光联AI Agent安全防护服务解决方案

光联世纪

AI时代的新安全挑战

AI Agent删光Meta AI总监邮箱

Meta AI总监把最火的AI智能体OpenClaw接上了自己的工作邮箱。由于OpenClaw需要「压缩上下文」来处理大的信息量，在压缩的过程中，OpenClaw把之前设定的「未经批准不得操作」这条指令给忘了，疯狂删除邮件，喊停三次全部无视。

问题点

关键指令遗忘：大语言模型的底层机制决定上下文窗口有限，信息会被压缩，而被压缩掉的，可能恰好是最重要的那条安全指令。

AI Agent恶意Skill供应链投毒事件

2026年1月底，OpenSourceMalware（安全研究组织）称已在ClawHub上发现至少至少14个恶意'skills'，这些skills伪装成加密货币交易或钱包自动化工具，试图向用户系统投递恶意软件。

2月初，Koi Security（为OpenAI、Fortune Top50公司提供安全服务）称已在OpenClaw中发现341个恶意Skill。

Date	Event
Jan 27 – 29, 2026	Initial wave of at least 14 malicious "skills" uploaded to ClawHub.
Jan 27 – Feb 1, 2026	Second, larger wave published, bringing the total to over 230 malicious packages across ClawHub and GitHub.
Feb 1 – 2, 2026	Koi Security completes a full audit of all 2,857 skills on ClawHub, identifying 341 malicious entries—335 of which belong to the single ClawHavoc campaign.
Feb 2 – 3, 2026	Widespread reporting by security firms and community researchers. OpenClaw creator introduces a user-reporting feature in response.

问题点

① **缺乏Skill安全检查：**Agent Skill可被利用进行提示注入、RCE等攻击，使得Agent产生非预期操作。

② **Agent未校验Skill来源可信：**公开不等于可信，Agent应用需校验第三方源安全性。

企业AI落地面临的四大风险

AI Agent风险提示

关于OpenClaw安全应用的风险提示

原创 CNCERT 国家互联网应急中心CNCERT 2026年3月10日 19:02 北京

近期，OpenClaw（“小龙虾”，曾用名Clawdbot、Moltbot）应用下载与使用情况火爆，国内主流云平台均提供了一键部署服务。此款智能体软件依据自然语言指令直接操控计算机完成相关操作。为实现“自主执行任务”的能力，该应用被授予了较高的系统权限，包括访问本地文件系统、读取环境变量、调用外部服务应用程序编程接口（API）以及安装扩展功能等。然而，由于其默认的安全配置极为脆弱，攻击者一旦发现突破口，便能轻易获取系统的完全控制权。

前期，由于OpenClaw智能体的不当安装和使用，已经出现了一些严重的安全风险：

1.“提示词注入”风险。网络攻击者通过在网页中构造隐藏的恶意指令，诱导OpenClaw读取该网页，就可能导致其被诱导将用户系统密钥泄露。

2. “误操作”风险。由于错误的理解用户操作指令和意图，OpenClaw可能会将电子邮件、核心生产数据等重要信息彻底删除。

3.功能插件（skills）投毒风险。多个适用于OpenClaw的功能插件已被确认为恶意插件或存在潜在的安全风险，安装后可执行窃取密钥、部署木马后门软件等恶意操作，使得设备沦为“肉鸡”。

4.安全漏洞风险。截止目前，OpenClaw已经公开曝出多个高中危漏洞，一旦这些漏洞被网络攻击者恶意利用，则可能导致系统被控、隐私信息和敏感数据泄露的严重后果。对于个人用户，可导致隐私数据（像照片、文档、聊天记录）、支付账户、API密钥等敏感信息遭窃取。对于金融、能源等关键行业，可导致核心业务数据、商业机密和代码仓库泄露，甚至会使整个业务系统陷入瘫痪，造成难以估量的损失。

https://mp.weixin.qq.com/s/4Ds8wa_iSgSvnNwH2DfaDw

AI Agent风险总结

权限失控：高危操作执行

默认以高权限运行，易被诱导执行危险系统命令（如 rm -rf /），可能导致系统崩溃或永久性损坏。

注入攻击：提示词注入攻击

网络攻击者通过网页中隐藏恶意指令，诱导OpenClaw读取该网页，导致其被诱导将用户系统密钥泄露。

网络逃逸：横向移动风险

可利用网络工具扫描、探测内部网络，成为攻击者横向移动的跳板，威胁整个局域网安全。

数据泄露：核心机密窃取

具备广泛的文件访问能力，可能被利用窃取API密钥、数据库密码等核心机密，造成商业损失。

企业AI Agent安全防护全景架构

光联携手华为，以“云、网、安”一体化架构为基础，为企业构建可视、可控、可管、可运营的AI安全运营服务体系，助力AI安全落地与规模化应用。

华为AI安全运营三层体系

智能运营：
AI对抗AI

可视 → 可控 → 可管

网络围栏
【边界NGFW】

第一层
受控环境与最小权限

切断与公网的一切联系，白名单访问公网资源、内网模型和必要服务，严格遵循最小权限原则，防止系统层面的越权操作。

Agent围栏
【Agent安全网关】

第二层
统一认证、授权、审计

所有流量必经Agent安全网关，实现统一的身份认证、权限控制和全链路审计，杜绝非法访问；模型围栏，过滤模型输入输出不合规内容及提示词注入攻击。

主机围栏
【AIDR】

第三层
Agent异常行为监管

通过部署主机及服务器智算EDR，实时监控和审计AI Agentw在终 端上的所有行为，及时发现异常。

华为AI Agent安全防护方案整体架构

FNET 光联世纪

AI Agent资产可视攻击面可视	AI Agent认证和访问控制	AI Agent威胁防御 (Skills带毒、恶意行为分析)	AI Agent行为可追溯可审计	AI Agent数据防泄密

安全运营	AI Agent资产和风险面可视 (Agent资产、模型、漏洞)	Agent安全态势	关联分析&AI分析	图DB溯源	SOAR响应

供应链&部署安全

- 供应链藏毒扫描
- 环境限定
(容器/VM/沙箱
/网络隔离)
- 安全配置检查
- 应用限定
(非关键)
- 模型限定
(内部模型)

Agent围栏 (Agent安全网关)

					数据窃取
提示词注入	外部输入消毒	高危/越权操作拦截	API异常访问检测	主题保持	内部敏感信息过滤
命令注入攻击检测	透明代理	工具调用异常校验		意图识别	

网络围栏 (NGFW)

资产可视和暴露面	Agent零信任	边界入侵	Skill/模型文件杀毒	恶意外联行为	数据窃取
AI Agent应用识别	NAC	入侵检测	网络病毒检测	AI C2检测（含加密流）	网络数据防泄密
AI Agent网络访问控制	远程ZTNA			恶意域名/URL过滤	

主机围栏 (AIDR)

资产可视和暴露面	Agent零信任	Skill/模型文件杀毒	关键目录&文件保护	权限控制	动态行为检测	数据窃取
AI Agent发现	身份认证	静态签名检测	Memory/会话保护	敏感进程启动	动态脚本行为	窃密木马检测
漏洞发现和管理	Posture检查	AI静态分类	配置保护	敏感文件操作	恶意指令执行	窃密工具检测
			凭证保护	网络访问	窃密/远控木测	敏感数据识别&防窃密
				系统调用	内存扫描	

E2E行为链审计（内部通信、进程链行为观测；意图偏离、关联&风险评估）

OS/虚拟机/容器

CPU 鲲鹏硬件加速

NPU/GPU 昇腾硬件加速

安全防护：实现AI Agent可视、可控、可管

方案价值

全面可视

全程可控

全链可管

AI 应用审计大屏

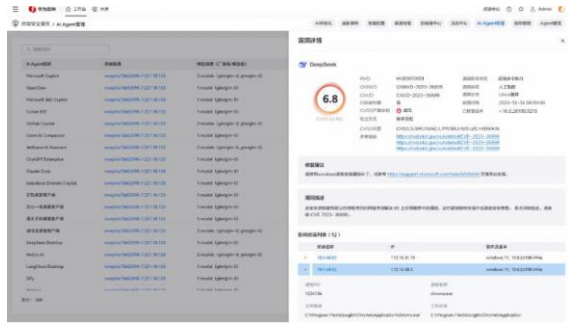
AI Agent资产和漏洞可视

AI资产可视

基本信息

AI Agent名称	Microsoft Copilot	安装包类型	Install.exe
安装路径	C:/Program Files	主程序	/usr/libexec/postfix/master
配置文件路径	/etc/docker/daemon.json	官网最新版本	v1.2.3
最近访问时间	2026-03-15 10:00:00	模型名称	ResNet50/Mistral
父进程	init	当前进程	process123
若干子进程	子进程1	进程状态	进行中
漏洞信息	3		

AI应用漏洞可视



光联一体化服务体系

光联不仅帮助客户“建好网络”，更持续帮助客户“用好网络、管优网络”。从建设到运营，全生命周期服务，成为客户值得信赖的长期合作伙伴

标准交付体系，保障项目高质量落地

光联通过标准化交付流程，专业团队保障，及资源保障能力，确保项目高质量落地，为长期稳定运行打下坚实基础。

标准化交付流程

01 需求调研
- 业务需求梳理
- 现状评估分析
- 项目目标明确

02 方案设计
- 方案架构设计
- 技术方案输出
- 实施计划明定

03 POC测试
- 方案可行性验证
- 性能与风险评估
- 优化完善方案

04 项目实施
- 设备部署与配置
- 系统集成与调试
- 进度与质量管控

05 项目验施
- 功能测试验证
- 性能测试验证
- 客户验收确认

06 完整交付
- 交付文档输出
- 知识转移培训
- 项目完整交付

全流程可视化管理，关键节点严格把控，确保交付质量与进度

交付能力保障

专业团队保障	认证工程师团队，丰富项目经验，多领域专家协同交付
标准化流程体系	基于最佳实践的标准化流程和模板，确保交付质量一致性
项目管理机制	PMO全过程管理，风险识别与管控，确保项目按时高质量交付
多厂商协同能力	深度合作生态，跨厂商方案集成能力，保障复杂项目顺利交付
文档与知识沉淀	规范文档输出，知识库沉淀，助力客户长期稳定运营

交付资源保障能力

219 个
全球PoP节点
- 全球资源统一调度
- 海外项目快速部署
- 多区域协同交付保障

支撑全球化项目交付能力

2100+
国内服务点
- 属地化实施能力
- 快速上门支持
- 提升交付效率

缩短项目交付周期

5000+
认证工程师
- 多厂商集成能力
- 复杂场景实施经验
- 全国协同交付

提升复杂项目成功率

16 个
国内备件库
- 快速备件响应
- 降低故障恢复时间
- 提升业务连续性

保障长期稳定运行

光联骨干网：全球覆盖+安全智能，支撑高质量交付

依托全球骨干网与安全能力，为交付提供稳定、快速、安全的网络保障，确保项目高质量落地与稳定运行。

全球骨干网络资源分布

- 全球 219+ PoP 节点
  覆盖五大洲核心区域
- 国内 10ms 海外 20ms
  低时延直连体验
- 99.99% 可用性保障
  高可用网络运行
- 多运营商资源
  骨干直连，灵活调度 最优路径保障

● 核心 PoP 节点 ● 扩展 PoP 节点

中性网络 | 自有管理 | 全球互联 | 快速调度 | 灵活接入

骨干网安全能力

接入安全传输

SD-WAN、IPv6、SRv6、IPsec 等多安全传输技术，支持端到端加密与动态路由优化，保障传输安全与稳定。

智能威胁感知与流量分析

整合全球流量监测与智能分析能力，实时识别异常行为，结合 AI 与威胁情报，精准发现并阻断潜在风险。

安全服务体系

统一堡垒机、AAA认证、身份访问管理、终端可信等能力，提供“权限+审计”一体化安全运营。

全球协同防御

依托全球安全防护能力与 DDoS 防御资源，覆盖全球网络边缘，提供多层防护，确保业务连续性。

- 120+ 国家覆盖
- 330+ anycast 节点
- 377Tbps+ DDoS 防护容量

为交付体系提供坚实保障（五大维度）

全球快速部署

- 覆盖 73 级别以上机房
- 多运营商接入
- 最快响应交付

弹性灵活扩展

- 网络弹性带宽扩展
- 云资源弹性调度
- 按需扩展，快速响应

网络与专线能力

- 代理运营商资源（专线/猫打/裸光纤）
- 最后一公里接入资源丰富
- 代理国际出口（合规）

云与 IDC 能力

- 多云生态互联（华为云、阿里云等）
- 自建 / 合作 IDC 资源中心
- 云上专线接入与云间加速

交付伙伴能力

- 全国驻场专业交付团队
- 国内外服务商协同合作
- 认证工程师与专业实施团队保障

智能运维体系，主动保障业务稳定运行

光联依托7x24专业运维体系，通过NOC网络运维、SOC安全运营、COC云托管运维协同联动，实现从监控、分析、响应到优化的闭环保障，为客户业务稳定运行保驾护航。

NOC 网络运维

- 网络质量监控
- 链路状态监控
- 性能保障
- 故障响应
- 巡检与健康检查

保障网络稳定运行

SOC 安全运营

- 威胁检测
- 安全告警分析
- 安全事件处置
- 风险发现与防护
- 安全策略优化

持续降低安全风险

COC 云托管运维

- 云资源监控
- 云资源管理
- 成本优化
- 平台保障
- 容量规划

提升云资源使用效率

AI运维（AIOps）智能运营

- 智能告警降噪
- 故障根因分析
- 异常预测预警
- 智能决策与处置
- 运维知识沉淀

提升运维效率与质量

智能运维能力支撑

统一运维门户（Portal）
集中展示，便捷访问，运营可视化

IT服务管理（ITSM）
标准流程，自动派单，服务全流程管理

AI智能助手（AI Agent）
智能分析，辅助决策，提升运维效率

知识库与经验库
沉淀经验，快速复用，持续提升运维能力

专业团队 + 标准流程 + 智能平台

7×24小时专业服务

7×24小时	5000+	>98%	>99.95%
不间断运维	认证工程师	问题及时处理率	服务可用性保障

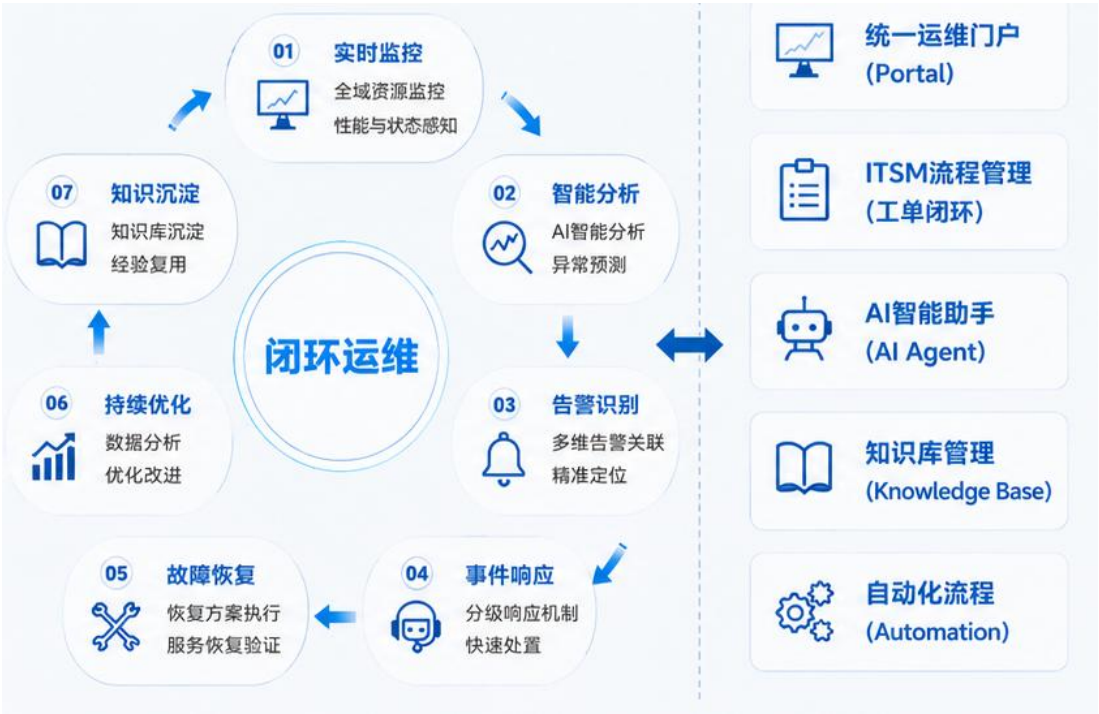
光联以AI运维能力（AIOps）与7×24持续运营能力为客户提供更智能、更高效、更稳定的运维保障，让网络、安全与云持续稳定支撑业务发展。

运维服务中心，让服务真正落地

光联通过AIOps能力、标准化流程与7×24持续运维机制，构建智能运维闭环能力，持续保障客户网络、安全与云稳定运行。

NSOC网安融合运维中心

标准运维流程体系



FNET 光联世纪

THANK YOU

感谢聆听 期待合作

关注微信公众号

关注视频号