

AI 领域数据合理使用的司法实践与产业启示

生成式人工智能的爆发式发展，使数据成为驱动产业创新的重要生产要素，而数据合理使用边界的模糊性，已成为制约人工智能产业合规发展的瓶颈。此类纠纷的本质是数据合理使用规则的适用分歧，不仅关乎相关法律条款的落地，更直接影响人工智能产业的合规发展。赛迪研究院建议建立健全数据合理使用规则体系、完善数据产权保护体系、强化协同共治，为人工智能产业合规发展、数据要素高效利用提供指引。

一、人工智能数据权益的核心知识产权界定需要厘清两个问题

（一）数据产权权属界定

一是权利主体与内容划分不明。数据流转涉及收集者、整理者、原始权利人等多主体，各主体的权利范围如何界定，缺乏统一标准。二是数据类型的权属差异模糊。从形态看，独创性汇编作品与非独创性数据集合的保护路径未明确区分；从来源看，个人信息数据、企业经营数据、公共数据的产权边界亟待厘清。三是非独创性数据集合争议突出。实践中此类数据的权属证明难度大，如数据堂案中双方争议点为非独创性数据的权利归属，因我国缺乏此类数据权利认定的明确依据，导致案件审理中权属判定难度较大，这也直接暴露了相关规则空白，进而影响数据的合法流转与利用。

（二）数据合理使用边界

数据合理使用的争议也是 AI 场景下数据权益纠纷的首要焦点。一是 AI 训练使用他人数据的合理使用认定标准不明。对于 AI 训练行为是否符合合理使用范畴，尚未细化明确判断要素，包括使用行为是否具备转化性、是否会对权利人的市场利益造成实质性损害等核心考量维度，导致实践中裁判标准不统一。二是商业与科研用途的边界划分不足。现行规则未对商业性 AI 训练与科研性 AI 训练的合理使用范围作出明确差异化界定，导致在实践中难以兼顾技术创新的发展需求与知识产权人的合法权益

保护。三是非法来源数据的合理使用争议。针对盗版数据、未经授权获取的敏感数据等非法来源数据，其是否存在合理使用的豁免情形，目前仍存在争议。

二、当前人工智能合理使用数据面临的制度困境

（一）制度存在空白，具体规则不明确

一是人工智能专项规制体系缺失。我国尚未出台针对人工智能数据合理使用的专项立法，《生成式人工智能服务管理暂行办法》虽明确训练数据需合法合规，但对数据来源合法性界定、流转环节违法性传导等主要问题未作细化，仅依托《著作权法》《反不正当竞争法》等传统法律构建多层保护体系，存在明显规制缺口。二是非独创性数据保护标准模糊。相关不正当竞争纠纷案如“刷宝 APP”不正当竞争纠纷案虽界定了非独创性数据集合的经营性利益保护规则，但缺乏统一的权利认定标准与救济路径，导致企业对数据投入的预期稳定性不足。三是全链条责任划分不明。AI 模型训练、内容输出等环节的侵权责任边界，以及避风港规则在 AI 平台的适用范围仅通过个案探索，未形成系统性规则，增加企业合规不确定性，加剧企业法律风险顾虑。

（二）制度弹性不足，平衡创新与保护的能力有限

一是规则包容空间不足。现有制度以著作权优先，反不正当竞争法兜底为核心，未明确转化性使用、小额使用等豁免情形，对商业性 AI 训练的合理使用预留空间狭窄。二是创新激励效能

受限。规则刚性过强导致企业因担心侵权风险放缓研发进度，难以兼顾技术创新需求与知识产权人合法权益的平衡。

（三）跨域适配性弱，制约企业国际化布局

一是规制体系差异显著。我国数据合理使用制度与欧美主流框架存在区别，美国实行逐案裁量的合理使用机制，欧盟坚持先授权后训练刚性规则，我国则缺乏明确细化的适配标准，企业难以精准对接不同法域规则。二是跨域衔接机制不完善。虽然我国已建立基础衔接框架，但尚需加强数据要素市场国际合作，建立数据授权、合规互认等合作机制，提高衔接的针对性和高效性，促进企业出海。三是企业出海数据合规成本攀升。我国 AI 企业出海需同时适配多法域数据合规要求，运营成本大幅增加，部分企业因难以应对规制差异，被迫收缩国际市场布局。

三、主要法域合理使用相关司法实践比较

（一）我国无专门规则，数据合理使用边界有待司法实践探索明确

我国暂无针对 AI 数据合理使用的专门立法规则，司法实践中依托《著作权法》《反不正当竞争法》等现有框架兜底调整，通过典型案例逐步明晰裁判边界。最高人民法院指导性案例某科技有限公司诉某文化传媒有限公司不正当竞争纠纷案具有重要司法意义。

该案中，某文化传媒有限公司未经许可，擅自抓取搬运某科技公司甲 APP 的短视频、用户信息及评论等数据，在其乙 APP 展示传播致内容同质化，某科技公司起诉后，法院认定其构成不正当竞争并判令担责。该案明确了平台对其投入人力、物力培育形成的、具有经营价值的数据集享有合法保护权益，同时界定了未经许可抓取、搬运该类数据并造成实质性损害的行为，属于不正当竞争这一侵权认定标准，为司法探索数据合理使用边界提供重要参考。

（二）美国以逐案裁量为主，配套政策强化创新激励

美国以《版权法》第 107 条合理使用四要素为核心，实行逐案判断模式，采用转化性使用和来源合规的判定标准。2025 年 6 月加州北区联邦法院审理的 Anthropic 案是训练数据合理使用的典型判例，AI 初创公司 Anthropic 使用海量书籍训练大模型 Claude，原告以其未经许可使用盗版书籍构建训练数据集为由，指控其构成版权侵权。

法院针对 Anthropic 的两类行为分别作出裁定，清晰界定合理使用边界。一是通过合法渠道购买书籍用于训练的行为，因具备高度转化性被认定为合理使用。这种训练并非复制、传播或替代原作，而是让模型通过学习文本逻辑形成独立的文本生成能力，类似人类阅读书籍积累素材后进行创作，契合版权法促进创造力发展的核心目的。二是存储并意图使用 700 万册盗版书籍的

行为，被明确认定为版权侵权，不享有合理使用豁免。此外，美国通过《AI 行动计划》将 AI 数据合理使用规则优化纳入政策议程，在国防军事领域推动 AI 与数据合规深度结合，形成司法个案指引与政策相协同，实现创新与规制的平衡。

（三）欧盟坚持先授权后训练，以立法主导构建合规体系

欧盟坚持先授权后训练刚性原则，仅为科研用途的合理使用预留有限例外，2025 年 11 月德国慕尼黑法院审理的德国音乐版权协会（GEMA）诉 OpenAI 案是训练数据使用边界的典型判例。该案核心围绕模型训练数据的授权义务展开，原告 GEMA 指控 OpenAI 在训练大模型时，未经授权使用受版权保护的歌词作品，模型生成内容时存在记忆并复刻歌词片段的行为，构成版权复制侵权。

该案中法院支持了原告诉求，认定商业性 AI 训练属于对版权作品的实质性使用，必须事先获得权利人授权，即便模型未完整复制原作，仅提取片段信息也需履行授权程序；同时明确，欧盟《数字单一市场版权指令》《人工智能法案》构建的全链条规制体系，优先于个案裁量。欧盟凭借成熟市场规模，使该裁判逻辑辐射全球，全球 AI 企业进入欧盟市场均需遵循先授权后训练规则。其背后的深层原因，可能是欧盟在全球人工智能竞争中相对落后，目前尚未拥有能与中美头部企业抗衡的成熟大模型。

四、启示与建议

（一）借鉴欧美司法实践经验，建立健全 AI 数据合理使用规则体系

借鉴欧美实践经验，进一步明确使用目的、数据性质等核心考量维度，增设数据合理使用场景例外情形，划定商业与公益使用边界。统筹权利人权益与使用者利益平衡，明确侵权认定与免责要件，完善数据合理使用的司法判定与监管规范。

（二）完善数据产权保护体系，激活数据要素价值

一是明确多元数据的产权归属规则。个人信息数据强化知情同意原则并明确企业有限使用权，企业数据以实质性投入确认整理者权益，公共数据明确政企使用边界并鼓励合规二次利用。二是推进数据确权登记制度落地。深入推广数据知识产权登记试点，明确登记证的法律效力、适用范围及查询流程。三是构建数据产权流转保障机制。完善交易规则与风控体系，明确数据流转中的权利义务转移标准；引入区块链技术实现数据流转全流程追溯，保障产权人权益，降低数据滥用风险。

（三）强化协同共治，构建产业良性生态

一是构建全流程合规管控体系。鼓励、指导企业建立数据来源审查、版权核验机制，留存投入凭证与合规记录。二是强化数据合规能力建设。鼓励行业协会针对核心版权数据、普通经营数据、公共数据数据制定差异化的合理使用规范，加强数据使用合

规行业自律。指导企业部署数据脱敏、侵权识别、区块链追溯等措施，适配 AI 全链条合规需求。三是适配跨域经营合规要求。跟踪美欧数据合理使用制度态势，针对不同法域规制特点，联合发布跨域合规指引，指导企业动态调整数据使用策略。

本文作者：赛迪研究院 张驰 刘新颜 王磊 刘欣亮 马涛
联系方式：16601115928
电子邮件：zhangch@ccidthinktank.com

赛迪智库

面向政府 · 服务决策

奋力建设国家高端智库

思想型智库 国家级平台 全科型团队
创新型机制 国际化品牌

《赛迪专报》《赛迪要报》《赛迪深度研究》《美国产业动态》《赛迪前瞻》
《赛迪译丛》《国际智库热点追踪周报》《工信舆情周报》《国际智库报告》
《新型工业化研究》《工业经济研究》《产业政策与法规研究》《工业和信息化研究》
《先进制造业研究》《科技与标准研究》《工信知识产权研究》《全球双碳动态分析》
《中小企业研究》《安全产业研究》《材料工业研究》《消费品工业研究》《电子信息研究》
《集成电路研究》《信息化与软件产业研究》《网络安全研究》《未来产业研究》

**思想，还是思想，才使我们与众不同
研究，还是研究，才使我们见微知著**

新型工业化研究所（工业和信息化部新型工业化研究中心）
政策法规研究所（工业和信息化法律服务中心）
规划研究所
产业政策研究所（先进制造业研究中心）
科技与标准研究所
知识产权研究所
工业经济研究所（工业和信息化经济运行研究中心）
中小企业研究所
节能与环保研究所（工业和信息化碳达峰碳中和研究中心）
安全产业研究所
材料工业研究所
消费品工业研究所
军民融合研究所
电子信息研究所
集成电路研究所
信息化与软件产业研究所
网络安全研究所
无线电管理研究所（未来产业研究中心）
世界工业研究所（国际合作研究中心）

通讯地址：北京市海淀区万寿路27号院8号楼1201 邮政编码：100846
联系人：王乐 联系电话：010-68200552 13701083941
传 真：010-68209616 电子邮件：wangle@ccidgroup.com