

# AI 行业日报

模型与工具能力 · 产业基础设施 · 应用商业化 · 研究开源 · 资本监管 | Data Source: aihot.virxact.com

API耗时: 0s 精选条目: 40 条 焦点: 8 条 快讯: 0 条

## Executive Summary

Qwen系列在具身智能领域实现重大突破，发布Qwen-RobotNav、Qwen-RobotManip和Qwen-RobotWorld三大模型，分别在导航、操作和环境交互方面达到SOTA水平。MiniMax开源428B参数的M3模型权重，采用MSA架构降低长上下文计算成本。Cartesia推出Sonic 3.5和Ink 2实时语音模型，在流式处理性能上表现领先。字节跳动的Seedance 2.0 Mini视频生成模型成本降低50%，API定价显著下降。

DeepSeek完成74亿美元融资，估值超过500亿美元，投资结构显示国资参与度提升。Anthropic企业订阅份额达41%，首次超越OpenAI的39.5%，但面临特朗普政府出口管制禁令。Salesforce以36亿美元收购Fin强化AI客服能力。微软因GitHub算力短缺转向AWS采购资源，Copilot Cowork考虑集成DeepSeek V4降低成本。Google Cloud发布OKF v0.1规范，推动AI Agent知识表示标准化。

后续需关注Qwen机器人系列的商业化落地节奏、MiniMax M3在开源社区的应用情况、DeepSeek与微软合作的具体条款。算力供需平衡、API定价策略调整、AI Agent治理框架建立将成为关键变量。Anthropic与五角大楼关系修复进展以及OpenAI在军用AI市场的竞争态势值得持续跟踪。

## 重点 今日核心进展

### ★ 1. Qwen-RobotNav: 面向智能体导航系统的可扩展导航模型

Qwen: [Blog Retrieval \(API\)](#) · 昨天 08:00 · 模型与工具能力

Qwen发布Qwen-RobotNav，基于Qwen3-VL在15.6M样本上训练，统一了视觉语言导航、目标导航、目标跟踪、自动驾驶和具身问答五个领域，无需修改架构即可在推理时切换任务模式和观察参数。模型在多项基准取得SOTA: VLN-CE RxC成功率76.5%，HM3Dv2目标导航75.6%（仅RGB），EVT-Bench跟踪率90.0%，NAVSIM PDMS 9

[能力进展](#) [基础设施](#) [新发布](#)

<https://qwen.ai/blog?id=qwen-robotnav>

### ★ 2. Qwen-RobotManip: 对齐解锁机器人操作基础模型的规模化能力

Qwen: [Blog Retrieval \(API\)](#) · 昨天 08:00 · 模型与工具能力

Qwen-RobotManip是通义千问基于Qwen-VL的视觉-语言-动作(VLA)基础模型，引入覆盖表示、运动和行为三维度的统一对齐框架。仅使用开源机器人数据集和人演示视频，构建约38,100小时预训练语料，涵盖15种机器人形态。在LIBERO-Plus达91.4%，RoboTwin-C2R Hard达69.4%，RoboCasa365 Composite-Unseen

[能力进展](#) [基础设施](#) [新发布](#)

<https://qwen.ai/blog?id=qwen-robotmanip>

### ★ 3. MiniMax 开源 M3 模型权重及 MSA 技术论文

公众号: [MiniMax \(稀宇科技\)](#) · 昨天 22:40 · 模型与工具能力

MiniMax上周五开源了428B总参数、23B激活参数的M3模型权重，同步发布MSA(MiniMax Sparse Attention)技术论文，该架构显著降低长上下文计算成本。M3是首个从预训练阶段就进行文本、图像等多模态交错混合训练的开源模型。发布两周后，M3在Artificial Analysis综合智能指数、GDPval-AA排行榜均获开源模型第一，Code Ar

[能力进展](#) [基础设施](#) [新发布](#)

<https://mp.weixin.qq.com/s/AW6L89QZkwN-jD27hQ84ww>

### ★ 4. DeepSeek 完成首轮外部融资，估值超 500 亿美元

The Decoder: [AI News \(RSS\)](#) · 14 小时前 · 产业与基础设施

中国AI初创公司DeepSeek完成首轮外部融资，募资超500亿元人民币（约74亿美元），估值超500亿美元。投资结构特殊：多数投资者将资金投入CEO梁文锋管理的有限合伙企业，无投票权且锁定期五年；仅国资AI基金直接投资并保留投票权。创始人梁文锋个人投入约200亿元，腾讯和宁德时代为主要外部投资者。梁文锋表示优先基础AI研究与AGI开发，将继续开源。Deep

[能力进展](#) [基础设施](#) [监管/资本](#)

<https://the-decoder.com/deepseek-takes-outside-money-for-the-first-time-at-a-50-billion-valuation>

### ★ 5. 小米 MiMo Claw 正式版发布: 旗舰模型+金山办公, 全新订阅服务上线

公众号: [小米 MiMo](#) · 10 小时前 · 应用与商业化

小米推出云端轻量化Claw类产品MiMo Claw正式版，搭载与OpenClaw框架深度适配的MiMo-V2.5-Pro旗舰模型。该模型原生兼容MCP工具调用协议，内置百万级超长上下文，支持单会话千次以上连续工具调用；依托MTP三层解码架构，在OpenClaw标准Agent工作流中吞吐效率提升约3倍。ClawEval测试中任务达标率(Pass3)达63.8%

[能力进展](#) [基础设施](#) [新发布](#)

<https://mp.weixin.qq.com/s/N5ac768a8LkhEjOVpkR1dQ>

## ★ 6. Google Cloud 推出 OKF v0.1: 供应商中立的 Markdown 规范, 为 AI 智能体提供结构化上下文

MarkTechPost (RSS) · 16 小时前 · 应用与商业化

Google Cloud 发布 Open Knowledge Format (OKF) v0.1, 一种供应商中立的 Markdown 规范, 为 AI 智能体提供结构化上下文知识。OKF 将知识表示为带 YAML 前置元数据的 markdown 文件目录, 每个概念对应一个文件, 通过 `type`、`title`、`description` 等少量保留字段实现互操作。无需专有服务、SDK 或运行时, 目

能力进展 基础设施 新发布

<https://www.marktechpost.com/2026/06/16/google-cloud-introduces-open-knowledge-format-okf-a-vendor-neutral-markdown-spec-for-giving-ai-agents-curated-context>

## ★ 7. 成本砍半, 字节跳动推出 Seedance 2.0 Mini 视频生成模型

IT之家 (RSS) · 22 小时前 · 模型与工具能力

字节跳动火山引擎旗下火山方舟体验中心于 6 月 15 日上线 Seedance 2.0 Mini 视频生成模型, 计划近期开放 API。该模型比 Seedance 2.0 Fast 快 2 倍, 输出质量相当。图生视频定价 0.023 元/千 tokens, 视频生视频 0.014 元/千 tokens, 720P 规格下单秒生成成本约 0.5 元, 较 Seedance 2.0 标准版降低约一半。模型面向

能力进展 新发布

<https://www.ithome.com/0/964/672.htm>

## ★ 8. Qwen-RobotWorld: 具身智能体的无界世界

Qwen: Blog Retrieval (API) · 昨天 08:00 · 模型与工具能力

Qwen-RobotWorld以语言为统一动作接口, 采用双流Multimodal Diffusion Transformer (MMDiT) 架构, 将Qwen2.5-VL作为动作编码器。在4个基准测试中取得顶尖成绩, 统一20余种机器人形态, 基于860万跨场景训练对和1300多项操作技能。语言接口标准化500多种动作类别, 支持操作、自动驾驶、室内导航的联合训练。还支持Scene2Robot人类到机器人转

能力进展 基础设施

<https://qwen.ai/blog?id=qwen-robotworld>

## 能力 模型与工具能力

### 1. Qwen-RobotNav: 面向智能体导航系统的可扩展导航模型

Qwen: Blog Retrieval (API) · 昨天 08:00

Qwen 发布 Qwen-RobotNav, 基于 Qwen3-VL 在 15.6M 样本上训练, 统一了视觉语言导航、目标导航、目标跟踪、自动驾驶和具身问答五个领域, 无需修改架构即可在推理时切换任务模式和观察参数。模型在多项基准取得 SOTA: VLN-CE RxR 成功率 76.5%, HM3Dv2 目标导航 75.6% (仅 RGB), EVT-Bench 跟踪率 90.0%, NAVSIM PDMS 9

能力进展 基础设施 新发布

<https://qwen.ai/blog?id=qwen-robotnav>

### 2. Qwen-RobotManip: 对齐解锁机器人操作基础模型的规模化能力

Qwen: Blog Retrieval (API) · 昨天 08:00

Qwen-RobotManip 是通义千问基于 Qwen-VL 的视觉-语言-动作 (VLA) 基础模型, 引入覆盖表示、运动和行为三维度的统一对齐框架。仅使用开源机器人数据集和人演示视频, 构建约 38, 100 小时预训练语料, 涵盖 15 种机器人形态。在 LIBERO-Plus 达 91.4%, RoboTwin-C2R Hard 达 69.4%, RoboCasa365 Composite-Unseen

能力进展 基础设施 新发布

<https://qwen.ai/blog?id=qwen-robotmanip>

### 3. MiniMax 开源 M3 模型权重及 MSA 技术论文

公众号: MiniMax (稀宇科技) · 昨天 22:40

MiniMax 上周五开源了 428B 总参数、23B 激活参数的 M3 模型权重, 同步发布 MSA (MiniMax Sparse Attention) 技术论文, 该架构显著降低长上下文计算成本。M3 是首个从预训练阶段就进行文本、图像等多模态交错混合训练的开源模型。发布两周后, M3 在 Artificial Analysis 综合智能指数、GDPval-AA 排行榜均获开源模型第一, Code Ar

能力进展 基础设施 新发布

<https://mp.weixin.qq.com/s/AW6L89QZkwN-jD27hQ84ww>

### 4. 成本砍半, 字节跳动推出 Seedance 2.0 Mini 视频生成模型

IT之家 (RSS) · 22 小时前

字节跳动火山引擎旗下火山方舟体验中心于 6 月 15 日上线 Seedance 2.0 Mini 视频生成模型, 计划近期开放 API。该模型比 Seedance 2.0 Fast 快 2 倍, 输出质量相当。图生视频定价 0.023 元/千 tokens, 视频生视频 0.014 元/千 tokens, 720P 规格下单秒生成成本约 0.5 元, 较 Seedance 2.0 标准版降低约一半。模型面向

能力进展 新发布

<https://www.ithome.com/0/964/672.htm>

### 5. Qwen-RobotWorld: 具身智能体的无界世界

Qwen: Blog Retrieval (API) · 昨天 08:00

Qwen-RobotWorld以语言为统一动作接口, 采用双流Multimodal Diffusion Transformer (MMDiT) 架构, 将Qwen2.5-VL作为动作编码器。在4个基准测试中取得顶尖成绩, 统一20余种机器人形态, 基于860万跨场景训练对和1300多项操作技能。语言接口标准化500多种动作类别, 支持操作、自动驾驶、室内导航的联合训练。还支持Scene2Robot人类到机器人转

能力进展 基础设施

<https://qwen.ai/blog?id=qwen-robotworld>

## 6. Cartesia 发布 Sonic 3.5 与 Ink 2 实时语音模型

X: [Testing Catalog \(@testingcatalog\)](#) · 17 小时前

Cartesia 推出 Sonic 3.5 和 Ink 2 两个模型，作为单一实时语音栈，分别负责文本转语音和语音转文本。Ink 2 在 Artificial Analysis 的流式语音转文字排行榜上排名第一。Sonic 3.5 在实时文本转语音中位列榜首，首音频延迟约 82ms。Cartesia 成为目前唯一同时拥有 #1 听与说模型的提供商。

能力进展 新发布

<https://x.com/testingcatalog/status/2066773392527655252>

## 产业 产业与基础设施

### 1. DeepSeek 完成首轮外部融资，估值超 500 亿美元

The Decoder: [AI News \(RSS\)](#) · 14 小时前

中国 AI 初创公司 DeepSeek 完成首轮外部融资，募资超 500 亿元人民币（约 74 亿美元），估值超 500 亿美元。投资结构特殊：多数投资者将资金投入 CEO 梁文锋管理的有限合伙企业，无投票权且锁定期五年；仅国资 AI 基金直接投资并保留投票权。创始人梁文锋个人投入约 200 亿元，腾讯和宁德时代为主要外部投资者。梁文锋表示优先基础 AI 研究与 AGI 开发，将继续开源。Deep

能力进展 基础设施 监管/资本

<https://the-decoder.com/deepseek-takes-outside-money-for-the-first-time-at-a-50-billion-valuation>

### 2. Anthropic 5月企业AI订阅份额首超OpenAI，特朗普政府禁令反促采用量创新高

TechCrunch: [AI \(RSS\)](#) · 1 小时前

Anthropic 5月企业AI订阅市场份额达41%，首次超越OpenAI (39.5%)。公司刚完成650亿美元融资、估值9650亿美元，并因首次盈利季度秘密提交IPO。特朗普政府以出口管制为由要求Anthropic禁止非美国人访问最新模型Mythos 5及Fable 5，导致两款模型下架。Ramp首席经济学家指出，类似争议（如3月被国防部列为供应链风险）反而推动Anthropic企业采用量创纪录

能力进展 监管/资本 新发布

<https://techcrunch.com/2026/06/16/anthropics-latest-feud-with-the-trump-admin-may-actually-help-it-sales-data-suggests>

### 3. 五角大楼将大部分日常AI工作流从Anthropic转移，目标9月前完全切断

X: [阿易 AI Notes \(@AYi\\_AInotes\)](#) · 23 小时前

五角大楼宣布已将超2/3日常AI工作流从Anthropic转移，目标9月前清零。起因是年初五角大楼要求Anthropic签署协议允许Claude用于大规模监控和全自动武器，CEO Dario Amodei以模型不可靠为由拒绝。五角大楼将其列为“供应链风险”，起诉未果。OpenAI调整立场获得订单。Polymarket预测6月底前和解概率仅9%。此事件凸显AI公司需在原则与政府合作间抉择。

能力进展 新发布

[https://x.com/AYi\\_AInotes/status/2066679835607412846](https://x.com/AYi_AInotes/status/2066679835607412846)

### 4. 微软旗下GitHub遭遇AI算力短缺，转面向AWS寻求支持

Hacker News 热门 ([buzzing.cc 中文翻译](#)) · 19 小时前

微软旗下GitHub面临AI算力短缺，微软因此转向亚马逊AWS寻求计算资源支持。原文来自Hacker News热门讨论，标题为“Microsoft turns to AWS as GitHub faces AI capacity crunch”。

能力进展 基础设施

<https://runtimewire.com/article/microsoft-github-aws-ai-capacity-crunch>

### 5. Salesforce以36亿美元收购AI客服平台Fin

TechCrunch: [AI \(RSS\)](#) · 昨天 22:34

Salesforce宣布以36亿美元收购AI客服平台Fin（前身为Intercom）。Fin提供可跨实时聊天、WhatsApp、短信、电话、Slack等多渠道解决客户问题的AI智能体。Salesforce计划利用Fin的技术和团队增强其企业级Agentforce平台，该平台允许企业构建自定义AI智能体以自动化任务。交易预计在Salesforce 2027财年第四季度（即2027年初）完成。Fin联

能力进展 监管/资本

<https://techcrunch.com/2026/06/15/salesforce-acquires-ai-customer-service-platform-fin-for-3-6b>

### 6. 微软考虑为 Copilot Cowork 集成 DeepSeek V4

X: [Kim \(@kimmonismus\)](#) · 6 小时前

微软正考虑为 Copilot Cowork 提供微软托管的 DeepSeek V4 版本，作为更便宜的模型选项。Copilot Cowork 将放弃无限定价，转向按使用量计费，原因是成本过高（用户每周执行数百项任务导致费用激增）。若采用 DeepSeek，该模型将是可选的、经过微调与安全防护，并完全托管于 Azure。Axios 报道称微软已微调了一个可用模型，最终决定待定。

能力进展 监管/资本

<https://x.com/kimmonismus/status/2066946013026263110>

### 7. 美国司法部援引国家安全为xAI未经许可的燃气轮机辩护

The Decoder: [AI News \(RSS\)](#) · 10 小时前

美国司法部在一份驳回诉讼的动议中称，xAI的聊天机器人Grok对军事行动至关重要，以此为其在密西西比州Southaven的Colossus 2设施运行未经许可的燃气轮机辩护。NAACP已提起诉讼，指控xAI的燃气轮机数量从4月的27台增至57台，导致氮氧化物排放飙升111%。国防部首席数字与人工智能官Cameron Stanley表示，Grok是支持机密和绝密网络军事任务的四款AI模型之一，包括近

能力进展 监管/资本

<https://the-decoder.com/doj-invokes-national-security-to-defend-xais-unpermitted-gas-turbines-in-naacp-lawsuit>

## 8. SpaceX 以 600 亿美元股票收购 AI 编程公司 Cursor

TechCrunch: AI (RSS) · 13 小时前

SpaceX 在历史性 IPO 数天后，同意以 600 亿美元股票收购 AI 编程初创公司 Cursor，旨在帮助其围绕 xAI 构建的 AI 部门追赶主要 AI 实验室。此前 Cursor 正接近完成一轮 20 亿美元融资，估值 500 亿美元，投资方包括 Andreessen Horowitz、Thrive 和 Nvidia。SpaceX 在 IPO 期间向投资者表示，其 AI 产品可寻址市场

基础设施 监管/资本

<https://techcrunch.com/2026/06/16/spacex-to-acquire-cursor-for-60b-in-stock-days-after-blockbuster-ipo>

## 9. Fable 遭美国政府封禁，TechCrunch 质疑真正原因并非模型越狱

Hacker News 热门 (buzzing.cc 中文翻译) · 7 小时前

美国政府对 Anthropic 的模型 Fable 实施封禁，但 TechCrunch 发文质疑，实际原因可能并非此前认为的“模型越狱”问题。该文章在 Hacker News 引发讨论，获得 103 个点赞。

能力进展

<https://techcrunch.com/2026/06/15/the-us-governments-anthropic-models-ban-was-never-about-an-ai-jailbreak>

## 10. Nvidia 加入 AI 债务热潮，发行 200 亿美元债券

The Decoder: AI News (RSS) · 昨天 23:43

Nvidia 计划通过自 2021 年以来的首次债券发行筹集至少 200 亿美元，消息援引知情人士透露。此举标志着 Nvidia 加入 AI 领域的债务融资热潮。

基础设施 监管/资本

<https://the-decoder.com/nvidia-joins-ai-debt-boom-with-20-billion-bond-sale>

## 11. 马斯克：AI 将达 Stockfish 级编码

X: Elon Musk (@elonmusk, xAI) · 10 小时前

AI 将实现 Stockfish 级别的编码和通用计算机使用

<https://x.com/elonmusk/status/2066880262668247091>

## 应用 应用与商业化

### 1. 小米 MiMo Claw 正式版发布：旗舰模型+金山办公，全新订阅服务上线

公众号: 小米 MiMo · 10 小时前

小米推出云端轻量化 Claw 类产品 MiMo Claw 正式版，搭载与 OpenClaw 框架深度适配的 MiMo-V2.5-Pro 旗舰模型。该模型原生兼容 MCP 工具调用协议，内置百万级超长上下文，支持单会话千次以上连续工具调用；依托 MTP 三层解码架构，在 OpenClaw 标准 Agent 工作流程中吞吐效率提升约 3 倍。ClawEval 测试中任务达标率 (Pass3) 达 63.8%

能力进展 基础设施 新发布

<https://mp.weixin.qq.com/s/N5ac768a8LkhEjOVpkR1dQ>

### 2. Google Cloud 推出 OKF v0.1：供应商中立的 Markdown 规范，为 AI 智能体提供结构化上下文

MarkTechPost (RSS) · 16 小时前

Google Cloud 发布 Open Knowledge Format (OKF) v0.1，一种供应商中立的 Markdown 规范，为 AI 智能体提供结构化上下文知识。OKF 将知识表示为带 YAML 前置元数据的 markdown 文件目录，每个概念对应一个文件，通过 `type`、`title`、`description` 等少量保留字段实现互操作。无需专有服务、SDK 或运行时，目

能力进展 基础设施 新发布

<https://www.marktechpost.com/2026/06/16/google-cloud-introduces-open-knowledge-format-okf-a-vendor-neutral-markdown-spec-for-giving-ai-agents-curated-context>

### 3. Meta 在 Facebook 上线 "AI Mode"，基于平台公开信息合成答案

TechCrunch: AI (RSS) · 昨天 02:30

Meta 宣布在 Facebook 推出 "AI Mode" 搜索功能，利用 Meta AI 从公开帖子 (含群组 and Reels) 提取信息并合成答案，用户可用自然语言提问获得摘要。同时新增视频拼贴剪辑、过渡效果及 AI 照片预设 (可更换服装、发型和配饰)，体育迷可在 Stories 中点击 "AI Edit" 虚拟穿上队服。这些更新延续了此前动态头像、Marketplace 自动回复和创作者 AI 助手的部

能力进展 新发布

<https://techcrunch.com/2026/06/15/metas-new-ai-mode-on-facebook-pulls-from-public-info-across-its-platforms>

### 4. Subagent：让模型把琐碎任务委托出去

OpenRouter: Announcements (RSS) · 昨天 08:00

OpenRouter 推出 openrouter: subagent 服务器工具，允许前沿模型在生成过程中将独立的琐碎任务 (如文档总结、结构化数据提取、文本重格式化) 委托给更小、更便宜、更快的 worker 模型执行，从而节省前沿模型的 token 消耗。

能力进展 新发布

<https://openrouter.ai/blog/announcements/subagent-server-tool>

### 5. OpenRouter 新增免费模型 gpt-oss-20b 和 Gemma4 26B

X: OpenRouter (@OpenRouter) · 昨天 02:16

OpenRouter 上新增免费容量，由 @eigenlabs 的 Darkbloom 提供：gpt-oss-20b 和 Gemma 4 26B。今天就开始使用这些模型吧 ↓

能力进展 新发布

<https://x.com/OpenRouter/status/2066585705581797616>

## 6. Copilot Cowork 全球正式可用，支持多模型

X: [Satya Nadella \(@satyanadella\)](#) · 8 小时前

Copilot Cowork 现已全球正式可用，并支持多模型！每个组织都可以让长期运行的智能体处理复杂的多步骤任务，基于你组织的独特知识和专有技术。 <https://www.microsoft.com/en-us/microsoft-365/blog/2026/06/16/copilot-cowork-is-now-generally-available/?v=15>

能力进展

<https://x.com/satyanadella/status/2066911399494963335>

## 7. 教育部"阳光志愿"信息服务系统全新升级上线：智能筛选志愿，数据权威可信

IT之家 (RSS) · 22 小时前

教育部"阳光志愿"信息服务系统今日全新升级上线，依托招生、学籍、就业等海量官方数据，免费为考生和家长提供志愿填报服务。系统支持31个省区市本专科普通批次志愿筛选，输入高考成绩、位次及个性化条件即可快速生成合理参考方案。AI助手"智慧小招"24小时在线解答政策规则。平台数据由高校直接报送、官方核验，真实可靠。同时推出专业倾向测评和21项生涯测评工具，帮助学生认清特长、规划未来。

监管/资本 新发布

<https://www.ithome.com/0/964/675.htm>

## 8. AI 版支付宝官宣开启邀测：右滑打开"阿宝"，官方放出 100 个邀请码

IT之家 (RSS) · 21 小时前

支付宝今日开启 AI 版邀测，用户可右滑进入新版界面，在对话框或语音中输入指令，由"阿宝"助手代办事项。以查询公积金为例，阿宝自动匹配对应小程序和服务入口，用户点击确认即可完成，将多步跳转折叠为一句指令。所有涉及资金变动和支付的环节均需用户本人确认。首批放出 100 个邀请码。

能力进展

<https://www.ithome.com/0/964/691.htm>

## 9. Grok for PowerPoint 发布：在 Microsoft PowerPoint 内直接生成和编辑幻灯片

xAI: [News \(网页\)](#) · 昨天 08:00

xAI 于 6 月 16 日发布 Grok for PowerPoint，作为免费 Microsoft 365 插件上线。用户无需离开应用即可利用 Grok 将大纲转为完整幻灯片，进行内容研究、撰写、排版，并支持添加单张幻灯片、调整样式主题、重构章节。插件还能调用 Grok 连接器，从邮件或 SharePoint 中获取信息。该插件同样适用于 Word 和 Excel。

新发布

<https://x.ai/news/introducing-powerpoint-addin>

## 10. Midjourney V8.1 推出 Draft mode 草稿模式与新功能预览

Midjourney: [Updates \(RSS\)](#) · 2 小时前

Midjourney V8.1 的 Draft mode 草稿模式每次生成24张低分辨率低质量图片。用户可对任意图片点击"Vary"，将其渲染为全质量、全分辨率版本。草稿任务消耗的快速小时数减半。

新发布

<https://updates.midjourney.com/draft-mode-for-v8-1-and-new-feature-previews>

## 研究 研究与开源进展

### 1. OpenAI 发布 Deployment Simulation 方法：通过模拟部署预测模型发布前行为

OpenAI: [官网动态 \(RSS\)](#) · [排除企业/客户案例](#) · 昨天 08:00

OpenAI 近日发布 Deployment Simulation 方法，通过在隐私保护下重放历史对话、用新候选模型重新生成回复，模拟模型上线后的实际表现。在多个 GPT-5-series Thinking 部署中，该方法比传统评估更准确地估计了不良行为频率，发现新型对齐问题，并降低模型识别测试的风险。它还能扩展至涉及工具使用的智能体场景。传统评估存在覆盖不足、选择偏差和模型可识别测试等局限，而

能力进展 新发布

<https://openai.com/index/deployment-simulation>

### 2. 公开聊天数据能否预测真实世界AI失调?

OpenAI: [Alignment 研究博客 \(RSS\)](#) · 6 小时前

OpenAI利用WildChat公开数据集（2023年4月至2024年5月收集的100万条对话）模拟模型部署，预测GPT-5.1、GPT-5.2、GPT-5.4在真实生产环境中的不良行为率。与私有生产数据对比发现，WildChat模拟的平均预测误差约3倍；但对技术性和智能体型失调的预测精度下降。研究验证了公开数据集作为外部审计工具的可行性。

能力进展 新发布

<https://alignment.openai.com/validating-public-evals>

### 3. Anthropic: 智能体编码中专业知识回报持续存在

Anthropic: [Research \(发表成果 · 网页\)](#) · 4 小时前

Anthropic 基于约40万次 Claude Code 交互会话（2025年10月至2026年4月）分析发现：人类主导规划决策（做什么），Claude 主导执行决策（怎么做）。领域专业知识越强，模型每次指令完成的工作量越多。各类职业完成任务的成功率与软件工程师平均相近；领域专家成功率更高，但与中级用户差距不大。七个月间调试会话占比下降近一半，使用转向端到端智能体任务（部署运行代码、分析数据、编

能力进展

<https://www.anthropic.com/research/claude-code-expertise>

### 1. 免费LLM API比较：速率限制、模型与真实成本（2026）

OpenRouter: Announcements (RSS) · 昨天 01:00

13个平台提供免费LLM API，含永久免费层与试用额度。OpenRouter拥有20+免费模型，单密钥无需信用卡；Groq以约320 tokens/秒运行Llama 3.3 70B；Google AI Studio支持1M上下文；Mistral实验层约10亿token/月但需同意数据训练；Cerebras约1M token/天；GitHub Models提供GPT-4o、Claude 3.5 S

能力进展 基础设施 新发布

<https://openrouter.ai/blog/tutorials/free-llm-apis-compared>

### 2. 智能体AI治理：你的API密钥就是护栏

OpenRouter: Announcements (RSS) · 昨天 00:00

智能体AI使用量将在未来两年急剧上升，但治理滞后--仅五分之一企业拥有成熟治理模型。典型事故：销售智能体重试失败调用后自行升级到GPT-5.5，一夜消耗200美元。API路由位于智能体与模型之间，是强制执行预算上限、模型白名单、提供方准入和请求日志的理想点。IBM报告指出97%遭遇AI安全事件的组织缺乏AI访问控制。最小可行治理方案：为每个智能体 workflow 分配独立API密钥，在API密钥层面实现预算

能力进展 监管/资本 新发布

<https://openrouter.ai/blog/insights/agenic-ai-governance>

### 3. WorkBuddy日活飙升至行业第二的3-4倍，非技术用户涌入

公众号: 数字生命卡兹克 · 11 小时前

从3月至今，WorkBuddy日活用户数已达行业第二名的3-4倍，用户不再限于开发者，大量HR、运营、行政等非技术岗位也在使用。其企业版和项目功能进一步扩展了Agent办公场景。同期，Trae Work、QoderWork、Kimi Work等产品纷纷改名或出新，争夺市场。腾讯云认为这可能是十年一遇的机遇。

能力进展 基础设施

<https://mp.weixin.qq.com/s/iAgbUzPrsbpquv4s2XiH0g>

### 4. 人工智能是否已经让自助类非虚构书籍销声匿迹了？

Hacker News 热门 (buzzing.cc 中文翻译) · 1 小时前

2026年Q1美国成人非虚构书籍销量同比下降9%，自助类下跌26.3%，仅手工艺/爱好/古董/游戏和宗教两个子类别增长。一位出版了多本《纽约时报》畅销书的作者透露，其五本书的印刷版年销量从2022年基准连续下滑：2023年-5%，2024年-13%，2025年-46%，2026年年化跌幅达-57% vs 2025年；若持续，2026年销量将比2022年减少约80%。所有格式在2025年下半年环比上

能力进展

<https://tim.blog/2026/06/12/has-ai-already-killed-nonfiction>

### 5. 毕业生陷入AI检测荒诞循环：手写摘要被判99%AI率，AI写部分0%

公众号: 数字生命卡兹克 · 22 小时前

当前毕业生面临论文AIGC率检测荒诞困境。学生手写摘要被判99%AI率，纯AI写部分却为0%。学校要求AIGC率不超40%，学生用Claude反复修改并花上百元检测费（维普20元/篇，知网/万方2元/千字符），最终降至36.1%。答辩时老师要求改回学术表达，AI率回升至37.21%。同一论文在不同平台检测结果差异巨大（48%、44%、59%）。部分平台提供降重收费服务，少数学校已改用AI使用声明

能力进展

<https://mp.weixin.qq.com/s/DRzNllojGLO-iqqPyJ7f8Q>

### 6. 项目负责人揭秘为何苹果 AI 版 Siri 姗姗来迟：推倒重来，彻底重构

IT之家 (RSS) · 昨天 07:16

苹果AI版Siri迟迟未上线，项目负责人迈克·罗克韦在WWDC技术分享会上透露，去年团队曾做出在原有Siri基础上小幅改良、新增工具调用的可运行版本，但因无法达到产品愿景，最终选择推倒重来，完整从零重构系统，依托全新大模型搭建。重构后的Siri拥有独立应用程序，原生支持多模态交互，隐私保护贯穿底层架构，并覆盖iPhone、iPad、Mac、Apple Watch、Vision Pro、CarPl

能力进展

<https://www.ithome.com/0/964/620.htm>

### 7. GitHub Copilot CLI 初学者指南：常用斜杠命令概览

GitHub Blog · 昨天 04:15

GitHub Copilot CLI 为初学者提供了常用斜杠命令的概述，帮助用户通过命令控制终端中的 AI 智能体。

能力进展

<https://github.blog/ai-and-ml/github-copilot/github-copilot-cli-for-beginners-overview-of-common-slash-commands>

### 8. OpenAI 的领先优势正在快速缩小

Gary Marcus: The Road to AI We Can Trust (RSS) · 2 小时前

评论认为 OpenAI 正面临多重危机：缺乏护城河导致市场领先地位下滑；最大投资者微软持续疏远，近期甚至公开考虑将主要产品外包给中国；亏损速度远超预期，年亏损额以 8 倍增长。华盛顿方面可能打压 Anthropic，但也可能反而帮助其崛起，而 Elon Musk 成为另一个潜在的竞标者。

新发布

<https://garymarcus.substack.com/p/openais-lead-is-dwindling-fast>

## 9. Skydio CEO Adam Bry: 硅谷不应为无人机使用画红线

The Verge: AI (RSS) · 昨天 22:00

Skydio是美国最大的无人机制造商，主攻公共安全、军事、能源、基建巡检等企业市场。CEO Adam Bry表示，特朗普政府去年底禁止中国产无人机后，廉价消费级无人机几乎消失，Skydio产品成为主要替代方案。公司认为无人机正从工具转向自主基础设施--通过机库、远程操控和软件整合实现规模化应用，AI在其中扮演关键角色。访谈还涉及Skydio与军方合作的态度，以及自主技术如何带动公司扩张。

监管/资本

<https://www.theverge.com/podcast/949195/skydio-ceo-adam-bry-autonomous-drones-china-red-lines-military>

---

## 10. Meta 解散工程部门引发热议

Hacker News 热门 (buzzing.cc 中文翻译) · 5 小时前

6月16日，一篇标题为 "Why is Meta destroying its engineering organization?" 的博客文章出现在 Hacker News，获得110个点赞。文章指出 Meta 正在解散其工程组织，引发业界广泛讨论。具体原因和后续影响尚未明确。

<https://newsletter.pragmaticengineer.com/p/why-is-meta-destroying-its-engineering>

---