“匿名化内容提供者”，使创作者既参与技术进程又免于显性侵权风险。例如Stable Diffusion模型使用的数据集源自LAION-5B的公开网络数据，数据被包装为“公共知识资源”，从而淡化个体版权主张；其二，开发者可能通过经济回报进一步确保内容生产者接受开发者为他们准备的角色。为实现技术合法性与创作生态的平衡，采取了版权分成及订阅分成等经济补偿策略以建立契约关系。这种合作模式为版权拥有者提供了公平的报酬机制，确保其知识产权成果能够在市场流通中实现价值转化，构建了可持续的合作关系网络。两种策略共同作用，实现开发者主导的网络进一步动态稳定。

（三）招募：异质行动者的差异化招募

“招募”是通过特定设备或机制，将一组相互关联的角色分配给参与者的过程，描述此阶段就要分析其中的多边谈判、实力考验和利益分配以及成功所需的技巧^[2]。在此阶段，开发者作为网络核心行动者，通过定义角色、分配职责与构建协作关系，将版权方、用户、法律机构、学术界、企业客户等其他行动者纳入网络，并确保其接受预设角色，从而推动技术落地，确保得到多方支持。首先，版权方因权属主张成为网络稳定性最脆弱的节点，最有可能发生背叛，开发者通过动态收益分配机制降低其退出风险，固化其“授权内容提供者”角色。在稳定版权方的同时，开发者对其他行动者则采取差异化策略：首先，开发者推动“技术中立性”或“合理使用”原则的司法适用，甚至游说政策制定者，法律机构会转译为“创新支持者”而非“侵权追责者”，使之成为支持新科技革命的群体。例如OpenAI在面临《纽约时报》诉讼时援引“合理使用”原则，主张模型训练属于知识转化而非内容复制。其次，在用户方面，开发者通过强调公共利益属性，用户被定义为“技术受益者”，使其接受AI服务的便利性优先于关注版权争议，强化AI服务的公共产品属性，从而使公众利益与技术资本诉求形成共识；最后，开发者主动与高校与研究机构建立合作，将学术界转译为“协同创新者”，形成支持人工智能大模型开发的利益联盟，增强开发的学术公信力。例如，北京大学与字节跳动共建“豆包大模型系统软件联合实验室”^[3]，构建产学研协同育人新模式，将学术公信力注入商业实践。

（四）动员阶段的动态稳定机制与脱嵌风险

“动员”是行动者“用来确保相关集体的所谓代理人能够正确地代表这些集体，而不被后者背叛的关键策略，是行动者网络系统稳定的阶段。代理人既是群体的“代言人”，又是群体间的“协调者”。在此阶段，网络建构者通过一系列策略，确保这些代理人准确传达所属群体的立场与需求，并在行动者网络中保持忠诚。开发者通过法律授权、技术协议等制度化手段，将出版商、社交平台、艺术家等原创内容生产者的代理人将后者嵌套至网络中，明确各方权责；同时，利用透明化技术工具实现版权方收益的可追踪，从而提升了代理人的话语权与忠诚度。通过赋权与制衡，开发者将版权方、艺术家和作者等群体深度纳入行动者网络，构建了稳态延续的网络。以Adobe的大模型Firefly为例，其通过透明的数据

管理机制缓解艺术家对其作品未经授权抓取的担忧；同时，Firefly提供了明确的授权机制和补偿方案，使艺术家能够按照授权条款参与AI训练过程。当作品被用于训练或生成内容时，艺术家会获得合理收益。这种合法性共建框架为生成式大模型的版权治理提供了可复用的动员路径。然而，仍需注意的是，庞大的训练数据拥有海量权利主体，开发者无法与所有创作者个体直接对话，转而与版权聚合商协商，这会导致难以做到百分百的准确授权。例如，《金融时报》与OpenAI签署内容授权协议，Reddit也与谷歌签署了类似的交易。此类协议虽然保障了出版商获取经济利益，但这些巨额收入与内容生成者却没有直接关系，其经济回报仍为零。这种未征求内容生成者意见的交易可能会引发大众的不满情绪，从而威胁到网络稳定性。针对此，应建立集体管理组织解决授权难题：为出版商以及所有做出贡献的主体建立专门的人工智能训练作品集体管理组织，以便在前者作品用于商业用途时可以有效地加入并获得报酬。

历经转译的四个阶段，开发者推动构建了更具韧性的行动者网络。从行动者网络的视角看，为缓解人工智能大模型所涉及的训练数据相关的权益冲突，开发者需注意以下几点：首先，应迎合其他行动者的利益诉求。开发者需综合考虑多元行动者的权益，如版权方的署名权和修改权、原创个体的著作权、以及平台用户的数据所有权等。其次，重视“物”的能动作用。将技术、数据、协议等“物”视为行动者，借助其力量实现版权方、开发者、平台用户等实体的利益共享。为此，需建立有效的经济补偿与激励机制，构建量化价值交换系统，例如分层分成协议、数据贡献代币化等；并通过技术手段增强数据训练透明度，让参与者能够知晓数据使用情况。

最后，科学事实的形成是群体协作的结果。科技进步并非孤立发生，而是通过多个行动者的共同努力和协作逐步构建而成。这意味着整个协作网络中，各方群体的贡献都不可被忽视。尤其是除开发者之外的其他行动者，他们的参与对于整个网络的稳定与成功具有不容忽视的作用。

结语

我国人工智能大模型领域已出现过多起训练数据相关的纠纷和争议。作为技术主导者，开发者须摒弃“技术至上主义”的单向逻辑，主动承担社会责任，保障版权方的合法权益。开发者应构建包容性协商机制、实施透明化操作并探索弹性利益分配模式，既彰显对人类知识成果的尊重，又确保创作者持续供给训练资源的积极性，从而推动人工智能大模型行动者网络由“脆弱联结”向“韧性生态”进化。

参考文献

- [1] LATOUR B. Reassembling the social: An introduction to actor-network-theory[M]. Oxford: Oxford Press, 2007: 70-74
- [2] Callon M. Some elements of a sociology of translation: Domestication of the scallops and the fishermen of St. Brieuc Bay[M]//Law J E. Power, action, and belief: A new sociology of knowledge. London: Routledge, 1986: 196-223.
- [3] 北京大学-字节跳动成立“豆包大模型系统软件联合实验室”，聚焦AI系统软件关键技术问题[EB/OL]. (2024-12-12) [2025-02-14]. <https://www.ithome.com/0/817/209.htm>