

GPT-6 Astra网络安全触顶、英伟达129亿美元购Hugging Face、Fable 5.1降价

头部企业动态

- 英伟达9月3日宣布以约129.3亿美元收购Hugging Face:向股东支付约119亿美元,另设最高约10亿美元员工留任股权;预计2027年上半年交割,承诺平台保持开放、不强制使用英伟达算力。 A · 83% 已核实 · 链接探活失败
- OpenAI 9月3日发布GPT-6 Astra,系其Preparedness框架下首个达到网络安全Critical级的模型;API定价每百万输入10美元/输出50美元,上下文约105万,逐步向Plus/Pro/企业及Azure、Bedrock开放。 A · 96% 已核实
- OpenAI 9月3日同步推出Daybreak for Frontline Defenders,承诺10亿美元补贴额度,面向水电、电网、地方政府、社区银行、开源项目等一线防御方,计划约6个月内先美国后盟国使用。 B+ · 88% 多源一致
- Anthropic 9月1日宣布企业前沿安全护栏(EFS):零数据留存结合滥用检测,监测数据存于客户自有云(S3/Azure Blob/GCS)与密钥之下,免费、今秋分阶段上线,与100余家客户共建。 B+ · 90% 多源一致

模型发布

- Anthropic 9月1日发布Claude Fable 5.1与Mythos 5.1(同一模型、不同护栏):Terminal-Bench 55.8%、科研代理52.6%;缓存读取价降75%至每百万0.25美元,典型负载约省25%,代理负载最多省45%。 A · 98% 已核实
- 谷歌9月2日发布Gemini 3.8 Flash与3.8 Flash Cyber:HLE-Verified 54.9%,CWE-Bench修补47.2%,Chrome安全团队称正确补丁为最佳商用模型2.6倍;年内优惠价每百万输入0.75/输出3.75美元。 A · 97% 已核实
- 阿里通义9月2日更新旗舰Qwen3.8-Max-0902,CodeArena前端编程榜1691分登顶,超Claude Opus 5 Max与Kimi K3 Max;均价每百万token 5美元,已接入千问API、千问办公、Qoder与千问App。 B+ · 90% 多源一致
- 微软AI 9月3日发布语音识别模型MAI-Transcribe-2:覆盖60种语言,FLEURS平均词错率5.2%居首,称比GPT-Transcribe快10倍;首发价每音频小时0.10美元,较前代降72%,经Foundry与MAI Playground提供。 B+ · 73% 多源一致 · 链接探活失败

热点与趋势

- 网络安全成为本周前沿模型主战场:OpenAI Astra首触Critical级并配套Daybreak,谷歌以Fairwind计划限量开放Flash Cyber,Anthropic以Mythos 5.1走可信访问;三家同步转向"能力分级+受控发放"模式。 B+ · 90% 多源一致
- 美国参议员桑德斯与众议员卡萨尔九月初宣布将提出《禁止人工超级智能法案》,拟永久禁止开发部署超级智能、暂停前沿AI直至联邦监管到位,并设内阁级AI监管机构,引述夏季"失控AI代理"入侵事件。 B · 63% 待复核 · 链接探活失败
- 国内大厂从"抢入口"转向"进 workflow":阿里千问办公上线一月、9月4日宣布用户破3000万且企业用户过半;字节8月25日发布豆包工作并打通飞书,腾讯WorkBuddy生态发布会跟进。 B · 75% 待复核
- AI算力融资持续升温:数据中心运营商Crusoe 9月3日据报完成逾30亿美元融资,估值300亿美元,10个月内估值涨约2倍;由Atreides与Valor领投,客户含Meta、微软、OpenAI。 B · 76% 同源引述

主动剔除(未核实,不外发)

- DeepSeek V5将于9月发布:仅来自匿名泄露帖,DeepSeek无公告、无API模型串、无技术报告,主动剔除。
- 月之暗面以500亿美元投前估值洽谈Pre-IPO轮并最快6个月内赴港上市:仅IT之家/澎湃转述消息人士,无公司确认。
- DeepSeek推进700亿元人民币融资、估值3000亿元:证券时报等引述传闻,未见一手来源。
- 英伟达拟向Nscale投资约20亿美元参与其pre-IPO融资:仅单一聚合源提及,未能交叉核实。
- OpenAI副总裁称GPT-6 Astra"进入AGI时代"及其"击败Claude Fable 5"的表述:仅新浪汇总/Al Jazeera单源转述,未见官方原话。

信息源清单(可审计)

条目	等级	可信度	探活	来源
英伟达9月3日宣布以约129.3亿美元收购Hug	A	83%	✗死	sec.gov/Archives/edgar/data/0001045810/0001045810...
OpenAI 9月3日发布GPT-6 Astra	A	96%	✓活	deploymentsafety.openai.com/gpt-6-astra...
OpenAI 9月3日同步推出Daybreak	B+	88%	✓活	theregister.com/security/2026/09/04/openai-com...
Anthropic 9月1日宣布企业前沿安全护栏	B+	90%	✓活	anthropic.com/news/enterprise-frontier-safegua...
Anthropic 9月1日发布Claude F	A	98%	✓活	anthropic.com/claude-fable-and-mythos-5-1...
谷歌9月2日发布Gemini 3.8 Flash	A	97%	✓活	blog.google/innovation-and-ai/models-and-resea...
阿里通义9月2日更新旗舰Qwen3.8-Max-	B+	90%	✓活	qbitai.com/2026/09/483101.html...
微软AI 9月3日发布语音识别模型MAI-Tra	B+	73%	✗死	venturebeat.com/infrastructure/microsoft-ais-m...
网络安全成为本周前沿模型主战场:OpenAI A	B+	90%	✓活	deepmind.google/models/gemini/cyber/...
美国参议员桑德斯与众议员卡萨尔九月初宣布将提出《	B	63%	✗死	sanders.senate.gov/press-releases/news-sanders...
国内大厂从"抢入口"转向"进 workflow":阿里千问办	B	75%	✓活	k.sina.com.cn/article_7857201856_1d45362c00190...
AI算力融资持续升温:数据中心运营商Crusoe	B	76%	✓活	techcrunch.com/2026/09/03/crusoe-reportedly-ra...

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=24 WebFetch=22 成本≈\$2.6069787499999997。自动生成,源已分级核验。