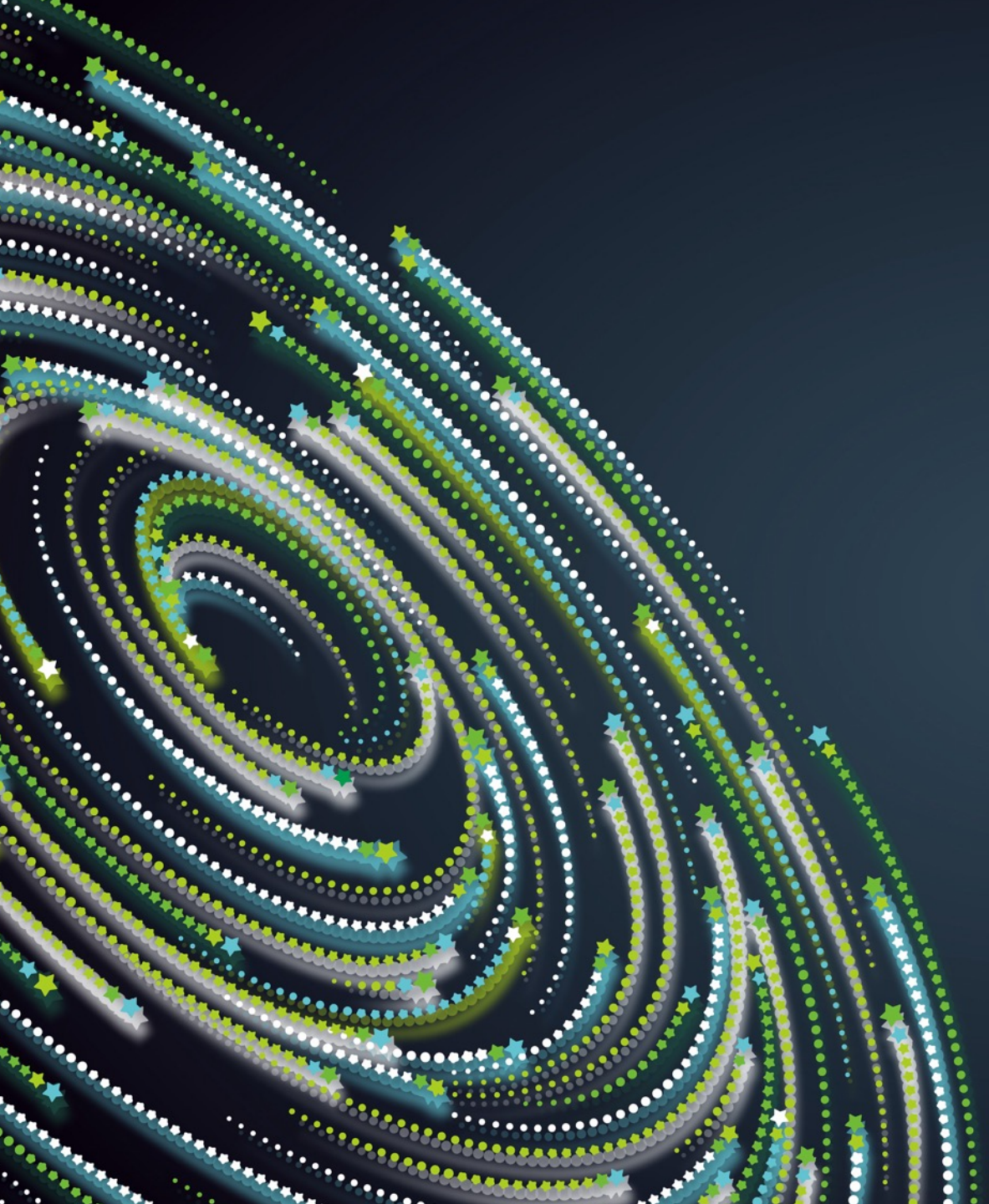
信息·趋势·感悟
THE POWER OF COMMUNITY STARTS WITH YOU
美国2026RSAC热点研讨
暨第十八届信息安全高级论坛
INFORMATION SECURITY FORUM 2026



从大模型到智能体

网络安全产品趋势与思考

绿盟科技 宫智

2025 vs 2026

RSAC 2026

RSAC 2025

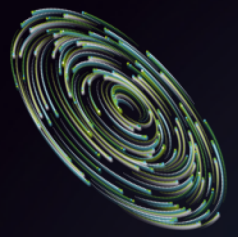
从“AI 辅助”到“AI 自主”、从“模型安全”到“智能体安全”

产品创新风向标：RSAC创新沙盒

2026 Geordie AI
智能体安全
2025 ProjectDiscovery
ASM / 漏管
2024 Reality Defender AI
深度伪造检测与治理
2023 HiddenLayer AI
模型安全
2022 Talon Cyber Security
零信任/安全浏览器
2021 Apiiro
应用安全/代码风险治理
2020 Securiti.ai
数据隐私合规

2019 Axonius
资产安全管理 / ASM
2018 BigID
数据安全治理 / 数据隐私
2017 UnifyID
身份认证 / 行为生物识别
2016 Phantom
安全编排自动化 SOAR
2015 Waratek
运行时应用保护 RASP
2014 RedOwl Analytics
用户行为分析 UEBA / 内部威胁
2013 Remotium
移动安全 / 虚拟移动基础设施
2012 Appthority
移动应用安全

2011 Invincea
高级威胁防护 / 沙箱隔离
2010 Altor Networks
虚拟化安全 / 云网络安全
2009 AlertEnterprise
工业安全 / 物理 - 逻辑安全融合
2008 Reconnex
数据泄漏防护 DLP
2007 Yoggie Security Systems
硬件安全 / 终端安全
2006 Imperva
WAF / 数据库安全
2005 Sourcefire
入侵防御 / 网络安全



创新沙盒冠军纵览（2005–2026）

2005-2010

边界安全

边界防护、硬件安全
数据库安全、DLP

2011-2015

应用与终端

移动安全、UEBA
高级威胁、内部威胁

2016-2019

安全治理

SOAR、数据安全
身份认证、资产治理

2020-2022

零信任合规

零信任访问、隐私合规
云原生安全、远程办公

2023-2026

AI原生安全

AI模型安全、深度伪造
攻击面管理(ASM)

防护对象升级

终端/网络 → 数据 →**AI模型/智能体**

防御范式演进

被动防御 → 零信任 →**智能自治**

技术风格变革

单点产品 → 平台化 →**AI原生自动化**

趋势变化：安全重心从 “端边” 转向 “数据” ，26年强化 “身份” 与 “自主” 。这标志着网络安全从**被动防御**向**主动**再到**智能**的范式转移，Agent 不再只是辅助工具，而是能够自主决策、自主响应的安全主体

RSAC 2026：AI 进化的分水岭

自主响应

Autonomous

减少人机交互，增加逻辑闭环。AI Agent 能够独立完成威胁检测、分析、响应的全流程，无需人工介入即可实现秒级处置。

量子准备

PQC

面对"现在截获，未来破解"的防御升级。采用晶格密码等后量子加密技术，为量子计算时代的安全威胁提前布局。

非人身份

NHI

针对 Agent 和机器账号的治理。企业需要为数以万计的 AI Agent 建立身份管理体系，防止越权操作和横向移动。

Agentic Era

智能体时代正式开启

Agentic SOC：从 Copilot 到 Agent

01

Chatbot

对话框

- 被动响应用户查询
- 基于规则的知识问答
- 无自主行动能力

02

Copilot

建议

- 主动提供分析建议
- 辅助决策但不执行
- 需要人工确认操作

03

Agent

执行

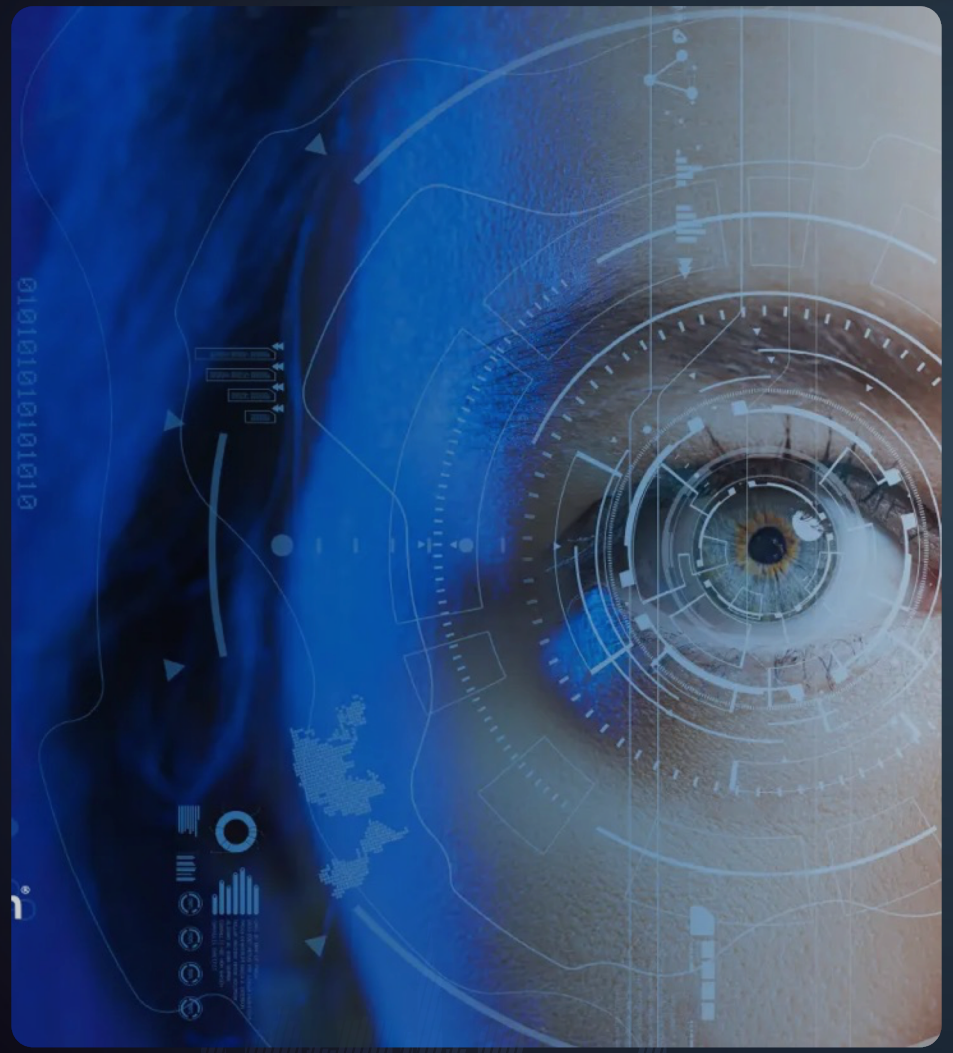
- 自主执行调查流程
- 自动封禁与修复
- 闭环响应无需人工

技术细节与核心突破

2026 年的 SOC 不再是对话助手，而是能自主执行调查、封禁、修复的数字安全分析师。

核心突破：决策逻辑的可解释性 (Explainable AI)

非人身份治理 (NHI)：治理"影子智能体"



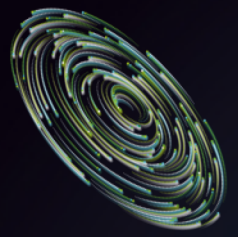
核心挑战

企业内部存在数以万计的 AI Agent，它们像"影子智能体"一样在系统中运行，缺乏统一的身份管理和权限控制。

解决方案

为每个 AI Agent 颁发“数字 DNA”，建立完整的身份生命周期管理，确保智能体不能绕过权限约束。

- 身份认证
- 权限管控
- 行为审计



商业巨头：平台化整合与生态锁定

营收规模50亿美元以上 | 增长率15%-25% | 平台化战略主导

业务模式：平台化战略

全栈整合
构建端到端统一平台，覆盖网络、端点、云、身份全层级

订阅转型
SECaaS模式深化，订阅收入占比超70%，增强收入可预测性

生态开放
Marketplace战略，700+第三方集成，构建网络效应

产品组合特点

Palo Alto Networks 三平台架构
STRATA（网络边界）+ PRISMA（云安全）+ CORTEX（检测响应）
通过共享威胁情报形成有机整体

Fortinet Security Fabric 生态系统
自研ASIC安全处理器，整合防火墙、SD-WAN、端点安全等，统一FortiOS操作系统管理

托管检测与响应服务向"AI主导、专家监督"的自主安全运营转变

技术优势

AI/ML深度集成
统一于AI驱动架构

硬件创新壁垒
Fortinet自研ASIC芯片

云原生安全
整合云安全完整生命周期

未来发展方向

01 **生成式AI自主代理：** 从"AI辅助分析师"向"AI自主代理"跃迁

02 **量子安全密码学：** NIST PQC标准产品化加速，IBM、Palo Alto布局

03 **边缘计算安全：** 5G安全、OT安全场景验证向规模化部署演进

代表性厂商

- CISCO（Splunk）
- Palo Alto Networks
- IBM Security
- Fortinet

高增长挑战者：云原生与数据驱动

营收规模10-50亿美元 | 增长率30+%

业务模式：云原生架构

云原生设计
微服务、容器化、DevOps持续交付，弹性扩展与快速迭代

AI-First
机器学习嵌入产品架构底层，行为检测替代签名，自适应进化

数据网络效应
全球遥测汇聚，Crowdsourced情报，客户越多→数据越多→检测越强

产品特点：架构创新

单代理架构
CrowdStrike Falcon轻量级代理(<100MB内存，<1% CPU)，所有功能云端动态加载

自动化响应
SentinelOne ActiveEDR毫秒级自动隔离、终止进程、回滚文件，勒索软件零损失

全球分布式
Zscaler 150+ PoP节点，零硬件纯云服务，消除VPN回传延迟与单点故障

实时威胁图谱
Threat Graph日处理数万亿事件，图计算关联分析，一次检测全球免疫

细分赛道领导地位

CrowdStrike	端点安全领导者 | 29% YoY
Zscaler	SASE先驱 | 35%+ YoY
SentinelOne	AI自主安全 | 40%+ YoY
Wiz	云安全领导者 | 20%+ YoY

发展趋势

- **自主安全运营**
  AI代理独立完成威胁狩猎、漏洞修复
- **AI代理安全**
  治理AI代理构建、访问与行为，Noma、Zenity布局
- **运行时优先**
  从静态配置分析转向实时行为检测，Upwind 95%误报降低

垂直领域专家：深耕与专业化壁垒

营收区间1-10亿美元 | 增长率50%-100%+

业务模式：极致专业化

身份安全专家

Okta： 7500+应用集成，开发者体验优先

CyberArk： 特权访问管理领导者，DevOps集成

云安全专家

Upwind： 运行时优先，95%误报降低，4000%收入增长

Wiz： Google 320亿美元收购，CNAPP领导者

邮件安全专家

Abnormal Security： 云邮件安全独角兽，545%搜索增长

竞争策略

云原生架构 + 快速部署 + 轻量级集成 + 体验

- 云原生
- 快速部署
- 弹性扩展
- 体验优先

产品特点：专业壁垒

领域深度

在特定安全领域建立不可替代的专业能力，如Mimic的勒索软件内核级防御

开箱即用

轻量级部署，分钟级上线，降低客户采用门槛，如Tailscale的mesh VPN、Tines的无代码自动化

无缝集成

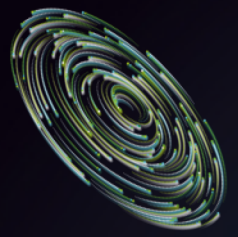
与主流平台深度集成，如Okta的7500+应用连接器、Snyk的IDE/CI/CD原生嵌入

发展趋势

AI安全治理： AITRiSM品类形成，治理AI代理全生命周期

深度伪造检测： Reality Defender检测音频/视频/图像/文本深度伪造，RSA创新奖

供应链安全： Chainguard零信任不可变基础设施，1700+可信容器镜像



新锐创新者：前沿首创与快速验证

营收规模5000万美元以下 | 超100%爆发式增长 |

业务模式特征

前沿技术场景首创性
定义和创造新市场，在成熟厂商响应前建立先发优势

轻量级部署与快速价值实现
云原生SaaS交付、无代理架构、预配置策略模板

新兴攻击向量专项防护
填补空白定位，避开成熟厂商直接竞争

发展路径选择

技术验证→规模化转化
从POC到生产环境，建立可重复销售流程

战略融资与生态合作
顶级风投、云厂商战略投资、平台生态集成

并购退出或独立IPO
以色列企业偏向并购，Wiz案例激励独立发展

前沿产品创新

Zenity AI代理安全 全球首家
意图感知分析，追踪代理完整执行路径，识别恶意或意外后果

Zama 同态加密 隐私计算
全同态加密商业化，使开发者能在不暴露数据情况下构建隐私保护应用

量子安全探索 前沿技术
后量子密码学（PQC）、量子密钥分发（QKD）应对量子计算威胁

产品规划启示：几个路径的战略选择

商业巨头

核心策略

平台整合与生态开放，通过并购快速补强能力短板

关键能力

全栈产品矩阵、统一管理编排、全球服务网络

适用场景

大型企业全生命周期安全、复杂混合环境统一治理

挑战者

核心策略

云原生架构与AI-first创新，数据网络效应构建护城河

关键能力

轻量级部署、自动化响应、实时威胁情报

适用场景

云优先战略企业、分布式办公环境、快速数字化转型

专家

核心策略

细分赛道深耕与专业化壁垒，开发者优先体验设计

关键能力

领域深度、开箱即用、无缝集成

适用场景

垂直领域需求、与现有安全栈互补

新锐

核心策略

在前沿领域探索突破，定义新市场机会

关键能力

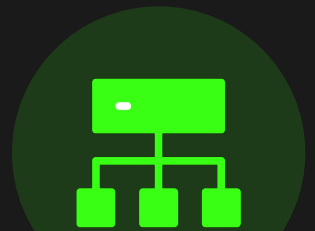
技术首创性

适用场景

特定安全痛点解决、快速价值实现

未来趋势：AI原生安全、零信任架构、运行时检测将成为三大路径的共同演进方向。

时代变迁：技术、需求与威胁的演进



阶段一：信息化时代

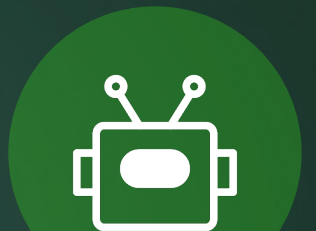
Web 1.0 / 2.0 时代

- **技术：** 静态网页、UGC、移动互联
- **需求：** 基础连接、信息发布合规
- **威胁：** 病毒、蠕虫、DDoS、网页篡改

阶段二：数字化时代

Web 3.0 时代

- **技术：** 云计算、大数据、IoT、移动优先
- **需求：** 业务上云、数据价值挖掘
- **威胁：** APT攻击、数据泄露、供应链攻击



阶段三：智能化时代

Web 4.0 时代

- **技术：** 生成式AI、数字孪生、元宇宙
- **需求：** AI赋能业务、极致体验、数据主权
- **威胁：** AI攻击、深度伪造、模型投毒

核心洞察：安全防护能力必须与技术演进速度同步升级，从被动防御转向主动智能防御

理念演进：从被动防御到主动韧性

边界防御 (Perimeter)

核心： 城堡护城河模型，构建坚固边界抵御外部入侵。
时代： 网络边界清晰
局限： 无力应对内部威胁及云时代边界模糊问题。

零信任 (Zero Trust)

核心： 永不信任，始终验证。默认环境不安全，逐次鉴权。
时代： 云原生/移动化
局限： 高度依赖身份体系，难防未知AI攻击。

韧性安全 (Resilience)

核心： 接受风险，快速恢复。构建能吸收冲击的主动防御体系。
时代： 超连接/智能化
突破： 从“防得住”转向“打不垮、恢复快”。

安全思维跃迁：被动防御 → 动态验证 → 主动适应与恢复

产品迭代：从工具到平台再到智能

01 传统安全工具

代表：杀毒软件 / 防火墙 / IDS

功能单一，基于特征库被动响应，防护范围有限

模式：软件授权 / 硬件买断

02 平台化安全产品

代表：SIEM / SOAR / XDR

多源数据整合，统一安全视图，支持自动化响应流程

模式：订阅制(SaaS) / 平台服务费

03 AI原生安全产品

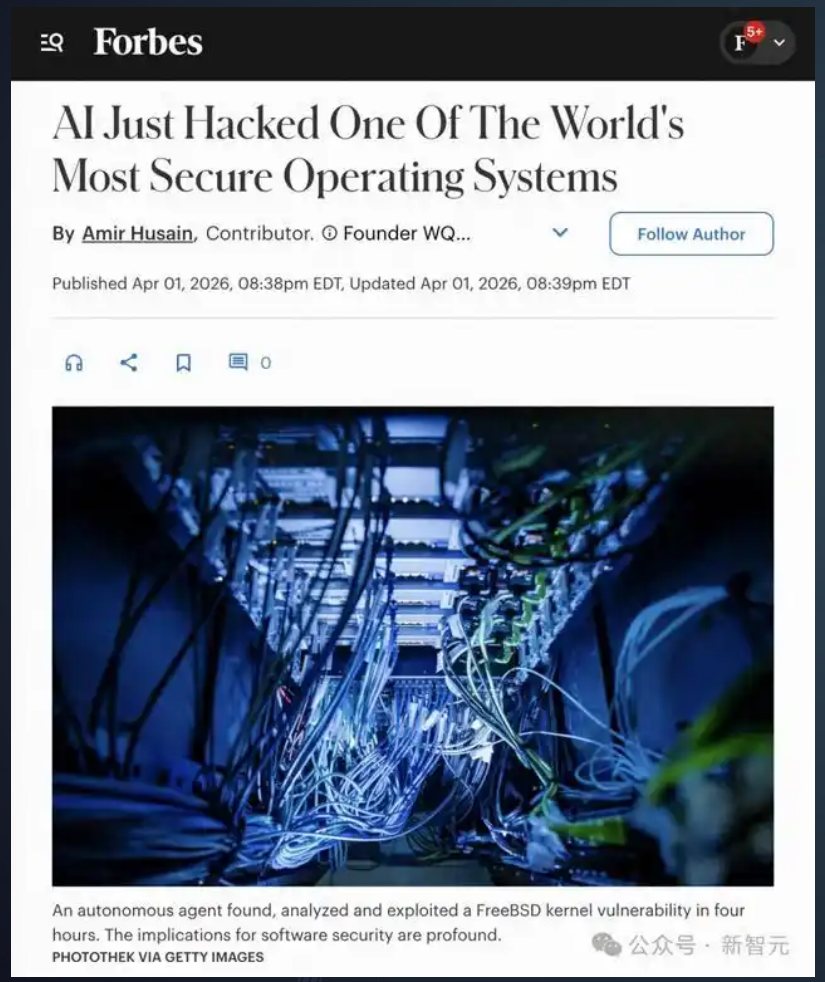
代表：AI Copilot / Autonomous SOC

AI驱动主动预测，自主响应威胁，实现深度人机协同

模式：API调用 / 成果付费 / 增值服务

核心洞察：产品演进不仅是技术升级，更是商业模式从“一次性”向“持续价值”的变革。

从旧金山到北京



Forbes

AI Just Hacked One Of The World's Most Secure Operating Systems

By Amir Husain, Contributor. Founder WQ...

Follow Author

Published Apr 01, 2026, 08:38pm EDT, Updated Apr 01, 2026, 08:39pm EDT

An autonomous agent found, analyzed and exploited a FreeBSD kernel vulnerability in four hours. The implications for software security are profound.

PHOTOTHEK VIA GETTY IMAGES

图片来自 公众号 新智元

从"集成"到"融合" (Native AI)

传统集成

通过外部 API 调用 AI 能力，AI 作为外挂模块存在，响应延迟高，无法实时阻断。

- ✕ 高延迟，无法实时响应
- ✕ 依赖外部服务稳定性
- ✕ 数据需要出域传输

原生融合

在网关和终端部署针对安全微调的 SLM (小语言模型)，实现 "In-line AI" 实时阻断。

- ✓ 毫秒级响应，实时阻断
- ✓ 本地部署，数据不出域
- ✓ 深度集成，无缝体验

齿轮完全咬合

Native AI 架构示意

安全网关内置 SLM

↑↓

终端 Agent 本地推理

↑↓

实时威胁检测与阻断

构建数字人团队，从人工驱动到智能自治

技术进阶：从固化流程到灵活编排，基于多智能体平台打造数字人团队，实现按需编排、即时组装的灵活性

主流智能体局限

本质问题： 现有安全智能体本质是「固化流程的自动化」，嵌入现有平台后难以适配不同行业的差异化需求。

场景局限： 电信与金融等不同行业的资产管理差异巨大，流程一旦改变，底层代码调整极其困难。

灵活性缺失： 无法灵活组合各种手段（平台能力、探针能力、自编写能力）完成业务场景诉求。

运营专家定义多智能体协作

运营专家可根据不同业务环境和客户场景， **定义各种Agent角色** ，并为这些Agent精准适配、挂载专属调用的安全工具或API接口以及其协同方式。

按需编排
即时组装能力

异构兼容
多厂商多平台

数字人团队架构

安全运营经理
架构的"大脑"
接收用户需求，拆解安全运营任务，精准调度麾下专业智能体，把控全局进度

威胁狩猎专家
主动出击，收集情报，深挖高隐蔽性威胁

分析研判工程师
过滤噪音，精准锁定高风险行为

暴露面评估师
监控资产变化，确保持续收敛

流程文档工程师
报告生成与归档，SLA规则检查

底层能力： 灵活调用异构全域探针、设备与平台（各种探针、各种平台、客户CMDB、ASM系统等）

实践效果：六维安全运营自治体系

以某单位数据库暴露面治理为例，展示数字人完整闭环能力

案例背景与痛点

某单位**数据库暴露面极大**，挖矿、勒索组织时刻盯着；爆破、未授权入侵后植入动作经常变化，躲避规则检测。

痛点： 人分析日志→扩展调查→写规则→受害地址持续验证监控，跨平台跨业务部门工作繁杂，精力难以聚焦琐碎事务。

六维自治能力矩阵

- **自动化运营** — 释放基础人效
- **自动化扩展** — 灵活挂载新设备
- **自主收敛** — 资产风险持续降级
- **自主报告** — 脱离人工撰写
- **工作流闭环** — SLA强制流转
- **持续迭代升级** — 规则库越用越强

五步闭环流程

1. **检查数据库操作日志**
   自动接入并全面检索数据库海量操作记录
2. **分析研判发现异常和风险**
   过滤噪音，精准锁定提权或敏感数据违规读取
3. **全面狩猎分析**
   追溯攻击源头，勾勒完整攻击链路与手法特点
4. **输出威胁检测规则**
   自动将狩猎经验转化为可机读、可下发的检测规则
5. **暴露面持续监控**
   针对攻击暴露的薄弱点，进行长期、周期性自动化监控

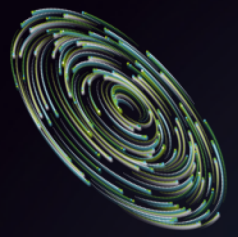
岗位任务自动拆解

威胁狩猎数字人
- ✓ 漏洞验证代码自主编写
- ✓ 检测规则自主编写
- ✓ 反测绘对抗自主绕过

分析研判数字人
- ✓ 任务自主规划
- ✓ 攻击链还原
- ✓ 关键证据提取

暴露面评估数字人
- ✓ 自主规则编写
- ✓ 自动添加监控
- ✓ 研判结论定性

流程工单数字人
- ✓ 定期工单跟踪
- ✓ 督促数字人工作
- ✓ SLA合规检查



数字团队重塑安全运营

自主可控

全面信创兼容
支持国产芯片与操作系统

开放兼容

多模型协同，异构设备灵活接入

持续进化

知识库越用越强
能力持续迭代升级

构建面向未来的**智能安全运营体系**

- 7×24H自主值守
- 全流程闭环
- 效率提升60%+

95%	91%	45%	60%+
AI降噪率	AI研判准确率	AI自主处置率	AI综合提效

AI辅助到AI主战 | 端到端自动化 | 数字工程师团队 | 全面信创兼容

AI智能体时代:机遇与风险并存

信息·趋势·感悟
THE POWER OF COMMUNITY STARTS WITH YOU
美国2026RSAC热点研讨
暨第十八届信息安全高级论坛
INFORMATION SECURITY FORUM 2026

Agentic AI 演进

AI正从**生成式应用**向**智能体形态**演进。2025年被称为"智能体元年",多模态智能体实现跨模态推理与决策,具备自主进化能力。

2023	2025	2026
生成式AI元年	智能体元年	系统级自主化

安全风险加剧

- 3天内爆发**2个高危RCE漏洞**
- 频繁更名导致**供应链投毒攻击**
- 大量资产**缺失有效防护**

OpenClaw 现象级爆发

典型案例

GitHub星标增长
27.2万
史上星标最多的非聚合类项目

全网资产规模
20万+
从1月底5.3K快速增长

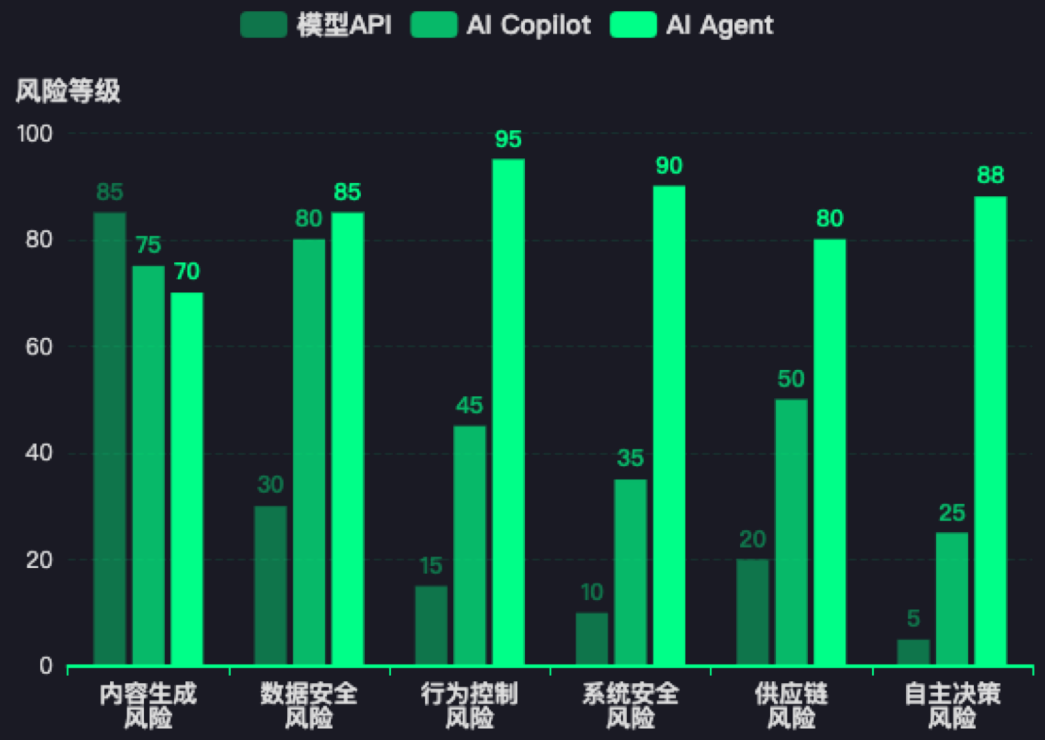
中国部署情况

全球部署排名	Top 1
主要云厂商	阿里云
关基机构暴露	多家境内机构

核心洞察: 智能体技术正在快速普及,但安全防护措施严重滞后,亟需建立系统化的安全管控体系。

AI安全风险演进:攻击面持续扩大

从模型API到AI Agent,风险的广度和深度显著增加



01 模型API阶段

AI以开放API形式提供基础能力,风险主要围绕模型内容的生成与控制

模型幻觉 | 非合规输出 | 提示词注入

02 AI Copilot阶段

AI深度集成到软件平台,能够接触用户数据和系统环境

数据泄露 | 隐私风险 | RAG投毒

03 AI Agent阶段

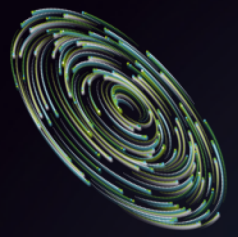
Agent具备自主规划、决策和执行能力,风险扩展到行为和系统控制层面

恶意工具调用 | 自主行为失控 | 系统提示词泄露

风险演进趋势

随着AI自主性的增强,攻击面从内容层扩展到数据层、行为层和系统层,呈现指数级增长态势

↑ 300%+



一体化智能体安全防护体系

从内容安全到行为安全,从静态评估到动态控制

信息·趋势·感悟
THE POWER OF COMMUNITY STARTS WITH YOU
美国2026RSAC热点研讨
暨第十八届信息安全高级论坛
INFORMATION SECURITY FORUM 2026

三层防护架构

01 资产与风险识别层
AI资产发现 · 空间测绘 · 风险识别

02 Skills安全测评层
静态分析 · 动态检测 · 权限控制

03 行为监测与控制层
输入输出围栏 · 工具审计 · Token管控

治理理念转变

- 内容安全 → **行为安全**
- 静态评估 → **动态控制**
- 外挂防护 → **内生治理**
- 单点防护 → **体系协同**

AI资产与风险识别

资产主动发现
网络空间测绘与AI资产指纹识别

空间测绘能力
全网分布热力图到具体IP定位

核心价值: 消除影子AI资产盲区,让风险无处遁形

Skills全方位测评

静态分析
识别恶意代码模式

动态检测
沙箱隔离环境运行

权限控制
最小化权限原则

核心理念: 从"盲目信任来源"走向"严格验证行为"

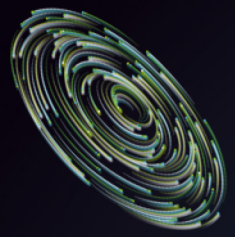
行为监测与控制

输入输出围栏

工具调用审计

Token资源监测

异常行为熔断



构建可信AI生态，守护智能体安全

资产发现与测绘

消除影子AI盲区
全网资产可视

Skills全方位测评

动静态结合检测
权限最小化控制

行为监测与控制

实时异常检测
毫秒级响应熔断

Token资源管控

算力消耗监测
防止恶意滥用

统一闭环:检测 · 测绘 · 管控

将游离于管控之外的碎片化智能体,转化为**集中治理、动态可控**的安全AI生态圈。实现从"内容审查"向"智能体行为管控"的跨越,确立纵深防御与实时响应架构。

供应链安全的升维：从 SBOM 到 ABOM

SBOM
Software Bill of Materials

传统软件物料清单，列出代码依赖和开源组件。

- 源代码依赖
- 开源组件版本

ABOM
AI Bill of Materials

必须列出 AI 基座版本、数据合规证明及 RLHF 安全对齐记录。

- AI 基座版本
- 训练数据来源
- RLHF 安全对齐

2026趋势展望：AI-Native与集体智能

01. AI-Native 架构主流化

AI成为安全底层基础设施，产品具备自主学习与自适应能力，从“附加功能”转为“原生基因”。

02. 攻防对抗白热化

攻击方利用AI自动化挖掘漏洞，防御方引入AI红队进行实时对抗，“以AI制AI”成为关键策略。

03. 数据安全成核心竞争力

数据主权意识觉醒，隐私计算与数据脱敏技术广泛应用，安全能力直接决定企业数据资产价值。

04. 安全即服务 (SECaaS)

安全服务场景化、智能化按需提供，企业运营模式从“自建自管”转向“订阅外包”，效率大幅提升。

产品发展思路与建议

价值驱动转型

围绕客户业务价值设计，解决实际痛点；提供场景化方案，拒绝功能堆砌。

构建开放生态

开放API赋能合作伙伴，共建解决方案；积极参与行业标准制定，引领生态共赢。

极致体验升级

简化部署与操作流程，降低门槛；打造可视化态势感知与智能报表，直观易用。

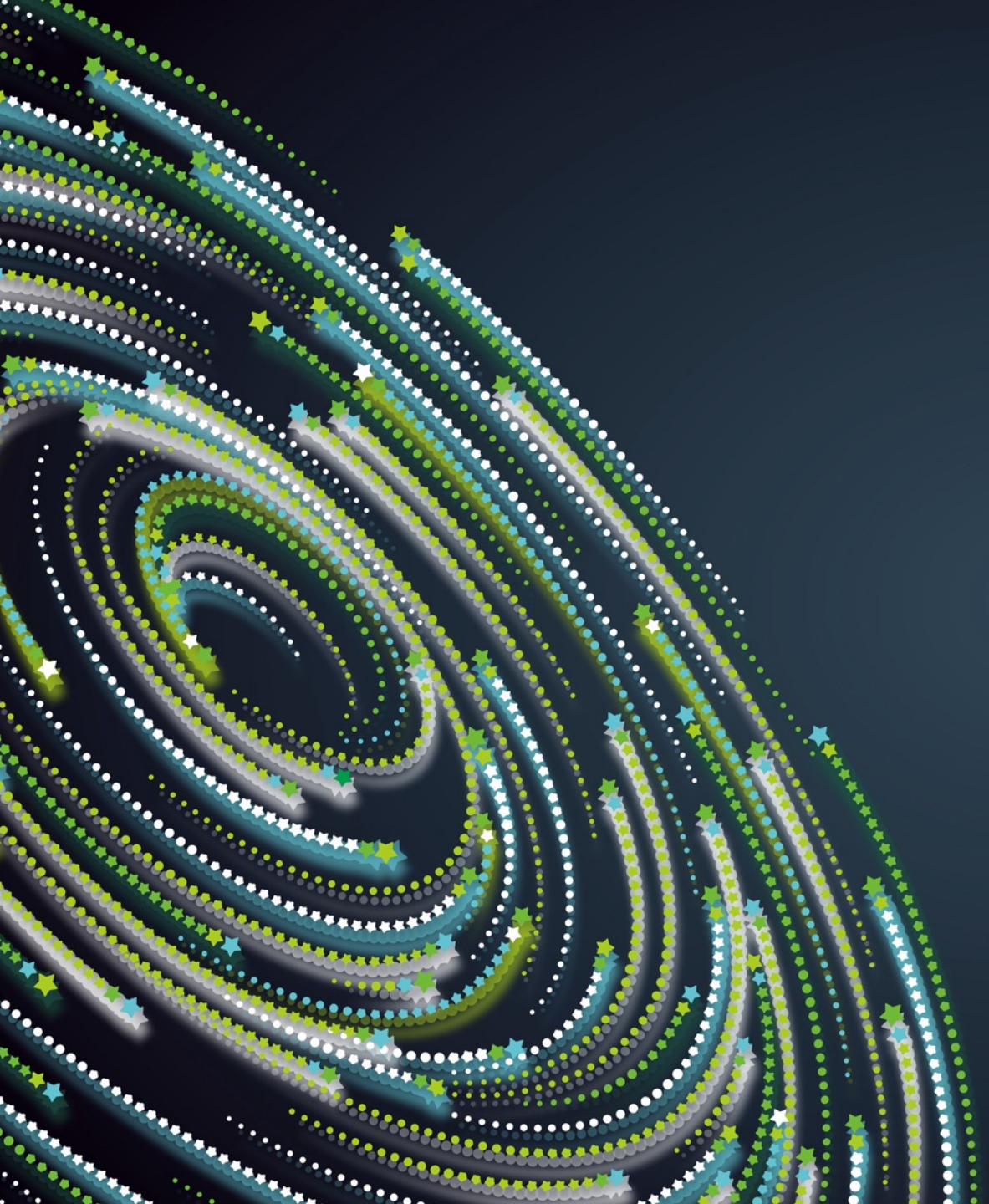
信息·趋势·感悟

THE POWER OF COMMUNITY STARTS WITH YOU

美国2026RSAC热点研讨

暨第十八届信息安全高级论坛

INFORMATION SECURITY FORUM 2026



智能体时代
安全不是阻碍速度的摩擦力
而是
保障业务疾驰的推进器