OpenClaw应用的安全防控

■主持人
黄冬生　杨浦区人民检察院副检察长
■嘉　宾
车延楠　上海科学技术交流中心高新技术发展处处长
顾　伟　徐汇区人民检察院检察官
房慧颖　华东政法大学副教授
张　勇　华东政法大学教授

开源AI智能体OpenClaw（俗称“养龙虾”）的应用，可能存在数据安全、网络安全、个人信息泄露隐患等风险。面对这些风险，该如何从技术、行为、社会监管等维度协同发力，加强安全防控？

AI生图

从技术研发、产业生态与用户使用等多维度统筹考量

车延楠：OpenClaw应用的风险与防范，需要立足发展，从技术研发、产业生态与用户使用等多个维度统筹考量。

在技术层面，不追求绝对化的封闭限制。可以通过权限管控、动态监测等手段，实现可控创新。在保障技术活力的同时，把越权操作、数据泄露、恶意利用等风险控制在可控范围。

在产业层面，持续强化外部监管，同时引导人工智能头部企业切实履行自律责任。

头部企业应秉持“能力越大、责任越大”理念，在平台治理、技能（Skill）审核、权限管控、风险预警等方面主动提高标准，从源头提升产品与服务的安全性、可靠性。

在个人层面，我对OpenClaw的应用持积极拥抱的态度。普通用户应主动学习，知晓OpenClaw便利，明晰风险。

在使用OpenClaw时，遵循最小权限原则，按需逐项开放而非全权授权，自行建立完整操作日志，确保每一步AI操作授权都可回溯。

预防监管重点应放在使用行为与应用场景

顾伟：OpenClaw的风险本质上集中在应用环节，预防监管重点应放在使用行为与应用场景上。具体来说：

一是完善行业治理。行业协会可为OpenClaw这类AI智能体设立明确准入门槛与底线标准。如针对AI自主执行可能引发的法律风险，明确AI智能体参与工作的责任条款，建立从人类到AI智能体再到行为的完整审计追踪；明确AI智能体应用的边界与流程，制定清晰使用规范，明确哪些任务可以自主执行、哪些须人工审核。

二是强化多部门协同监管。监管部门定期开展网络安全与数据安全检查，对相关企业、平台未采取必要技术隔离措施、未履行安全管理义务的，引导帮助完善制度设计和落实好管理责任。结合检查中发现的问题，梳理风险点，推动网信、工信、公安等部门加强开源AI智能体安全监管，提出立法建议，促进技术发展与安全规范同步。

三是强化司法引导。发布典型案例清晰界定合法与违法的行为边界，为技术研发、平台运营及普通用户提供明确行为指引。

例如，就检察机关而言，可以探索刑事检察与公益诉讼协同联动模式，针对OpenClaw滥用导致的群体性信息泄露、公共利益受损等问题，强化检察保护力度，制发典型案例，为打击恶意利用技术实施的违法犯罪行为提供参考。

实现技术、行为、社会三个层面的系统性治理

房慧颖：可以将OpenClaw风险纳入制度化治理框架，明确使用规则与责任归属，实现技术、行为、社会三个层面的系统性治理。

首先，在技术层面，聚焦最小权限和强制隔离。对于使用者来说，不让OpenClaw默认接触与任务无关的本地文件、账户凭证和高危系统，尤其不能直接赋予金融、政务、核心业务系统的高权限。对于运营者来说，应当把关键操作的二次确认、风险输出过滤、插件准入审查、异常熔断和权限分级等，作为运营安全最低标准。其次，在行为层面，把规则讲清楚、把边界划明白。对使用者而言，避免向AI智能体输入与任务无关的敏感信息，避免模糊、歧义和高风险指令。对运营者而言，建立实名核验、权限审批、异常监测、风险预警和审计复盘机制，使每一次高风险调用都有迹可循。把OpenClaw纳入AI使用规程，明确哪些业务可以用、哪些数据不能碰、哪些操作须人工复核、哪些场景原则上不得部署。最后，在社会层面，对OpenClaw这类智能体遵循分类分级、协同共治路径。开源社区应当承担漏洞披露和版本修补责任，平台和云服务商应当承担准入提醒和生态审查责任，行业协会应当尽快形成面向高风险行业的专项操作指引，监管部门则应当把网络安全、数据安全、人工智能治理和刑事打击衔接起来，形成“准入—预警—追溯—处置”的协同治理链条。

张勇：可以进一步聚焦授权、数据、评估与追责四个关键环节，形成闭环治理体系。一是明晰授权筑牢防线。可以围绕授权机制构建系统化风险防控体系：明确授权边界、规范授权流程、强化权限最小化配置，对高敏感操作强制设置二次确认，从源头压实权限合规责任。二是推进数据分类分级管理。可以按照数据敏感程度实施差异化权限与保护措施，实现精准管控。三是强化全流程风险评估。在技术研发、部署应用、运行操作等各阶段开展风险识别与研判，根据评估结果及时采取预警、限制、阻断等对应处置措施，做到早发现、早干预，从源头防范风险升级为违法犯罪行为。四是主体责任追责时确立优先顺序。优先运用民事追责、行政处罚等前置手段，将大量应用风险在行政监管和行业自律层面化解。鼓励企业主动开展内部安全评估与合规自查，把风险解决在前端。

（发言整理：杨浦区人民检察院　肖凤　嘉定区人民检察院　曹俊梅）