

OpenAI发布图像模型新版本并宣布千禧年难题解答

中文内部早报 · 5分钟版

今日总览

OpenAI成为今日焦点，发布了ChatGPT Images 2.5图像模型并宣布其内部AI系统解决了Navier-Stokes千禧年大奖难题。同时，OpenAI向所有付费用户全面推送Astra模型。

Agent和Coding领域，Meta的Muse智能体产品开放更多用户体验，用户反馈显示其设计和速度表现良好但存在命名不够直观等问题。GPT-6 Astra操控Blender的教程也已发布。

模型与多模态方面，Runway发布Adobe插件可在Premiere Pro和After Effects内直接生成编辑内容。Anthropic分享了通过Claude Platform降低成本并提升性能的三个方法。

大厂与产业链板块，Dwarkesh Patel研究显示预训练进步主要来自数据改进而非模型改进。Anthropic据报道签署高达5170亿美元算力合同。

监管安全与研究工具领域，OpenAI发布关于对齐与监测困境的长文，指出链式思维监控效果正随模型能力提升而减弱。Berkeley RDI发布开源平台CUA-Lite面向计算机使用智能体。

头条

今日头条

1. OpenAI发布ChatGPT Images 2.5图像模型

OpenAI · 09/09 08:00 · 前沿模型与多模态

OpenAI发布ChatGPT Images 2.5图像模型，生成延迟比Images 2.0降低最多50%，细节、编辑精度、参考照片保真度和多轮编辑一致性均有提升。
2. Meta智能体产品Muse开放体验，官方回应用户好评

X: Alexandr Wang (Scale AI 创始人/Meta 首席 AI 官) (@alexandr_wang) · 09/09 07:48 · 智能体、编程与工作流

Meta的Muse智能体产品开放更多用户试用，入口为https://muse.ai/join，Meta首席AI官Alexandr Wang称团队为产品倾注心血并感谢用户反响。用户评价称其设计、速度和浏览器等智能体流程表现出色，具备Instagram等Meta产品原生集成优势，但也指出soul.md等命名对普通用户不直观、feed内容相关性不足等问题。
3. Dwarkesh Patel研究：预训练进步主要来自数据改进

Dwarkesh Patel: Podcast & Blog · 09/09 08:00 · 大厂与产业链

Dwarkesh Patel发布实验分析，在最高1e19 FLOPs的算力预算下训练2019至2025年各年度代表性模型配方与数据语料，发现数据改进带来12.0x算力效率提升，模型改进为3.7x，数据贡献约为模型的3.24倍。
4. Runway发布Adobe插件，可在Premiere Pro和After Effects内直接生成与编辑

Runway 新闻 · 09/09 08:00 · 前沿模型与多模态

Runway发布Runway Plugins，新面板可直接嵌入Premiere Pro和After Effects，在时间线内生成图像和视频、重绘片段并放置结果。Edit Studio可基于已有片段用Aleph 2按原始时长重新渲染，插件免费下载（macOS和Windows），生成功能面向所有付费计划并消耗现有额度。
5. Anthropic讲解用Claude Platform降低成本并提升性能的三个方法

X: Claude Devs (@ClaudeDevs) · 09/09 08:00 · 前沿模型与多模态

Anthropic团队文章指出，优化prompt cache命中率、清除升级到前沿Claude模型后的提示词反模式、校准effort三个手段可在不牺牲性能的情况下降低成本。

模型

前沿模型与多模态

1. OpenAI 发布新动态：《Funding grants for new research into AI and teen development》

OpenAI 新闻 · 09/08 17:00

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. OpenAI 宣布以内部 AI 系统给出 Navier–Stokes 千禧年问题解答

OpenAI · 09/09 08:00

OpenAI 宣布其内部 AI 系统给出 Navier–Stokes 存在与光滑性问题的解答，证明初始光滑的流体可在有限时间内形成奇点，并附证明文稿与 Lean 形式化验证。
3. OpenAI 发布新动态：《The Work Now Within Reach》

OpenAI 新闻 · 09/08 21:00

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。

4. OpenAI 发布新动态：《Supporting independent journalism in Ukraine》

OpenAI 新闻 · 09/07 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。

智能体 智能体、编程与工作流

1. OpenAI 智能体团队宣布求解 Navier-Stokes 千禧年大奖难题并向独立证明者致贺

X: OpenAI (@OpenAI) · 09/09 01:23

OpenAI 宣布由一组智能体使用一个能力显著超过 GPT-6 Astra 的下一代模型给出 Navier-Stokes 千禧年大奖难题的解，该问题关注三维光滑流体运动的描述是否会失效，已悬置约 90 年。

2. OpenAI 发布新动态：《How GPT-5.6 Sol helps run quantum computing experiments》

OpenAI 新闻 · 09/09 01:00

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。

3. OpenAI 发布新动态：《Research acceleration: The view inside OpenAI》

OpenAI 新闻 · 09/06 16:00

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。

4. GPT-6 Astra操控Blender保姆级教程：三种玩法与踩坑实录

公众号：数字生命卡兹克 · 09/09 08:00

作者数字生命卡兹克发布GPT-6 Astra操控Blender的小白教程，介绍前期安装、官方MCP与Computer Use插件配置，并实测三种玩法。Computer Use花约4小时搭出天坛祈年殿，但耗掉200美刀Pro会员近半额度；MCP更快，可完成摩托车建模与组装动画，但复杂任务会单次运行超时，可拆分步骤或用CLI执行Python脚本解决。

产业 大厂与产业链

1. OpenAI 推出 AI 产业部署新进展

X: OpenAI (@OpenAI) · 09/09 08:00

据X: OpenAI (@OpenAI)，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。

2. NYU 数学家指控 OpenAI 在千禧年难题竞赛中不正当竞争，Bubeck 否认

TechCrunch AI · 09/09 08:00

NYU 数学教授 Tristan Buckmaster 与 Anthropic 数学家 Levent Alpöge 公布针对 Navier-Stokes 存在与光滑性这一千禧年难题的三项证明初步结果，并称其研究进展信息被泄露给 OpenAI，对方随后用大量算力沿其独特路线追赶证明。

3. Anthropic 据报道签约高达 5170 亿美元算力协议，锁定至少 14.8 GW 算力

The Decoder: AI News · 09/08 02:12

据 The Information 报道，Anthropic 在十一个月内签署了价值高达 5170 亿美元的算力合同，自 2025 年 10 月以来锁定至少 14.8 GW 算力，并计划自建数据中心。

4. TechCrunch 推出 AI 产业部署新进展

TechCrunch AI · 09/09 05:10

据TechCrunch AI，TechCrunch 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。

开源 开源项目与开发者工具

1. Hugging Face探讨AI安全边界：拒绝话题子集而非整个话题

Hugging Face Blog · 09/08 22:23

Hugging Face博客发布文章《Safety for Whom? Refusing the Right Subset of a Topic, Not the Whole Topic》，探讨AI系统在处理敏感话题时的安全机制设计原则。文章强调应精确拒绝特定子集内容而非完全屏蔽整个话题领域，以平衡安全性和功能性需求。

1. OpenAI 长文阐述对齐与监测困境，称 CoT 监控能力正在减弱

OpenAI · 09/06 17:00

OpenAI 发布长文《An Alien Mind》，回溯 2023 年 RLSlow 项目中确认推理模型可扩展训练的起点，并系统阐述目标对齐与价值对齐的区分。文章指出链式思维监控的效果正随着模型能力提升而逐步减弱，GPT-6 Astra 在对齐上显著优于 GPT-5.6 Sol；作者预期进展可能持续走向机器递归自我改进（RSI），呼吁自愿放缓扩展、建立第三方安全门槛并加强国际协调。

2. Simon Willison 评 OpenAI 用未发布模型求解 Navier-Stokes 千禧年大奖难题之争

Simon Willison 博客 · 09/09 07:55

OpenAI 用未发布模型在约88小时内给出 Navier-Stokes 存在与光滑性问题（七大千禧年难题之一）的解答，并通过 GPT-6 Astra 完成17小时 Lean 形式化验证，全程发送 490万条消息、消耗约3000亿输出 token。

3. OpenAI 更新 AI 治理与安全动态

OpenAI 新闻 · 09/08 08:00

据OpenAI 新闻，OpenAI 更新 AI 治理与安全动态。这类进展主要影响数据、版权、安全和合规部署边界。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. Berkeley RDI 发布开源平台 CUA-Lite，面向计算机使用智能体

Berkeley RDI: Blog (AI 安全与评测) · 09/07 00:00

Berkeley RDI 推出开源平台 CUA-Lite，为计算机使用智能体提供三个标准化抽象。其环境接口下运行 15+ 个覆盖桌面、浏览器和移动端的基准，并提供含 30k+ 可验证任务的免 VM 桌面沙箱；统一监督数据格式已转换 10+ 个公开 CUA 数据集；每模型一个 harness 在评测、SFT 和 RL 间共享，支持 14 个模型族。

2. Sam Altman 回应与 Anthropic 研究员的 Navier-Stokes 证明发布争议

X: Sam Altman (@sama) · 09/09 08:00

Sam Altman 发文回应与 Anthropic 一方围绕 Euler/Navier-Stokes 证明发布的争议，称对方只有 Euler 结果、双方协调失败，并称对方以抄袭指控相威胁。

3. Fortune 报道 OpenAI 多次修改 GPT-6 Astra 基准测试数据，部分成绩大幅变化

IT之家 · 09/06 14:46

据 Fortune 报道，OpenAI 自 9 月 3 日发布 GPT-6 Astra 公告以来多次修改评测基准数据：Astra 幻觉率曾从 4.2% 降至 2% 后又改回。