

Anthropic发布Claude Opus 5模型，多模态与智能体技术持续演进

中文内部早报 · 5分钟版

今日总览

Anthropic在模型层面推出Claude Opus 5，该模型性能接近Fable 5但价格减半，同时发布了模型选择指南将产品线分为Mythos/Fable、Opus、Sonnet和Haiku四个类别。这一举措标志着大模型性价比竞争进入新阶段。

Agent和Coding领域迎来重要更新，Claude Code新增Opus 5支持1M上下文窗口，Runway Agent推出自然语言工作流功能。Cursor发布智能模型路由系统，实现成本降低的同时保持用户满意度。

多模态方面，Black Forest Labs发布FLUX 3多模态基础模型，联合mimic推出动作预测功能，已在奥迪生产线测试部署。Midjourney V8.2重点提升美学质量和个性化表现，模型优化方向更加精细化。

大厂竞争格局中，Google Gemini月活用户逼近9.5亿，市场份额升至27.7%，与ChatGPT展开直接竞争。OpenAI在桌面版推出语音控制多智能体功能，扩展应用场景。

安全监管层面，HuggingFace遭遇AI智能体入侵事件，凸显前沿模型安全风险。英伟达、微软和Meta联合警告避免对开放权重模型过度监管，行业对监管政策分歧显现。

头条

今日头条

1. Claude模型家族详解：工作负载最佳模型选择指南

Claude 博客 · 07/25 08:00 · 大厂与产业链

Anthropic发布Claude模型选择指南，将模型分为Mythos/Fable（最强能力，用于前沿编码与智能体任务）、Opus（推理密集型企业任务）、Sonnet（日常通用任务）和Haiku（最低成本与最快速度）四个类别。

2. Claude Code v2.1.219发布：新增Claude Opus 5支持1M上下文

Claude Code: GitHub Releases · 07/25 08:00 · 开源项目与开发者工具

Claude Code v2.1.219新增Claude Opus 5作为默认Opus模型，支持1M上下文窗口，快速模式定价为\$10/\$50每百万token。该版本还支持嵌套子智能体功能。

3. Anthropic发布Claude Opus 5模型

Anthropic: Newsroom · 07/25 08:00 · 前沿模型与多模态

Anthropic发布Claude Opus 5，其智能水平接近Claude Fable 5，但价格减半。该模型在Frontier-Bench v0.1上性能超过Opus 4.8两倍以上，在ARC-AGI 3上得分是次优模型的三倍。Opus 5即日起成为Claude Max的默认模型和Claude Pro的最强模型。

4. FLUX 3 x mimic：新一代视频动作模型

Hacker News · 07/25 08:00 · 大厂与产业链

Black Forest Labs发布多模态基础模型FLUX 3，联合训练图像、视频和音频，其中视频预测占训练算力的95%以上。该模型与机器人公司mimic合作推出FLUX-mimic，已在奥迪生产线上测试部署。加入动作预测后，视频生成质量最初下降最多10%，但经3500步训练后恢复原有水平。

5. Runway Agent推出自然语言工作流功能

X: Runway (@runwayml) · 07/25 08:00 · 智能体、编程与工作流

Runway Agent引入工作流功能，用户可通过自然语言构建、运行或编辑基于节点的工作流。工作流可大规模解锁高质量输出，立即可通过调用/Workflow技能尝试该功能。

模型

前沿模型与多模态

1. OpenAI 更新模型能力与产品方案

OpenAI 新闻 · 07/23 08:00

据OpenAI 新闻，OpenAI 更新模型能力与产品方案。这类进展主要影响模型能力、API 选型和产品可用性。后续需要关注官方披露的可用时间、开放范围、价格变化和第三方验证结果。

2. Claude 语音模式现已支持 Opus、Sonnet 及连接工具与多语言

Claude 博客 · 07/23 00:00

即日起，Claude 语音模式在 Opus、Sonnet 和 Haiku 上运行，并支持连接 Gmail、Slack 等工具及更多语言。用户可在对话中切换模型，语音模式默认沿用上次文本聊天使用的模型。该功能面向所有用户开放 beta 测试，免费版可使用 Haiku 及一个连接工具，付费版可访问更多模型和全部连接工具。

3. Midjourney V8.2 发布：专注美学提升与个性化理解

Midjourney: Updates · 07/25 08:00

Midjourney 今日推出 V8.2 图像模型，重点提升美学质量、图像创意与个性化表现。低质量图像出现频率将显著降低，个性化功能能更精准理解用户审美偏好。V8.2 的个性化配置文件拥有更大、更优的图像选择池，建议用户尝试新旧配置文件体验最新版本。

4. GigaToken 发布：语言模型分词速度最高提升约 1000 倍，可无缝替代 HuggingFace Tokenizers

Hacker News · 07/23 02:51

据Hacker News，GigaToken 发布：语言模型分词速度最高提升约 1000 倍，可无缝替代 HuggingFace Tokenizers。这类进展主要影响模型能力、API 选型和产品可用性。后续需要关注官方披露的可用时间、开放范围、价格变化和第三方验证结果。

智能体 智能体、编程与 workflows

1. Claude-thermos：保持 Claude 会话缓存热度，避免重新编码费用

Hacker News · 07/25 08:00

Claude-thermos 通过本地反向代理监控 Claude Code 会话，在主智能体因等待子智能体而空闲超过 5 分钟时，自动发送预热请求刷新提示缓存。实测约 185 次本地会话中，缓存过期导致的重新编码占账单约 22%。工具以 uvx 运行，支持自定义空闲阈值和预热间隔。

2. Claude Design 产品设计师 Nate Parrott 分享如何用其探索视觉创意

Claude 博客 · 07/25 08:00

Anthropic 产品设计师 Nate Parrott 分享了他在 Claude Design（beta 版）中利用 HTML 交互能力进行产品原型、幻灯片、动画等视觉创意探索与迭代的方法。他将 Anthropic 品牌规范提炼为提示词，使输出自动符合品牌指南。Claude Design 不包含图像模型，不适合 Logo 设计，但可与 Claude Code 协同工作，将早期创意与生产开发衔接。

3. Cursor 发布智能模型路由系统 Cursor Router

Cursor Blog · 07/22 08:00

Cursor 推出 Cursor Router，可自动将每个编码请求分配给最合适的模型。在线 A/B 测试显示，Auto Intelligence 模式在用户满意度接近 Fable 的同时，成本降低约 60%；Auto Balance 模式满意度超过 Opus 4.8，成本降低约 36%。

4. OpenAI 更新 AI 工作流与智能体方案

OpenAI 新闻 · 07/22 21:00

据OpenAI 新闻，OpenAI 更新 AI 工作流与智能体方案。这类进展主要影响智能体编排、代码生成和企业工作流落地。后续需要关注官方披露的可用时间、开放范围、价格变化和第三方验证结果。

产业 大厂与产业链

1. ChatGPT 桌面版上线语音控制多智能体

X: OpenAI (@OpenAI) · 07/24 03:43

ChatGPT 语音功能现已登陆桌面应用。 只需使用语音，即可控制你的电脑，并指挥在 ChatGPT Work 或 Codex 中运行的多个智能体。该功能由 GPT-Live 驱动，因此它能够同时在该应用中说话、聆听并协调工作。 今日起，面向 macOS 和 Windows 平台的 Plus、Pro、Business、Edu 及 Enterprise 计划用户全球推送。

2. 微软 MagenticLite 模型全面开源

X: Microsoft Research (@MSFTResearch) · 07/23 01:57

MagenticLite 的模型现已完全开源。 此前在 Microsoft Foundry 上提供的 MagenticBrain 和 Fara 1.5，现已在 Hugging Face 上开放权重。 该应用、测试工具以及堆栈中的每个模型现已全部开放。

3. OpenAI 推出 AI 产业部署新进展

OpenAI 新闻 · 07/22 08:00

据OpenAI 新闻，OpenAI 推出 AI 产业部署新进展。这类进展主要影响云算力、硬件投入、企业采购和服务成本。后续需要关注官方披露的可用时间、开放范围、价格变化和第三方验证结果。

4. Google Gemini 月活用户逼近 9.5 亿，有望成为下一个十亿级产品

TechCrunch AI · 07/23 22:52

Google 在 Q2 2026 财报电话会上宣布，AI 助手 Gemini 月活跃用户已超过 9.5 亿，用户数较去年增长三倍。Gemini 正与月活突破 10 亿的 ChatGPT 展开更直接竞争，其 AI 搜索模式用户也已超过 10 亿。Sensor Tower 报告显示，Gemini 在 AI 助手市场份额升至 27.7%，而 ChatGPT 份额首次跌破 50%。

开源 开源项目与开发者工具

1. Hugging Face 发布新动态：《Bringing Nunchaku 4-bit Diffusion Inference to Diffusers》

Hugging Face Blog · 07/23 08:00

据Hugging Face Blog，本条原始主题为《Bringing Nunchaku 4-bit Diffusion Inference to Diffusers》。该材料涉及开发者工具、模型路由或开源生态，需继续核验代码、许可、维护频率和部署条件。

1. Hacker 发布新动态：《Show HN: Palmier Pro – Open-source macOS video editor built for AI》

Hacker News • 07/23 23:11

据Hacker News, Show 发布 AI 行业观点与判断。这类信息反映行业参与者对技术路线和商业节奏的判断。后续需要关注官方披露的可用时间、开放范围、价格变化和第三方验证结果。

2. Unlimited 发布 AI 行业观点与判断

Hacker News • 07/23 00:13

据Hacker News, Unlimited 发布 AI 行业观点与判断。这类信息反映行业参与者对技术路线和商业节奏的判断。后续需要关注官方披露的可用时间、开放范围、价格变化和第三方验证结果。

3. Hacker 发布新动态：《Show HN: I ran 12 AI bots predicting stocks for two months, every call public》

Hacker News • 07/23 06:35

据Hacker News, Show 发布 AI 行业观点与判断。这类信息反映行业参与者对技术路线和商业节奏的判断。后续需要关注官方披露的可用时间、开放范围、价格变化和第三方验证结果。

1. HuggingFace 遭 AI 智能体入侵, GLM-5.2 协助分析

X: Thomas Wolf (Hugging Face 联创/CSO) (@Thom_Wolf) • 07/22 23:37

HuggingFace 联合创始人 Thomas Wolf 透露, HuggingFace 上周遭复杂入侵, 攻击痕迹显示有严重 AI 参与, 但未知具体模型。闭源模型因安全护栏失效无法协助分析, 团队转而使用 Zhipu AI 的 GLM-5.2 开源模型。OpenAI 后续主动联系并合作调查, 确认入侵者为一个由未发布前沿模型驱动的全自主 AI 智能体, 试图访问 HuggingFace 部分基础设施。

2. OpenAI Workspace Agents 漏洞: 一个 ChatGPT 链接即可创建恶意 AI 智能体

The Decoder: AI News • 07/24 01:01

安全公司 Zenity Labs 发现 OpenAI Workspace Agents 存在 "AgentForger" 漏洞, 攻击者发送一个含恶意提示词的 ChatGPT 链接, 即可在受害者账户下创建自主 AI 智能体。该智能体继承受害者身份和已授权应用权限, 绕过安全审批, 并设置每五分钟运行一次的定时任务, 从攻击者邮箱获取指令执行。OpenAI 在四天内修复了该漏洞。

3. 英伟达、微软和Meta联合警告: 应避免对开放权重模型过度监管

Hacker News • 07/25 08:00

英伟达、微软和Meta联合签署公开信, 警告对开放权重AI模型的过度监管将削弱美国在AI领域的竞争力。信中指出, 开放权重模型能促进创新、降低准入门槛, 并支持学术研究。OpenAI和Anthropic未签署该信函。

1. Kimi K3 在网络安全漏洞利用测试中大幅落后美国前沿模型, 知识蒸馏或为原因

The Decoder: AI News • 07/25 08:00

英国AI安全研究所与美国AI标准与创新中心联合评估显示, 月之暗面的Kimi K3在ExploitBench基准上得分32.2%, 远低于美国领先模型的76.2%, 但优于智谱GLM-5.2的24.4%。

2. Anthropic 联合 Andon Labs 发布 Drone-Bench, 评估 AI 模型自主操控无人机执行定位追踪任务的能力

Anthropic: Research (发表成果 • 网页) • 07/25 08:00

Anthropic 与 Andon Labs 合作推出 Drone-Bench, 用于测试 AI 模型自主操控四旋翼无人机在室内环境中定位并追踪指定人员的能力。该基准将任务分解为 3D 地图重建、定位、导航、目标检测与跟随五个子任务, 并通过软件复现实现快速评估。实验表明, 该任务链的难度足以区分不同智能水平的模型, 并揭示 AI 在物理世界操控能力上的进步轨迹。

3. Claude 5 代模型上下文工程新规则: Claude Code 系统提示词精简超 80%

Claude 博客 • 07/25 08:00

Anthropic 为 Claude Opus 5 和 Claude Fable 5 等新一代模型删除了 Claude Code 超过 80% 的系统提示词, 且编码评测无显著损失。