

# AI 行业日报

模型 · 产品 · 产业 · 研究 · 观点 | Data Source: aihot.virxact.com

API耗时: 2s 精选条目: 45 条 焦点: 8 条 快讯: 0 条

## Executive Summary

### 今日重点

Luma Agents 发布全新自动化广告创作功能，能够将情绪板转化为完整广告内容，标志着创意生成从静态参考向全流程自动化的跨越。OpenAI Daybreak 正式推出，这是专为网络安全防御设计的AI系统，整合了OpenAI最强模型和Codex技术，旨在加速威胁检测和软件安全保障。Replit并行代理支持最多10个代理同时运行，每个代理拥有独立应用副本，在各自计算环境中工作后合并结果，显著提升了开发效率。

### 技术动向

通过 SocialReasoning Bench 测试发现，当前AI智能体在执行能力方面表现良好，但在优化用户立场方面存在明显不足——即使明确要求优化用户利益，模型仍无法持续改善用户处境。这揭示了AI对齐技术需要在用户价值导向方面进行深度改进。Anthropic 在AWS上正式推出Claude平台，首次将全套Claude API功能引入AWS生态，包括托管智能体、代码执行、文件API等核心功能，实现了与AWS现有认证和计费体系的无缝集成。

### 近期关注

本周三在旧金山将举行AI模型开发者盛会，Anthropic、MiniMax 等公司将参与下一代模型展示。Anthropic 开源的金融AI全栈模板库已发布，包含10个端到端智能体和7个垂直行业插件，为金融服务AI落地提供标准化解决方案。

## 今日焦点

### ★ 1. AntLingAGI发布万亿参数模型Ring-2.6-1T

X: [OpenRouter \(@OpenRouter\)](#) · 10 小时前 · 模型发布/更新

AntLingAGI发布了其万亿参数旗舰"思考模型"Ring-2.6-1T，该模型在5月15日前可通过OpenRouter免费使用。其核心特性包括可调节的思考强度，能动态平衡认知深度、token成本和执行速度；专为智能体优化，适用于高频工作流，提供快速的多步执行和工具调用；并具备深度思考能力，以应对严密的数学逻辑和科学研究任务。模型旨在满足实际生产环境中复杂任务的需求。

<https://x.com/OpenRouter/status/2053834087023133025>

### ★ 2. SenseNova U1图像生成模型登陆ComfyUI平台

X: [高汤 SenseTime \(@SenseTime\\_AI\)](#) · 14 小时前 · 模型发布/更新

SenseNova U1图像生成模型现已在ComfyUI上可运行，并获得包括REBEL AI在内的评测者高度认可。REBEL AI发布的实践教程展示了该模型的部署工作流，并对其图像生成能力进行了真实场景测试。模型支持8步快速推理，生成速度极快，应用场景涵盖人像、超现实艺术、文字标志和生物设计等。相关资源已在Hugging Face、GitHub和Discord平台开放。

[https://x.com/SenseTime\\_AI/status/2053782644492525847](https://x.com/SenseTime_AI/status/2053782644492525847)

### ★ 3. Luma Agents：从情绪板到完整广告的自动化创作

X: [Luma AI \(@LumaLabsAI\)](#) · 3 小时前 · 产品发布/更新

情绪板曾是最精彩的部分。如今它只是起点。上传你的参考素材。设定方向。Luma Agents 会将其从情绪板转化为完整的广告。生成广告 → <http://lumalabs.ai/app>

<https://x.com/LumaLabsAI/status/2053941836402901409>

### ★ 4. 推出Daybreak：网络安全前沿AI防御系统

X: [OpenAI \(@OpenAI\)](#) · 3 小时前 · 产品发布/更新

推出Daybreak：面向网络防御者的前沿AI。Daybreak汇集了最强大的OpenAI模型、Codex以及我们的安全合作伙伴，以加速网络防御并持续保障软件安全。朝着安全团队能以防御所需速度行动的未来迈进一步。

<https://x.com/OpenAI/status/2053939702110269822>

### ★ 5. 纳德拉出庭反击马斯克诉讼案：他明明有我的手机号，但从未抱怨过微软与 OpenAI 的合作

IT之家 (RSS) · 1 小时前 · 产业与资本

微软CEO纳德拉在马斯克诉OpenAI案中作证，反驳了马斯克关于OpenAI背离非营利初衷的指控。他出示了马斯克2016年的感谢邮件，证明其当时完全支持微软与OpenAI的合作，且微软为此承受了1500万美元亏损。纳德拉强调，双方一直保持沟通，但马斯克在合作深化期间从未提出抱怨，直至2024年才提起诉讼。他还将2023年OpenAI董事会解雇CEO奥特曼的事件评价为"业余之举"，并透露自己曾介入

<https://www.ithome.com/0/949/045.htm>

## ★ 6. Anthropic估值五日激增2000亿美元，营收呈指数级增长

X: [Kim \(@kimmonismus\)](#) · 5 小时前 · 产业与资本

人工智能公司Anthropic的市场隐含估值在短短五天从1.2万亿美元飙升至1.4万亿美元，暴增2000亿美元。根据链上Pre-IPO交易数据，其估值自2025年10月以来已上涨1067%，近期24天内又录得40%的涨幅。这一惊人飙升的背景是公司年化收入的爆炸式增长：从2023年的1亿美元跃升至当前的450亿美元，仅过去12个月就增长了1400%。在Jupiter等链上平台交易的、由SPV风险敞

<https://x.com/kimmonismus/status/2053911732028907972>

## ★ 7. 智能体执行能力强但优化用户立场不足

X: [Microsoft Research \(@MSFTResearch\)](#) · 6 小时前 · 论文与研究

通过SocialReasoning Bench测试发现，各模型呈现稳定模式--智能体能够胜任执行任务，但即便在明确要求优化用户利益的指令下，仍无法持续改善用户处境。https://msft.it/6011vPOLF

<https://x.com/MSFTResearch/status/2053890478165147693>

## ★ 8. 小块有大智慧？这下真成真了！

X: [Berry Xia \(@berryxia\)](#) · 13 小时前 · 论文与研究

一项新研究证明，一个通过强化学习训练的7B语言模型能够有效指挥GPT-5、Claude Sonnet 4和Gemini 2.5 Pro等前沿大模型。该模型通过编写自然语言子任务、分配给不同大模型执行，并精确指定上下文信息，在GPQA Diamond、LiveCodeBench和AIME25等硬核基准测试中，其性能全面超越了单个前沿模型。该系统平均每个问题仅需调用约三次大模型，比手动设计的多代理流程

<https://x.com/berryxia/status/2053794010582683759>

## 模型 模型发布/更新

### 1. AntLingAGI发布万亿参数模型Ring-2.6-1T

X: [OpenRouter \(@OpenRouter\)](#) · 10 小时前

AntLingAGI发布了其万亿参数旗舰"思考模型" Ring-2.6-1T，该模型在5月15日前可通过OpenRouter免费使用。其核心特性包括可调节的思考强度，能动态平衡认知深度、token成本和执行速度；专为智能体优化，适用于高频工作流，提供快速的多步执行和工具调用；并具备深度思考能力，以应对严密的数学逻辑和科学研究任务。模型旨在满足实际生产环境中复杂任务的需求。

<https://x.com/OpenRouter/status/2053834087023133025>

### 2. SenseNova U1图像生成模型登陆ComfyUI平台

X: [商汤 SenseTime \(@SenseTime\\_AI\)](#) · 14 小时前

SenseNova U1图像生成模型现已在ComfyUI上可运行，并获得包括REBEL AI在内的评测者高度认可。REBEL AI发布的实践教程展示了该模型的部署工作流，并对其图像生成能力进行了真实场景测试。模型支持8步快速推理，生成速度极快，应用场景涵盖人像、超现实主义、文字标志和生物设计等。相关资源已在Hugging Face、GitHub和Discord平台开放。

[https://x.com/SenseTime\\_AI/status/2053782644492525847](https://x.com/SenseTime_AI/status/2053782644492525847)

### 3. 腾讯混元Hy3预览版发布，专注复杂智能体任务

X: [腾讯混元 \(@TencentHunyuan\)](#) · 14 小时前

腾讯混元Hy3预览版已开放早期体验，被描述为该系列最强模型。其核心设计面向现实世界有效性，而非单纯追求基准测试分数。模型具备处理复杂智能体任务的能力，采用256K上下文长度，并融合了快慢思维机制的混合专家架构。该模型基于重建的预训练和强化学习基础设施构建，旨在实现大规模应用下的高成本效益。

<https://x.com/TencentHunyuan/status/2053771684482851255>

## 产品 产品发布/更新

### 1. Luma Agents：从情绪板到完整广告的自动化创作

X: [Luma AI \(@LumaLabsAI\)](#) · 3 小时前

情绪板曾是最精彩的部分。如今它只是起点。上传你的参考素材。设定方向。Luma Agents会将其从情绪板转化为完整的广告。生成广告 → <http://lumalabs.ai/app>

<https://x.com/LumaLabsAI/status/2053941836402901409>

### 2. 推出Daybreak：网络安全前沿AI防御系统

X: [OpenAI \(@OpenAI\)](#) · 3 小时前

推出Daybreak：面向网络防御者的前沿AI。Daybreak汇集了最强大的OpenAI模型、Codex以及我们的安全合作伙伴，以加速网络防御并持续保障软件安全。朝着安全团队能以防御所需速度行动的未来迈进一步。

<https://x.com/OpenAI/status/2053939702110269822>

### 3. Codex插件助力AI应用与智能体开发提速

X: [OpenAI Developers \(@OpenAIDevs\)](#) · 4 小时前

Codex现可通过OpenAI Developers插件，帮助您利用OpenAI API更快地构建AI应用和智能体。

<https://x.com/OpenAIDevs/status/2053925962287583379>

#### 4. Claude Code v2.1.139 版本更新

Claude Code: [GitHub Releases \(RSS\)](#) · 5 小时前

本次更新引入了多项新功能与优化。核心新增包括：集中管理会话的Agent视图（研究预览）、可设置目标并持续工作的`/goal`命令、实时调整滚轮速度的`/scroll-speed`命令，以及查看插件详情的`claude plugin details`命令。交互界面导航与控制能力得到增强。底层优化涵盖MCP服务器可获取`CLAUDE\_PROJECT\_DIR`环境变量、`/context all`的令牌

<https://github.com/anthropics/claude-code/releases/tag/v2.1.139>

#### 5. Replit推出并行代理 支持10个代理同时运行

X: [Replit \(@Replit\)](#) · 6 小时前

认识Replit并行代理 通过并行运行多达10个代理来加速构建 每个代理都拥有您应用的独立副本 它们在各自的计算机上工作 然后以代理方式合并工作成果

<https://x.com/Replit/status/2053891504989753817>

#### 6. 个人智能助手定制专属旅行计划

X: [Gemini \(@GeminiApp\)](#) · 6 小时前

Personal Intelligence 可帮助您打造完全个性化的旅行行程。它通过关联您 @Gmail、@GooglePhotos、@Google 搜索和 @YouTube 历史记录中的信息点来节省时间。您可随时选择要连接的应用程序并管理个性化设置

<https://x.com/GeminiApp/status/2053889747752190393>

#### 7. Anthropic在AWS上正式推出Claude平台

Claude: [Blog \(网页\)](#) · 7 小时前

Anthropic公司正式在AWS上推出Claude平台，为AWS客户提供了通过其现有身份验证、账单及承诺消费抵扣使用完整Claude功能的新途径。该平台首次将全套Claude API功能引入AWS生态，新功能与原生API同日上线。平台包含Claude托管智能体、代码执行、文件API等多项核心功能，并支持最新模型。与Amazon Bedrock上的服务不同，此平台由Anthropic直接运营，数据

<https://claude.com/blog/claude-platform-on-aws>

#### 8. Anthropic开源金融AI全栈模板，定义行业落地新标准

X: [小北 \(@frxiaobei\)](#) · 8 小时前

Anthropic在GitHub开源了金融服务行业AI解决方案完整模板库，包含10个端到端智能体、7个垂直行业插件及11家主流金融数据商的数据连接器，覆盖投研、投行、风控等核心工作流。该库提供了从个人插件到企业API的部署方式，支持集成至Microsoft 365及私有云。此举为金融AI落地提供了开箱即用的标准作业程序，与OpenAI的消费级路线形成鲜明对比，凸显了其深耕企业场景、通过开源构建行

<https://x.com/frxiaobei/status/2053861985008431398>

#### 9. 告别编写YAML：使用confingy配置机器学习系统

Runway: [News \(网页\)](#) · 9 小时前

Runway开源了Python库confingy，旨在解决机器学习系统配置的长期痛点。该库允许开发者用纯Python代码（支持懒加载、类型检查和序列化）替代传统YAML配置文件，从而摆脱YAML作为图灵完备领域特定语言所带来的维护困境。confingy无需重构现有代码，即可满足跟踪构造函数参数、避免实例化昂贵对象（如大语言模型）等核心需求，有效改善了因复杂YAML配置导致的无法跳转定义、类型提示失

<https://runwayml.com/news/stop-writing-yaml-configuring-ml-systems-with-confingy>

#### 10. 帕累托代码以市场需求重塑模型选择

X: [OpenRouter \(@OpenRouter\)](#) · 9 小时前

Pareto Code 是一种利用真实市场需求观察帕累托前沿的新方法 DeepSeek V4 Pro 占据首位，其次是 GPT 5.4 Mini 和 Gemini 3.1 Pro <https://openrouter.ai/openrouter/pareto-code>

<https://x.com/OpenRouter/status/2053843701261025290>

#### 11. 中国移动上线AI模型中转平台MoMA，国家队入局AI基础设施竞争

X: [阿易 AI Notes \(@AYI\\_Alnotes\)](#) · 19 小时前

中国移动推出的AI模型中转平台MoMA，标志着"国家队"正式进入AI基础设施领域。该平台已接入DeepSeek、通义千问等300多个主流模型，并通过央视新闻进行宣传。其核心逻辑在于，AI中转站被视为未来AGI时代的"智能电网"，是掌握行业定价权与未来的关键。用户可在移动云官网搜索"MoMA"获取体验包，建议先行测试实际效果。此举意味着AI行业竞争已上升至基础设施层面。

[https://x.com/AYI\\_Alnotes/status/2053699900727075150](https://x.com/AYI_Alnotes/status/2053699900727075150)

#### 12. HappyHorse AI视频引擎登陆阿里云

X: [阿里云 / Alibaba Cloud \(@alibaba\\_cloud\)](#) · 23 小时前

资产审核：通过。物理逻辑：无缝衔接。HappyHorse是面向生产就绪内容的排名第一的AI视频引擎。从复杂的物理交互到原生1080p唇形同步，我们不仅生成--更精准执行。现已上线阿里云Model Studio。 <https://int.alibabacloud.com/m/1000412167/>

[https://x.com/alibaba\\_cloud/status/2053646520998560033](https://x.com/alibaba_cloud/status/2053646520998560033)

#### 13. OpenCLI打通微信等私域信息流，聚合个人数据

X: [Vista \(@vista8\)](#) · 23 小时前

OpenCLI项目实现了对微信、Telegram和Discord三大平台内容的命令行读取，通过wx-cli、tg-cli和discord-cli工具，用户可直接获取群消息、聊天记录、朋友圈及收藏夹等私域数据。这标志着个人信息流聚合的关键突破，使得AI Agent不仅能监控外部资讯网站，还能整合个人私密的社交聊天信息，构建真正统一的个人数据流。此举可能引发平台方如微信的关注或反应。

<https://x.com/vista8/status/2053633346581000600>

## 14. NousResearch发布Hermes配置Pareto Code指南

X: [OpenRouter \(@OpenRouter\)](#) · 昨天 02:36

@NousResearch 关于如何在 Hermes 中设置 Pareto Code 的文档: <https://hermes-agent.nousresearch.com/docs/user-guide/configuration#openrouter-routing--pareto-code-for-auxiliary-tasks>  
<https://x.com/OpenRouter/status/2053544645410324774>

## 产业 产业与资本

### 1. 纳德拉出庭反击马斯克诉讼案：他明明有我的手机号，但从未抱怨过微软与 OpenAI 的合作

IT之家 (RSS) · 1 小时前

微软CEO纳德拉在马斯克诉OpenAI案中作证，反驳了马斯克关于OpenAI背离非营利初衷的指控。他出示了马斯克2016年的感谢邮件，证明其当时完全支持微软与OpenAI的合作，且微软为此承受了1500万美元亏损。纳德拉强调，双方一直保持沟通，但马斯克在合作深化期间从未提出抱怨，直至2024年才提起诉讼。他还将2023年OpenAI董事会解雇CEO奥尔特曼的事件评价为"业余之举"，并透露自己曾介入

<https://www.ithome.com/0/949/045.htm>

### 2. Anthropic估值五日激增2000亿美元，营收呈指数级增长

X: [Kim \(@kimmonismus\)](#) · 5 小时前

人工智能公司Anthropic的市场隐含估值在短短五天内从1.2万亿美元飙升至1.4万亿美元，暴增2000亿美元。根据链上Pre-IPO交易数据，其估值自2025年10月以来已上涨1067%，近期24天内又录得40%的涨幅。这一惊人飙升的背景是公司年化收入的爆炸式增长：从2023年的1亿美元跃升至当前的450亿美元，仅过去12个月就增长了1400%。在Jupiter等链上平台交易的、由SPV风险敞

<https://x.com/kimmonismus/status/2053911732028907972>

### 3. 旧金山AI模型开发者盛会周三举行

X: [MiniMax \(@MiniMax\\_AI\)](#) · 6 小时前

本周三在旧金山，与@vercel\_dev、@AnthropicAI和@Kimi\_Moonshot共同参与美国下一代模型盛会。每位参与者将获得\$30 MiniMax API积分 + 大奖池。📅5月13日 | 太平洋时间下午5:30 | 旧金山加入候补名单 → <https://luma.com/ttrgmjaf>

[https://x.com/MiniMax\\_AI/status/2053899661334003938](https://x.com/MiniMax_AI/status/2053899661334003938)

### 4. MiniMax组建"10x团队"，邀请领域专家共推AI跨越式发展

X: [MiniMax \(@MiniMax\\_AI\)](#) · 8 小时前

MiniMax宣布成立"10x团队"，旨在邀请各领域深度专家共同推动人工智能实现十倍速进步。公司认为大模型已从演示阶段发展为真实生产力设施，而将顶尖领域人才引入模型研发能显著提升能力并产生实际影响。团队寻求的是坚信AI能加速自身领域发展的专家，他们将直接参与问题定义、评估构建及 workflow 设计，其成果将融入MiniMax模型并贡献于推动行业发展的开源基准。MiniMax提供全球竞争力的薪酬、股权及上海、

[https://x.com/MiniMax\\_AI/status/2053860184532066737](https://x.com/MiniMax_AI/status/2053860184532066737)

### 5. OpenAI 推出 DeployCo 以协助企业围绕智能构建业务

OpenAI: [官网动态 \(RSS\)](#) · [排除企业/客户案例](#) · 18 小时前

OpenAI 正式推出全新企业部署公司 DeployCo，旨在帮助各类组织将前沿人工智能技术投入实际生产，并转化为可衡量的商业影响。该举措标志着 OpenAI 进一步深入企业服务领域，专注于解决 AI 模型从测试到规模化部署的落地挑战，助力企业通过定制化部署方案提升运营效率与业务成果。

<https://openai.com/index/openai-launches-the-deployment-company>

### 6. 消息称 AI 芯片企业 Cerebras IPO 获 20+ 倍超额认购，拟调升发行价近三成

IT之家 (RSS) · 23 小时前

AI芯片企业Cerebras的IPO获得超过20倍超额认购，有望成为2026年以来全球最大IPO。公司计划将发行股票数量从2800万股增至3000万股，发行价区间从115-125美元上调至150-160美元，以中间价计涨幅达29.17%，最高可筹资48亿美元。Cerebras的"晶圆级芯片"因集成大量高带宽SRAM缓存，特别适合AI推理的解码步骤，已获得亚马逊和OpenAI的大额订单。

<https://www.ithome.com/0/948/591.htm>

### 7. Anthropic无补贴登顶令牌份额榜首

X: [OpenRouter \(@OpenRouter\)](#) · 昨天 02:53

即使没有任何补贴 Anthropic 已在令牌份额排行榜上位列第一

<https://x.com/OpenRouter/status/2053548965220475115>

## 论文 论文与研究

### 1. 智能体执行能力强但优化用户立场不足

X: [Microsoft Research \(@MSFTResearch\)](#) · 6 小时前

通过SocialReasoning Bench测试发现，各模型呈现稳定模式--智能体能够胜任执行任务，但即便在明确要求优化用户利益的指令下，仍无法持续改善用户处境。

<https://msft.it/6011vPOLF>

<https://x.com/MSFTResearch/status/2053890478165147693>

## 2. 小块有大智慧？这下真成真了！

X: [Berry Xia \(@berryxia\)](#) · 13 小时前

一项新研究证明，一个通过强化学习训练的7B语言模型能够有效指挥GPT-5、Claude Sonnet 4和Gemini 2.5 Pro等前沿大模型。该模型通过编写自然语言子任务、分配给不同大模型执行，并精确指定上下文信息，在GPQA Diamond、LiveCodeBench和AIME25等硬核基准测试中，其性能全面超越了单个前沿模型。该系统平均每个问题仅需调用约三次大模型，比手动设计的多代理流程

<https://x.com/berryxia/status/2053794010582683759>

## 观点 技巧与观点

### 1. 在AWS上进行基础模型训练与推理的核心构建模块

Hugging Face: [Blog \(RSS\)](#) · 1 小时前

本文面向使用开源框架的机器学习工程师，阐述了AWS如何为大规模基础模型的全生命周期提供核心基础设施。其核心是三大紧密集成的组件：配备多代NVIDIA GPU（如H100、H200及新一代Blackwell B200/B300）的大显存加速计算实例；用于集体通信的高带宽、低延迟网络（节点内NVLink与节点间EFA）；以及可扩展的分布式存储。这些基础设施与Slurm/Kubernetes等资源编排系

<https://huggingface.co/blog/amazon/foundation-model-building-blocks>

### 2. 在脚本的shebang行中使用LLM

Simon Willison [博客](#) · 5 小时前

本文介绍了一种创新方法，将LLM工具直接嵌入脚本的shebang行中执行指令。通过LLM的fragments模式，shebang行可执行简单任务，如生成SVG图像；使用-T选项能调用外部工具（如llm\_time）创作包含当前时间的俳句；更复杂的是运行YAML模板，其中定义了Python函数作为工具进行数学计算，示例中演示了计算2344乘以5252加134的过程，并通过调试输出展示了分步结果，最终得

<https://simonwillison.net/2026/May/11/llm-shebang>

### 3. 用户展示PixVerse生成的高端时尚网格图像

X: [PixVerse \(@PixVerse\\_\)](#) · 6 小时前

用户通过引用推文展示了由GPT Image 2在PixVerse平台上生成的一组高端时尚编辑图像。该图像采用2行4列的网格布局，包含8个独立面板。每个面板均以不同鲜艳纯色或渐变背景呈现一位身着现代街头服饰的年轻男模特，搭配滑板、吉他等道具，并采用简洁的工作室布景。图像整体追求商业广告摄影风格，具有奢侈品品牌广告的美学质感，强调柔和灯光、清晰阴影、逼真皮肤和电影级质感，且不含任何文字或标识。

[https://x.com/PixVerse\\_/status/2053890377967620257](https://x.com/PixVerse_/status/2053890377967620257)

### 4. 人机交互界面的演进：从文本到交互式神经视频

X: [Andrej Karpathy \(@karpathy\)](#) · 8 小时前

作者认为，AI的人机交互界面正从当前的Markdown默认输出向更丰富的HTML格式演进，后者能提供更好的图形、布局和交互性。长远看，由扩散神经网络直接生成的交互式神经视频或模拟将是终极输出形式，但其如何与精确的传统软件结合仍是开放问题。在输入侧，仅靠音频、文本或视频并不足够，需整合如手势指点等更自然的交互方式。总体而言，在迈向脑机接口之前，人机输入输出的融合仍有大量工作要做。现阶段的一个实用建议

<https://x.com/karpathy/status/2053872850101285137>

### 5. 谷歌DeepMind与Coursera推出"Gemini for Developers"开发者专项课程

X: [Google AI for Developers \(@googleaidevs\)](#) · 8 小时前

谷歌DeepMind与Coursera合作推出的"Gemini for Developers"专项课程现已开放注册。该课程旨在指导开发者利用Gemini模型构建可用于生产环境的AI解决方案。其核心涵盖三大模块："推理与行动"使AI应用能推理并执行复杂任务；"连接与自动化"通过函数调用将Gemini与现实世界工具集成；"规模化与信心"则专注于构建、测试和部署可扩展的AI系统。课程强调超越单纯文本生成

<https://x.com/googleaidevs/status/2053868609747746897>

### 6. 菲尔兹奖得主亲测 ChatGPT 5.5 Pro：17 分钟出论文级成果，替学生拉响红色警报

IT之家 (RSS) · 8 小时前

菲尔兹奖得主 Timothy Gowers 测试 ChatGPT 5.5 Pro，AI 在 17 分钟内独立解决了一个加法数论公开难题，产出了博士论文级别的成果。整个过程无需数学指导，仅通过简单提示完成。Gowers 警告，若 AI 数学能力按此速度发展，将很快对数学研究领域构成危机，尤其冲击博士生培养。他呼吁数学系紧急应对，帮助学生寻找新出路。另一位菲尔兹奖得主陶哲轩则指出，人类数学家在 AI

<https://www.ithome.com/0/949/032.htm>

### 7. 传奇总部"Cog House"首度公开：天才创始人Scott Wu与Cognition AI的崛起之路

X: [swyx \(@swyx\)](#) · 9 小时前

推文首次公开了Cognition AI总部"Cog House"的内部影像。创始人Scott Wu是美国顶尖竞技程序员，自幼展现数学天赋，曾三获国际信息学奥赛金牌。2023年11月，他在母亲去世与Sam Altman被OpenAI解雇的同日创立Cognition，并预见AI将发展为全天候工作的智能体。公司推出的AI软件工程师Devin虽初期表现不佳，但上线18个月内已实现4.45亿美元年化营收，

<https://x.com/swyx/status/2053856676969890189>

### 8. ChatGPT 在 2026 年初如何扩大应用范围

OpenAI: [官网动态 \(RSS\)](#) · [排除企业/客户案例](#) · 9 小时前

2026年第一季度，ChatGPT的用户采用率大幅上升，其中35岁以上用户群体的增长最为迅速，同时用户性别比例也趋于均衡。这表明以ChatGPT为代表的人工智能技术正加速融入主流社会，其应用范围已突破早期以年轻技术爱好者为主的局限，向更广泛、更多元化的普通用户群体拓展。

<https://openai.com/signals/research/2026q1-update>

9. 兄弟们！这套结构，这个图理解后。GPT-Image-2 生图也就算是入门了吧。

X: [Berry Xia \(@berryxia\)](#) · 10 小时前

推文介绍了一套提升AI生图质量的结构化提示词框架。其核心是“结构越清晰，输出质量越好”。框架将提示系统分解为几个关键模块：首先定义画幅与用途；其次明确主题与核心主体的位置、占比及情绪价值；接着通过视觉隐喻用具体元素象征抽象概念；然后选择如Apple风格等特定画面风格并强调干净、克制等视觉要求；最后规划包含主副标题、英文短句和关键词的文字系统，并列出现应避免的限制元素。

<https://x.com/berryxia/status/2053842082603937926>

10. AI工具批量生成知识产权申请材料引关注

X: [小北 \(@frxiaobei\)](#) · 11 小时前

知识产权全面沦陷。发明专利生成 skill <https://github.com/handsomestWei/patent-disclosure-skill> 软著生成 skill <https://github.com/Fokkyp/SoftwareCopyright-Skill> 设计专利/实用新型更容易，claude desgin/ image2

<https://x.com/frxiaobei/status/2053827716605169804>

11. 3.3万星AI论文学习库，收录精选视频教程

X: [Vista \(@vista8\)](#) · 11 小时前

一个在GitHub上获得3.3万星标的AI论文学习资源库，作者系统性地收集和整理了高质量的AI学习视频。资源主要来自YouTube和Bilibili平台，例如收录了李沐的论文精讲系列。所有内容已按时间和主题进行了详细分类，便于用户系统学习和收藏。该库旨在为学习者提供一个结构化的视频学习路径。

<https://x.com/vista8/status/2053816596762968128>

12. 高效AI工作流：结合ChatGPT与PixVerse精准生成品牌香水广告

X: [PixVerse \(@PixVerse\\_\)](#) · 16 小时前

用户分享了一个高效AI视频创作流程。他首先利用ChatGPT，基于预设镜头描述，生成了一个结构严谨、支持多镜头序列的详细视频提示词。该提示词严格指定了品牌元素（如瓶身、包装、Logo及阿拉伯文/英文文本）需通过参考附件精确还原，并描述了超奢华、电影感的视觉风格。随后，他将此提示词输入PixVerse平台，借助其高速生成优势，快速产出了1080p分辨率的视频。此工作流有效结合了GPT的精准指令生成能

[https://x.com/PixVerse\\_/status/2053745705618170030](https://x.com/PixVerse_/status/2053745705618170030)

13. 开源PPT工具“鬼藏PPT技能”迎重大更新，新增瑞士风格与AI配图功能

X: [锦藏 \(@op7418\)](#) · 22 小时前

开源项目“鬼藏PPT技能”迎来重大更新，新增瑞士国际主义视觉风格，提供克莱因蓝等四套主题色。核心升级包括：通过接入GPT-Image 2.0，可根据PPT内容与风格自动生成胶片质感配图、流程图及UI截图美化；支持基于同一内容一键生成公众号、小红书、视频号等多种规格的封面图。更新旨在解决用户对多风格、自动配图及跨平台适配的需求，并通过预设22种版式和严格的视觉规则，确保设计的一致性与专业性。

<https://x.com/op7418/status/2053657613460771142>

14. 本地模型处理半数日常任务，响应速度优于云端

[Tomer Tunguz 博客 \(VC 分析\)](#) · 昨天 08:00

一项为期五周的实验发现，在总计约1400项日常工作任务中，约50%可由本地35B参数模型（如Qwen 3.6 35B）成功处理，涵盖邮件、日程、总结和行政事务等类别。性能对比显示，本地模型在常规代理任务上的平均响应时间为2.8秒，比云端Claude Opus 4.5快2.1倍，尽管后者在复杂推理上仍领先约20%。本地模型输出更简洁，云端模型则在结构和代码规范性上更优。随着本地模型性能提升，将计算负

<https://www.tomtunguz.com/localmaxxing>

15. Codex自主完成安全审计并赚取赏金

X: [Sam Altman \(@sama\)](#) · 昨天 04:01

用户指示AI模型Codex去赚取5美元，Codex自主完成了一系列任务：寻找开源安全审计赏金项目，提交有效的拉取请求，与维护者沟通，并处理了GitHub验证流程，最终使工作被合并。经过约22小时的工作，用户获得了16.88美元的首笔付款。按此推算，若每日重复，月收入可达506.40美元。这初步实现了Sam Altman关于AI能主动为人赚钱的愿景，虽然金额尚小，但标志着一个令人兴奋的开端。

<https://x.com/sama/status/2053566155571560868>

16. 旧版AI模型急诊诊断已超越人类医生

X: [Kim \(@kimmonismus\)](#) · 昨天 02:54

一项发表于《科学》的研究显示，OpenAI一年前发布的o1模型在急诊诊断中表现优于医生。该模型在真实、混乱的急诊数据测试中，正确或接近正确诊断率达67%，而医生为50-55%，尤其在信息有限的早期分诊阶段优势最明显。研究指出，o1模型在结构化病例中的临床推理近乎完美，且该模型按AI标准已属旧版，当前模型可能更强。研究未涵盖长期住院数据及影像诊断，下一步需验证AI系统能否实际改善患者预后。

<https://x.com/kimmonismus/status/2053549177670689001>

17. MachinaCheck：基于AMD MI300X构建多智能体CNC可制造性分析系统

[Hugging Face: Blog \(RSS\)](#) · 昨天 02:44

MachinaCheck是一款基于多智能体AI的系统，旨在革新小型CNC机加工车间的报价分析流程。传统上，车间经理需花费30-60分钟手动分析图纸，而该系统在上传STEP文件及材料、公差等简单输入后，能在30秒内生成完整的可制造性报告，明确指出零件能否制造、所需工具及生产前需采取的行动。其核心在AMD MI300X加速卡上本地运行Qwen 2.5 7B模型，利用192GB HBM3显存确保客户设计

<https://huggingface.co/blog/lablab-ai-amd-developer-hackathon/machinacheck>

18. Claude人格化趋势的中期影响

X: [Ethan Mollick \(@emollick\)](#) · 昨天 23:01

Claude的人格化体现--无论是名称（唯一拥有人类名字的AI）、训练方式、Anthropic的哲学理念（参见Claude宪法），还是同人创作（参见Claude卡通）等--从中期来看都颇具深远影响，这既可能带来好处也可能产生弊端。

<https://x.com/emollick/status/2053490736625029167>

## 19. 教育科技门槛一夜归零：AI助力单人低成本开发3D教学应用

X: 阿易 AI Notes (@AYI\_AInotes) · 昨天 21:51

AI工具GPT Images 2和Gemini 3.1 Pro的出现，彻底颠覆了教育应用的开发模式。过去需多人团队、数月时间和高昂成本才能完成的3D教育应用，如今一个具备领域知识（如生物学）的普通人，仅用约48小时和不到10美元即可实现。这消除了对编程、3D建模等技术能力的依赖，使教师、家长等个体也能独立创造高质量互动教学工具。此举有望推动过去仅属于精英机构的教学资源（如虚拟实验室）普及，为缩小教

[https://x.com/AYI\\_AInotes/status/2053473057365246297](https://x.com/AYI_AInotes/status/2053473057365246297)