

AI 技术每日分析：Z.ai 揭开 Ox Alpha 身份并发布 GLM-5.3-Flash

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 8 月 27 日星期四

摘要

今天 AI 技术方向的新动态，明显从“谁又发了一个更大的模型”转向几条更具体的工程线：模型如何以更低成本进入真实工作流，Agent 如何在企业软件中被反复训练，通用 Agent 如何从“帮你做东西”走向“帮你把业务跑起来”，以及开放模型在安全任务上的成本效率如何变化。Z.ai 正式揭开 Ox Alpha 身份，确认此前匿名测试模型就是 GLM-5.3-Flash，并披露其在代码与 Agent 基准、成本和国产芯片承载方面的数据；Deep Cogito 完成 4300 万美元 A 轮融资，继续押注开放权重模型、强化学习后训练和递归式自我改进；Arga 获得 1000 万美元种子轮融资，用“企业软件数字孪生”解决 Agent 在 Salesforce、Workday 等复杂环境中无法大规模重复训练的问题；Runnable 完成 2100 万美元 A 轮融资，开始把 Agent 能力从网站和应用生成扩展到广告投放、SEO、社交媒体和客户获取；Fireworks AI 则用 CyberGym 测试 DeepSeek V4 Pro 等模型，显示安全 Agent 的评估正在从一次性代码题走向数十轮真实漏洞操作。AI 竞争的核心，正在从单次回答能力继续向训练环境、执行闭环、成本结构和可控性延伸。

目录

一、Z.ai 正式揭开 Ox Alpha 身份：匿名测试成为模型发布前的真实流量实验	3
二、Deep Cogito 融资 4300 万美元：企业开始重新重视“可拥有、可训练”的开放权重模型	4
三、Arga 用“企业软件数字孪生”训练 Agent：Salesforce 也需要一个可反复重置的训练场	5
四、Runnable 融资 2100 万美元：Agent 开始从“帮你搭网站”转向“帮你找客户”	5
五、Fireworks 把安全 Agent 放进 CyberGym：评估开始看“真实漏洞 + 长程任务 + 成本”	6
趋势观察	6
参考资料	7

今天 AI 技术方向的新动态，明显从“谁又发了一个更大的模型”转向几条更具体的工程线：模型如何以更低成本进入真实工作流，Agent 如何在企业软件中被反复训练，通用 Agent 如何从“帮你做东西”走向“帮你把业务跑起来”，以及开放模型在安全任务上的成本效率如何变化。Z.ai 正式揭开 Ox Alpha 身份，确认此前匿名测试模型就是 GLM-5.3-Flash，并披露其在代码与 Agent 基准、成本和国产芯片承载方面的数据；Deep Cogito 完成 4300 万美元 A 轮融资，继续押注开放权重模型、强化学习后训练和递归式自我改进；Arga 获得 1000 万美元种子轮融资，用“企业软件数字孪生”解决 Agent 在 Salesforce、Workday 等复杂环境中无法大规模重复训练的问题；Runnable 完成 2100 万美元 A 轮融资，开始把 Agent 能力从网站和应用生成扩展到广告投放、SEO、社交媒体和客户获取；Fireworks AI 则用 CyberGym 测试 DeepSeek V4 Pro 等模型，显示安全 Agent 的评估正在从一次性代码题走向数十轮真实漏洞操作。AI 竞争的核心，正在从单次回答能力继续向训练环境、执行闭环、成本结构和可控性延伸。

一、Z.ai 正式揭开 Ox Alpha 身份：匿名测试成为模型发布前的真实流量实验

8 月 26 日，Z.ai 正式发布 GLM-5.3-Flash，并确认此前在 OpenRouter 和 OpenCode 上以“ox-alpha”匿名运行的模型就是它的预发布测试版。官方称，匿名测试期间模型一度成为当周最受欢迎模型之一，流量全部由中国 AI 芯片承载。

这条消息的重要性不只是“谜底揭晓”。匿名预览正在形成一种新的模型发布方式：先把模型放进真实开发者流量中，用免费或低价方式观察代码、Agent、长上下文和工具调用的实际表现，再正式发布品牌与权重。与封闭 Benchmark 相比，这种方式更接近真实生产流量，也能更早暴露速度、兼容性和行为稳定性问题。

Z.ai 披露，GLM-5.3-Flash 在其引用的 Artificial Analysis Intelligence Index v4.1.1 上得到 57 分，折扣后单任务成本约 0.045 美元；在 DeepSWE v1.1 和 AutomationBench 等代码、Agent 基准上相较 GLM-5.2 提升明显。需要注意，这些数字一部分来自厂商发布，仍应结合第三方复测看待，但它已经说明一个趋势：开放模型的竞争正在从“能不能追上闭源模型”变成“在可接受质量下，谁能把每次任务的成本压得更低”。

二、Deep Cogito 融资 4300 万美元：企业开始重新重视“可拥有、可训练”的开放权重模型

Deep Cogito 8 月 26 日宣布完成 4300 万美元 A 轮融资，由 TQ Ventures 领投，Benchmark、Nexus Venture Partners、Atreides Management、South Park Commons 以及客户兼战略投资方 Zscaler 参与。公司累计融资超过 5600 万美元。

Deep Cogito 的定位很明确：重点不是做一个面向普通用户的聊天产品，而是做后训练和自我改进系统，让企业能够拥有自己的开放权重模型，并用专有数据和业务结果继续训练。公司创始团队曾参与 Google AI Search，其技术路线强调大规模强化学习、后训练和递归式自我改进。

这类公司的价值在企业 AI 进入第二阶段后开始上升。企业如果长期依赖外部闭源 API，虽然上线快，但在成本、数据控制、模型行为可解释性和定制空间上都会受制于供应商。开放权重模型加上企业自己的后训练体系，提供了另一种路径：让模型更贴近某个垂直业务，同时把推理和数据留在自己可控的基础设施中。

三、Arga 用“企业软件数字孪生”训练 Agent：Salesforce 也需要一个可反复重置的训练场

Arga 在 8 月 26 日宣布获得 1000 万美元种子轮融资，由 General Catalyst 领投，Box Group、Emergence、Gradient 和 SV Angel 参与。它解决的是企业 Agent 非常具体的训练难题：真实的 Salesforce、Workday、邮件系统并不适合被 Agent 反复“试错”。

传统强化学习环境可以把任务跑几次，每次失败后重新初始化。但企业软件有复杂的权限、Webhook、对象关系和状态变化，很难一键恢复。Arga 因此选择完整复制企业应用结构，构建带权限体系和状态变化的“数字孪生”，让 Agent 可以在不破坏生产系统的情况下反复训练。

这类基础设施可能会成为企业 Agent 规模化的关键环节。代码 Agent 进步快，一个重要原因就是代码世界天然拥有 Git、测试、容器、回滚和沙箱；业务软件过去缺少同等级的训练基础设施。Arga 实际上是在给企业软件补一套类似“CI/CD + 强化学习环境”的底座。

四、Runable 融资 2100 万美元：Agent 开始从“帮你搭网站”转向“帮你找客户”

印度创业公司 Runable 在 8 月 26 日完成 2100 万美元 A 轮融资，由 Susquehanna Venture Capital 和 Nexus Venture Partners 共同领投。公司最初做浏览器和 Agent 基础设施，后来转向面向小企业的一体化 Agent 平台。

Runable 目前可以通过自然语言生成网站、应用、演示文稿，并把数据库、身份认证、支付、部署和分析等基础能力一起打包。新一轮产品重点开始向“增长”侧扩展，包括广告投放、社交媒体、SEO、AI 搜索结果优化和客户触达。

TechCrunch 披露，Runnable 注册用户约 170 万；公司称 3 月启动付费后 3 周达到约 200 万美元年化收入水平，但目前仍存在负毛利，原因之一是补贴用户 AI 推理成本。这恰好说明 Agent 商业化的现实难题：产品越能真正替用户执行工作，Token、浏览器、工具调用和外部服务成本越高。未来 Agent 公司真正的壁垒，很可能既包括能力，也包括单位任务经济性。

五、Fireworks 把安全 Agent 放进 CyberGym：评估开始看“真实漏洞 + 长程任务 + 成本”

Fireworks AI 在 8 月 26 日发布对 DeepSeek V4 Pro 0813 的安全 Agent 评估结果。其 CyberGym 测试包含 840 个真实代码库中的对抗性安全任务，平均每个任务约 89 轮交互，并与 Kimi K3、GPT-5.5、Claude Opus 4.8 等模型进行对比。

Fireworks 称，DeepSeek V4 Pro 在这组测试中没有出现拒答，并在“每次成功的成本”上表现突出。由于测试由推理平台和基准提供方发布，这些结果应视为厂商实验数据，而不是独立安全结论；同时，CyberGym 面向的是授权安全测试，不代表对现实攻击行为的鼓励。

更值得关注的是评估方式本身。过去安全模型经常用单轮问答或静态代码题评估，现在开始转向真实代码库、长任务链和多轮工具执行。对企业而言，安全 Agent 的价值也会越来越用“发现一个真实问题要花多少钱、需要多少轮、能不能稳定复现和修复”来衡量。

趋势观察

今天几条动态共同指向一个变化：AI 行业的工程重心正在向“模型之外”移动。匿名模型测试让真实流量成为发布前评估场，Deep Cogito 押注企业自己的后训练体系，Arga 给企业软件复制训练环境，Runnable 把

Agent 拉到业务结果和收入环节，Fireworks 把安全评测拉到长程真实任务。模型能力仍然重要，但能否被训练、被重复执行、被低成本调用、被审计和被安全控制，正在决定 AI 系统是否真正进入生产。

参考资料

1. Z.ai | 《GLM-5.3-Flash: Frontier Intelligence, Flash Cost》 | 2026-08-26 | 核验 Ox Alpha 身份、基准与成本数据。

<https://z.ai/blog/glm-5.3-flash>

2. Business Insider | 《Mystery solved: Chinese lab Z.ai says it's behind the Ox Alpha model...》 | 2026-08-26 | 独立补充匿名测试背景。

<https://www.businessinsider.com/ox-alpha-model-made-by-china-z-ai-2026-8>

3. The Wall Street Journal | 《Deep Cogito Aims to Put Companies in Control of Their Own AI》 | 2026-08-26 | 核验 Deep Cogito 融资和企业开放模型路线。

<https://www.wsj.com/cio-journal/deep-cogito-aims-to-put-companies-in-control-of-their-own-ai-50ddf44b>

4. Business Wire / Yahoo Finance | 《Deep Cogito Raises \$43M Series A...》 | 2026-08-26 | 补充融资参与方与公司技术定位。

<https://au.finance.yahoo.com/news/deep-cogito-raises-43m-series-170000457.htm>

5. TechCrunch | 《Arga is building a better way to train enterprise AI agents》 | 2026-08-26 | 核验 Arga 融资与企业软件数字孪生训练环境。

<https://techcrunch.com/2026/08/26/arga-is-building-a-better-way-to-train-enterprise-ai-agents/>

6. TechCrunch | 《Runnable hits \$21M to bet AI agents can go from building businesses to growing them》 | 2026-08-26 | 核验 Runnable 融资、用户与产品方向。

<https://techcrunch.com/2026/08/26/runable-hits-21m-to-bet-ai-agents-can-go-from-building-businesses-to-growing-them/>

7. Business Wire / Yahoo Finance | 《Runnable Raises \$21M Series A to help small businesses run, build and grow》 | 2026-08-26 | 补充公司官方口径。

<https://finance.yahoo.com/small-business/articles/runable-raises-21m-series-help-120000916.html>

8. Fireworks AI | 《DeepSeek V4 Pro is Redefining Security Agent Economics》 | 2026-08-26 | 核验 CyberGym 安全 Agent 评估数据。

<https://fireworks.ai/blog/DeepSeek-V4-Pro-Security>

9. GitHub Community | 《remote-agent - an open-source free AI agent for self-hosted AI automation & workflows》 | 2026-08-25/26 | 补充本地、自托管 Agent 与策略审计长尾趋势。

<https://github.com/orgs/community/discussions/205799>

10. LLM Releases | 模型发布追踪页 | 2026-08-26 更新 | 用于交叉检查近期开放模型发布节奏。

<https://www.llm-releases.com/>

联系我们，请扫描二维码



新质生产力工作委员会
官方公众号



工业智能算网
gyznsw.cn

新质生产力工作委员会：

中国高技术产业发展促进会新质生产力工作委员会，专注于推动工业人工智能、智能制造、数字化转型等前沿技术发展，为企业提供政策解读、技术咨询和产业对接服务。

工业智能算网：

专注于工业人工智能、新质生产力、工业软件 CAE、智能制造等前沿技术。提供每日动态分析、技术趋势解读、解决方案分享，推动工业智能化转型。

网站地址：<https://gyznsw.cn>