“AI龙虾”掀热潮 智能体如何系好“安全带”？

▶ 本报记者 罗晓燕

近期，一款名为OpenClaw（昵称“龙虾”）的开源AI智能体，在科技圈乃至大众视野中爆热。然而，就在网民热捧之际，国家互联网应急中心、工业和信息化部网络安全威胁和漏洞信息共享平台、国家安全部等密集发声，揭示其严重的安全风险，并发布详细使用指南，提醒用户务必审慎对待。

当人工智能（AI）真正拥有了操作电脑的能力，人们该如何系好“安全带”？对此，多位业内人士在接受记者采访时表示，以OpenClaw为代表的AI智能体，目前仍处于发展初期阶段，亟需从底层技术架构上完善安全机制，并同步强化国家监管制约。

2026年3月11日，在云南省蒙自市，用户在开源AI智能体“龙虾”手机网页版浏览。新华社发（薛莹莹/摄）

2026年3月11日，在云南省腾冲市一家手机专卖店，工作人员在和同伴交流开源AI智能体“龙虾”的操作体验。新华社发（赵辉/摄）

一场养“龙虾”狂欢

OpenClaw究竟有何魔力？有分析人士认为，它的魅力在于其前所未有的“高操作权限”——能像人一样操作电脑、执行任务，将AI从“陪聊”推向“做事”的新阶段。

三呆科技AI架构师布丰观察到，不同人群对“龙虾”的想象截然不同。“有人看到的是企业办公智能化的效率潜力，有人看到的是服务自身的个人助理潜能。”他表示，这种“靠谱的工作伙伴”在提高泛办公场景效率上潜力巨大，甚至衍生出“数字分身”概念——当“你”下班离开电脑后，让数字化的“你”继续工作。

布丰告诉记者，这波“养龙虾”热潮中还掺杂明显的FOMO（错失恐惧症）情绪。“尽管有过热成分，但‘龙虾’热作为首个人类和AI关系里程碑式的时刻，已展现出人类愿意以更加对等的姿态看待AI，这是很重要的。”

360集团创始人周鸿祎在近期举行的“龙虾安全媒体交流会”上表示，“龙虾”代表一种新的智能体范式。传统智能体更像是能力预设的工具，而“龙虾”更像是具备自主能力的进化体。“未来用户面对的可能不再是‘养一个智能体’，而是管理大量AI的智能体。”周鸿祎认为，如果能通过“龙虾”的普及驱动智能体在政府机构和企业落地，让公众接受智能体理念，对中国AI产业发展是非常大的贡献。

“野蛮生长”下的安全隐患

OpenClaw持续走热，其安全问题备受关注。不少用户更是从付费安装“龙虾”到付费卸载“龙虾”。

近期多个官方部门针对“龙虾”安全问题发布风险提示。工业和信息化部网络安全威胁和漏洞信息共享平台3月10日发布《关于防范OpenClaw开源AI智能体安全风险的预警提示》，针对这款热门工具存在的安全隐患，明确作出风险警示并提供相关防范指引。国家互联网应急中心同步发布《关于OpenClaw安全应用的风险提示》，进一步明确OpenClaw在权限配置、插件使用等方面存在安全风险，为各类用户规范使用该工具提供安全指引。3月17日，国家安全部发布《“龙虾”（OpenClaw）安全养殖手册》，就开源AI智能体OpenClaw的使用风险提出专业指引。

“开源的AI产品一般不具备传统企业产品的成熟度，‘龙虾’是个上限很高同时下限也很低的技术产品。”布丰表示，当FOMO情绪高涨时，人们看到的是新奇。但当Token（词元）账单砸下来且大大超出预算，再叠加对运作机制的不了解，用户的安全恐惧就会以另一种形式爆发。

“争议来得迅猛，是因为技术跨界突破了传统的安全防线。”北京威努特技术有限公司（以下简称“威努特”）产品经理任道鑫表示，过去的大模型被关在对话框“笼子”里，而OpenClaw直接拥有了操作系统的交互权。

“‘龙虾’当前最致命的安全漏洞并非单一的缺陷，而是底层架构对‘执行层’的无条件信任。”任道鑫表示，例如近期频发的OpenClaw默认绑定到0.0.0.0（监听所有网络接口）导致公网全量暴露，以及臭名昭著的ClawJacked（WebSocket跨域劫持漏洞），这些现象本质上都指向同一个“致命伤”——原生应用在赋予AI极高权限的同时，完全没有匹配相应的身份认证与行为拦截机制，如同让一个没有交通规则意识的驾驶员直接把车开上高速公路。

“这是典型的‘全链路失控’。”任道鑫表示，一是模型侧的“幻觉与意图误判”，导致正常指令被曲解为删除操作；二是执行侧的权限“裸奔”，系统没有对敏感操作（如文件删除、高危命令）进行二次确认或权限隔离；三是生态侧的供应链“投毒”，正如近期安全通报指出的，公开市场（如ClawHub）上存在大量带有恶意代码的技能插件。而威努特在近期推出的WinClaw（安全龙虾）中内置安全引擎和白盒化插件规则，正是为了精准切断这3条失控链条而设计的。

周鸿祎认为，AI产业正进入大模型+智能体的“双线”进化阶段。“龙虾”印证了这一趋势：大模型不断提升认知能力，智能体持续强化执行能力。随着AI从“会回答问题”走向“能动手干活”，相关安全问题自然浮出水面。

“‘龙虾’作为新生事物，不应因其存在潜在风险而被简单否定，而应在发展中逐步建立安全边界。普通用户在使用智能体时，要特别注意账户与资金安全，不要向智能体泄露密码、口令等敏感信息。”周鸿祎提醒道。

技术与制度需双管齐下

面对OpenClaw的原生安全隐患，国内已涌现出多款国产“龙虾”。中国信息通信研究院专家此前提示，尽管“龙虾”智能体已经更新到最新版本，能修复已知的安全漏洞，但并不意味着完全消除安全风险。

通过云化部署、沙箱隔离等方式，能否将“危险工具”真正转化为“可靠工具”？任道鑫认为，沙箱隔离等技术手段确实是现阶段防范原生隐患的有效途径之一。但需要明确的是，无论是云化部署还是端侧本地化，核心在于“行为管控的精细度”。

“国产化改造不能只是简单套壳，必须深入到智能体的执行逻辑内核中去。未来，威努特会针对不同的用户场景需求，持续探索最优部署与隔离形态，以确保在极致的算力释放与隐私安全之间找到平衡点。”任道鑫进一步表示，开源OpenClaw面临的是“结构性缺陷”，单纯依赖打“补丁”属于头痛医头、脚痛医脚，无法根治。只要其多层架构（网关层、智能体层、执行层）之间依然缺乏内生的信任校验机制，“零日漏洞”（0-day）就会层出不穷。“长期看，安全治理必须是‘原生技术体系重构’与‘国家制度规范’的双管齐下。”

周鸿祎认为，OpenClaw等AI智能体目前仍处于发展初期阶段，普遍面临使用门槛偏高、结果稳定性不足等问题，其底层安全机制仍有待进一步完善。如若缺乏有效管控，让智能体随意与外部系统交互，或在公开环境中执行复杂任务，可能导致用户密码、API密钥等敏感信息被诱导泄露。此外，OpenClaw虽然支持通过外部技能包扩展能力，但部分技能包来源复杂，如果缺乏审核机制，可能面临被植入恶意代码的安全隐患。

码上读报 扫码阅读全文

竞逐太空计算赛道，中国抢先“落子”

前不久，马斯克申请发射100万颗卫星构建太空数据中心的新闻引发热议，太空计算迎来前所未有的关注。

“抢占太空算力赛道是支撑我国航天事业高质量发展、保障国家太空安全的战略举措。”今年全国两会期间，全国人大代表、北京航空航天大学教授张涛提交了有关建设太空算力星座的建议。他认为：“太空计算是必争的战略赛道，短期投入大、长期回报高。它既是算力瓶颈的终极解法，也是国家安全、数字经济与太空强国的核心支撑。”

在太空轨道上部署计算能力，这个曾被视为“天方夜谭”的构想，已成为各国抢占未来制高点的新赛道。

全球正掀起一股算力上天的浪潮，中国已抢先“落子”。

北京邮电大学联合多家单位共建的“天算星座”已于2021年10月启动。

2025年5月，国星宇航、之江实验室携手发射全球首个太空计算卫星星座，标志着我国首个整轨互联的太空计算星座正式进入组网阶段。

《科技日报》2026. 3. 19 管晶晶

制造业数字化转型纵深推进

当前，我国制造业数字化转型覆盖广度明显提升，进入规模化普及阶段。中国信息通信研究院（以下简称“信通院”）前不久发布的《制造业数字化转型发展报告（2025年）》显示，截至2025年12月，全国规模以上工业企业开展数字化改造比例达89.6%，数字化设备普及率达到57.7%；累计建成3.5万余家基础级、8200余家先进级、500余家卓越级、15家领航级智能工厂。

业内专家认为，随着政策体系日益健全，应用持续深化，数字技术和产品供给水平明显提升，数字基础设施支撑坚实有力，制造业数字化转型整体步伐不断加快。

信通院信息化与工业化研究所工业发展研究部主任袁媛认为，在全球经济增长放缓、我国制造企业竞争压力加剧的背景下，新一代信息技术正从创新范式、生产方式到组织模式，对制造业进行系统性重塑，助力企业从“拼价格”转向“拼效率、拼能力”。一方面，数字技术催生“内生增长”新动能。另一方面，数字技术给企业开辟“价值创造”新赛道。

《经济日报》2026. 3. 19 李芃达