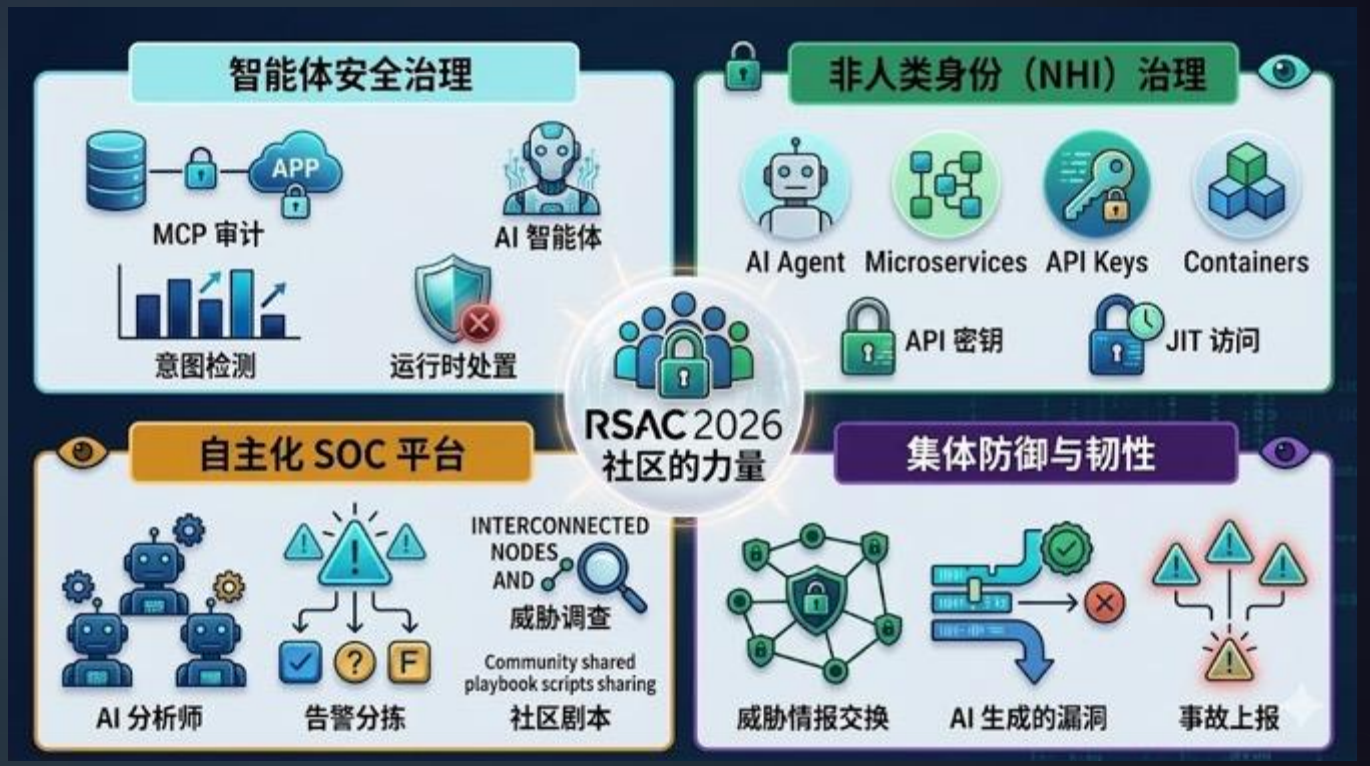
从RSAC看网络安全技术和行业发展

绿盟科技 叶晓虎

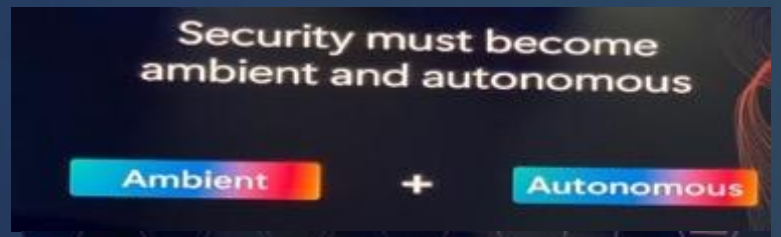
Power of Community

35 RSAC Conference



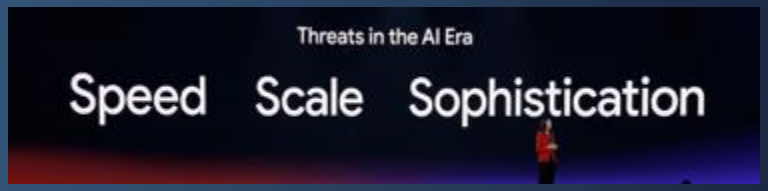
Opening Keynote

Microsoft 微软



cisco 思科

Google 谷歌



从参展商看行业的变化

- 本届大会共有约 **649 家** 参展商。相比 2024 年的 600 家和 2025 年的 650 家左右，展商规模保持稳定，但**初创企业**的更迭率极高；
- 有 **261 家** 是在 2020 年以后成立的，这意味着近 40% 的展商是疫情后崛起的新兴势力。

核心分类	展商数量 (约)	占比	行业观察
威胁检测与情报 (TDR/TI)	117	18%	依然是基本盘，但全部集成了 AI 分析。
AI 安全 (Security for AI)	115	17.7%	增长最快，专注 LLM 防护和 Agent 治理。
安全运营 (SecOps)	101	15.5%	重点在于“自主化 SOC”和自动化响应。
数据安全 (Data Security)	95	14.6%	侧重于 AI 模型训练中的隐私保护（DSPM）。
身份与访问管理 (IAM)	91	14.0%	核心议题转向“非人类身份 (NHI)”管理。
应用安全 (AppSec)	83	12.8%	关注 AI 生成代码的安全审计。
云安全 (Cloud Security)	82	12.6%	关注云原生环境下的 AI 基础设施安全。
其他 (GRC/网络/端点等)	<60	-	传统品类增长乏力，多向 AI 概念靠拢。

不同区域的行业组团


ITALIA
ITALIA


Germany
Germany


KOREA
KICA
KOREA
KICA


SINGAPORE
ALICE
ABSOLUTE


NSFOCUS
AUTONOMOUS SECURITY OPERATIONS
24*7 ASOC
Innovation
ERED ENGINE
Detection Agent
Smart Triage Agent
Investigation Agent
Threat Hunting Agent
Response Agent

不同的发展策略

互联网巨头

Microsoft cisco

构建端到端生态体系

核心优势足生态整合与全栈技术覆盖，技术路线聚焦"身份-生态-运维"的深度协同。

初创阵营

Geordie

聚焦AI安全细分场景创新

填补巨头生态短板，代表资本重点流入的新兴技术基座。

安全厂商

paloalto NETWORKS CROWDSTRIKE

将Agent安全无缝融入现有防御栈

加速向平台模式靠拢，聚焦"防护-运营-治理"一体化，争夺智能体时代的SOC主导权。

Innovation Sandbox

冠军

Geordie AI
AI 智能体实时治理与意图修复

Token Security
针对非人身份 (NHI) 的权限与合规管理

Humanix
针对语音/文字冒充的自然语言威胁检测

Charm Security
AI 模拟反诈，拦截复杂的社交工程攻击

治理/身份

研发安全

人为因素

基础设施

ZeroPath
AI 原生漏洞自动修复，取代传统扫描

Clearly AI
自动化威胁建模与安全设计评审

Glide Identity
去密码化、设备级信任的身份协议

Crash Override
工程供应链追踪与 SLSA 自动化合规

Fig Security
自动化修复 SecOps 流程的逻辑失效点

Realm Labs
AI 推理透明度监控与决策风险审计

一些关键技术

01 Beam 引擎

技术原理：注入式治理 (Context Injection)。不同于暴力阻断（Kill Connection），Beam 引擎在**毫秒级** 内向 AI 上下文窗口注入"确定性约束"。

业务不中断

引导 AI "自我修正"，而非强行关机

自愈性

让安全成为 AI 业务流的原生组成部分

02 足迹捕捉

全方位可视化识别

L1 静态层

扫描代码库，发现隐藏的 Agent 框架（如 LangChain）

L2 流量层

识别 API 行为指纹，区分"人类"与"机器人"流量

L3 端点层

监控浏览器插件及本地运行的自主 AI 实例

03 IBAC 技术

从 RBAC 演进到 IBAC

RBAC："它有权限吗？"

（AI 账号通常权限极高）

↓

IBAC："它应该这么做吗？"

实时对比 Agent 当前操作与初始任务目标的匹配度

场景示例：拦截试图访问工资表的"翻译智能体"

智能体时代安全的原则

行业坐标

PARADIGM SHIFT

2026年是"智能体安全元年"，安全重心发生根本性转变：从传统的"防护模型"转向"治理行为"。这标志着AI安全进入了一个全新的发展阶段。

核心挑战

BUSINESS BOTTLENECK

AI Agents拥有高特权，能够自主执行复杂任务。传统的拦截式安全机制已成为业务发展的瓶颈，严重制约了AI生产力的释放。

冠军启示

NEW ERA BEGINS

Geordie AI的夺冠标志着"意图驱动"与"低摩擦"安全时代的到来。这一创新方案证明了安全可以与业务高效协同，而非相互制约。

安全范式转变

旧	→	新
防护模型 基于规则的特征匹配 被动防御		**治理行为** 基于意图的实时监控 主动引导

关键洞察

- 传统安全机制无法理解AI的**推理逻辑**，导致误报严重
- 企业需要**“无感安全”**产品，释放AI生产力
- **意图驱动**成为AI安全的新金标准

核心结论： 在AI时代，最好的安全是让业务跑得更快，而不是拉紧手刹。

智能体摩擦力: Agentic Friction

定义：安全机制对自动化流程的干扰程度

高摩擦

传统方案

频繁报错

安全系统过度敏感，正常业务被频繁拦截

手动干预

需要人工审核和放行，严重影响效率

生产力损失

抵消 AI 生产力红利，安全成为业务阻力

低摩擦

Geordie AI

软性引导

通过上下文注入引导AI自我修正，而非强制阻断

持续运行

业务不中断，AI Agent保持高效运行

生产力释放

释放 100% AI 生产力，安全成为业务助力

核心结论：在 2026 年能提供 “无感安全” 的产品更具溢价权。100% AI生产力释放

智能体安全治理的关键技术

01 NHI 2.0 身份治理

ID-as-a-Proxy

将 Agent 视为**"第三类身份"**（非人非机器），建立动态指纹与信誉评分，彻底消除"影子智能体"

02 微秒级动态权限

JIT + ABAC

最小权限原则演进为**"即时任务权限 (JIT)"**，权限仅在子任务执行瞬间生效，动作偏移即刻撤销

03 MCP 深度协议审计

Context-Aware

在 Anthropic MCP 标准层建立**"智能体防火墙"**，监控模型与外部数据源的连接，拦截逻辑诱导

04 "HR模型"全生命周期

Lifecycle Governance

引入入职审计、行为轨迹实时监控与**一键挂起 (Kill Switch)** 机制，像管理员工一样管理 AI

智能体安全实践

大模型
安全智链社区

AI 红队自动化测试

非法OpenClaw安装

提示词注入

Skill投毒

非法外联与横向越权

智能体资产发现　智能体安全围栏

Skill安全检测　智能体行为管控

智能体防护

无侵入式对接应用

透明代理/网关监听

AI网关

MCP网关

对外调用
API网关

AI安全一体机

https://github.com/NSF-AIGuard/NSF-ClawGuard

安全运营演进的趋势：Agentic SOC

01

从"副驾驶"向"自主智能体"

自主调查链与动态剧本生成，AI从助手进化为具备自主意图的数字分析师。

AUTONOMOUS AGENTS

02

多智能体协作网格

专业化智能体分工与协同审计机制，构建可信赖的Agent Mesh架构。

SECURITY AGENT MESH

03

超大规模语义遥测与溯源

自然语言调查与攻击故事线自动生成，取证时间从小时级缩短至秒级。

SEMANTIC TELEMETRY

04

预测性威胁狩猎与数字免疫

攻击路径仿真与社区联动防御，实现从“响应”到“预测”的战略转变。

PREDICTIVE HUNTING

自主化率成为SOC重要评价指标

专业化智能体分工架构

未来的SOC不由一个"全能AI"统治，而是由一群 专业化智能体 组成的协作网格，每个Agent负责特定领域任务

告警分拣Agent

ALERT TRIAGE

- 实时分析海量告警流
- 智能去重与优先级排序
- 误报率降低85%+

深度溯源Agent

INVESTIGATION

- 跨域数据关联分析
- 攻击链重建与归因
- 威胁情报自动关联

自动化修复Agent

REMEDIATION

- 自动隔离受感染资产
- 策略自动调整加固
- 修复效果持续验证

自主调查链

AI能够理解告警背后的攻击者意图，无需人类干预即可自主调用各类安全工具进行跨域取证。

- **端点取证 (EDR)：**自动提取进程行为、文件操作、注册表变更等关键证据
- **网络取证 (NDR)：**分析流量异常、C2通信、横向移动等网络层威胁
- **身份取证 (IAM)：**追踪凭证滥用、权限提升、异常登录等身份威胁

动态剧本生成

突破传统SOAR的预设剧本限制，AI根据威胁的实时演进轨迹动态生成并执行响应路径。

传统SOAR		AI-SOAR
预设固定剧本	VS	实时动态生成
无法应对零日		自适应零日攻击
人工配置依赖		自主决策执行

零日漏洞应对能力：AI通过行为分析识别异常模式，无需已知签名即可检测并响应未知威胁

AI数字安全运营团队

核心：运营专家对角色、职责、协同的定义

绿盟安全运营专家 可根据不同的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、挂载其专属调用的安全工具或 API 接口以及其协同的方式

运营经理（大将军）—— 架构的"大脑"

接收长官（用户）的需求，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，是保障产品极大程度自动化的核心指挥中枢

调度

威胁狩猎专家

主动出击，对外收集情报，对内深挖事件库的高隐蔽性威胁，将分析成果转化为规则

研判工程师

过滤噪音，精准锁定高风险行为。核心原则：坚决不为失败的尝试浪费精力。

暴露面评估师

监控资产变化，确保持续收敛，防止已修复的漏洞死灰复燃。

流程文档工程师

主导报告生成与归档，负责 SLA 规则维度的检查和强制度量。

调用

底层能力：灵活调用异构全域探针、设备与平台

绿盟UTS | 绿盟IDPS | 友商SOC | 友商探针 | 客户CMDB | ASM系统 | + 按需自动适配

检测规则 自主编写

自主规则编写 自动添加

漏洞验证代码 自主编写

反测绘对抗 自主绕过

复盘机制 自我改进

缺席的美国政府和Crypto Panel

NSA

CISA

FBI

RSAC 2025

Jen Easterly

Claude带来的冲击

Claude @claudeai

Introducing Claude Code Security, now in limited research preview.

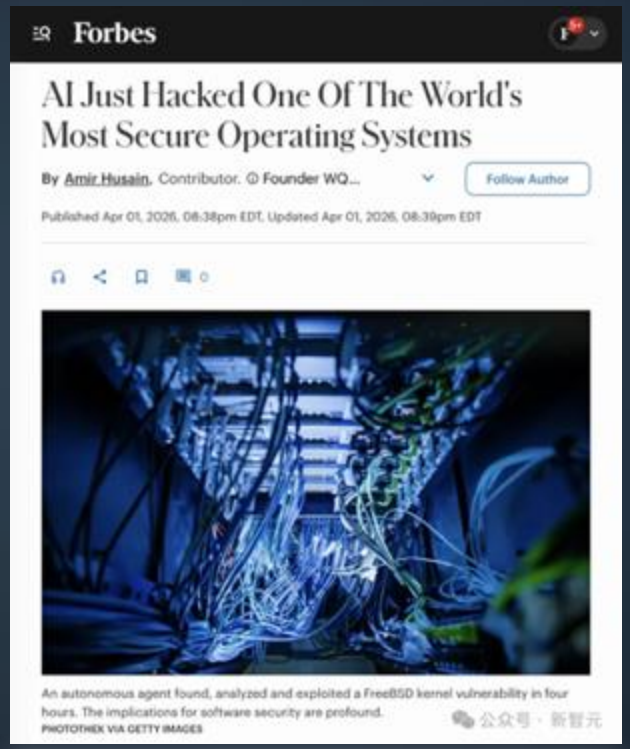
It scans codebases for vulnerabilities and suggests targeted software patches for human review, allowing teams to find and fix issues that traditional tools often miss.

Learn more: anthropic.com/news/claude-co...

介绍 Claude 代码安全，现处于有限研究预览阶段。

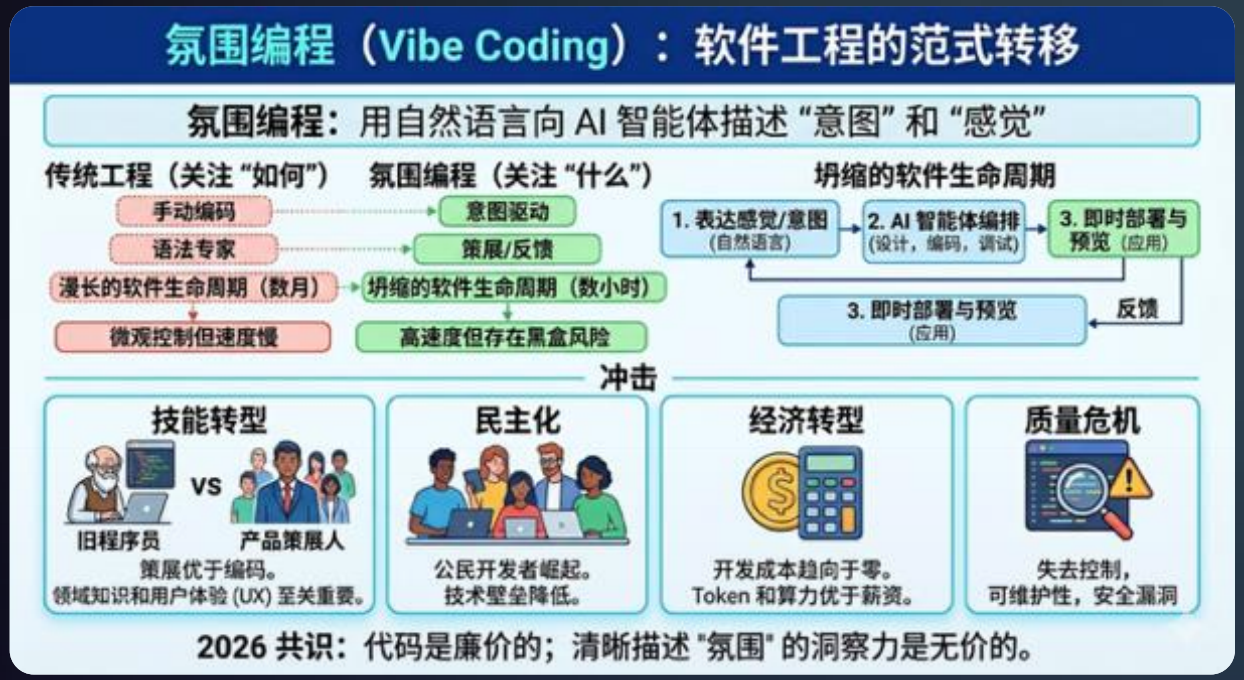
它扫描代码库漏洞，并建议针对性的软件补丁供人工审核，帮助团队发现并修复传统工具常忽略的问题。

了解更多： anthropic.com/news/claude-co...



- Claude模型能力超出其他模型
- 闭源模型目前比开源模型能力强
- 使用国外模型带来新安全风险，需要有应对方法
- 垂域模型在专业领域的应用上具有优势
- 大参数模型在推理规划能力上更优，但成本过高，需要轻量化。

AI时代下安全系统的实现架构



拥抱AI改造产品研发流程体系

Phase 1 • Vibe Coding

定义：Prompt-and-Pray 模式，AI 全权生成，人工零干预。
需求：自然语言，无结构化约束。
质量：无保障，技术债指数级增长。
痛点：交付速度假象，中期维护爆炸。

Phase 2 • Spec-Coding（SDD）

定义：以结构化规格文档驱动 AI 生成，Spec 先于代码存在。
需求：接口契约 / 数据模型 / 异常处理。
质量：Spec 即验收标准，人工手动对照。
价值：输出变为可预期，形成协作基准。

Phase 3 • Harness Coding

定义：人类主导架构标准，AI 在工程化护栏内高效执行。
需求：Spec + 架构决策(ADR) + 边界定义。
质量：流水线级自动化门禁 (Lint/SAST/Test)。
价值：全链路可控，可规模化复制。

实践主题：构建 AI 原生产品研发体系，打造 AI Coding 端到端工程能力

领域纵深

产品经理	SE架构师	研发	测试	攻防研究	公共支撑
产品管理	需求设计	开发实现	质量测试	安全研究/运营	平台/度量/流程

试点探索：根据团队和产品的性质，端到端实践选择的编程范式

（探索AI原生编程范式，积累AI编程工程能力，总结编程范式的实践地图，总结Harness Engineering，推广优化）

流程运转：试点实践 → 岗位纵深探索 → 案例实践总结 → 横向复制验证 → 沉淀复用与循环开拓

价值积累：人才&分享交流 —— 落地工程能力 —— 总结编程范式实践路径

新时代的网络安全企业

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

构建面向未来的智能安全防护体系

拥抱AI

AI不是替代，而是增强

全流程AI赋能

在需求分析、架构设计、代码开发、测试验证等各个环节充分使用AI工具，提升效率与质量

工程范式革新

深度改造软件工程方法论和工作流程，建立AI驱动的DevSecOps新范式

安全原子能力

原子化是智能化的基础

开放性架构

构建标准化、模块化的安全原子能力，支持灵活组合与快速集成

智能体对接

实现与安全智能体的深度对接，支持自动化威胁检测与响应

数据与知识

数据是AI时代的石油

SKILLS体系

建立结构化的安全知识库，沉淀攻防经验、漏洞案例、处置方案

高质量数据

积累标注精准、场景丰富的高质量安全数据，支撑AI模型训练

AI安全新战场

攻防对抗进入智能时代

业务价值导向

安全始终为业务服务，在保障安全的前提下最大化业务价值

新威胁新战法

应对AI带来的新攻击面，发展AI驱动的防御方法与对抗技术

以AI赋能安全，以安全护航AI，构建智能时代的安全防护新范式

谢谢聆听