AI 技术每日分析：CrowdStrike 拟收购 Fiddler AI 与 Dreadnode，布局智能体安全

中国高技术产业发展促进会新质生产力工作委员会

博雅云创 & 中科创新驱动

2026 年 11 月 4 日

摘要

网络安全企业 CrowdStrike 同日宣布拟收购 Fiddler AI 与 Dreadnode，把智能体运行观察、自动化攻击测试和应急处置纳入统一安全平台；新发表的 Strategist 研究用可复用代码策略优化智能体推理流程，在 30 个基准任务中有 24 项超过最强对照，平均推理费用降低 39%。欧洲方面，Bull 与 Schwarz Digits 签署合作意向，计划结合欧洲制造的 AI 服务器与 STACKIT 云平台提供本地司法管辖范围内的计算服务。美国新安全中心则提出“主权智能体”风险框架，推动长期自主系统的安全治理讨论。

Contents

1. CrowdStrike 宣布两项收购：智能体运行监测与攻击测试进入安全平台 **2**
2. Strategist 论文发表：把推理方法写成可复用的代码蓝图 **2**
3. Bull 与 Schwarz Digits 签署合作意向：欧洲主权 AI 云走向服务器与平台协同 **3**
4. CNAS 发布“主权智能体”报告：长周期自主系统的控制权成为安全议题 **4**
趋势观察 **4**
参考资料 **4**

1. CrowdStrike 宣布两项收购：智能体运行监测与攻击测试进入安全平台

11 月 3 日，CrowdStrike 宣布拟收购 AI 可观测性企业 Fiddler AI 及专注自动化 AI 安全测试的 Dreadnode。公司计划将前者能力整合为 Falcon Observe，跟踪智能体的输入输出、工具调用和运行行为；后者将强化模型与智能体在提示词注入、数据泄露和工具滥用等方面的主动测试。CrowdStrike 同日还公布面向智能体的安全管理能力，包括必要时中止风险执行流程。收购仍需完成约定程序，部分集成功能属于后续产品计划，而非已经覆盖全部企业客户的正式部署。来源为 11 月 3 日公司公告和产品博客。[AI-1][AI-2][AI-3]

智能体安全正在从“模型回答是否合规”扩展到“软件正在代表谁执行什么操作”。一个能读写仓库、调用浏览器和提交工单的智能体，其风险与普通聊天窗口明显不同。对企业开发团队来说，模型接入日志、外部工具权限、异常动作拦截以及可追溯的审计链开始构成同一套工程需求。Fiddler 与 Dreadnode 分别从运行观察和对抗性评估切入，两类小型技术公司的能力被大型平台并购，也反映出 AI 安全产品正在出现更细的专业分工。

2. Strategist 论文发表：把推理方法写成可复用的代码蓝图

11 月 2 日公布、11 月 3 日进入技术社区视野的论文《Strategist: Evolving Reasoning with Agents》，由 Taiming Lu 等研究者提出自动演化推理策略的元智能体。系统接到新任务后，先从已有蓝图库中检索相关策略，再编写、运行、评价候选推理程序，把有效经验保存供后续任务复用。论文在 30 个基准上进行测试：其中 24 项超过最强对照方法，平均准确率为 74.7%，对照为 74.0%，推理成本下降 39%。这是论文设定中的

实验结果，尚需要独立复现和真实生产任务验证。[AI-4]

过去，团队通常通过调整提示词或人为编排工作流来提高智能体表现，不同项目之间很难复用调试经验。Strategist 尝试把检索、分解、验证与重试规则转化为可执行资产，使一个任务中积累的推理组织能力转移到下一项任务。实验报告同时给出了质量与成本指标：部署时不宜只比较单次答案得分，还应观察搜索次数、额外模型调用、工具执行时间和错误恢复能力。对于代码智能体和研究助手，策略复用可能降低反复试验的工程成本。

3. Bull 与 Schwarz Digits 签署合作意向：欧洲主权 AI 云走向服务器与平台协同

11 月 3 日，法国超级计算机厂商 Bull 与德国 Schwarz 集团技术子公司 Schwarz Digits 宣布达成初步合作协议，拟共同发展受欧洲司法管辖的 AI 计算服务。按公布的分工，Bull 负责服务器制造及相关技术，Schwarz Digits 通过 STACKIT 提供云计算、存储和 AI 服务入口，双方还将探索联合设计数据中心服务器，并向客户开放 Bull 的 AI 软件能力。合作交易的财务条件尚未公开，联合服务的具体上市和验收时间亦有待确定。[AI-5][AI-6]

协议涉及硬件制造、计算资源交付和云平台运营。欧洲公共机构与受监管企业采购 AI 服务时，除了模型能力，也关注数据所在地、跨境访问规则、芯片与服务器供应链以及平台运维责任。Bull 的计算设备与 STACKIT 客户渠道相互补充，但未来能否形成稳定市场，仍取决于实际算力性能、价格、能源效率和企业应用迁移工作量。中小 AI 云供应商的竞争空间，正在沿着行业合规和区域市场继续细分。

4. CNAS 发布“主权智能体”报告：长周期自主系统的控制权成为安全议题

11 月 3 日,美国新安全中心(CNAS)发表研究报告《Sovereign Agents》,讨论一种在较长周期内持续规划、获得资源并维持执行能力的自主智能体可能引发的治理问题。报告把一般按用户授权完成任务的持续性智能体，与可能绕开原有控制边界、形成自身行动连续性的“主权智能体”风险情景进行区分，建议研究更完善的权限限制、身份验证、监测和终止机制。这是一份前瞻性政策研究，并未宣称已发现具备完整自主生存能力的现成商业系统。[AI-7][AI-8]

随着长任务智能体能够连续修改文件、调度多个代理、访问云资源甚至管理账户，如何确认其行为始终符合人的授权范围，成为系统设计问题。技术团队可从短期操作入手：为代理身份设置期限，限制资源预算，对敏感工具采用二次授权，对跨任务持久状态保留清理与回滚机制。报告提醒业界，智能体能力越强，软件权限、基础设施治理与事故响应越需要同时成熟。

趋势观察

本期 AI 动态分别发生在安全工具、推理方法、区域算力和治理研究四条线上。模型能力继续提升，企业更重视的是如何让长时间运行的软件可管理、可复用、可核算，并能在出现偏差时及时恢复控制。

参考资料

1. CrowdStrike 官方博客 | CrowdStrike to Acquire Fiddler to Deliver Falcon Observe for AI Agent Observability | 2026-11-03 | Fiddler 拟收购及智能体运行观察方案

   https://www.crowdstrike.com/en-us/blog/crowdstrike-to-acq

uire-fiddler/

2. CrowdStrike 官方博客｜ CrowdStrike to Acquire Dreadnode to Bring Automated AI Security Testing to the Falcon Platform ｜ 2026-11-03 ｜ Dreadnode 拟收购、AI 攻击测试能力
https://www.crowdstrike.com/en-us/blog/crowdstrike-to-acquire-dreadnode/

3. CrowdStrike 官方博客｜ CrowdStrike Advances AI Security and Safety from Endpoint to Cloud ｜ 2026-11-03 ｜智能体安全产品与行动控制的企业公告
https://www.crowdstrike.com/en-us/blog/

4. arXiv ／ Taiming Lu 等｜ Strategist: Evolving Reasoning with Agents ｜ 2026-11-02 ｜代码蓝图机制、30 项基准、推理成本实验
https://arxiv.org/abs/2611.03618

5. Reuters ｜ Schwarz Digits and Bull partner up to build European AI computing services ｜ 2026-11-03 ｜合作主体、STACKIT 平台与服务器制造分工
https://www.reuters.com/business/retail-consumer/schwarz-digits-bull-partner-up-build-european-ai-computing-services-2026-11-03/

6. Bull ／ GlobeNewswire ｜ Bull and Schwarz Digits: A Franco-German Alliance for Europe's Digital Sovereignty ｜ 2026-11-03 ｜合作意向及欧洲主权算力方案原始公告
https://www.globenewswire.com/news-release/2026/11/03/3394200/0/en/bull-and-schwarz-digits-a-franco-german-allianc

e-for-europe-s-digital-sovereignty.html

7. Center for a New American Security | Sovereign Agents: The Pathway to Irreversible Loss of Control | 2026-11-03 |主权智能体定义、风险场景及控制建议
   https://www.cnas.org/publications/reports/sovereign-agents
8. Axios | AI sovereign agents security threat | 2026-11-03 |研究报告的外部报道与行业讨论
   https://www.axios.com/2026/11/03/ai-sovereign-agents-security-threat