

国产开源模型竞速升温;DeepSeek推视觉Agent;Claude全线加水印

头部企业动态

- Anthropic自8月2日起在全线Claude产品输出中嵌入机器可读文本水印,以符合欧盟AI法案第50(2)条,并于8月14日发官方博客《How Claude's text watermark works》解释原理。 **A • 96% 已核实**
- OpenAI于8月1日宣布未公开发布的下一代前沿模型Astra在内部解决了数学与理论计算机科学的10个长期开放问题(含非sofic群存在性构造、球堆积新界)。 **B- • 80% 同源引述**
- Anthropic于8月16日21:58 UTC起发生大规模服务中断,Claude.ai、Claude Code、Claude Cowork多项服务受影响,官方已确认故障。 **B • 60% 待复核 • 链接探活失败**
- OpenAI于8月17日任命Dali Rajic为首席营收官(CRO),被外界视为商业化与后续资本运作提速的信号。 **B • 71% 待复核**

模型发布

- DeepSeek于8月13日正式发布DeepSeek-V4-Pro(0813)正式版,强化Agent能力(Terminal Bench 2.1达87.9),支持1M上下文,并自8月16日起实施峰谷分时计费,谷时价格为峰时一半。 **A • 97% 已核实**
- 智谱8月14日发布新一代基座模型GLM-5.3,编程能力较GLM-5.2提升50%,在Terminal Bench 3.0等基准居开源第一,CyberGym网络安全评测84.5%;API已于8月19日上线,权重定于8月28日开源。 **B+ • 91% 多源一致**
- 阿里8月3日发布迄今最大模型Qwen3.8-Max:2.4万亿参数MoE、1M上下文、原生多模态,经阿里云Model Studio API全球开放,并宣布近期释出开源权重版本。 **B+ • 90% 多源一致**
- Google DeepMind于8月13日将Gemini 3.7 Flash转正式版(GA),定位'最智能的编码与Agent主力模型',在软件工程与Agent工作流上有大幅提升,优惠定价持续至2026年底。 **A • 96% 已核实**

热点与趋势

- 8月被业内称为'大模型疯狂八月':20天内11+款模型密集发布,GLM-5.3(8月28日开权重)、Qwen3.8-Max开源在即,中国厂商以开源高端模型对标Claude/GPT,开源竞速成为主叙事。 **B- • 64% 多源一致 • 链接探活失败**
- DeepSeek于8月21日推出实验模型V4-Flash-Vision-Exp,为文本旗舰V4 Flash补上视觉理解,多模态Agent基准(Terminal Bench 2.1达83.9)显著提升,Bloomberg称其对标Anthropic Opus 4.8。 **A • 96% 已核实**
- 低价与分时定价加速AI推理'逼近零价'竞赛:Axios 8月1日报道DeepSeek廉价模型冲击行业定价,而其8月中旬又引入峰谷计费(峰时输出\$3.96/百万token),推理定价从固定价走向电力式浮动价。 **B • 61% 待复核 • 链接探活失败**
- Agent基础设施成新战场:Cloudflare于8月下旬发布面向AI Agent的浏览器运行时Kitesurf,跑在Workers平台上,CPU/内存开销约为Chromium的1/3至1/7,通过23.5万+项Web平台测试。 **B • 72% 待复核**

主动剔除(未核实,不外发)

- OpenAI'正筹备史上最大IPO'的说法仅见聚合媒体转述,无SEC文件或官方确认,剔除。
- 月之暗面Pre-IPO估值传闻互相矛盾(知乎称500亿美元、另有事实核查帖称300亿美元),无一手源,剔除。
- ChatGPT'突破10亿活跃用户'仅见聚合站单源,OpenAI官网403未能核实,剔除。
- 字节Seed'8月5日发布SeedRealtime全双工大模型'仅见搜狐自媒体账号单源,未找到官方发布页,剔除。
- Meta Muse Glimmer(Apache 2.0开源)与Muse Spark 1.2权重开源消息仅见聚合站,未能定位Meta官方一手源,剔除。

信息源清单(可审计)

条目	等级	可信度	探活	来源
Anthropic自8月2日起在全线Claude	A	96%	✓活	anthropic.com/news...
OpenAI于8月1日宣布未公开发布的下一代前沿	B-	80%	✓活	imfounder.com/science-tech/ai/ai-updates-augus...
Anthropic于8月16日21:58 UTC	B	60%	✗死	bleepingcomputer.com/news/artificial-intellige...
OpenAI于8月17日任命Dali Rajic	B	71%	✓活	kraviona.com/blog/latest-ai-news-august-2026...
DeepSeek于8月13日正式发布DeepSe	A	97%	✓活	api-docs.deepseek.com/updates/...
智谱8月14日发布新一代基座模型GLM-5.3,	B+	91%	✓活	stdaily.com/web/gdxw/2026-08/14/content_564134...
阿里8月3日发布迄今最大模型Qwen3.8-Ma	B+	90%	✓活	scmp.com/tech/article/3362738/alibabas-ai-mode...
Google DeepMind于8月13日将Ge	A	96%	✓活	ai.google.dev/gemini-api/docs/changelog...
8月被业内称为'大模型疯狂八月':20天内11+	B-	64%	✗死	zhuanlan.zhihu.com/p/2071707711306216661...
DeepSeek于8月21日推出实验模型V4-F	A	96%	✓活	api-docs.deepseek.com/updates/...
低价与分时定价加速AI推理'逼近零价'竞赛:Ax	B	61%	✗死	axios.com/2026/08/01/deepseek-model-cheap-ai-p...
Agent基础设施成新战场:Cloudflare	B	72%	✓活	aiagentstore.ai/ai-agent-news/this-week...

*评分:A 官方一手抓原文 / B+ ≥2独立源 / B 单家权威或未复核 / B- 同源引述 / C 二手(不采)。本期 WebSearch=10
WebFetch=6 成本≈\$1.7034780000000005。自动生成,源已分级核验。*