连马化腾朋友圈都在提“安全龙虾”

网络安全专家：盲目跟风养虾有风险，龙虾必须“圈养”，不能“放养”

3月27日至29日，2026全球开发者先锋大会在上海徐汇西岸举行，现场的“龙虾养殖场”“龙虾派对”等活动掀起热潮，以红色龙虾为图标的开源软件OpenClaw成为全场焦点。活动中，不仅有AI爱好者踊跃参与云端及本地“小龙虾”部署，不少银发一族也在工作人员指导下体验安装、“养龙虾”，现场甚至出现参会者抱电脑排队装机的热闹场景。

这只AI“龙虾”从极客小众玩法走向全民实操，成为AI智能体走进大众生活的鲜活缩影，但随之而来的疑问也备受关注：这股热潮究竟是资本虚火，还是能真正推动效率革命？其安全部署的边界又该如何界定与把控？

“龙虾智能体的价值，在于实现了AI从‘对话建议’到‘规划执行’的跃迁，但热潮之下，权限集中带来的安全隐患、黑灰产诈骗等风险更不容忽视，普通用户务必警惕。”北京众安天下科技有限公司创始人、中国计算机行业安全协会数据安全产业专家委员会委员杨蔚表示。

在接受新闻晨报独家专访时，杨蔚说：“最近，国家安全部推出了龙虾OpenClaw安全养殖手册，提醒广大用户理性辨别、规范使用。连马化腾的朋友圈都在提‘安全龙虾’。我想说，新技术值得探索，但安全必须先行。对普通用户，我的建议是，可以尝试，但必须‘圈养’，绝不能‘放养’。”

从“只说不做”的顾问，到“能想会干”的执行者

“龙虾之所以能吸引这么多人饲养，是因为大家都认为，它比豆包、DeepSeek等AI工具更能干活。”刚刚开始养虾的上海爷叔老段告诉新闻晨报记者。

深耕网络安全领域多年的白帽安全专家杨蔚，揭示了龙虾吸引力的缘由。“当我们打开豆包、DeepSeek，让它们帮我们写一篇文章、搭一个方案框架、找一份行业资料，甚至构思一条视频脚本时，我们其实一直停留在AI的舒适区——对话与建议。它们就像博学的顾问，能给你思路、给你素材、给你方向，但始终需要你亲自动手，把这些建议落地成具体的成果。”

而龙虾智能体的出现，恰恰打破了壁垒，让AI从只说不做的顾问，变成了能想会干的执行者，实现了从对话建议到规划执行的跨越。

它显著降低了门槛，让不懂编程的普通人，也能拥有专业开发团队的自动化能力，释放出惊人的数字杠杆效应。曾经，跨软件、复杂流程的自动化，是程序员的专属领域——要写代码、调接口、做调试，普通人只能望而却步。但现在，龙虾智能体让“自然语言 = 自动化指令”，只要你能清晰地描述需求，它就有望代劳大部分操作，将复杂工作流程大幅简化。

它还让那些“只可意会、不可言传”的隐性知识不再“藏在肚子里”，而是能被记录、被传承、被放大。“我们都有这样的体验：很多工作的高效，依赖的不是书面上的规则，而是个人多年积累的经验、习惯和直觉——这些隐性知识，就像老匠人的独门绝技，只能靠口传心授。而龙虾恰恰提供了一种能力，能把这些经验，转化为可记录、可优化、可复用的数字技能。”

杨蔚说，龙虾智能体的价值，从来不是对现有AI工具的简单升级，而是一次能力的根本性重构。人们养虾，是因为它让AI从“纸上谈兵”走向“落地实干”，也让自动化走进普通人的工作生活。

养虾风险的关键，藏在“全权委托”里

养虾热潮之下，杨蔚更加关注那些被热闹遮蔽的隐患。

“这只AI‘龙虾’的核心风险，恰恰源于它最核心的能力本身。”杨蔚直言，龙虾智能体本质上是一个必须拥有高权限才能工作的自动化执行体，它的“能力”与“风险”从来都是一体两面，无法分割。

他打了一个形象的比方：“这就好比你聘用了一位能力超强，但背景、意图都没经过任何验证的‘数字全能助理’，你放心地把自己的电脑、手机，甚至所有线上账户的操作权限，全都交给了它。”

风险的关键，就藏在这份“全权委托”里——权限的高度集中，让AI智能体成为了一把“双刃剑”。杨蔚进一步解释，当前龙虾智能体技术迭代速度飞快，功能不断升级，但在安全架构搭建、权限的最小化节制，以及操作全流程的审计监管等方面，还处于相对早期的阶段，尚未形成完善的防护体系。

得知不少用户不做任何安全防护直接“养虾”，杨蔚直言“不可以”。“普通用户如果缺乏基本的防护措施，就盲目跟风部署、使用，后果不堪设想。”他警示道，用户的各类数字资产都将面临显著风险，小到日常使用的游戏账号、社交媒体账号，大到与资金安全紧密相关的网上银行账户，都可能因为这过高的权限的泄露，陷入安全危机。

龙虾必须“圈养”，绝不能“放养”

热潮之下，不少普通用户出于好奇或工作需求，也想尝试部署、使用龙虾智能体。那么，在享受其便捷性的同时，普通用户必须遵守的最低限度安全准则是什么？

杨蔚给出了明确答案。“我认为三个核心准则是‘隔离、最小权限与全程可监督’。”杨蔚强调，这是普通用户尝试使用龙虾智能体的底线，缺一不可，具体可分为三条必须严格恪守的安全要求。

第一条是环境物理隔离，这是不可妥协的前提。“绝对不要在存有敏感数据的日常主力电脑或手机上安装。”杨蔚提醒道，这里的敏感数据包括工作文件、财务信息、通讯记录，以及涉及敏感商业相关的文件和数据等。他建议，用户应使用一台功能单一、无重要数据的专用设备，若条件有限，也可在主力机上通过完全独立的虚拟机运行，从根源上隔绝敏感数据与智能体的接触。

第二条是权限最小化。杨蔚解释，用户在安装与配置龙虾智能体时，需严格审视并限制其可访问的目录、应用和网络权限，只授予其完成具体任务所必需的最低权限，杜绝“过度授权”。“例如，若仅用于整理公开资料，就应禁止其访问浏览器历史、通讯录及所有社交、金融类应用，包括涉及内容发布的社交应用等。”他举例说明，通过权限管控，可大幅降低龙虾带来的安全风险。

第三条是操作可监督。“在运行期间，尤其是在初期，要时刻关注龙虾的操作。”杨蔚表示，对于重要或不可逆的操作，比如删除文件、发送邮件、执行安装、修改系统设置等，必须为智能体设置“二次确认”的交互指令，确保最终执行权始终掌握在用户自己手中。同时，要避免在无人监督时，让智能体执行涉及数据修改、对外通信等敏感任务，且需确保其关键操作的日志清晰、可查，以便后续进行审计与回溯，及时发现并规避潜在风险。

对于希望降低部署门槛和风险的普通用户，可以考虑在天翼云、阿里云、腾讯云等主流云平台上进行尝试。“这类云服务通常能提供基础的环境隔离，相当于为‘龙虾’准备了一个独立的‘云上鱼缸’，使其操作与您本地的个人电脑物理隔离。但请注意，这主要解决了‘环境隔离’问题，‘权限最小化’与‘操作可监督’等安全责任，仍需用户根据云平台指引进行认真配置和持续关注‘龙虾’本身的安全风险。”杨蔚强调。

“幽灵操作”现身，诈骗陷阱需警惕

龙虾不“圈养”，不仅会带来安全隐患，甚至可能出现让人费解的“幽灵操作”。

近日，上海市民姜女士分享了自己的亲身经历：刷手机时偶然发现，自己在某知名生活服务平台的账号，竟在一条陌生帖子下评论了“有wifi吗？”而评论发布的时间，她早已进入梦乡。

平台客服调查后表示，其账户仅在本人手机上登录，排除了盗号可能，推测是误触了平台的“猜你想评”自动评论功能。

尽管这个事件可能与“龙虾”无关，但杨蔚指出，这恰恰说明了自动化操作可能带来的不可控后果。“如果没有把‘龙虾’圈养好，这类幽灵事件，它也可以干。”

更值得警惕的是，随着龙虾智能体的热度攀升，不法分子已开始借这一新技术实施犯罪。“任何有热度的新技术，都可能被黑灰产迅速‘蹭热点’，利用公众的信息差实施犯罪。”杨蔚提醒大家，当前需重点防范以下四类骗局。

首先，需高度警惕“付费安装与教学”骗局，这是当前最为常见和典型的直接诈骗形式之一。龙虾智能体本身是开源免费的，不存在“内部渠道”“付费破解版”，任何声称提供这类服务，或是收取高价提供“一对一远程安装指导”的，都是骗局。

“其危害远不止骗取学费那么简单。”他警示说，在所谓的“远程协助”过程中，用户的设备控制权会完全暴露，不法分子可趁机植入木马、盗取账号密码，甚至威胁资金安全、泄露隐私，极端情况下还可能遭遇赎金勒索，造成严重损失。

第二，警惕“AI自动赚钱”投资陷阱，这类骗局最具蛊惑性，潜在损失也最大。不法分子常将其包装成“AI全自动炒股”“躺赚项目”等，还会展示伪造的高收益截图，吸引公众投入资金。杨蔚揭露，这类骗局本质上是资金盘或传销，更隐蔽的是，不法分子还会诱导用户购买天价“定制版软件”或“API密钥套餐”，实为“杀猪盘”的变种，一旦入局，损失难以挽回。

第三，警惕“恶意技能插件”供应链攻击。龙虾智能体可通过安装“技能”（Skills）扩展功能，但社区内的技能缺乏严格审核，这给了不法分子可乘之机。他们会制作伪装成“效率工具”“赚钱秘籍”的恶意插件，用户一旦安装，这些插件就可能在后台悄悄窃取浏览数据、本地文件，甚至将用户的设备变为攻击他人的“肉鸡”，危害极大。

最后，警惕“API密钥盗刷”风险。运行龙虾智能体需配置大模型API密钥，这就相当于一张预存了额度的数字信用卡，一旦被恶意技能窃取，就可能遭遇“盗刷”，产生无法追索的巨额费用，给用户带来经济损失。

文/晨报记者 何雅君
图/国家安全部微信公众号 受访方