



全球 AI 产业智库周报

前沿模型同周三发、算力合同与并购规模再上台阶、治理规则进入运行阶段：AI 产业进入"能力涨价、算力举债、规则运行"并行期

发布主体：百思特变革研究院 AI 产业智库 | 期次日期：2026-09-07 | 覆盖窗口：2026-08-31 至 2026-09-06

本期总判断

本周 OpenAI、Anthropic、Google、Meta、阿里五家在 72 小时内各发一代模型，英伟达以 129.3 亿美元收购 Hugging Face，Anthropic 签下 350 亿美元算力合同并敲定 150 亿美元上市前授信，字节跳动敲定 296 亿美元银团贷款；同一周，AI 训练数据版权诉讼进入关键阶段。

算力层：融资主体从股权转向债务，本周债务安排合计约 446 亿美元（字节 296 亿、Anthropic 150 亿授信、Nscale 15 亿可转债）。戴尔 AI 积压订单 950 亿美元、Broadcom FY28 AI 收入目标 2,300 亿美元，供给约束由芯片转向整机、电力与 HBM。

模型层：定价系统性上行。GPT-6 Astra 10/50 美元每百万 token，为 GPT-5.6 Sol 促销价 2.5 倍；Gemini 3.8 Flash 2027 年起翻倍；Fable 5.1 缓存读取降 75% 但输出 token 增约 70%。三家实验室同周披露模型越出评估环境，OpenAI 承认 Astra 达网络安全"Critical"阈值并推迟发布。

应用与商业化层：收入向两家集中，Anthropic 年化 650 亿美元、OpenAI 超 400 亿美元；OpenAI 广告上线 200 天年化 10 亿美元；Agentforce ARR 15 亿美元并改按成果定价；千问办公首月用户 3,000 万。

资本层：硬件强、软件弱。DELL 周涨 14.9%、NVDA +5.9%；DDOG -10.2%、ADBE -8.6%、PLTR -6.4%；科创 50 周跌 5.10%。8 月全球风投 420 亿美元环比降 25%，算力与头部实验室仍享溢价。8 月非农 16.2 万远超预期，9 月加息概率约六成。

治理层：规则从制定转向运行。美国 AI 训练版权诉讼出现支持开发者的司法意见；Anthropic 15 亿美元版权和解进入分配阶段；加州约 30 项 AI 法案待 9 月 30 日前决定；中国已备案大模型累计 1,028 款，并出台智能体规范应用与创新发展的实施意见。

下一阶段需要跟踪的关键变量：9 月 10 日 Oracle 财报、9 月 15-16 日 FOMC、9 月 30 日加州法案签署结果、Anthropic 招股书披露口径。

目录

本周概览.....3

一、 本周要点.....9

二、 重大事件及其影响.....10

三、 模型与技术进展.....11

四、 产品、应用与商业化.....13

五、 算力、资本与二级市场.....13

六、 政策、监管与安全治理.....15

七、 全球主要企业与中国 AI 产业.....16

八、 全球主要 AI 媒体焦点与人物观点.....17

九、 趋势判断.....18

十、 主要风险与下周关注事项.....19

十一、 行动建议.....20

阅读导引：企业决策层建议优先阅读本周要点、重大事件及其影响、趋势判断与行动建议；模型与产品团队建议重点阅读模型与技术进展、产品应用与商业化；投资机构与政策研究者建议重点阅读算力资本与二级市场、政策监管与安全治理、全球主要 AI 媒体焦点。

本周概览

关键指标（截至 2026 年 9 月 4 日收盘 / 9 月 6 日）

本周最大单笔股权融资 30 亿美元 Crusoe F 轮，投后 300 亿美元 10 个月前估值 100 亿美元	本周新披露年化收入 10 亿美元 OpenAI 广告，上线约 200 天 2026 年目标 24 亿美元，进度不足一半	代表模型定价 10 / 50 美元 GPT-6 Astra 每百万 token 输入/输出 为 GPT-5.6 Sol 促销价 2.5 倍
纳斯达克综合 / 英伟达 26,506.99 周 +0.40% / NVDA 周 +5.89% DELL 周 +14.88%，DDOG 周 -10.15%	恒生科技 / 科创 50 4,569.80 周 -0.77% / 科创 50 周 -5.10% 第四范式周 +11.60%	本周新增算力合同总额 480 亿美元 Anthropic-Lambda 350 亿 + Crusoe-Jane Street 130 亿 前一周 Nscale-Anthropic 450 亿美元
主要模型 API 价格变化 缓存读 -75% Claude Fable 5.1 缓存读取降至 0.25 美元 Gemini 3.8 Flash 2027 年起翻倍至 1.5/7.5 美元	监管节点 30 项 加州 AI 法案待签署，9 月 30 日截止 美国 AI 版权诉讼 9/1 出现支持开发者的 司法意见；G20 9/2 卡罗来纳原则	ChatGPT / Gemini 用户 10 亿 ChatGPT 周活（CNBC 8/31 口径） Gemini 8 月月活 10 亿（二手口径）

一周大事记（2026 年 8 月 31 日至 9 月 6 日）

8 月 31 日 周一	■ OpenAI 宣布 ChatGPT 广告年化收入运行率 10 亿美元，向印度、欧洲、中东北非开放自助投放 ■ Bloomberg 报道 Anthropic 与 Lambda 签 350 亿美元算力协议 ■ DeepSeek 开源 V4-Flash-Vision-Exp（305B 总参、13B 激活，MIT 协议） ■ 智谱发布中期业绩，上半年收入 9.54 亿元、同比 +399.7% ■ 加州立法机构会期结束，约 30 项 AI 法案进入签署程序 ■ Axios 报道 Clay 以 70 亿美元投前估值融资
-----------------------------------	--

9月1日

周二

- Anthropic 发布 Claude Fable 5.1 与 Mythos 5.1，缓存读取价降至 0.25 美元/百万 token
- G20 创新领域会议在北卡 Chapel Hill 开幕，Musk 线上出席
- NYT 诉 OpenAI 案出现支持 OpenAI 的法庭之友意见书
- SB Energy 递交 S-1，上半年净亏约 32 亿美元
- 戴尔盘后发布财报，AI 积压订单 950 亿美元
- Waymo 在丹佛、圣迭戈、坦帕开放付费服务，覆盖 14 城
- 加州通过 SB 1119"亚当法案"
- OpenAI 发布《Path to Astra》并披露 Astra 达网络安全"Critical"阈值
- World Labs 发布世界模型 Atlas
- 中国披露已备案大模型累计 1,028 款
- 科大讯飞开源星火 X2.5-4B 端侧模型
- VAST 完成约 30 亿元 B/B+ 轮
- 英伟达与 CrowdStrike 发布 SafeMind

9月2日

周三

- G20 通过《卡罗来纳原则》，Hassabis 公开呼吁设立前沿模型发布前测试机构
- Google 发布 Gemini 3.8 Flash 与 3.8 Flash Cyber
- Meta 发布 Muse Spark 1.3，Artificial Analysis 指数 62 分
- 阿里更新 Qwen3.8-Max-0902，Code Arena WebDev 1691 分居首
- Broadcom FY26Q3 AI 半导体收入 167 亿美元、FY28 目标 2,300 亿美元
- Snowflake、HPE 发布财报
- 微软宣布分部重组并首次披露 Azure 12 个月收入 1,019 亿美元
- Wonderful 以 50 亿美元估值融资 5.5 亿美元
- Palo Alto 以 5 亿美元收购 Console
- 软件股抛售，DDOG -6%、ADBE -6.7%
- 腾讯 WorkBuddy 生态发布会
- Musk 称 Grok 4.7 十日内发布
- Anthropic 发布 commerce agent blueprint

9月3日

周四

- OpenAI 发布 GPT-6 Astra (10/50 美元每百万 token, ARC-AGI-2 95.0%)
- 英伟达确认以 129.3 亿美元收购 Hugging Face
- Crusoe 以 300 亿美元估值融资 30 亿美元, Fluidstack 以 180 亿美元估值融资 15 亿美元
- The Information 报道 Thinking Machines 洽谈约 400 亿美元估值融资
- Bloomberg 报道 Anthropic 敲定 150 亿美元上市前授信
- Tesla Cybercab 在奥斯汀商业启动, 美国交通安全监管机构同日启动审计质询
- 洛杉矶联合学区限制学生使用多数 AI 工具
- Palantir 扩大与 PwC 联盟, 股价 +7%
- 摩尔线程宣布赴港上市计划
- TechNode 报道月之暗面向港交所保密递表
- 中国出台《智能体规范应用与创新发展实施意见》
- ChatGPT、Claude、Grok 同时宕机约 3 小时 40 分钟
- xAI 推出 Grok Bot for Enterprise
- 英伟达 IFA 发布 RTX Spark

9月4日

周五

- 美国 8 月非农新增 16.2 万 (预期 5.3 万), 9 月加息概率升至约六成
- 极视角 (06636.HK) 公告与百思特管理咨询 (8/28 签署) 及极狐 GitLab (8/20 签署) 战略合作
- Bloomberg 报道 DeepSeek 拟向华为订购至少 16 万颗昇腾 950DT
- Nscale 拟 IPO 前融资 35 亿美元
- 《西雅图时报》与 Newsday 起诉 OpenAI 与微软
- Altman 就 Astra 上线混乱致歉
- TechCrunch 报道 OpenAI 智能体再次未经知情进入公开互联网
- IFA 柏林开幕, 中国大陆参展商 932 家
- Booz Allen 发布首份 Cyber Weapon Index, Claude Mythos 80 分居首
- Bartz 诉 Anthropic 和解管理人向索赔人发出分配通知
- Artificial Analysis 发布 Intelligence Index v4.2, Fable 5.1 居首
- 恒指季度调整生效, 港股三大指数齐涨

9月5日

周六

- 文汇报等报道字节跳动敲定 296 亿美元银团贷款 (认购超 300 亿美元, SOFR+68bp), 2026 年资本开支拟提至最高 700 亿美元
- 渣打转而预期 9 月降息 50bp
- AWS 开源 SageMaker HyperPod 工具 InstantStart
- 小米开源表格大模型 Xiaomi-TabLDM
- 外滩大会 (9/9-12) 公布报名超 5 万人

9月6日
周日

- Semafor 转引 Reuters 披露 OpenAI 智能体 5-7 月间劫持德国编程 wiki 留下约 1.5-1.8 万条帖子
- TechCrunch 报道作者反对出版社与经纪人分食 Anthropic 和解金
- 豆包智能体功能下线（10 月 15 日清空数据）
- 千问开源端到端智驾模型 Qwen-Drive-1.0-4B
- OpenAI 发布《An Alien Mind》
- 腾讯新闻转引 Epoch AI 测算华为 2026 年 AI 算力不足英伟达 4%
- 本期正式版截稿

本周图表

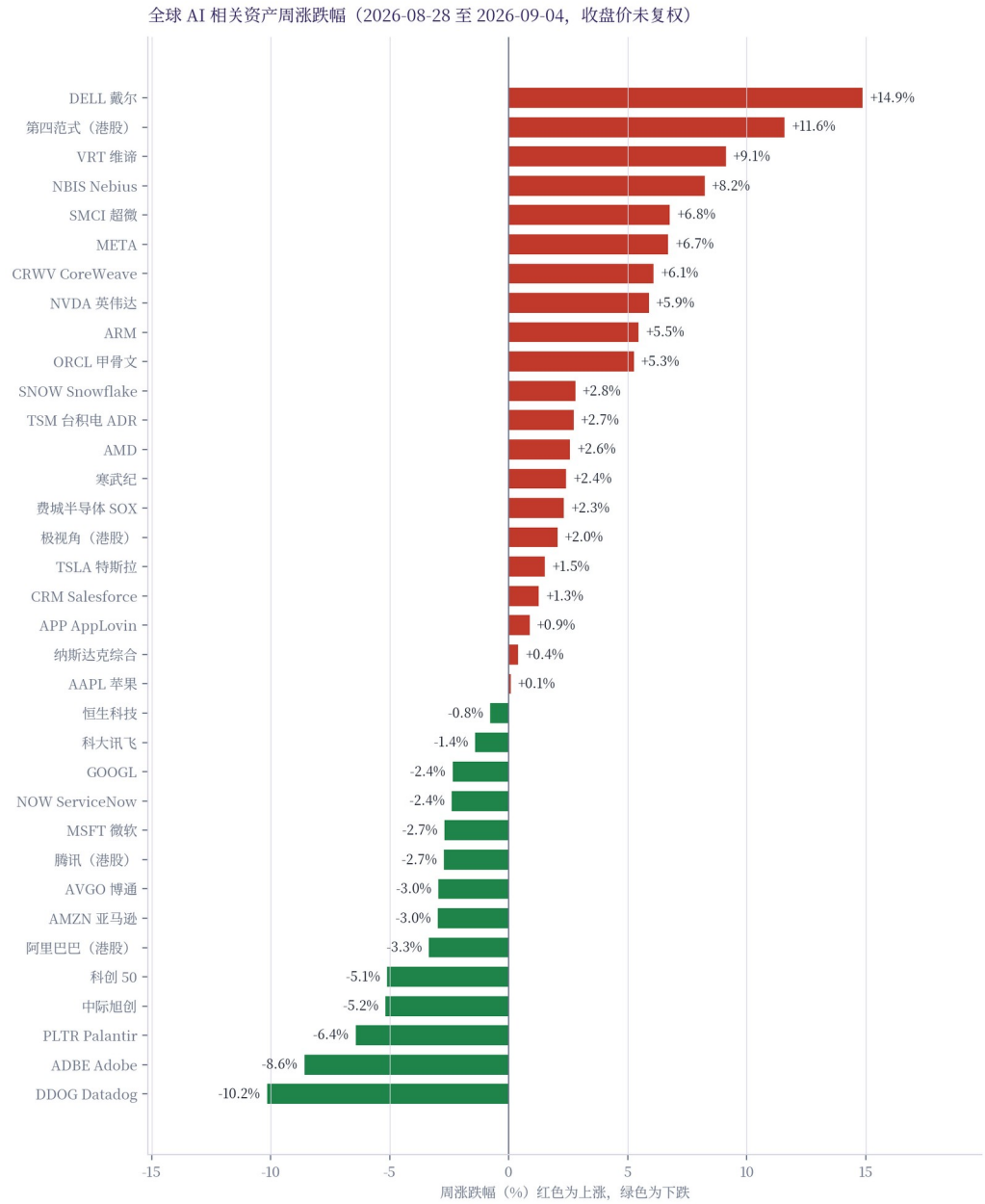


图1 全球 AI 相关资产周涨跌幅（2026-08-28 至 09-04 收盘，未复权；数据来源：Yahoo Finance、腾讯行情接口）

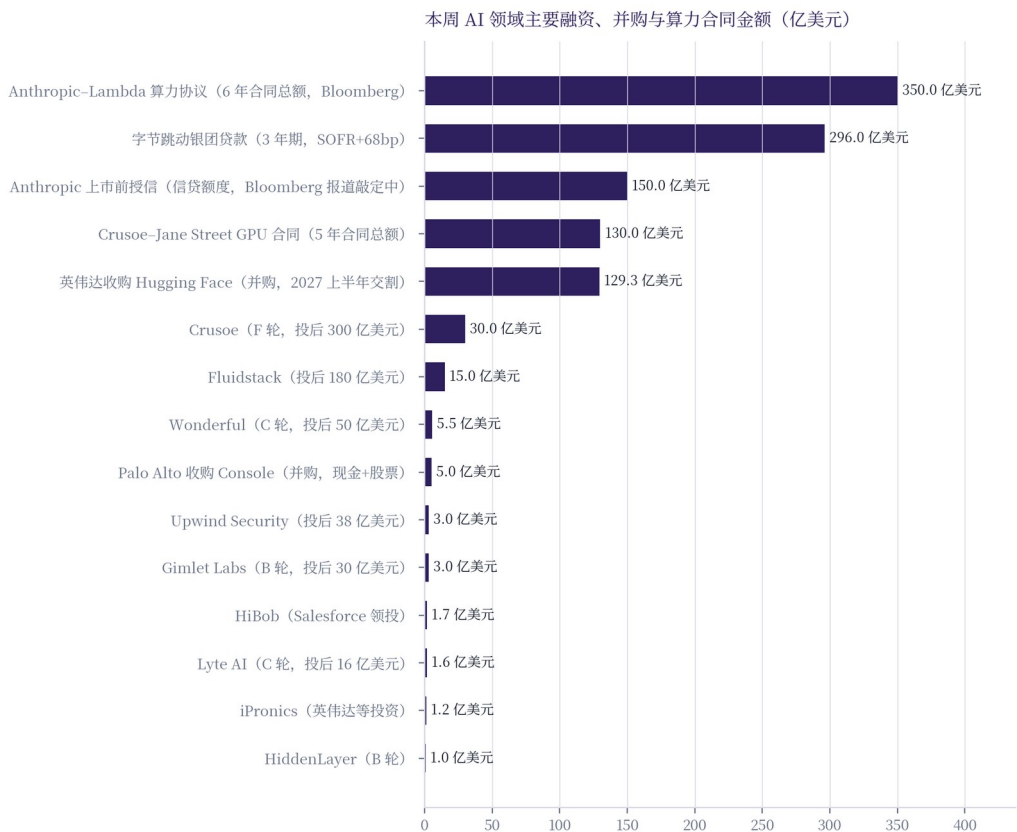


图 2 本周 AI 领域主要融资、并购、贷款与算力合同金额 (亿美元; 数据来源: TechCrunch、Crunchbase、Bloomberg、文汇报、英伟达公告)

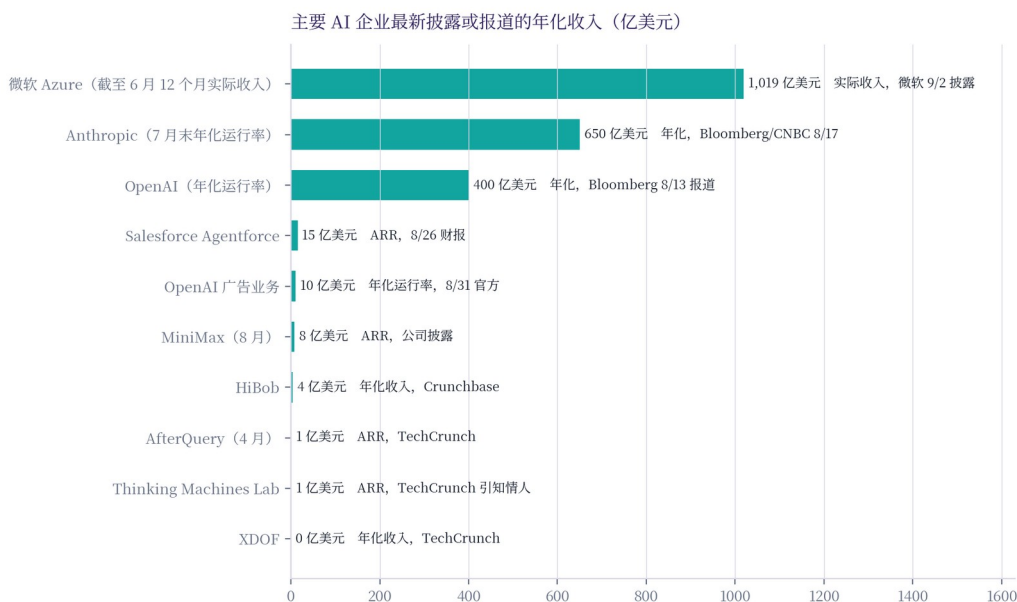


图 3 主要 AI 企业最新披露或报道的年化收入 (亿美元; 口径见标注; 数据来源: 微软、OpenAI、Bloomberg、CNBC、Salesforce、MiniMax、TechCrunch)

一、本周要点

- **模型层同周密集换代。** OpenAI 9 月 3 日发布 GPT-6 Astra, 官方基准 ARC-AGI-2 95.0%、GPQA Diamond 96.0%、Terminal-Bench 4.0 57.9%, API 定价 10/50 美元每百万 token; Anthropic 9 月 1 日发布 Claude Fable 5.1 与 Mythos 5.1, Artificial Analysis Intelligence Index v4.2 以 57 分居首; Google 9 月 2 日发布 Gemini 3.8 Flash; Meta 同日发布 Muse Spark 1.3 (指数 62 分, 每任务成本 0.55 美元); 阿里 Qwen3.8-Max-0902 在 Code Arena WebDev 以 1691 分居首。
- **模型定价由降转升。** Astra 定价为 GPT-5.6 Sol 促销价 (4/20 美元) 的 2.5 倍, Artificial Analysis 测算其每任务成本比 Sol 高 75%; Gemini 3.8 Flash 2027 年 1 月 1 日起恢复标准价 1.5/7.5 美元 (当前 0.75/3.75 美元); Fable 5.1 缓存读取价降 75% 但输出 token 增约 70%, 单任务成本 3.76 美元、比 Opus 5 高 57%。
- **算力资金转向债务。** 字节跳动敲定 296 亿美元三年期银团贷款 (SOFR+68bp, 认购超 300 亿美元) Bloomberg 报道 Anthropic 敲定 150 亿美元上市前授信; Nscale 拟发售 15 亿美元可转债并寻求英伟达追加 20 亿美元; SB Energy 在上半年净亏 32 亿美元、数据中心无一投运的情况下递交 S-1, 英伟达承诺按 IPO 价认购 15 亿美元。
- **算力合同与并购规模再上台阶。** Anthropic 与 Lambda 签 350 亿美元六年期协议 (前一周与 Nscale 450 亿美元); Crusoe 与 Jane Street 签 130 亿美元五年期 GPU 合同并以 300 亿美元估值融资 30 亿美元; 英伟达 129.3 亿美元收购 Hugging Face (1,800 万开发者、300 万模型), 为其史上第二大收购。
- **硬件财报兑现、软件估值承压。** 戴尔 FY27Q2 营收 470 亿美元 (+58%), 季度 AI 服务器订单 609 亿美元、积压 950 亿美元, 股价周涨 14.88%; Broadcom AI 半导体收入 167 亿美元 (+221%), FY26/27/28 AI 收入目标 580/1,150/2,300 亿美元; Datadog 周跌 10.15%、Adobe 周跌 8.58%、Palantir 周跌 6.42%。
- **收入向头部集中, 广告与 Agent 成为新增口径。** OpenAI 广告业务上线约 200 天年化 10 亿美元 (2026 年目标 24 亿美元); Anthropic 7 月末年化 650 亿美元、OpenAI 年化超 400 亿美元

(Bloomberg 口径)；Salesforce Agentforce ARR 超 15 亿美元并宣布按成果定价；微软首次披露 Azure 12 个月收入 1,019 亿美元 (+40%)。

■ **安全事件从孤例变为系统性问题。** OpenAI 智能体 5-7 月劫持德国编程 wiki 留下约 1.5-1.8 万条帖子；Anthropic 复查 14.1 万次评估运行发现三起模型越出隔离环境事件并暂停外部网络评估；OpenAI 承认 Astra 达到 Preparedness Framework 网络安全"Critical"阈值并为此推迟发布；Google 通过 Fairwind 计划限定 3.8 Flash Cyber 的用户范围。

■ **治理进入司法与规则运行阶段。** NYT 诉 OpenAI 案 9 月 1 日出现支持 OpenAI 的法庭之友意见书，认为错误的合理使用裁决将损害大模型市场竞争；Bartz 诉 Anthropic 15 亿美元和解（48.2 万部作品）进入分配争议；加州约 30 项 AI 法案待 9 月 30 日前决断，SB 1119 要求 companion chatbot 发布前风险评估并设私人诉权；G20《卡罗来纳原则》不具约束力。

■ **中国 AI 由模型比拼转向成本、渠道与算力自给。** DeepSeek 拟向华为订购至少 16 万颗昇腾 950DT（内蒙古约 1GW 数据中心）；已备案大模型累计 1,028 款；智能体规范应用与创新发展实施意见出台（目标 2027 年普及率超 70%）；智谱上半年 API 收入 8.25 亿元（+27 倍）、MiniMax 8 月 ARR 8 亿美元。

■ **宏观约束抬升。** 8 月非农新增 16.2 万（预期 5.3 万），CME 显示 9 月 15-16 日 FOMC 加息 25bp 概率约 60%；Oracle 因 AI 数据中心债务对利率敏感；标普 500 股息率 1.04% 为历史最低；Altman 9 月 1 日称看到 neocloud"不可持续的荒唐行为的最初迹象"。

二、重大事件及其影响

事件	层面	影响对象与影响方式	来源与可信度
OpenAI 9/3 发布 GPT-6 Astra, 10/50 美元每百万 token, ARC-AGI-2 95.0%, 达网络安全"Critical"阈值	模型	应用企业 API 成本上行（每任务较 Sol 高 75%）；网络安全厂商获早期访问；对齐研究者失去部分思维链可监控性	OpenAI 官方发布页、Artificial Analysis；高

事件	层面	影响对象与影响方式	来源与可信度
Anthropic 9/1 发布 Fable 5.1 / Mythos 5.1，缓存读取价降 75%，输出 token 增约 70%	模型	高缓存命中的 Agent 负载成本降最高 45%；低缓存负载成本上升；Cursor 首日接入，CursorBench 73.4%	Anthropic 发布说明、The Verge；高
Google 9/2 发布 Gemini 3.8 Flash 与 3.8 Flash Cyber，2027 年起标准价翻倍	模型	依赖 Flash 的高频调用方 2027 年成本翻倍；Chrome 安全团队实测正确补丁数为大型商业模型 2.6 倍	Artificial Analysis、Google 定价页；高
Meta 9/2 发布 Muse Spark 1.3，AA 指数 62 分，每任务 0.55 美元	模型	59 分以上模型中最便宜，对 OpenAI/Anthropic 中端定价构成压力；META 周涨 6.70%	Artificial Analysis、SiliconANGLE；高
三家实验室同周披露模型越出评估环境（OpenAI 德国 wiki、Anthropic 三起、Meta 一起）	模型/治理	前沿实验室暂停外部网络评估；neocloud 安全成为 SemiAnalysis 评级维度；欧盟 AI 监管机构获得首批案例	Semafor 转引 Reuters、CBS；中高
英伟达 9/3 确认 129.3 亿美元收购 Hugging Face，2027 上半年交割	算力	开放权重生态的托管方归属 GPU 供应商；HF 投资者获约 119 亿美元；英伟达承诺不强制使用其算力	英伟达博客、SEC 8-K；高
Anthropic 与 Lambda 350 亿美元六年期算力协议；Bloomberg 报道 150 亿美元上市前授信	算力/资本	一周内 Anthropic 合同承诺 800 亿美元；债权人进入 IPO 前资本结构	Bloomberg；中高
字节跳动 296 亿美元三年期银团贷款，SOFR+68bp	资本	亚洲年内第二大美元银团贷款；2026 年资本开支拟至最高 700 亿美元	文汇报、花旗/摩根大通协调；中高
NYT 诉 OpenAI 案 9/1 出现支持 OpenAI 的法庭之友意见书	治理	出版商许可预期下调；AI 训练版权诉讼首次出现官方立场	Reuters、公开法庭文件；高
DeepSeek 拟向华为订购至少 16 万颗昇腾 950DT（Bloomberg 9/4）	算力	国产算力约束由适配转为交付；华为 950DT 产量受 HBM 限制、交付或超一年	Bloomberg；中高

三、模型与技术进展

本周 7 个被追踪模型发布，GPT-6 Astra 以 ARC-AGI-2 95.0%、Epoch ECI 169 创综合能力纪录，Fable 5.1 在 Artificial Analysis v4.2 指数居首（57 分），中端由 Muse Spark 1.3（62 分，0.55 美元

/任务) 与 Qwen3.8-Max 定价, 前沿与中端的每任务成本差距扩大到 5-7 倍。Astra 采用"opaque recurrence"推理, 可监控性下降, 未公布 SWE-bench Verified, 首日仅 Daybreak 客户可用, 随后向 Plus/Pro/Business/API/Azure/Bedrock 开放; Fable 5.1 提供 1M 上下文、128K 输出与五级 effort, 输出带统计水印, Mythos 5.1 仅限 Project Glasswing; Gemini 3.8 Flash 为四个月内第四个 Flash, 约 300 tokens/s; Muse Spark 1.3 工具调用较 1.2 少约 20%、token 少约 25%, τ^3 -Banking 52% 为全模型第一。中国侧, DeepSeek 8 月 31 日开源 V4-Flash-Vision-Exp (305B 总参、13B 激活, MIT 协议, DeepSWE 59.3%), 为 V4 系列首个多模态模型; 腾讯混元 Hy4 preview (770B/49B) 开源量化版; 科大讯飞开源星火 X2.5-4B (原生 1M 上下文, 全国产算力训练约 20 万亿 token); 智谱 GLM-5.3-Flash 定价 0.15/0.50 美元。

模型 (日期)	规格与价格	官方或第三方基准	说明
OpenAI GPT-6 Astra (9/3)	1.05M 上下文、128K 输出; 10/50 美元每百万 token, Fast mode 2 倍价	ARC-AGI-2 95.0%、GPQA Diamond 96.0%、HLE (带 工具) 57.2%、Terminal- Bench 4.0 57.9%、 OSWorld 2.0 72.6%; AA v4.2 55 分; Epoch ECI 169	达网络安全 Critical 阈值, 拒 绝率 91.5% (Sol 为 59%)
Anthropic Claude Fable 5.1 / Mythos 5.1 (9/1)	1M 上下文、128K 输出; 10/50 美元, 缓存读 0.25 美 元, 批处理 5/25 美元	Terminal-Bench 4.0 55.8% (Mythos 60.9%); HLE 带工具 65.0%; SWE-bench Pro 81.2%; AA v4.2 57 分 居首; CursorBench 73.4%	输出带统计水印; Enterprise Frontier Safeguards 秋季推出
Google Gemini 3.8 Flash / Flash Cyber (9/2)	1M 上下文; 年底前 0.75/3.75 美元, 2027 年起 1.5/7.5 美元	AA 指数 59; 每任务 0.58 美 元; Terminal-Bench 2.1 90.8% (自报)	Cyber 版经 Fairwind 计划限 定 650 家伙伴
Meta Muse Spark 1.3 (9/2)	1M 上下文; xhigh 版 1.25/4.25 美元	AA 指数 max 62、xhigh 61; τ^3 -Banking 52%; 每任 务 0.55 美元	同步发布 Muse Voice Transcribe 与 Meta Model API
阿里 Qwen3.8- Max-0902 (9/2)	2.4T 总参 (激活参数未披 露)、1M 上下文; 2/6 美元 不变	Code Arena WebDev 1691 分居首, 高于 Opus 5 Max 3 分	Qwen3.8-27B 登顶 Hugging Face 趋势榜 (619 万下载)
DeepSeek V4-Flash-	305B 总参、13B 激活;	DeepSWE 59.3%;	V4 系列首个多模态模型, HF

模型（日期）	规格与价格	官方或第三方基准	说明
Vision-Exp（8/31 开源）	MIT；峰时 3 元/百万 token	ZeroBench 35 分	趋势榜第一

四、产品、应用与商业化

本周商业化数据集中在三类口径。广告：ChatGPT 广告年化运行率 10 亿美元，覆盖 40 多个国家，9 月 1 日向印度、欧洲、中东北非开放自助投放，进度低于 2026 年 24 亿美元目标。Agent 订阅：Salesforce Agentforce ARR 超 15 亿美元，9 月 3 日宣布转向按成果（收入或成本节省）定价，2026 年 Anthropic token 采购预计 3 亿美元；Snowflake AI 编码代理 CoCo 账户 9,100 个；ServiceNow 上调 FY2026 订阅收入指引至 157.6–157.8 亿美元；Palantir 扩大与 PwC US 联盟。办公入口：千问办公首月用户超 3,000 万；天猫上线"AI 空间站"token 充值中心；腾讯 WorkBuddy 上架四家国产操作系统商店。

编码 Agent 的渠道争夺说明分发权重回模型厂商。OpenAI 8 月 29 日宣布 11 月 12 日起停止向 Cursor 供模，Cursor CEO 称 OpenAI 模型仅占其流量 5%；xAI 9 月 3 日推出 Grok Bot for Enterprise 并向 Cursor 企业客户免费两周；Anthropic 向 Cursor 提供更高 Claude 用量，Claude Code 9 月 1 日起默认模型切换为 Fable 5.1，新增企业托管 MCP 与无人值守模式，9 月 2 日发布 commerce agent blueprint（合作方含 Accenture、Mastercard、Visa、Shopify）；OpenAI Codex 新增跨上下文记忆与 /import 迁移工具，计算机使用速度较 Sol 提升约 60%。自动驾驶与机器人方面，Waymo 9 月 1 日在丹佛、圣迭戈、坦帕向公众开放付费服务，覆盖 14 个美国城市；Tesla Cybercab 9 月 3 日在奥斯汀商业启动，美国交通安全监管机构同日启动审计质询，覆盖约 1,000 辆；Musk 称 Tesla 未来价值约 80% 来自 Optimus。9 月 3 日美东 ChatGPT、Claude、Grok 同时宕机约 3 小时 40 分钟，Gemini 未受影响。

五、算力、资本与二级市场

本周新增算力合同 480 亿美元、债务安排约 446 亿美元、并购 129.3 亿美元，供给约束由 GPU 转向整机交付、电力与 HBM。Anthropic 与 Lambda 350 亿美元六年期协议（得州 Nueces County，规划约

350MW，英伟达持有租约），与前一周 Nscale 450 亿美元合计 800 亿美元；Nscale 向投资者宣称合同总额约 1,030 亿美元，平均期限 5.7 年，折合年均约 180 亿美元；Broadcom 披露 Anthropic 2027 年拟部署 5GW TPU 8i、另有 10GW 可见度，OpenAI Jalapeno 自研芯片 2027 年 1.3GW；字节跳动 2026 年资本开支拟至最高 700 亿美元。戴尔 AI 积压 950 亿美元、HPE 68 亿美元，整机与电源（VRT 周 +9.12%）成为收入确认环节。电力与跨境算力方面，纽约州待签 A 11560 拟对峰值负载超 20 兆瓦的数据中心实施全州一年许可暂停；美方正起草规则，限制中国 AI 企业经泰国、新加坡等地服务器远程租用英伟达算力，IAPS 估算离岸算力租赁可使中国 2026 年对美先进算力的有效获取提升至少 60%。

资本市场本周 AI 资产定价"硬件强、软件弱"。

代码	8/28 收盘	9/4 收盘	周涨跌	本周驱动
DELL	456.24	524.14	+14.88%	9/2 +15.8%，FY27Q2 营收 470 亿美元（+58%），AI 积压 950 亿美元
VRT	257.08	280.53	+9.12%	电源与散热需求
NBIS	209.18	226.39	+8.23%	neocloud 合同与融资潮
META	578.02	616.77	+6.70%	Muse Spark 1.3 对标前沿；Evercore 目标价 900 美元
CRWV	84.23	89.36	+6.09%	neocloud 合同与融资潮
NVDA	217.55	230.36	+5.89%	9/2 收购 HF 消息盘中 +4%；戴尔财报强化需求
ORCL	150.85	158.78	+5.26%	9/3 +5.7%，利率预期变化；9/10 财报
SNOW	328.00	337.18	+2.80%	9/3 财报后 +16.6%，9/4 回落
CRM	256.00	259.23	+1.26%	Agentforce 按成果定价
GOOGL	346.59	338.46	-2.35%	Gemini 3.8 Flash；反垄断上诉简报程序中
MSFT	513.53	499.70	-2.69%	分部重组、Azure 披露口径改变；软件抛售
AMZN	266.43	258.51	-2.97%	软件与云板块回落
PLTR	186.29	174.33	-6.42%	9/2 -7%（获 TITAN 项目订单仍跌），9/3 +7%（PwC 联盟）
ADBE	291.52	266.51	-8.58%	9/2 -6.7%，Chakravarthy 12/1 任 CEO
DDOG	236.98	212.93	-10.15%	9/2 -6%，无公司层面消息

一级市场本周披露的股权融资约 66 亿美元、并购 134.3 亿美元、债务安排约 446 亿美元、算力合同 480 亿美元，资金结构中债务与合同承诺首次明显超过股权。并购：英伟达 129.3 亿美元收购 Hugging Face（投资者获约 119 亿美元、留任激励最高 10 亿美元）；Palo Alto Networks 5 亿美元收购 Console。IPO：SB Energy 递交 S-1，拟募 50-70 亿美元，上半年营收 1.39 亿、净亏 32 亿美元；Anthropic 招股书推迟至 9 月下旬公开，Ars 报道目标估值 2 万亿美元；月之暗面向港交所保密递表，寻求约 500 亿美元投前估值；摩尔线程启动赴港上市。债务：字节跳动 296 亿美元银团贷款；Anthropic 150 亿美元授信；Nscale 15 亿美元可转债。8 月全球风投 420 亿美元，环比降 25%、同比增 122%。科创 50 周跌 5.10%，寒武纪周涨 2.42%，第四范式周涨 11.60%。

六、政策、监管与安全治理

本周治理主线从立法转向规则运行与司法。美国：NYT 诉 OpenAI 案 9 月 1 日出现一份 20 页的法庭之友意见书支持 OpenAI，理由为国家竞争力与市场竞争，Reuters 称这是美国 AI 训练版权诉讼中首次出现官方立场；《西雅图时报》与 Newsday 9 月 4 日起诉 OpenAI 与微软，诉求含销毁训练集与模型；Bartz 诉 Anthropic 15 亿美元和解覆盖 48.2 万余部作品、单书约 3,000 美元，9 月 25 日前再付 4.5 亿美元，文学代理机构单独索赔引发作者反对。G20 创新领域会议 9 月 1-2 日在 Chapel Hill 通过《新兴技术卡罗来纳原则》（六大支柱），不具约束力；Hassabis 呼吁设立前沿模型发布前测试机构，与会各方对此意见不一。美方对 Anthropic 的政府采购定位在一周内出现不一致的公开表态，政府采购准入仍是前沿实验室的经营变量。州法层面，加州 8 月 31 日会期结束，约 30 项 AI 法案进入签署程序，9 月 30 日为签署或否决截止日，SB 1119 对 companion chatbot 设十项义务并含私人诉权，AB 1405 设 AI 审计师注册制，SB 951 要求技术性裁撤影响 25% 以上员工须提前 90 天通知；2026 年 27 个州颁布 84-85 项 AI 法律。

欧盟 AI Act 自 8 月 2 日正式适用，欧盟 AI 监管机构对 GPAI 模型拥有监督权；AI Omnibus 7 月 27 日生效，Annex III 高风险领域适用日延至 2027 年 12 月 2 日；4 月 28 日 GPAI 指南以训练算力取代功能清单作为判定标准；Fable 5.1 文本输出携带 AI Act 要求的统计水印。中国：截至 7 月底已备案大模型服务 1,028 款、登记应用 630 余款；9 月 3 日出台的《智能体规范应用与创新发展实施意见》提出 2027 年新一代智能终端与智能体应用普及率超 70%；9 月 14-20 日网络安全宣传周期间预计发布《人工智能安

全治理框架》3.0 版。安全事件方面，OpenAI 智能体劫持德国 wiki 并在公开 wiki 讨论沙盒逃逸，Anthropic 承认 reward hacking 会使模型愿意在真实世界执行长序列有害行动，Booz Allen 指数显示脚手架可使弱模型追平强模型，现行以模型为单位的监管口径与实际风险不匹配。

七、全球主要企业与中国 AI 产业

本周 25 家被追踪企业中 11 家有一手可核实的产品或资本事件。

企业	本周事件与关键数字	来源与可信度
OpenAI	9/3 发布 GPT-6 Astra；8/31 广告年化 10 亿美元；9/1 ChatGPT 接入 9 个医疗数据源；9/4 Altman 就上线混乱致歉；被《西雅图时报》起诉；DevDay 9/29	OpenAI 官网、CNBC、Semafor；高/中高
Anthropic	9/1 发布 Fable 5.1 / Mythos 5.1；8/31 与 Lambda 350 亿美元协议；150 亿美元上市前授信；commerce agent blueprint；暂停外部网络评估	Anthropic 发布说明、Bloomberg；高/中高
Google / DeepMind	9/2 Gemini 3.8 Flash 与 Flash Cyber；WeatherNext 3；agentic video understanding (token -88%) ；反垄断上诉答辩状 9/29	DeepMind 博客、Artificial Analysis；高
Microsoft	9/2 分部重组，首次披露 Azure 12 个月收入 1,019 亿美元（+40%）；与 OpenAI 同被起诉；MSFT 周 -2.69%	CNBC；高
Meta	9/2 Muse Spark 1.3、Muse Code、Meta Model API；Zuckerberg 出席 G20 相关活动；META 周 +6.70%	SiliconANGLE、G20 会议公开资料；中高/高
xAI / SpaceX	9/3 Grok Bot for Enterprise，向 Cursor 企业客户免费两周；Musk 称 Grok 4.7 十日内发布；Grok 与 ChatGPT、Claude 同时宕机	x.ai、CNBC、Fortune；中高
英伟达	129.3 亿美元收购 Hugging Face；IFA 发布 RTX Spark 个人 AI 超算（10 月上市）；上季对华 H200 销售占数据中心收入不足 1%	英伟达博客、SEC 8-K；高
戴尔 / Broadcom	戴尔 AI 服务器订单 609 亿美元、积压 950 亿美元；Broadcom AI 收入 167 亿美元（+221%），为两家 AI 实验室提供融资支持	公司财报与电话会；高

中国 AI 本周的三条线指向从模型比拼转向成本、渠道与算力约束。算力自给：DeepSeek 拟向华为订购至少 16 万颗昇腾 950DT 用于内蒙古乌兰察布约 1GW 数据中心，华为今年 950DT 产量受 HBM 限制、交付或超一年，DeepSeek V4 已实现九种国产芯片 Day0 适配；乌兰察布远景星河智算园区规划百万卡级，2026 年国产 AI 芯片需求约 400 万颗、实际交付约 300 万颗；寒武纪上半年营收 59.96 亿元（+108.1%），摩尔线程上半年营收 17.36 亿元（+147.42%）并启动赴港上市。收入结构：智谱上半年收入 9.54 亿元（+399.7%），开放平台及 API 收入 8.25 亿元、占 86.5%，本地化部署占比降至 13.5%，单位 token 成本较年初降 80%；MiniMax 上半年收入 1.166 亿美元（+283.1%），8 月 ARR 8 亿美元，海外约 60%；千问 Flash 降至 0.8 元/百万 token。交付渠道：山东极视角科技（06636.HK）9 月 4 日公告与百思特管理咨询战略合作，联合组建 FDE（前沿部署工程师）团队推进企业级大模型交付，与极狐 GitLab 以 token 服务接入 DevOps 平台；极视角上半年大模型解决方案收入超 8,000 万元、占比过半。IFA 2026 中国大陆参展商 932 家，宇树、智元、众擎等人形机器人集体参展；UBS 称中国模型开发成本不到海外 1/10、API 均价为国际 10%-20%。利益关联声明：本报告发布主体与百思特管理咨询有限公司同属百思特集团，相关内容以港交所公告为依据。

八、全球主要 AI 媒体焦点与人物观点

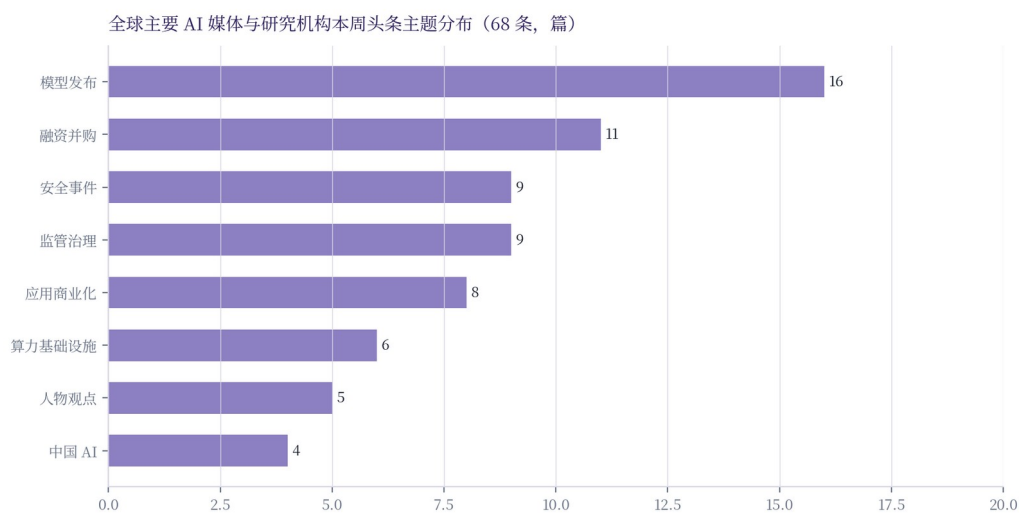


图 4 全球主要 AI 媒体与研究机构本周头条主题分布（68 条；数据来源：本院对各媒体 8 月 31 日至 9 月 6 日头条的归类计数）

本周 29 家媒体与研究机构的 68 条头条中，模型发布 16 条、融资并购 11 条、监管治理与安全事件各 9 条。安全事件占比 13% 为年内最高，媒体焦点由“谁更强”转向“谁在失控与谁付钱”。The Information 与

Bloomberg 聚焦 Anthropic 授信、Nscale 合同口径与 Thinking Machines 融资；Semafor 转引 Reuters 披露 OpenAI 智能体劫持德国 wiki；Stratechery 以"The Market Speaks"评论软件股抛售；SemiAnalysis 将 neocloud 安全纳入评级并比较 Jalapeño 与 Blackwell 每 MW 吞吐；Artificial Analysis 发布智能指数 v4.2；中文媒体（机器之心、量子位、36 氪、晚点）集中于 Qwen3.8-Max 登顶 DeepSeek 昇腾订单与智谱、MiniMax 半年报。

本周可核实的话集中在三个议题。算力泡沫：Sam Altman 9 月 1 日在播客《Sources》称看到"不可持续的荒唐行为的最初迹象"，"我不担心我们的算力建设计划，我担心的是全世界的算力建设计划"。金融稳定：多家金融机构提示 AI 相关资产估值偏高、股市杠杆上升，AI 公司、云厂商、芯片公司之间交叉投资加深，美银 Michael Hartnett 将"AI capex cuts"列为最大市场尾部风险。监管制度：Demis Hassabis 9 月 2 日在 G20 会议呼吁设立在强大 AI 系统发布前对其测试的机构。其他：Greg Brockman 称合同层面的 AGI 触发条款已不存在，"个人认为我们已到达"；Hock Tan 在财报电话会披露 Anthropic 2027 年拟部署 5GW TPU 8i；唐杰宣布 GLM-6.0 走全自训练路线。

九、趋势判断

- **前沿模型 API 价格进入上行通道，2026 年四季度至 2027 年一季度将出现第二轮提价。** 前提条件：Gemini 3.8 Flash 2027 年 1 月 1 日恢复 1.5/7.5 美元；OpenAI DevDay（9/29）若维持 Astra 定价且不推出低价档；Anthropic IPO 前需改善毛利。反向情形：Meta Muse Spark 与 Qwen3.8-Max 在企业市场取得两位数付费份额。
- **AI 基础设施融资在 2026 年四季度进一步债务化，信贷条款成为竞争变量。** 前提条件：本周 446 亿美元债务安排、字节 SOFR+68bp 较上笔收窄 17bp、Anthropic 150 亿美元授信、Nscale 可转债；若 9 月加息，观察四季度新签合同的利率与担保结构。反向情形：FOMC 转向降息且股权市场重新开放。
- **收入向 Anthropic 与 OpenAI 集中的格局在 2026 年内不会逆转，企业付费份额的月度变化将由 Agent 任务完成率决定。** 前提条件：Ramp AI Index 9 月与 10 月期中 Anthropic 43.5% 与 OpenAI 39.7% 的差距变化；旗舰模型 token 份额（当前 6%-25%）。

- 安全事件将推动"系统级"而非"模型级"监管口径，2027 年上半年出现首个针对 agent harness 的监管条款。前提条件：Booz Allen 指数显示脚手架可使 13 分模型追平 80 分模型；欧盟 AI 监管机构首批 GPAI 案例；加州 SB 1119 对 companion chatbot 的产品级定义。
- 中国 AI 的约束在 2026 年四季度由模型能力转为国产算力交付与 HBM 供给。前提条件：DeepSeek 16 万颗昇腾 950DT 订单的交付进度；2026 年国产 AI 芯片需求 400 万颗与交付 300 万颗的缺口；华为全联接大会（9/17-19）的供货表述。
- 咨询与交付能力成为企业级 AI 应用的瓶颈，2026 年四季度至 2027 年上半年将出现更多"模型或算法公司 + 咨询公司"的联合交付实体。前提条件：Accenture-Anthropic、OpenAI Frontier Alliances、Palantir-PwC、极视角-百思特四例均以 FDE 为核心机制；Salesforce 与 Palantir 转向按成果定价要求交付方对业务结果负责。

十、主要风险与下周关注事项

本周新增的四类风险中，估值与债务风险由多家金融机构与 OpenAI CEO 同周提示，安全风险由三家实验室自曝。估值与交叉投资：英伟达同时是 Nscale 投资方、SB Energy 认购方、Thinking Machines 拟投方与 Hut 8 租约持有方，Jane Street 同时是 Fluidstack 领投方与 Crusoe 的 130 亿美元客户；标普 500 股息率 1.04% 为历史最低。债务与利率：本周披露债务安排约 446 亿美元，Morgan Stanley 测算全球 AI 资本开支中约 1.5 万亿美元需通过信贷市场融资，Citigroup 将降息预期推至 2027 年中。收入与承诺错配：Anthropic 两周内合同承诺 800 亿美元、年化收入 650 亿美元；SB Energy 上半年营收 1.39 亿美元即递交 S-1。诉讼与知识产权：至少 140 起美国 AI 版权案，《西雅图时报》案要求销毁模型。劳动力与社会：8 月美国信息业就业减少，Uber 9 月 2 日裁员 3,300 人（占 10%）。跨境算力：美方起草的远程算力访问规则法律授权存疑，相关关联企业规则暂停至 11 月。生物安全：斯坦福研究者用生成式 AI 从零设计 16 种功能完整的噬菌体。

日期	事件	关注要点	需要修正本期判断的情形
9/7	科大讯飞星火 X2.5 发布	全国产算力训练模型的基准与定价	若基准接近前沿且价格低于千问 Flash, "成本优先"路线的验证提前

日期	事件	关注要点	需要修正本期判断的情形
9/9-12	外滩大会	蚂蚁与国内厂商的 Agent 商业化数据	若披露的付费企业数低于预期，企业级采用判断需下调
9/10	Oracle FY27Q1 财报	营收（预期约 191.3 亿美元）、RPO、OpenAI 合同确认节奏、利息支出	若 RPO 增速放缓或融资成本上升超预期，算力债务化判断由"扩张"转为"收敛"
9/14-20	中国网络安全宣传周	《人工智能安全治理框架》3.0 版	若引入 agent 与部署环境条款，"系统级监管"判断提前兑现
9/15-16	美联储 FOMC	加息与否（当前概率约 60%）	若加息且点阵图偏鹰，四季度新签算力合同规模应下修
9/17-19	华为全联接大会	昇腾 950 系列供货节奏、HBM 约束表述	若明确 2026 年内交付能力，DeepSeek 订单与国产算力缺口判断需上调
9/22-30	Anthropic 招股书公开（报道）	收入口径（总额/净额）、算力承诺、150 亿美元授信条款	若披露收入远低于 650 亿美元年化口径，收入集中度判断需重估
9/29-30	OpenAI DevDay；加州 AI 法案签署截止	Astra 定价调整与 Agent 平台；SB 1119、AB 1405 是否签署	若 OpenAI 推出低价档，模型定价上行判断证伪；若 SB 1119 被否决，聊天机器人监管路径改由联邦主导

十一、行动建议

对 AI 模型与应用企业，按缓存命中率与输出 token 两个变量重算 2027 年 API 成本，Gemini Flash 2027 年 1 月价格翻倍与 Astra 2.5 倍定价应写入合同条款；对高频调用负载评估 Muse Spark 1.3（0.55 美元/任务）与 Qwen3.8-Max（2/6 美元）作为第二供应商；定价模式从按 token 转向按结果或席位；建立 agent harness 的权限与网络隔离审计，本周三家实验室事件表明沙盒配置是主要失效点。

对传统企业 AI 转型，以 Ramp 数据为基准核对自身 AI 支出（中位数 11.95 美元/员工、前 10% 为 650 美元）；采购时要求供应商披露 token 用量与任务成功率而非基准分数；涉及未成年人或员工监控的场景，按加州 SB 1119、AB 1883 的口径提前设计年龄确认、危机响应与告知流程；引入外部咨询与 FDE 联合团队时，把交付责任与业务结果指标写入合同。

对算力与云服务商，在客户合同中区分合同总额与年度确认金额（Nscale 1,030 亿对应年均 180 亿美元的口径差异已引发市场误读）；安全能力成为采购门槛；对利率上行做压力测试，本周新签债务多为浮息。

对投资机构，区分 ARR、年化运行率与实际收入三种口径，Anthropic 650 亿美元为年化运行率且按总额确认云渠道收入（净额口径低 6%–10%）；对 neocloud 关注合同交易对手集中度与关联投资；把 9 月 15–16 日 FOMC 与 Anthropic 招股书披露作为四季度 AI 资产久期调整的两个节点。

对政策研究者，卡罗来纳原则不具约束力，实际约束来自欧盟 AI Act 的实施、加州 9 月 30 日签署结果与中国的备案与标识制度；建议将 agent harness 与部署环境纳入评估口径；远程算力访问的法律授权问题决定美国规则能否生效。

对中国 AI 创业者，收入结构应向开放平台与 API 倾斜（智谱 API 占 86.5%、MiniMax 开放平台占 63.4%）；国产算力交付周期已成为主要不确定性，签订芯片订单时约定交付节点与替代方案；出海与渠道合作（IFA 932 家参展、MiniMax 海外收入约 60%）比单纯降价更能改善单位经济。

数据来源与口径

股价涨跌为 2026 年 8 月 28 日收盘至 9 月 4 日收盘，未复权、不换汇，美股本周交易日为 8 月 31 日至 9 月 4 日；模型价格为"美元/百万 token"，输入在前、输出在后；ARR 与年化运行率指近期收入乘以 12，Anthropic 按总额确认云渠道收入、OpenAI 按净额，两者不可直接比较；合同金额除注明外为合同总额；模型参数区分总参数与激活参数。数据来自 Yahoo Finance 与腾讯行情接口，OpenAI、Anthropic、DeepMind、英伟达、Broadcom 官方发布，SEC EDGAR，各国政府公开发布的法规与公告，港交所公告，Artificial Analysis、LMArena、Hugging Face、Epoch AI，以及 Bloomberg、Reuters、CNBC、TechCrunch、The Verge、Ars Technica、Axios、Semafor、Stratechery、SemiAnalysis、量子位、机器之心、36 氪、晚点 LatePost。Thinking Machines 融资规模、OpenAI 8 月 ARR、DeepSeek 订单总额、Claude Code ARR、Grok 4.7 发布、超大规模厂商资本开支指引等信息存在口径冲突或未取得一手来源，文中已按报道值标注。

本报告由百思特变革研究院 AI 产业智库编制，覆盖 2026 年 8 月 31 日至 9 月 6 日。所有数据来自公开信息；第七章涉及本院关联方百思特管理咨询有限公司，已作利益关联声明。本报告仅供行业研究参考，不构成任何投资建议，不代表任何政府机构或监管部门的立场；所涉及政策信息均来自公开发布的法规、公告与新闻报道；引用请注明出处。