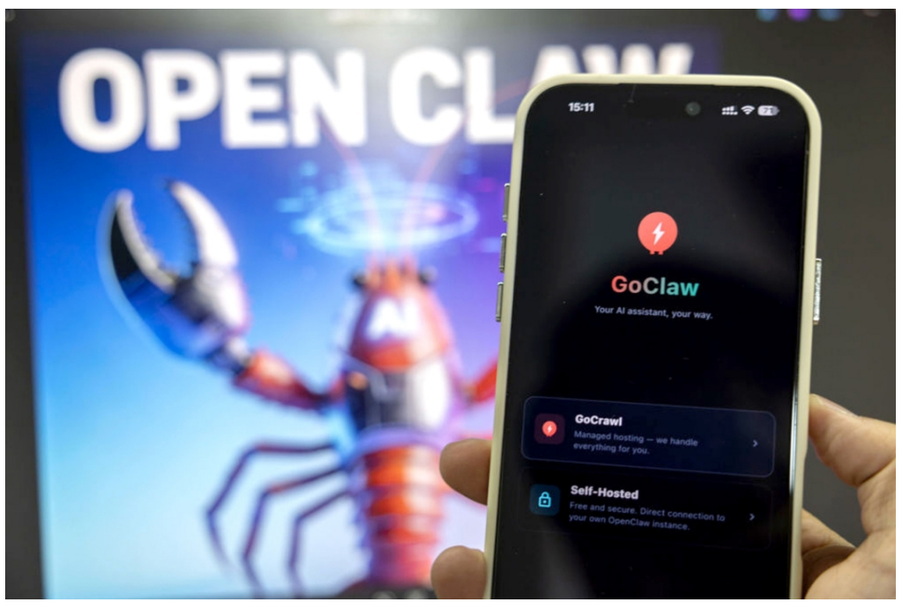
OpenClaw迅速走红暴露出一系列亟待解决的隐患

多国启动安全治理与风险防控行动

□ 本报记者 赵阳

近一段时间以来，一款名为OpenClaw的开源人工智能体在全球科技圈迅速走红。这一人工智能体可自主在现实场景中执行任务，代表用户采取行动，标志着人工智能正从“对话助手”向“行动助手”的加速跨越。但伴随技术快速普及，多重安全隐患也集中显现，包括中国在内，美国、欧盟、英国、日本、韩国等多国监管机构、网络安全专家及国际组织接连发出警示，启动行动型AI安全治理与风险防控行动。



图为开源AI智能体*OpenClaw*手机端页面。 新华社发(伊凡 摄)

风险集中暴露引发全球安全警报

“作为开源人工智能体领域的典型代表，OpenClaw凭借强大的自然语言交互与任务执行能力，在短时间内迅速吸引了广泛关注与应用。然而，其开放性与高权限特性，也使其成为网络安全领域不容忽视的焦点，暴露出一系列亟待解决的安全风险。”作为我国人工智能领域的专家学者，浙江大学光华法学院求是特聘教授熊明辉一直在关注着OpenClaw。

他在接受《法治日报》记者采访时，将OpenClaw的第三方技能包或依赖库形象地比喻成攻击者的“特洛伊木马”。“恶意行为者通过篡改软件包、植入后门，或利用已知漏洞进行渗透，一旦成功潜入企业内网，便可窃取核心商业数据、知识产权，甚至取得内部系统控制权，对企业数据安全构成系统性威胁。”

“设备与系统层面的安全隐患同样严峻。”熊明辉对此解释说，为完成复杂任务，OpenClaw在开发与运行时往往需要较高的系统权限。若配置存在疏漏，或软件自身存在未修补的漏洞，可能导致主机敏感信息（如密钥、配置）泄露。攻击者借此可劫持设备，将其纳入僵尸网络或用于发起进一步的内部攻击，严重威胁企业基础设施的稳定与安全。此外，在扮演个人助手时，OpenClaw处理通讯录、日程、聊天记录等大量敏感个人信息时，安全防护的薄弱环节可能导致这些数据被窃取与滥用。“更危险的是，在集成金融交易功能的场景下，一旦智能体被攻破，攻击者可能诱导其执行错误转账、越权查询等操作，直接造成用户财产损失，动摇数字经济的信任基石。”

也正因如此，熊明辉把OpenClaw看作是一把双刃剑，在带来效率革命的同时，也引入了多维度的安全挑战。

3月8日，我国工业和信息化部网络安全威胁和漏洞信息共享平台（NVDB）发布预警提示，监测发现OpenClaw开源AI智能体部分实例在默认或不当配置情况下存在较高安全风险，极易引发网络攻击、信息泄露等安全问题。

美国网络安全企业CrowdStrike在专项报告中指出，行动型AI智能体的安全风险远高于传统对话式AI，一旦被劫持，可直接完成用户设备接管、敏感信息窃取、核心数据篡改等操作，危害覆盖个人隐私、企业商业机密与公共机构信息。日本、韩国多家大型企业及金融机构紧急下发通知，全面禁止办公设备安装OpenClaw，严防核心业务数据泄露。英国信息专员办公室也发布风险提示，提醒公共服务机构谨慎使用行动型AI工具，避免公民隐私信息遭到非法窃取。

多方协同搭建立体安全防护网

面对OpenClaw带来的系统性安全风险，多国政府、科技企业、安全机构，已经或正在搭建应急安全防护体系。

美国联邦通信委员会、联邦贸易委员会联合发布临时监管规范，要求OpenClaw等行动型AI严格落实权限最小化、操作可追溯、高危行为人工审核三项核心要求。欧盟正式将OpenClaw纳入高风险AI监管范畴，强化数据合规与隐私保护要求。韩国、加拿大、澳大利亚等国快速跟进，发布针对性使用指引与安全规范。我国网信、工信等部门同步完善开源AI安全评估指南，强化供应链安全监管，压实开发者与使用者双重安全责任。

“在人工智能应用浪潮中，为坚守安全、可控、可信的底线，需构建多层次防护体系。”熊明辉把这种防护体系归纳为“四维”防线。第一道防线聚焦安全配置与权限管理。贯彻“最小权限”原则，确保AI智能体仅能访问完成任务必需的最小数据集与操作权限。第二道防线致力于供应链安全保障。AI生态具有开放性，第三方技能包、插件的引入扩大了潜在攻击面。需建立严格的审核机制，对引入组件开展全方位安全审查，确保来源可信、代码透明。第三道防线依赖监管与合规建设。行业需明确统一的安全标准与行为规范，为开发者和使用者提供清晰指引。在金融、能源等关键领域，企业应建立严格的内部合规框架，确保AI应用符合数据安全与隐私保护相关法律法规，将合规要求融入企业运营全流程。第四道防线融合前沿技术与人的因素。通过持续的安全培训与清晰的风险告知，提升参与者的安全意识与操作能力，让安全文化渗透到AI生命周期的各个环节。

“人工智能安全治理需具备全球视野。各国应加强合作，共同制定国际性安全标准与监管框架，建立威胁情报共享机制，携手应对网络风险。构筑‘技术防护、管理规范、合规约束、用户教育、全球协作’五位一体的纵深防御体系，才能在享受AI创新红利的同时，守住发展的安全底线。”熊明辉说。

凝聚国际共识构建长效安全体系

OpenClaw的全球走红，标志着人工智能正式进入行动时代，技术发展潜力巨大，但安全底线不容突破。构建长效治理体系，才能真正守住安全、可控、可信的发展底线。目前，多国政府、国际组织、科技界已发出明确声音，为行动型AI安全发展划定路径。

联合国教科文组织发布《2026国际人工智能安全报告》，确立跨国界行动型AI安全基线，推动监管互认与标准协同。国际标准化组织、国际电工委员会联合推进ISO/IEC27090专项标准，将于年内正式发布，成为全球首个聚焦AI系统网络安全防控的权威指南。欧盟表示将持续完善《AI法案》配套细则，强化行动型AI全生命周期监管；美国提出建立行动型AI安全认证体系，未通过认证的产品不得面向公众销售；我国加快推进开源供应链安全标准落地，明确开发者、平台、用户三方责任，建立穿透式追责机制，全球制度监管正逐步走向统一与规范。

国际合作方面，联合国多次发布联合声明，呼吁各国加强行动型AI安全协作，共同打击利用AI实施的网络犯罪、数据窃取与间谍活动。G7与金砖国家建立常态化安全沟通机制，开展跨境联合应急演练，提升协同处置能力。我国积极参与全球AI治理规则制定，推动形成多边、透明的治理体系，为全球行动型AI安全发展贡献中国方案与中国智慧。全球多国专家共同呼吁，行动型AI必须坚守伦理底线，划定不可接受风险清单，确保人类始终掌握最终控制权，让技术发展始终服务于人类共同利益。

从技术爆红到风险预警，从紧急应对到长效治理，OpenClaw引发的全球关注，深刻诠释了人工智能进入行动时代的机遇与挑战。国际社会以一致态度表明，安全是人工智能创新发展的前提与基石。唯有各国携手合作、政企协同、全民参与，严守安全底线、完善治理体系、凝聚国际共识，才能真正筑牢行动型AI安全防线。

是“合理使用”还是“版权侵权”

大英百科全书起诉OpenAI或重塑行业规则

□ 本报记者 王艺茗

大英百科全书公司及其子公司梅里亚姆—韦伯斯特公司近日在美国曼哈顿联邦法院起诉OpenAI，指控这家人工智能（AI）巨头滥用它们的参考资料训练人工智能模型。这场诉讼的核心争议在于，OpenAI未经授权使用大英百科全书近10万篇文章训练AI，究竟是应被允许的“合理使用”，还是必须禁止的“版权侵权”。

业内人士指出，这起案件远不止一起AI版权纠纷。从训练数据的“输入端”到生成内容的“输出端”、从传统的版权侵权到新兴的商标与署名权争议，大英百科全书正在向AI时代的“来源秩序”发起反击。

双方争论“合理使用”边界

本案的第一条争议，聚焦于AI模型训练阶段（数据输入）的数据获取行为是否构成版权侵权。据路透社报道，大英百科全书在3月13日提交的诉状中称，OpenAI利用近10万篇大英百科全书的在线文章、百科全书和词典条目，训练聊天机器人ChatGPT如何回应用户的提问，这种复制行为是“系统性、规模化”的。诉状将ChatGPT描述为对大英百科全书可信赖的高质量内容“搭便车”，将后者的内容价值转移至OpenAI，且未支付任何补偿。

OpenAI发言人3月16日在回应该诉讼时表示：“我们的AI模型旨在推动创新，其训练基于公开可获得的数据，并且符合‘合理使用’原则。”这是AI行业对抗版权诉讼的标准抗辩框架——他们认为将受版权保护的内容转化为训练数据，属于“转换性使用”，不应受到版权限制。

然而，本案的特殊之处在于AI所使用的大英百科全书内容的性质。与普通网页信息或新闻资讯不同，百科条目和词典释义经过严格的编撰、审核和更新流程，具有高度的原创性和权威性，本身就是具有稳定商业价值的版权产品。当AI模型吸收的是这类“高质量结构化知识体系”而非零散的互联网信息时，“转换性使用”的边界将会被重新审视。

值得注意的是，大英百科全书并非被动等待侵权的发生。诉状披露，该公司曾于2024年11月主动联系OpenAI探讨授权合作可能，但OpenAI方面“从未认真考虑授权”，尽管其已与其他同类出版商达成了授权协议。

事实上，目前的一些学术研究已经表明，在生成式人工智能时代，数据已经“不再局限于静态内容，而是渗透到AI生命周期的每个阶段——从塑造模型参数的训练样本，到驱动实际部署的提示词和输出结果”。这意味着，传统的“输入端”合规控制，可能已无法覆盖数据在模型内部持续发挥作用的全过程。大英百科全书发起诉讼的案件恰恰触及这一核心矛盾：即使承认训练阶段的复制是“转换性”的，但当这些内容通过模型输出不断被再利用时，权利人的控制权又该如何保障？

AI“记忆化”复现引声讨

如果说训练阶段的争议尚可在“转换性使用”的框架下辩论，那么大英百科全书所提出的与“输出阶段”相关的指控则直接将OpenAI推向版权侵权的传统禁区——复制。

大英百科全书在13日的诉状中附上详细对比证据，指控ChatGPT在回应用户请求时，生成与原作“逐字相同或高度近似”的内容。诉状明确指出：“ChatGPT复制了原告受版权保护内容的表达，含义和信息，并将其重新包装给消费者。ChatGPT没有添加任何新的表达、含义或信息。”OpenAI还通过人工智能生成相关内容摘要，“蚕食”了大英百科全书的网络流量。

这正是当前AI版权案件中最具争议也最核心的问题——模型究竟是在“抽象学习”，还是在特定条件下对原文进行“记忆化”复现？当用户输入“请给我《大英百科全书》中关于教育的文章”时，ChatGPT输出的内容与原版几乎完全一致。这种情况下，AI不再是被动地“学习”知识，而是主动地“提供”受版权保护的原文。

从数据保护的角度看，这一现象揭示了一个深层困境：数据一旦被纳入模型训练，其存在形式便发生了根本性转变，已经从独立的、可识别的作品，转化为分布式的、难以追踪的参数和权重。大英百科全书指出，虽然能够确认OpenAI使用了近10万篇文章，但“真正的复制范围只有OpenAI自己知道”。这种信息不对称，使得权利人在主张权利时处于天然的弱势地位。

提停止侵权及索赔要求

有媒体分析认为，大英百科全书起诉OpenAI案最具辨识度的创新之处，不在版权，而在商标与来源标注。诉状指控OpenAI不仅暗示自身获得授权复制大英百科全书的内容，还在AI产生的“幻觉”信息中不当引用大英百科全书，将事实上错误的陈述归因于这家拥有250多年历史的权威知识机构。

这里触及了一个超越版权法的深层问题：在AI时代，“来源可信度”和“品牌署名权威”究竟应当如何保护？对于百科、词典这类知识品牌而言，内容固然重要，但真正稀缺的是被社会长期承认的可信来源身份。如果AI生成了错误内容，却挂上大英百科全书的名字，损害的就不只是某个条目的点击量，而是品牌所代表的知识权威。

欧美媒体指出，此次诉讼是版权方针对科技公司未经许可使用资料训练AI系统发起的众多维权行动之一。大英百科全书去年曾对人工智能初创公司Perplexity AI提起类似的版权诉讼，该案目前仍在审理过程中。

业内人士指出，本案是继大英百科全书起诉Perplexity AI后，传统知识机构对AI时代“来源秩序”发起的一次关键反击。虽然OpenAI坚持“合理使用”抗辩，但本案很可能被并入纽约南区法院的多地区诉讼（MDL），与《纽约时报》等案件一并审理，最终判决可能重塑整个AI行业的游戏规则。

据悉，大英百科全书公司已在诉状中要求法院下令禁止OpenAI的侵权行为，同时向OpenAI索取数额不详的赔偿。目前来看，无论此案未来的判决结果如何，一个基本共识正在形成：AI的发展不能以消解“来源秩序”为代价，数据的保护也需要适配AI时代的技术特征。

立法动态

哈总统签署实施新宪法措施的法令

当地时间3月17日，哈萨克斯坦总统托卡耶夫签署实施新宪法措施的法令，新宪法将于7月1日正式生效。据哈萨克斯坦总统府网站消息，托卡耶夫表示，新宪法是建设强大、充满活力且具有竞争力的国家的重要基础，为维护国家主权和领土完整提供法律保障，进一步推动社会公平正义和法治建设，并为教育、科学、创新和文化发展创造条件。哈萨克斯坦宪法法院官网于2月发布新宪法草案，共11章95条。与现行宪法相比，草案修改内容包括将议会由两院制改为一院制、恢复副总统职位、设立人民理事会等。根据草案内容，新宪法将废除参议院，设立由各政党共145名议员组成的一院制议会“库鲁尔泰”，议员任期5年。

阿政府将在国会推动一系列新法案

当地时间3月17日，阿根廷政府在举行高层会议后宣布将在国会推动一系列新法案，如残疾人法案，并对刑法、私有财产法等相关法案进行修改。阿根廷内阁首席阿多尼希望在国会推动立法改革。他表示，首先将对刑法进行改革，“重点将放在加大刑罚力度上”。对私有财产法的修改则包括改革《征用法》《土地法》及为促进社会融合的土地所有权规范等。另据总统府消息，阿根廷政府还将推动对冰川法、残疾人法案进行修改。此外，还计划推动通过高校融资法。据报道，17日的高层会议还确定了递交相关议案的时间表，并已经就与政治盟友及各省长的协商策略达成一致，这对于在国会顺利推动通过以上新法案至关重要。

阿联酋议会通过传染病防治法草案

阿联酋联邦国民议会近日审议通过传染病防治法草案，旨在通过立法强化全国公共卫生防控体系。该草案对传染病病例报告、疫苗接种义务以及故意传播疾病等行为设定了明确法律责任，以加强医疗机构和社会公众在疫情防控中的义务。根据草案规定，公立和私立医疗机构中的医生、药剂师、药房技术员以及其他医疗从业人员必须严格履行传染病报告职责。若医务人员在知悉或怀疑出现传染病感染或死亡病例后未在8小时内向主管部门报告，将被处以3万至10万迪拉姆罚款。草案还对违反防疫规定和故意传播疾病的行为设定了严格惩罚。该法案的出台被视为阿联酋加强公共卫生治理和防疫法律体系的重要一步，旨在通过更严格的法律约束提升社会整体防疫意识与责任。

（本报记者 吴琼 整理）

俄美乌谈判意外推迟 和平进程再遇阻

□ 本报驻俄罗斯记者 史天昊

近期，备受国际社会关注的俄罗斯、美国、乌克兰三方谈判宣布推迟，原定举行的面对面磋商未能如期启动，具体重启时间尚未确定。这一消息一经公布，立即引发全球舆论高度关注。作为俄乌冲突缓和与政治解决的关键渠道，此次谈判推迟不仅意味着俄乌和平进程再度陷入停滞，更折射出美俄战略博弈、俄乌核心分歧以及地区局势复杂联动的深层矛盾。从谈判筹备受阻、各方立场对峙到未来局势走向，谈判推迟背后的多重因素，正深刻影响着欧洲安全格局与全球地缘政治稳定。

核心议程难达一致

此次俄美乌三方谈判推迟，并非临时突发状况，而是多方在前期沟通中矛盾累积、分歧难以调和的结果。据多方消息源证实，原计划举行的本轮磋商，是自冲突僵持以来，三方首次计划在同一平台就停火条件、安全保障、人道主义通道等关键议题展开直接对话，一度被外界视为危机转向政治解决的重要契机。然而，在临近会谈之际，三方却在日程安排、会谈地点、核心议程等基础性问题上难以达成一致，最终导致谈判被迫延后。

当地媒体在报道中称，三方在谈判举办地点上的争执最为突出。俄方坚持谈判应在中立第三国举行，以保证谈判环境的公平与安全，反对在西方主导的国家或地区开展对话；美方则倾向于在其认可的地点组织会谈，试图掌握谈判流程的主导权；乌方虽表态支持尽快启动对话，但在具体安排上更多与美方保持协同。此外，三方在谈判议程设置上同样存在明显分歧，俄方希望优先讨论乌克兰中立地位、安全保障等结构性问题，美方与乌方则更关注战场停火、领土现状、外部援助等现实议题，双方难以形成统一的谈判清单。

有分析认为，外部局势变化也是此次谈判推迟的原因之一。当前国际热点议题频发，美国难以集中精力推动俄乌对话；欧洲内部对和谈的态度存在分歧，部分国家支持快速缓和局势，另一部分则主张保持对俄施压。诸多因素叠加之下，原本被寄予厚望的三方谈判最终宣告推迟，为本就脆弱的和平前景蒙上阴影。

深层矛盾难以调和

“谈判推迟的表象之下，是俄美乌三方在核心利益与安全诉求上的根本对立，这也是导致对话难以推进的真正原因。”对于此次谈判的推迟，此间分析人士指出，从俄罗斯角度出发，其核心目标在于推动乌克兰保持中立地位、放弃加入北约，同时确保自身边境安全与地缘战略空间不受威胁。俄方多次强调，谈判必须建立在现实基础之上，任何脱离当前战场态势与安全格局的方案都不具备可行性。

美国作为斡旋方与乌克兰的主要支持者，其立场始终服务于自身全球战略布局。美方一方面表态支持外交解决冲突，另一方面持续向乌克兰提供军事与经济援助，不断强化对俄遏制态势。在谈判问题上，美国既希望通过对话管控危机规模、避免局势失控升级，又不愿轻易放松对乌支持、放弃对俄战略施压机会，这种摇摆性立场直接影响了谈判的推进节奏。

乌克兰则在领土完整、主权独立与外部安全保障问题上态度坚决。乌方明确表示不会在国家主权与领土问题上妥协，同时希望通过谈判获得长期安全承诺与国际社会支持。面对俄方提出的相关条件，乌方难以接受，而外部援助的持续输入，也在一定程度上降低了乌方在核心议题上让步的意愿。三方立场相向而行难度极大，使得谈判从筹备之初便面临难以逾越的障碍。

此外，美俄之间的战略博弈贯穿谈判始终。围绕欧洲安全秩序、地缘影响力、能源格局等核心利益，双方的竞争早已超出乌克兰危机本身。谈判推迟本质上也是双方在外交主导权、规则制定权上的一次角力，短期之内难以达成实质性妥协。

和平前景再添变数

随着俄美乌谈判推迟，乌克兰危机的未来走向再度变得不明朗，地区和平与稳定面临新的考验。从短期来看，谈判停滞意味着双方在战场上的对峙状态仍将延续，前线交火、远程打击等行动可能持续发生，人道主义局势难以得到根本缓解。对于欧洲而言，危机长期化将持续拖累能源供应、经济复苏与安全稳定，难民问题、能源价格波动等外溢效应将进一步显现。

乌克兰危机已进入第五个年头，危机的延宕已经对国际贸易、粮食安全、能源格局产生深远影响，此次谈判推迟无疑将加剧全球经济与安全领域的不确定性。广大发展中国家普遍呼吁各方保持克制，尽快重返谈判桌，通过外交途径化解分歧，避免冲突升级带来更大范围的危机。联合国及多个国际组织先后发声，敦促俄美乌摒弃对抗思维，为重启对话创造有利条件，避免和平机会一再流失。

分析人士指出，尽管当前谈判受阻，但政治解决仍是乌克兰危机唯一可行的最终出路。军事对抗无法带来持久和平，只有各方回归理性、相互尊重核心安全关切，才有可能打破僵局。未来，随着国际社会斡旋力度加大、三方立场逐步调整，谈判仍具备重启的可能性，但这一过程注定充满曲折与博弈。

此次谈判推迟，既是当前多方矛盾的集中体现，也为后续对话敲响了警钟。在全球安全形势日趋复杂的背景下，俄美乌三方能否克服分歧、重启磋商，不仅关系俄乌两国的前途命运，更深刻影响着欧洲安全秩序与全球战略稳定。国际社会仍在期待，各方能够以对话代替对抗、以协商化解分歧，让和平的曙光早日降临。

乌克兰危机延宕至今，已逾四年时间。图为当地时间3月16日，顿涅茨克地区曼古什镇，一栋邮局建筑在无人机袭击后部分坍塌。 CFP供图