“龙虾”（OpenClaw）安全养殖手册

OpenClaw（昵称“龙虾”）是一款开源AI智能体工具，上线不久便迅速成长为2026年度现象级“开源奇迹”。不少用户从付费安装“龙虾”，到付费卸载“龙虾”，养“龙虾”正在成为一场智能体的狂欢。但火热的“龙虾”在创新改变生活的同时，也存在原生风险。小安特别提示，广大用户要理性辨别、规范使用，以积极的心态和慎重的执行拥抱人工智能时代，让“龙虾”成为遵规守纪、产能高效的“数字员工”。

摸清“龙虾”的生产特点

“龙虾”智能体通过整合通信软件和大语言模型，依托高权限实现自主操作，成为其核心优势。

从“给出方案”到“落地执行”。“龙虾”不像大模型智能体通过问答提供咨询建议，而是可以通过聊天程序远程执行用户指令，自主完成任务。

从“固定功能”到“多种插件”。“龙虾”内置了大量技能插件，用户可直接下载使用，形成覆盖文件管理、邮件撰写、日历调度、网页浏览、定时任务等多场景的工具链。

从“普通工具”到“自我进化”。“龙虾”可以长期记忆用户使用记录，持续理解用户行为偏好，“越用越懂用户”，所以大家称之为“养龙虾”。

从“被动等待”到“主动服务”。“龙虾”可根据用户要求，主动感知外部情况，主动触发预警或执行动作，完成“夜间下达指令、晨间获取成果”的智能服务。

了解养“龙虾”的风险隐患

主机可能被接管。为实现“做事”能力，用户常赋予其最高系统权限，可能引发因 AI 误操作造成的数据损失。更严重的是，运行后可能被攻击者神不知鬼不觉获取设备管理权限，从而引发主机被远程操控，资源被非法占用等安全风险。

数据可能被窃取。部分用户缺乏数据安全意识，个人敏感数据交由“龙虾”处理，一旦被攻破，可能造成个人隐私泄露，带来财产与安全风险。

言论可能被篡改。“龙虾”智能体可在社交网络自主发声，一旦被攻击者接管，可能被用于生成和传播虚假信息、实施诈骗等不法活动。

技术可能有漏洞。“龙虾”缺乏专业维护与漏洞修复机制，攻击者可能通过恶意插件投毒等方式，诱导智能体突破权限管控，主动窃取本地设备的核心敏感信息，其隐蔽性远超传统木马程序。

“养虾人”必看安全指南

给自己的“龙虾”全面体检。检查控制界面是否暴露在公网、权限配置是否过高、存储的凭证是否已泄露、安装的插件来源是否可信等问题。对于严重安全风险，请立即采取隔离、下线等处置措施。

为自己的“龙虾”做好防护。必须遵循最小权限原则，严格限制智能体的操作范围。对存储的敏感数据必须进行加密，建立完整的操作审计日志，尽量在隔离环境（如专用虚拟机、沙箱）中运行“龙虾”，限制其对核心资源的访问。

让自己的“龙虾”老实好用。“龙虾”并非供人娱乐的数字宠物，而是能够自主执行任务、承担流程操作、持续学习成长的“数字员工”，养“虾”人应理性看待、规范使用，让其在合规、安全、可控的前提下成为提升治理效能，服务生产生活的数字化生产工具。

来源：国家安全部