

开源大模型使用端的知识产权风险及合规对策

韩 硕¹, 丁天曲²

(1. 湘潭大学知识产权学院, 湖南湘潭 411105; 2. 湘潭大学公共管理学院, 湖南湘潭 411105)

摘 要: 开源协议是开源大模型生态治理中的关键要素, 对其认识不足、履行不当既可能引发知识产权侵权风险, 也可能陷入投入产出失衡的“合规陷阱”, 亟需合规指引。梳理世界范围大模型领域的典型开源协议, 立足人工智能价值链理论识别因协议违约可能导致的知识产权风险, 运用“法律-技术-情报”交叉融合范式分析解决对策。数据预训练阶段的权利信息清理不足、开发使用阶段的协议条款及其差异认知不清、输出阶段衍生内容的控制缺位是风险主要致因。企业应建立“技术过滤+权责明示+数据合规”的全周期权利信息清理机制; 注重宽松开源协议、限制性宽松开源协议与感染性开源协议的类型学特征和风险特点, 分别设置合规管理体系; 从数据处理、模型训练、生成控制与用户监督4个维度对开源代码的输出进行合规管理。

关键词: 生成式人工智能; 开源生态; 开源协议; 风险治理; 企业合规管理

中图分类号: D922.29; G203

文献标识码: A

2025年1月, 杭州深度求索公司发布 DeepSeek-R1 开源大模型, 重构全球人工智能产业“闭源主导”格局^[1], 推动智能技术的普惠化进程, 为新型产业生态演化提供了系统性支撑^[2], 同时, 开源生态的勃兴虽预示我国智能产业正迈向认知范式重构新阶段, 但蕴含的系统性知识产权风险也需审慎评估。开源协议是开源知识产权生态治理中的关键要素^[3], 本质是权利人以其真实意思表示塑造的权利让渡契约, 协议条款对于使用、复制、修改和分发“开源资源”的社会公众具有法律约束力。截至2022年, 经 OSI 认证的开源协议已有 80 多种, 不同开源协议的声明义务、公开要求、保留机制、禁止行为、再分发条件等授权条款不尽相同^[4]。有调研指出, 条款表述的复杂性和不同协议之间的细微差异, 使得很多开发人员对开源协议授权标准的理解不够深入, 也未能充分认识到开源协议不履行、履行不充分或错误组合不兼容许可证的开源代码可能发生的法律后果, 开源协议的选择与开源资源的利用存在合规困境^[5]。加之人工智能对数据、算法、算力复用的多样化方式, 使得不同开源协议的潜在法律风险难以识别, 大模型领域因开源协议履行不当引发的争议屡见不鲜。DeepSeek、Llama 等开源大模型

的成功经验表明, 人工智能的技术创新与产业发展不可避免地要依赖于“开源资源库”。对开源协议的认识不足, 一方面可能因履行授权义务不当、对开源协议的责任传导效应认识不清, 引发知识产权侵权风险。另一方面, 过度履行协议义务可能引发合规成本超限, 形成投入产出失衡的“合规陷阱”, 为我国人工智能产业生态留下隐患。据此, 本文立足开源大模型使用端, 梳理 AI 领域主流开源协议条款, 进行类型化解读, 阐明其类型学特征与约束边界, 多维度解构开源协议履行不足可能带来的知识产权风险, 提出合规对策。

1 文献综述

1.1 研究现状

开源原指开源软件, 现在还包括开源硬件、开放数据、开放算法、开放标准、开放内容以及开放组织等^[6]。近年来, 国内开源生态起步, 但对规则运用能力较弱, 开源生态建设处于探索状态^[7]。开源生态效能方面, 大模型开源能够统领产业进步节奏、带动终端场景以及云端服务协同发展^[8]。避免重复开发的社会成本问题^[9], 帮助发现并修复漏洞与风

险^[10],是推动 LLM 研究和伦理部署最稳健的路径^[11]。开源生态建设方面,加强开源体系顶层设计^[12],开展政府协同^[13],通过财政补贴和税收减免加大支持力度^[14],创建国际化的开源协作平台、制定共享和使用规范、建立持续的创新和更新机制^[15]。开源生态风险方面,利用开源模型从事非法活动^[16],可能引发能源过度损耗、不正当竞争^[17]。违反开源协议产生版权风险^[18]。开源生态治理方面,是否开源应基于严格的风险分类与评估,高能力模型在初期可能过于危险^[19]。以敏捷治理替代“命令—控制型”监管,实施政府社区合作治理机制^[20]。“增量风险”机制更符合开源的技术应用特征^[21]。

1.2 研究评价

现有研究从开源效能、生态建设、风险识别、风险治理四个维度展开,系统阐明了人工智能开源生态创新效能和治理面向,证明本文研究具有理论意义与实践意义。局限在于:第一,多聚焦国家维度的宏观制度建构,对微观层面的企业风险识别与产业合规策略缺乏系统回应;第二,偏重对大模型开源的技术脆弱性与应用不可控性可能引致的风险进行分析,忽视了开源协议这一兼具技术规范与法律效力的开源生态关键要素,对不同开源协议的认知偏差或履行不当所衍生的知识产权风险及合规对策研究阙如。

2 开源许可协议类型化解读

广义上的开源许可协议被分为宽松型许可、感染性许可以及公共领域许可三类,前两类广泛适用于人工智能领域的开源实践。由于一些开源社区在宽松许可协议基础上增加额外限制,因此,本文将宽松许可协议进一步划分为“通用型”与“限制型”。

2.1 “通用型”宽松开源协议

MIT 开源协议唯一的义务要求是在修改后的代码或者发行包中声明原作者的许可信息。BSD 开源协议要求源代码的重新分发必须保留版权声明、条件列表、免责声明。BSD 3 - Clause License 增加“商标禁用”条款。Apache 2.0 开源协议共包括 7 个条款,主要义务包括:一是声明机制,二次授权软件或软件衍生品,必须向接受方提供协议副本;修改应进行标识并保留声明以及 Notice 文件。二是禁诉条款,用户提起诉讼所获授权将立即终止。三是责任限制,增加贡献者的免责声明,但允许贡献者自愿提供仅由其自身承担的担保责任。Mulan PSL 是中国首个开源通用协议,2020 年 2 月 14 日由 OSI 董事会批准通过。除未明确允许“贡献者”提供仅由其自身承担的担保责任外,其他内容与 Apache2.0 类似。

表 1 “通用型”宽松开源协议对比
Table 1 Comparison of “general - purpose” permissive open source licenses

许可协议	再分发要求	专利条款	商标条款	贡献者担保	代表性模型
MIT	版权声明、免责声明	无	未禁止	未明确	CLIP、Whisper、Wav2Vec 2.0、Coqui TTS、DeepSeek - R1
BSD	版权声明、免责声明	无	禁止未授权使用	未明确	BLIP - 2
Apache 2.0	许可证副本、标注修改、版权声明、Notice 文件	专利授权禁诉条款	禁止未授权使用	允许贡献者自愿承担	GPT - NeoX - 20B、GPT - J - 6B、BERT、T5、Falcon - 40B、SAM、VoxCeleb Encoder、RNNT、ChatGLM2 - 6B、ERNIE - ViLG
木兰 PSLv2	同 Apache 2.0	专利授权禁诉条款	禁止未授权使用	未明确	MindSpore Transformers 工具

2.2 感染性开源协议

感染性 (copyleft) 协议要求任何“源代码”的修改或再创作必须在同样协议下发布。GPL 协议是最具代表性的感染性许可协议,主体条款包括:第一条,GPL 协议的适用对象,只要包含了 GPL 授权代码的“全部或部分”,整个软件都必须受到 GPL 的约束。第二条,使用者的基本权限,使用者可以无限次

运行未修改的程序,进行衍生品的制作、运行和传播 (propagate)。但转让 (convey) 只能在“雇佣修改或者为运行作品提供设施”目的下进行,传播和转让的区别在于是否向第三方提供软件程序的副本。第三条,禁止版权人或再分发者利用技术保护措施 (TPMs) 和反规避法律。第四条,转让软件副本的条件,包括“明显并适当标示版权信息”“保留许可证

声明、免责声明”“向接收者提供许可证副本”。第五条、第六条共同组成传播行为的核心义务体系。第五条包括三款:第一款要求修改后的源代码传播须标注修改信息、完整保留许可证声明与免责条款、确保 GPL 协议适用于作品及其所有组合形态;第二款是交互界面例外规则,即当程序用户界面不具备展示功能时,可豁免前述义务;第三款为独立作品排除条款,GPL 协议效力仅及于被整合作品本身,不延伸至集合中具有独立功能的组成部分。第六条,非源代码的传播条件,第一款确立“源代码伴随义务”的四种方式,直接交付、承诺交付、网络对等访问及 P2P 传播告知机制;第二款引入系统库豁免条款,将操作系统运行必需的标准组件排除在源代码交付范围之外;第三款创设用户产品安装信息披露义务,要求附着于有形商品的目标代码,转让方必须提供技

术修改支持信息。第八条规定,违反开源协议将导致授权自动终止。第九条,豁免框架。对单纯运行程序、P2P 传输接收等非主动传播行为进行义务豁免。第十条针对企业并购等结构化交易场景,确立权利义务概括承受规则。

GPL 3.0 的强传染性仅适用于通过传统方式分发软件副本的场景。AGPL v3.0 在其基础上增设了网络服务条款(第十三条),规定若通过远程交互(如 SaaS 模式)向用户提供程序功能,即使未主动分发软件副本,仍需公开服务端源代码。LGPL v3.0 针对代码库调用场景进行了特殊设计,开发者以动态链接方式将 LGPL 授权的代码与私有代码结合使用,且无需公开私有部分的源代码,但若直接修改 LGPL 授权代码并静态链接分发,仍须履行开源义务。对比见表 2。

表 2 强传染性开源协议对比
Table 2 Comparison of strong copyleft open source licenses

协议类型	GNU GPL v3.0	GNU AGPL v3.0	GNU LGPL v3.0
许可证类型	强 Copyleft(强传染性)	网络服务专用强 Copyleft	弱 Copyleft(弱传染性)
主要适用场景	软件、操作系统、工具等	在线服务、网络应用	软件库、中间件等
修改后要求	必须在同一许可证下再许可	必须在同一许可证下再许可	修改后再发布时需同样许可证
源代码披露要求	修改、传播需提供完整源代码	网络交互需披露源代码	库本身适用披露要求
目标代码传播	必须附带或承诺提供源代码	同上	同上
系统库要求	可豁免包含在源代码中	同上	同上
用户产品	提供可安装、修改的说明	同上	同上
典型条款编号	0~6 条:范围、权限、传播		
	第 11 条:专利条款 第 13 条:与 AGPL 兼容条款	网络服务下源代码披露义务	弱传染性适用范围说明
是否允许商业使用	允许	允许(需开放源代码)	允许(非库部分可闭源)

2.3 “限制型”宽松开源协议

限制型宽松开源协议是以 Apache 2.0 为基础,增加其他限制条款完成的。非商用型协议禁止任何商业用途、军事用途或其他敏感用途,代表性的开源协议包括 Meta 的 LLaMA(1.0 版本)、清华的 GLM-130、Stable AI 的 DeepFloyd IF 等。商用型许可,在允许商业利用的同时,增加行为规范和用途限制条款。代表许可证包括 BigScience Open RAIL、CreativeML Open RAIL、Llama2、BAAI Aquila 等。

限制型宽松开源协议中较为典型的条款包括:Meta Llama2 的第二条是一个单独的行为限制条款,

它规定若提供的产品或服务在前一个日历月中的月活跃用户数超过 7 亿,必须经过 Meta 许可。Open RAIL-M 系列开源协议由 RAIL 制定,该协议禁止将开源模型用于一系列不当用途,如违法行为、仇恨言论、侵害他人权利等。以 BigScience RAIL License 1.0 为例,该协议在附件 A 中列出 10 项禁止内容,如禁止参与、促进、生成非法或不法活动或内容;参与、促进、煽动或促成对个人或群体的骚扰、虐待、威胁或欺凌;在就业、福利、住房等方面促成歧视或其他非法或有害行为等。与之类似的是 CreativeML Open 系列协议,它们都在附件中列出了禁止性义务。

表 3 “限制型”宽松开源协议对比
Table 3 Comparison of “restrictive” permissive open source licenses

许可协议名称	是否允许商业使用	行为限制	应用模型
Meta LLaMA 研究许可	不允许	仅限学术研究用途,不得用于商业化	LLaMA (v1)
GLM - 130B 特别许可	不允许	禁止商业、军用、非法用途	GLM - 130B
DeepFloyd IF	不允许	与 GLM - 130B 类似	DeepFloyd IF
Meta Llama 2 社区许可	有条件允许	附件明确行为限制条款	Llama 2 等
BigScience Open RAIL	允许	附件明确行为限制条款	BLOOM 176B
Creative ML Open RAIL	允许	附件明确行为限制条款	StableDiffusion

3 开源大模型使用端的知识产权风险

本文对开源大模型使用端的知识产权风险识别以“人工智能价值链理论”(AI value chain)为基础。“价值链”的概念最早由经济学家创造,用于阐明产品从生产至价值实现的“链式”路径^[22],后在技术治理领域得到广泛运用。例如,欧盟《人工智能法案》中,“价值链”概念多次出现,学界认为,它既能指代人工智能系统实现其预期功能的各环节,也强调人工智能价值链风险的整体治理,全生命周期治理、产业上下游风险协同治理、集中式治理是其关键理念^[23]。该理论对本文的可适用性在于:首先,全生命周期是对人工智能模型预训练、模型校准、模型微调以及模型的发布与输出等“功能实现”各环节的静态描述,从各环节入手,可以系统、全面审视风险。其次,产业上下游风险协同治理理念强调人工智能的系统风险蕴含了各主体间复杂行为及其累积效应^[24]。因而除自身行为外,还需关注、管理上下游主体的行为风险。最后,集中式治理理念阐明运用“价值链”理论的必要性。囿于社会治理成本,法律通常会确立特殊的责任承担方式。一类是,明确将责任承担归结于某一特定主体,例如《药品管理法》中以“药品上市许可持有人”为核心的 MAH 责任制度。另一类是所有主体之间的连带责任,例如《中华人民共和国民法典》中的产品责任制度。显然,集中式法律治理理念要求开源大模型使用者做最全面的风险识别与管理,以避免第三方行为带来的法律责任。

3.1 预训练阶段的知识产权风险:权利信息难以清理

开源生态发展初期,开源社区通过信息清理、技术筛查以及制度性设计有效降低“开源资源”中的

知识产权瑕疵可能引发的第三方权利冲突风险。人工智能技术的范式变革重构了开源生态的风险图谱。开源大模型通常利用开源数据库、公共互联网抓取的海量数据进行预训练,而后成为下游“服务大模型”的底层基础模型,例如,Stable Diffusion 便采用了预训练多模态模型 CLIP。不仅如此,Hugging Face 等开源平台还允许上传用户训练模型,供全球用户下载。对 Hugging Face 上的 7903 个项目进行实证研究表明,开源大模型的训练数据通常存在未经授权问题^[25]。训练数据的规模性、用户行为的隐蔽性与不可控性加之开源大模型的黑箱属性,使得传统权利清理机制在人工智能场景下面临困境。斯坦福大学人工智能研究所 2024 年研究数据显示,90% 的开源模型存在多层级数据递归引用现象。即便采取技术措施,大模型依然会输出版权侵权内容。法律责任角度,各国法律实践为支持开源大模型生态建设,普遍为开发者或发布平台设置了预训练阶段的豁免规则。美国法院在 Kadrey v. Meta Platforms 一案中,认定被告未经授权进行数据训练行为构成合理使用。欧盟《人工智能法案》(EUAIA)和《产品责任指令》为非高风险的开源模型设置了豁免规则。我国学界通说认为,应将预训练阶段的数据使用行为纳入合理使用制度,而将“输出端”的数据传播侵权聚焦于开源模型使用者(服务提供者)^[26]。模型参数与训练数据的非线性映射关系使传统清理技术可行性锐减、海量数据聚合产生“畸高”权利主体识别成本,加之各国对开源大模型开发者的豁免倾向和开源协议普遍设置的绝对免责条款,使得开源大模型预训练阶段知识产权侵权风险的管理责任由开发端向使用端转移。

3.2 开发使用阶段的知识产权风险:协议条款认识不足

开源协议种类繁多,对不同开源协议的复杂法

律含义及其差异感到“困惑”是技术开发者选择开源协议时的常见困难^[27]。这种认知不足在现实中也引发诸多纠纷,风险主要来源于以下方面:

一是宽松开源协议之间的义务差异。不同宽松开源协议的条款之间存在细微差异。若未能识别这些差异,可能因未全面履行义务构成违反开源协议。例如,Mit 与 Apache 2.0 相比,保留声明义务的履行范围不同。Apache 2.0 和 MulanPSL 2.0 相比,授权主体范围不同,对于开源软件的“修改”声明的范围亦不同。BSD 包含多个版本,3- Clause BSD 许可证较为特殊,它不包含专利许可条款,允许接收者修改原始作品的代码。Apache v1.1 在 BSD 3 的基础上又附加了其他条款^[28]。

二是限制型宽松开源协议之间的禁止性条款差异。不同限制型宽松开源协议的禁止性条款存在多方面差异。例如,Llama 2 允许 MAU <7 亿用户量级情况下的商业应用,而 GLM-130B 禁止任何形式的商业使用。又如,仅 BigScience Open RAIL-M 协议要求使用者对生成内容进行标注。限制性宽松协议还存在用户行为控制风险,表现为技术提供者对终端行为链的监管失效。即使开发主体已履行协议认知义务,用户通过 API 调用或微调实现的非预期应用仍可能导致上游模型部署者实质性违约。例如,Stable Diffusion 默认附带内容安全过滤器,但已有用户移除过滤器并尝试将 Stable Diffusion 用于生成色情或 Deepfake 内容^[29]。

三是感染性开源协议在大模型领域的触发边界模糊。开源促进基金会的报告显示,开源大模型的开发存在对“GPL 技术资源”广泛的隐性调用。衍生作品属于 GPL 协议的适用对象、使用行为属于知识产权的控制范围、被使用的知识产品满足开源定义是触发 GPL 传染性的前提条件。但源于传统开源软件的 GPL 协议与人工智能技术特性间存在系统性割裂,导致三个标准的判断并不清晰^[30]。开源模型使用者一方面因行为性质模糊可能承担违约风险,另一方面因缺乏确定性指引不得不采取过度保守的合规策略,开发效率受损与合规成本激增^[31]。第一,GPL 协议中的“运行”“修改”“传播”“转让”“衍生作品”等核心概念的设置以传统软件工程范式为底层逻辑,难以适配 AI 开发特性。例如,利用 GPL 代码进行数据训练,既因运用“数据挖掘”技术构成“运行”行为,也因程序代码的解构与重组,构成“修改”和“再创作”。第二,人工智能大

模型与传统软件的技术路径差异使得协议义务的“涵摄”关系断裂。衍生作品必须包含受保护作品表达的实质性部分,而机器学习将开源代码“解构—重组—生成”的技术闭环使得开源模型使用者难以监测模型参数与原始代码之间的实质性关联。第三,大模型领域的“开源”概念缺乏清晰界定。当前,世界范围内对于何为人工智能“开源”并未达成统一认知。实践中,开源大模型的开放程度呈现光谱式分布,不同大模型对文档、方法论、权重系数、模型架构或使用信息均存在不同程度、不同类型的公开。从法律、政策角度看,欧盟《人工智能法案》、开源促进会(OSI)的《开源 AI 定义草案 v0.0.6》对“开源”定义亦不相同。有专家认为,仅通过开源协议发布模型而不开放其他组件,不应被视为开源。综上,开源模型使用者需要采取恰当的管理措施消除 GPL 协议触发条件不清潜在的知识产权风险。

3.3 模型输出阶段的知识产权风险:管理不善引致开源代码传播

开源大模型通常需要使用受开源协议约束的代码进行训练。典型例证是 GitHub 的 Copilot 编程辅助工具,该工具通过数十亿字节公开源代码微调训练而成^[32]。大模型对训练数据的记忆特性会导致训练代码实质相似内容传播。使用开源代码进行模型预训练实质上是对开源代码的重用、修改和再分发,必须遵守开源许可证中的授权条件。实证研究表明,许可证信息在前述过程中往往是被忽略的,例如,开源开发者指控 Copilot 在输出代码的同时未附署名,也未包含 LGPL 许可证。OpenAI、GitHub 和 Microsoft 因此面临诉讼。一方面,与预训练阶段强调权利信息的清理与尽可能避免实质性相似内容输出的风险管理对策不同,大模型在代码生成场景下的风险应对,强调对开源代码修改、使用、再分发的合规、高效。另一方面,值得警惕是如果进行模型训练或软件开发的目的是实现专有许可模式收费的商业模式,GPL 等“感染性”开源协议所带来的污染效应是一大难题。

4 开源大模型使用端知识产权风险的合规对策

开源大模型使用端的知识产权风险管理,本质是要构建覆盖开源大模型价值链条的开源协议合规评估体系,系统应对预训练阶段、开发使用阶段与输

出控制阶段等 3 个维度的风险。

4.1 预训练阶段:强化数据合规全流程管理

AI 训练数据规模大、权利主体分散、权益结构复杂且呈现非结构化特征,依托开源社区的传统审查模式难以实现有效溯源,开源大模型使用者必须建构自主性权利治理框架。一是数据源头控制。优先选用来源可追溯、产权明确或经认证的标准化数据集进行预训练的开源大模型。部署前,剔除存在“在先权益”的重复数据,降低下游版权侵权风险。二是技术管理维度。一方面,综合运用水印技术、内容相似度检测、差分隐私和合成数据生成技术,对疑似侵权素材进行标记、隔离与实时过滤,消除原始数据中的版权表达残留。另一方面,利用人类反馈强化学习技术(RLHF)对开源大模型输出侵权内容予以否定性反馈。三是建立内部合规体系,开源模型使用者与服务模型部署者具有主体层面的同一性,从世界范围内的立法、案例及学说看,服务模型提供者的间接责任以未履行适当的注意义务为限,因此,部署严密且符合技术理性的合规体系,是开源模型使用者避免承担知识产权侵权责任的必要举措,应

在组织层面明确法律、技术、业务团队的职责分工,实现“技术过滤+权责明示+数据合规”的全流程风控。由于举证责任倒置,开源模型使用者还应当保存日志,为权利人提供核查工具。

4.2 开发使用阶段:协议认知与用户监测

开发使用阶段,开源模型使用者需对各类开源协议深入了解、正确认知,在此基础上,利用 Choose-ALicense、Licence Differentiator 等开源协议辅助选择工具,作出与自身商业偏好契合的正确选择。一是宽松开源协议的条款差异阐释。MIT 协议的声明保留条款最为宽松,它唯一的要求是在修改后的代码或者发行包中声明原作者的许可信息。BSD 协议的声明保留机制与 MIT 协议相同,但禁止未经许可使用开源主体、贡献者的商标或名义进行宣传。Apache 2.0 协议则要求“每一份”(不仅是整体)被修改的文件都必须保留软件的版权、专利、商标声明以及软件中所包含的 Notice 文件。Mulan PSL 开源许可协议则与 Apache2.0 类似。Apache2.0 和 Mulan PSL 包含了专利禁诉条款,“反咬一口”会引发“授权终止”。

表 4 “通用型”宽松开源协议差异概览

Table 4 Overview of differences in “general – purpose” permissive open source licenses

名称	开发者	保留声明	商标禁用	禁诉条款
MIT	麻省理工	整体履行	否	否
BSD	加州大学	整体履行	是	否
Apache 2.0	Apache 基金会	所有修改文件	是	是
Mulan PSL	中国开源云联盟	所有修改文件	是	是

二是限制型宽松开源协议的合规对策。应建立跨协议文本的对比审查体系,一方面,通过语义要件解构与场景适应性分析,精准识别异质性条款的规制边界,如 MAU 计量标准中的“活跃用户”等概念。另一方面,识别不同限制性宽松协议在禁止性条款差异,具体见下表:

强化用户行为监测部署,应通过核心义务矩阵将抽象条款转化为可操作的技术控制要件,例如在模型推理层部署协议合规引擎,利用实时意图分析算法阻断禁止性场景的指令响应;同时构建动态监测系统,整合用户行为日志的区块链存证与敏感内容生成审计模块,形成风险“识别—阻断—溯源”的闭环治理。

三是感染性开源协议的合规对策。感染性开源

协议的合规对策,关键在于清晰界定开源大模型领域“传染性”条款的触发条件。一般认为,GPL 协议的触发条件有三:第一,被使用的资源本身满足“开源”的定义。第二,“被传染者”属于 GPL 协议的适用对象。第三,开源大模型使用者的行为属于“开源资源权利人”的专有权范围。从“开源”定义看,开源促进会(OSI)的《开源 AI 定义草案 v0.0.6》较为详细,可为参照。人工智能系统开源包括:(1)数据:训练方法与技术、训练数据集的来源、范围与特征、数据的获取与筛选方式、数据标注程序与清洗方法、数据集的溯源信息;(2)代码:需要公开支持系统训练、部署与运行的完整代码。用于数据预处理的代码、用于训练、验证与测试的代码、支持性库、推理代码以及模型架构说明;(3)模型:需公开模型参

数,包括权重参数、训练中间阶段的检查点、最终优化器状态。从适用对象看,GPL 的适用对象是“受保护作品”“未修改的原程序”或“基于该程序的作品”。GPL 许可协议规定:“若某作品的可识别部分并未源自该程序,且可以合理地视为独立作品,那么在单独分发时,本许可证及其条款不适用于这些部分。”因此,企业需要清晰识别各种组件之间在“表达”层面的相互影响。例如,若被引用的开源组件

仅承担微小功能而未构成整个软件的主要部分,则无需将无关组件开源。从是否落入专有权范围角度看,感染性开源协议“触发”的前提是未经许可的使用、修改或再分发行为构成侵权。应建立完善版权法律服务队伍,具体判断:是否使用了开源技术的“表达”;修改或分发的“衍生品”与源代码是否存在实质性相似;是否符合权利限制规则的豁免框架。

表 5 限制性宽松开源协议禁止性条款差异概览

禁止性条款	LLaMA2	GLM - 130B	DeepFloyd IF	BigScience	CreativeML
非商业用途	×	✓	✓	✓	✓
军事用途	✓	✓	✓	×	×
非法行为	✓	✓	✓	✓	✓
危害公共利益	×	✓	×	×	×
伤害未成年人	✓	×	×	✓	✓
骚扰、霸凌、诽谤	✓	×	×	✓	✓
仇恨和歧视	✓	×	×	✓	✓
专业建议	✓	×	×	✓	✓
隐私与数据滥用	✓	×	✓	✓	✓
侵犯他人权利	✓	×	✓	×	×
恶意代码/病毒	✓	×	×	×	×
高风险应用	✓	×	✓	×	×
欺诈、虚假内容	✓	×	×	✓	✓
深度仿冒	✓	×	×	✓	×
未标明 AI 生成	×	×	×	✓	×
自动决策	×	×	×	✓	✓
司法执法用途	×	×	×	✓	✓
披露已知风险	✓	×	×	×	×
规避安全措施	×	×	✓	×	×
出口管制	×	×	✓	×	×

4.3 输出管理阶段:开源代码使用合规的全流程防控

鉴于开源大模型输出管理阶段风险应对需求的特殊性,企业需从数据预处理、模型训练、生成控制及用户监管四个维度构建风险管理框架。探索将合规义务融入大模型算法的自动化治理范式。数据预处理环节,通过强化数据清洗与许可协议识别策略,构建代码与开源许可证元数据关联架构,利用大模型的自回归特性将许可声明作为代码片段的后续标记进行联合训练,增强模型对代码与协议关联性的

语义理解。模型训练阶段,融入差分隐私训练机制降低代码记忆效应,或构建协议敏感性上下文学习模块,实现高风险协议内容的动态规避。生成控制环节,设置特殊的输出过滤范式,避免对开源协议保护内容的简单抑制,动态注入版权声明、原始许可证标识及衍生作品修改记录,同步满足 MIT 协议的署名要求与 GPL/LGPL 协议的传染性约束条件;模型部署前需开展系统性合规能力评估,重点验证其协议约束代码的识别与处置效能。用户监管层面,优化用户协议,禁止使用者实施未经授权的使用、修改

及传播行为;建立对抗性提示工程防御机制,阻断用户通过诱导式输入获取违反开源协议内容。

5 结论

本文以开源技术使用端的知识产权风险为视角,类型化解读人工智能领域典型开源协议,识别不同类型开源协议下的风险点和风险致因,给出合规对策。首先,根据开源协议“传染力”将其分为通用型宽松开源协议、限制型宽松开源协议和感染性协议三类,总结了各类协议的核心义务差异。其次,基于人工智能价值链理论全面审视风险致因,包括数据权利清理困难、协议理解与履行偏差、生成内容的控制错位等。本文建议“开源智能技术”使用企业设置贯穿数据采集及预处理、模型开发、模型训练、模型部署和服务应用全周期的合规监管对策,为开源大模型使用企业提供了合规指引。

参考文献:

- [1] 郭亚军,徐苑茜,梁艳丽,等. 从 ChatGPT 到 DeepSeek:生成式人工智能迭代对图书馆的影响[J]. 图书馆论坛,2025,45(7):140-149. Guo Y J, Xu Y X, Liang Y L, et al. From ChatGPT to DeepSeek: The impact of generative artificial intelligence iterations on libraries[J]. Library Tribune, 2025, 45(7): 140-149.
- [2] 王珏. 基于 DeepSeek 的未来产业生态构建[J]. 西北大学学报(哲学社会科学版), 2025, 55(3): 71-80. Wang J. Ecological architecture of future industries: Based on DeepSeek[J]. Journal of Northwest University (Philosophy and Social Sciences Edition), 2025, 55(3): 71-80.
- [3] 喻玲,邵滨. 开源社区知识产权治理模式及变革——基于 36 个开源社区使用协议的考察[J]. 科学学研究, 2024, 42(9): 1938-1945. Yu L, Shao B. Intellectual property governance model of open-source community and its reform: Based on 36 open-source community use protocols[J]. Studies in Science of Science, 2024, 42(9): 1938-1945.
- [4] Lin Y H, Ko T M, Chuang T R, et al. Open source licenses and the creative commons framework: License selection and comparison[J]. Journal of Information Science and Engineering, 2006, 22: 1-17.
- [5] Kapitsaki G M, Kramer F. Open Source License Violation Check for SPDX Files[C]. Proceedings of the 14th International Conference on Software Reuse. Cham: Springer Verlag, 2015: 90-105.
- [6] 孟小峰,黄匡时,梁玉成,等. 社会计算与社会智能: 机遇与挑战[J]. 中国科学基金, 2023, 37(3): 439-450. Meng X F, Huang K S, Liang Y C, et al. Social computing and social intelligence: Opportunities and challenges[J]. Bulletin of National Natural Science Foundation of China, 2023, 37(3): 439-450.
- [7] 王黎莹,赵春苗,王举铎,等. 知识产权与标准协同推进人工智能产业创新机制与路径优化[J]. 科学学与科学技术管理, 2024, 45(4): 52-67. Wang L Y, Zhao C M, Wang J D, et al. Research on the mechanism and path optimization of intellectual property rights and standards to promote artificial intelligence industry innovation[J]. Science of Science and Management of S. & T., 2024, 45(4): 52-67.
- [8] 王冲鹄,陈丝. 人工智能开源平台发展态势研究[J]. 信息通信技术与政策, 2018(8): 56-61. Wang C J, Chen S. Research on the development trend of open source platform in artificial intelligence[J]. Information and Communications Technology and Policy, 2018(8): 56-61.
- [9] Muzammil K. The importance of open source development in technology[EB/OL]. <https://www.aalpha.net/blog/importance-of-opensource-development-in-technology/>. 2024-08-28/2025-04-10.
- [10] Ferrandis C M, Lizarralde M D. Opening sourcing AI: Intellectual property at the service of platform leadership[J]. Journal of Intellectual Property, Information Technology and Electronic Commerce Law, 2022, 13(3): 224-246.
- [11] Manchanda J, Boettcher L, Westphalen M, et al. The Open Source Advantage in Large Language Models (LLMs) [EB/OL]. <https://arxiv.org/abs/2412.12004>. 2025-02-02/2025-04-10.
- [12] 张超,冯泽,陈凯华,等. 关于健全数字公共产品制度推进开源创新生态建设的若干思考[J]. 中国科学院刊, 2024, 39(11): 1944-1955. Zhang C, Feng Z, Chen K H, et al. Reflections on strengthening digital public goods institutions to promote construction of open source innovation ecosystem[J]. Bulletin of Chinese Academy of Sciences, 2024, 39(11): 1944-1955.
- [13] 谢新水. 智能跃迁、开源创新与主权 AI: DeepSeek 现象推动人工智能开源创新生态体系建设[J]. 电子政务, 2025(3): 40-48. Xie Y S. Intelligent leap, open source innovation, and sovereign AI: how the DeepSeek phenomenon propels the building of an open source innovation ecosystem for AI[J]. E-Government, 2025(3):

- 40-48.
- [14] 魏巍,曾铮,刘蕾.从 DeepSeek 突破看我国人工智能产业创新范式、挑战与应对[J/OL].经济纵横,1-13 [2025-04-23]. Wei W, Zeng Z, Liu L. Innovation paradigm, challenge, and response of China's AI industry from the breakthrough of DeepSeek[J/OL]. Economic Review Journal, 1-13 [2025-04-23].
- [15] 范从来,郑琼洁.人工智能:新质生产力的新质态及其发展战略[J].经济理论与经济管理,2024,44(11):1-14. Fan C L, Zheng Q J. AI: The new qualitative forms of new quality productive forces and its development strategy[J]. Economic Theory and Business Management, 2024, 44(11): 1-14.
- [16] Vake D, Šinik B, Vičič J, et al. Is open source the future of AI? A data-driven approach[EB/OL]. <https://arxiv.org/abs/2501.16403>. 2025-01-27/2025-04-10.
- [17] 周辉.开源人工智能模型的法律治理[J].上海交通大学学报(哲学社会科学版),2024,32(8):18-33. Zhou H. Legal governance of open source artificial intelligence model[J]. Journal of Shanghai Jiaotong University(Philosophy and Social Sciences), 2024, 32(8): 18-33.
- [18] Basanagoudar V, Srekanth A. Copyright conundrums in generative AI: Github copilot's not-so-fair use of open-source licensed code[J]. Journal of Intellectual Property Studies, 2023, 7(2): 58-68.
- [19] Seger E, et al. Open-sourcing highly capable foundation models: An evaluation of risks, benefits, and alternative methods for pursuing open-source objectives[EB/OL]. <https://arxiv.org/abs/2311.09227>. 2023-09-23/2025-04-10.
- [20] 苏宇,郭雨婷.人工智能开源生态的法律治理[J].宁夏社会科学,2024(5):119-130. Su Y, Guo Y T. Legal governance of the open source AI ecosystem[J]. Ningxia Social Sciences, 2024(5): 119-130.
- [21] 魏钰明,贾开,曾润喜,等. DeepSeek 突破效应下的人工智能创新发展与治理变革[J].电子政务,2025(3):2-39. Wei J M, Jia K, Zeng R X, et al. The innovation development and governance transformation of artificial intelligence under the breakthrough effect of DeepSeek[J]. E-Government, 2025(3): 2-39.
- [22] 迈克尔·E·波特.竞争优势[M].北京:中信出版社,2014:29-50. Porter M E. Competitive Advantage[M]. Beijing: CITIC Press, 2014: 29-50.
- [23] 李润生.人工智能价值链的法律类型[J].东方法学,2025(2):63-78. Li R S. The legal shaping of the artificial intelligence value chain[J]. Oriental Law, 2025(2): 63-78.
- [24] 谢尧雯.生成式人工智能价值链行政监管与侵权责任的匹配[J].政法论坛,2025,43(2):36-46. Xie X W. Matching administrative regulation and tort liability in the generative AI value chain[J]. Tribune of Political Science and Law, 2025, 43(2): 36-46.
- [25] Chakraborti M, et al. "Responsible AI in open ecosystems: Reconciling innovation with risk assessment and disclosure." [EB/OL]. arXiv preprint arXiv: 2409.19104 (2024). 2024-09-27/2025-04-10.
- [26] 王利明.生成式人工智能侵权的归责原则与过错认定[J].中国法律评论,2025(4):15-30. Wang L M. Principles of attribution and fault determination in generative artificial intelligence tort[J]. China Law Review, 2025(4): 15-30.
- [27] 吴欣,武健宇,周明辉,等.开源许可证的选择:挑战和影响因素[J].软件学报,2022,33(1):1-25. Wu X, Wu J Y, Zhou M H, et al. Selection of open source license: Challenges and influencing factors[J]. Journal of Software, 2022, 33(1): 1-25.
- [28] 王志强,伍胜,肖国强,等.开源许可证合规性研究[J].软件学报,2022,33(8):3035-3058. Wang Z Q, Wu S, Xiao G Q, et al. Study of open-source software license compliance[J]. Journal of Software, 2022, 33(8): 3035-3058.
- [29] Contractor D, McDuff D, Haines J K, et al. Behavioral use Licensing for Responsible AI[C]. Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, 2022: 778-788.
- [30] Benhamou Y. Open Source AI: Does the copyleft clause propagate to proprietary AI models? [J]. Revisiting the Definition of Derivatives in the AI-context, 2024(7): 1-12.
- [31] Henley M. Open source software: An introduction[J]. Computer Law & Security Review, 2008, 24(1): 77-85.
- [32] Please Note: citations are provided as a general guideline. Users should consult their preferred citation format's style manual for proper citation formatting[J]. Journal of Intellectual Property Studies, 2023: 58-68.

Intellectual property risks and compliance strategies on the user side of open – source large models

HAN Shuo¹, DING Tian – qu²

(1. Intellectual Property School, Xiangtan University, Xiangtan 411105, China;

2. School of Public Administration, Xiangtan University, Xiangtan 411105, China)

Abstract: The success of open – source large models such as DeepSeek and Llama demonstrates that technological innovation and industrial development in AI will inevitably rely on an open – source innovation ecosystem. Open – source licenses are a key element of governance in that ecosystem; in essence, they are contracts for the transfer of rights shaped by the rightsholder's genuine manifestation of intent, and their terms are legally binding on the public with respect to the use, reproduction, modification, and distribution of "open – source resources." Because open – source large models are widely used across the AI ecosystem, enterprises' insufficient understanding and improper performance may not only trigger intellectual – property (IP) infringement risks but also trap them in compliance dilemmas where investment and return fall out of balance; there is an urgent need for compliance guidance. At present, however, there is still a lack – both domestically and internationally – of systematic research on compliance strategies for open – source licenses at the user side of large models. This paper surveys representative open – source licenses in the large – model field worldwide, identifies IP risks that may arise from license breaches through the lens of the AI value chain, and analyzes countermeasures using an integrated "law – technology – information science" paradigm. In the data pre – training stage, open – source large models typically use massive datasets from open databases and the publicly accessible internet to perform pre – training, thereafter serving as the foundational base models for downstream "service – facing" large models. Because the relationship between model parameters and training data is a non-linear mapping that sharply reduces the feasibility of traditional cleansing techniques, and because the aggregation of massive data produces abnormally high costs for identifying rightsholders – combined with many jurisdictions' tendency to afford exemptions to open – source model developers and the absolute disclaimers commonly found in open – source licenses – the responsibility for managing IP infringement risks in the pre – training stage has shifted from the development side to the user side. Inadequate cleansing of rights – information in training data is the principal cause of IP infringement risk at the pre – training stage. During the development and use stage, differences exist both among different categories of open – source licenses and among clauses within the same category – for example, divergent obligations among permissive licenses, divergent prohibitions among weak – copyleft (restrictively permissive) licenses, and ambiguous trigger boundaries for strong – copyleft ("infectious") licenses in the large – model context. An unclear understanding of open – source licenses is the primary cause of IP infringement risk in this stage. At the output stage, open – source large models typically must rely on code subject to open – source licenses for training, and the model's propensity to memorize training data can lead to the dissemination of content that is substantially similar to the training code. A lack of control over derivative content is a principal cause of IP infringement risk at this stage. In the data pre – training stage, because AI training data are massive in scale, rightsholders are dispersed, entitlement structures are complex, and the information is often unstructured, traditional community – based review mechanisms struggle to achieve effective provenance tracing; users of open – source large models must therefore build an autonomous rights – governance framework. Enterprises should establish a full – lifecycle rights – information cleansing mechanism that integrates "technical filtering + explicit allocation of rights and responsibilities + data compliance." In the development and use stage, users of open – source models should develop a deep and accurate understanding of the various open – source licenses and, on that basis, use auxiliary license – selection tools – such as ChooseALicense and License Differentiator – to make choices aligned with their commercial preferences. They should also focus on the typological features and risk profiles of permissive licenses, weak – copyleft (restrictively permissive) licenses, and strong – copyleft ("infectious") licenses, and establish differentiated compliance management systems for each. In the output – management stage, in light of the particular risk – response needs associated with outputs from open – source large models, enterprises should construct a risk – management framework along four dimensions: data preprocessing, model training, generation control, and user oversight.

Key words: generative AI; open – source ecosystem; open – source licenses; risk governance; corporate compliance management