AI 技术每日分析

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 4 月 24 日

摘要

全球 AI 领域在底层安全机制、资本市场定价、政策监管走向及劳动力市场影响等方面发生了深远演变。Anthropic 专为网络安全开发的前沿模型 Mythos 遭遇未经授权的第三方访问，引发业界对超人类网络攻击能力的担忧；在二级资本市场，Anthropic 的隐含估值一度突破 1 万亿美元，短暂反超 OpenAI。与此同时，美国科技巨头第一季度的 AI 政策游说支出创下新高，劳动力市场则显示高薪和资深员工采纳 AI 的速度远超基层员工，职场“AI 鸿沟”迅速拉大。白宫最新备忘录还将“大模型蒸馏”上升为知识产权与国家安全议题。

Contents

1 前沿模型安全危机：Anthropic 彻查 Mythos 模型越权访问事件

1.1 事件背景与技术隐患

过去 24 小时内，备受瞩目的前沿 AI 安全事件爆发。据《金融时报》与彭博社披露，Anthropic 正在紧急调查一起针对其最新发布的高级模型“Claude Mythos”的越权访问事件。Mythos 是 Anthropic 于本月早些时候推出的一款具备极高网络安全能力的前沿模型。考虑到该模型在寻找和利用系统漏洞方面的速度和规模可能远超人类，Anthropic 仅将其试用权限开放给亚马逊、微软、苹果、思科和 CrowdStrike 等少数受信任的科技巨头，用于在正式向公众发布前协助检测和修复网络漏洞。然而，调查显示，有不良行为者利用 Anthropic 第三方外包测试人员的权限，成功潜入了用于模型开发的“供应商环境”并调用了 Mythos。

1.2 行业影响与治理挑战

尽管 Anthropic 官方声明目前没有证据表明越权活动蔓延到了供应商环境之外，但这起事件直接暴露了当前顶级 AI 实验室在第三方供应链及权限管理上的脆弱性。安全专家警告称，随着大语言模型能力的指数级跃升，一旦具备高级渗透测试和自动化代码生成能力的模型落入黑客之手，其发起的大规模网络攻击将使传统防御体系防不胜防。对于估值已达数千亿美元的 Anthropic 而言，如何在模型快速迭代测试与防止技术武器化之间取得平衡，将是其面临的最严峻考验。

2 资本市场历史性交锋：Anthropic 隐含估值破万亿，OpenAI 全面转向 B 端

2.1 估值倒挂背后的商业逻辑演变

根据 Forge Global 等二级市场交易平台的最新数据，Anthropic 的股票近期遭遇机构抢筹，其隐含估值一度达到约 1 万亿美元，在二级市场短暂超越了估值约 8800 亿美元的 OpenAI。Anthropic 在 2026 年 2 月完成 G 轮融资时的官方投后估值为 3800 亿美元，而近期二级市场给出的超额溢价，主要源于其在企业级市场的强劲表现以及 Claude 4.7 版本在复杂代码生成和 Agent 连续执行上的高可靠性。

2.2 OpenAI 的 B 端反击与“次级 AI 危机”隐忧

为了应对 Anthropic 在 B 端业务的步步紧逼，OpenAI 正在经历深度的内部战略调整。美联社报道称，OpenAI 已经全面将重心转向“高价值专业工作”领域的企业用户，甚至为此在资源优先级上放缓了部分面向消费者的娱乐级产品线。为了夺回在企业端的话语权，OpenAI 即将推出代号为“Spud”的新一代推理模型，主打更强的逻辑链条、对意图的深刻理解以及在生产环境中的极致可靠性。然而，市场同时警告，两家公司都

面临着高昂的推理算力成本和亏损压力，这种建立在巨大烧钱压力下的算力军备竞赛，正在演变为潜藏的“次级 AI 危机”。

3 科技巨头游说战升级：一季度 AI 政策游说支出创历史新高

3.1 北美科技巨头集体加码华盛顿游说

随着 AI 技术的快速渗透，美国联邦层面的监管政策正处于成型关键期，这直接触发了科技巨头游说支出的爆发式增长。据 Axios 综合联邦披露文件的数据显示，2026 年第一季度，硅谷核心企业在华盛顿的政策游说开销远超往年同期。Meta 以 710 万美元领跑，亚马逊与谷歌分别斥资 440 万美元和 290 万美元，Anthropic 与 OpenAI 作为纯粹的 AI 实验室，预算增幅最为惊人。

3.2 AI 政策博弈已蔓延至算力、能源与贸易

Anthropic 一季度支出达 160 万美元，重点议题为 AI 政府采购标准、国家安全风险防御以及供应链出口管制；OpenAI 支出创历史新高的 100 万美元，核心聚焦于 AI 训练与版权争议、云基础设施扩建及网络安全规范。与此同时，AMD、Nvidia 以及数据中心联盟的活动度也显著提高，说明 AI 政策的博弈场已经从单纯的算法伦理，全面蔓延至底层算力设施、电力能源、知识产权与国际贸易等宏观实体经济层面。

4 劳动力市场的新鸿沟与模型合规升级

4.1 高薪阶层采纳率远超基层，职场不平等加剧

《金融时报》联合多家机构对英美两国 4000 名职场人士进行的深度调研表明，AI 工具的采用率呈现出清晰的“倒金字塔”结构，收入最高、经验最丰富的员工采用 AI 的速度远远超过基层员工。超过 60% 的高薪

高管及资深专家每天都在工作中使用高级 AI 助手，而这一比例在低收入群体中仅为 16%。OpenAI 首席经济学家 Ronnie Chatterji 认为，AI 本质上是一种“熟练度补充工具”，更容易放大资深员工原有的生产力，而不是自动填平经验差距。

4.2 白宫直指“大模型蒸馏”式工业级知识窃取

白宫科技政策办公室主任 Michael Kratsios 签署的一份内部备忘录被媒体曝光，文件正式指责部分境外实体正在进行工业规模的美国前沿 AI 系统知识产权窃取行动。备忘录聚焦于一种被称为“模型蒸馏”的技术手段，即通过 API 大规模调用美国顶尖大模型，利用其高质量逻辑链输出来训练成本更低的本地模型。Anthropic 和 OpenAI 均公开反对这种策略，认为其严重违背 API 服务条款，并使竞争对手能够在缺乏庞大底层算力投入的情况下通过“捷径”快速抹平技术护城河。未来主流 AI 模型的 API 访问权限、身份验证和审计机制将面临更严苛审查。

5 参考文献

1. Financial Times (2026), *Anthropic investigating unauthorised access of powerful Mythos AI model*。
2. Bloomberg (2026), *Contractor access leads to Anthropic Mythos probe*。
3. Binance / Forge Global Market Reports (2026), *AI Trends: Anthropic's Valuation Surpasses OpenAI on Forge Global*。
4. Associated Press (2026), *ChatGPT maker OpenAI shifts its focus to business users amid Anthropic pressure*。

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