AI 安全与信任基础设施

CyberArmor.AI 产品家族

“你们有 AI 安全网关吗？” 有：CyberArmor.AI AI Trust Gateway（AI 信任网关）在线检查 AI 双向流量，按客户策略拦截、脱敏或警告，每项决策写入已签名审计记录。与普通网关不同，同样的检查还运行在笔记本电脑、浏览器、手机以及客户自己的应用中，不经过网关的流量同样受到保护。

Trust Gates：决策关口

URL 信任网关 URL Trust Gate
链接与网页在加载前完成检查
试点

上下文信任网关 Context Trust Gate
内容在 AI 读取前检查隐藏指令与投毒
试点

MCP 信任网关 MCP Trust Gate
工具投毒、冒充工具、工具定义被篡改；发现每台机器上的 MCP 服务器
路线图（工具投毒） · 试点（服务器发现）

提示词攻击载荷防护 Promptware Defense
像恶意软件一样运作的提示词，直接与间接：网页、文档、邮件或工具描述中的隐藏指令在 AI 执行前被拦截
生产（直接注入） · 试点（间接触发）

付款完整性信任网关 Payment Integrity Trust Gate
资金只流向已批准的去向：由指定审批人回拨核实，并核对付款实际去向与批准去向是否一致
试点（付款回拨验证）

智能体信任网关 Agent Trust Gate
AI 编程智能体（Claude Code、Codex）：抓取的网页与读取的文件先经检查
概念验证

AI 检测与响应（AIDR）及核心引擎

AI 信任网关（托管） AI Trust Gateway
托管在线检查 AI 双向流量
试点（响应侧策略）

AI 检测 AI Detection
提示词注入、敏感数据与危险输出；对身份、金融、健康与密钥数据进行脱敏
生产

AI 响应 AI Response
主机隔离、终止进程、防火墙联动
试点

运行时决策 API Runtime Decision API
一次调用：扫描、策略、裁决

策略引擎 Policy Engine
租户策略、工件与名单
生产

AI 路由器 AI Router
AI 服务提供商路由与目录
生产

智能体身份 Agent Identity
智能体注册与委托链
生产

恶意软件与勒索软件防护 Malware & Ransomware Defense
特征检测、隔离、勒索软件拦截
试点 · 概念验证（终止勒索进程）

防护运行位置

终端代理 Endpoint Agent
macOS、Windows、Linux：本地代理、文件、剪贴板、补丁
试点 · 生产（签名安装包）

浏览器扩展 CyberArmor.AI Protect
Chrome、Edge、Firefox、Safari、Opera
试点

iOS 应用 iPhone, iPad & Watch
设备端过滤、Safari 扩展、MDM 按应用检查
概念验证 · 路线图（托管应用代理）

Android 应用 Android
VPN 过滤、工作资料清单
概念验证

开发与办公插件 Developer & Office add-ons
Office 插件；VS Code、Cursor、Kiro、Visual Studio
试点

SDK 开发工具包 CyberArmor.AI SDKs
将检查能力嵌入您自己的 AI 应用：Python、Go、Java、Node.js、.NET、C/C++、PHP、Ruby、Rust；LangChain、LlamaIndex、OpenAI 与 Anthropic 封装
试点

运行时应用自我保护 CyberArmor.AI RASP
在运行中的应用内部：检查发往 AI 的流量、识别提示词注入、检查外发请求、报告实际加载内容；同样支持 9 种语言
试点

无代理集成 Agentless Integrations
防火墙 EDL、SIEM、Syslog/EDR 接入、M365
试点

机器人安全 Robotics (ROS 2)
话题与服务防护、执行器策略、急停
试点

证据、治理与信任

已签名审计记录 Signed Audit Trail
HMAC-SHA256 记录，ML-DSA-87 检查点
生产

合规与报告 Compliance & Reports
45 个框架、44 个策略包、附证据的证明
生产

AI 物料清单（A-BOM） AI Bill of Materials
AI 清单，结合 KEV 与 EPSS 漏洞数据
生产

影子 AI 发现 Shadow AI Discovery
正在使用的 AI 服务、应用与 MCP 服务器
试点

媒体来源验证 Media Provenance
C2PA 内容凭证
生产

身份与访问 Identity & Access
SSO、MFA、目录信息补充、后量子密钥
生产

AI 输出验证 AI Output Verification
结构、算术、引用、与来源一致性
路线图

控制台 客户门户（企业安全团队）· 合作伙伴控制台（MSP 与经销商）· 运营控制台（CyberArmor.AI 员工）

同样包含 设备端本地 AI 模型 · 沙箱引爆分析 · 输入时即时脱敏 · 银行官网以外的银行卡 PIN 保护 · 设备未在防护时主动报告 · 补丁管理 · 剪贴板脱敏 · 防火墙拦截名单 · 中国 AI 服务（Kimi、豆包、通义千问、文心一言等）识别与检查 · 客户数据留在客户所在国家 · 后量子密码 · 中英文培训视频与指南

成熟度以 CyberArmor.AI 状态页（cyberarmor.ai/status）公布为准，2026 年 10 月 29 日。CyberArmor.AI 产品与功能名称正在进行商标审查。