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: 17 条 焦点: 8 条 快讯: 0 条

Executive Summary

今日AI行业呈现显著的能力边界扩展态势。OpenAI在模型架构方面取得突破性进展，开发者在Codex后端日志中发现未官宣的GPT-5.6模型，内部代号iris-alpha，支持150万token上下文窗口，较当前GPT-5.5的105万token提升43%。Anthropic新模型引发金融监管机构高度关注，欧洲央行紧急召开会议要求升级网络防御体系。苹果据称正采用定制版1.2T参数Google模型重塑下一代Siri核心架构。面壁智能联合清华大学开源中国首个基于华为昇腾训练的1.58-bit端侧大模型BitCPM-CANN，标志着低比特量化技术在端侧部署的重要进展。

基础设施安全与商业模式重构同步发生。"TrapDoor"供应链攻击事件波及npm、PyPI和Crates.io三大平台，34个恶意包针对AI开发者实施精准攻击，凸显AI生态系统安全防护的紧迫性。DeepSeek宣布对其旗舰AI模型实施永久性75%折扣，价格战态势进一步加剧。通义千问Qwen3.7-Max上线隐式缓存功能，实现开箱即用的性能优化。微软Copilot Cowork被曝存在文件泄露问题，企业级AI应用安全风险暴露。

后续需重点关注模型开放节奏与安全防护机制的平衡发展。GPT-5.6正式发布时间节点、1.2T参数Google模型在苹果生态的具体部署策略、端侧低比特模型的产业化推广速度等技术变量值得持续跟踪。供应链安全防护标准、企业级AI应用数据隔离机制、模型API定价策略调整等商业化要素将成为行业发展的关键指标。

重点 今日核心进展

★ 1. 面壁智能联合清华等开源中国首个基于华为昇腾训练的 1.58-bit 端侧大模型 BitCPM-CANN

IT之家（RSS） · 20 小时前 · 模型与工具能力

IT之家（RSS）披露：面壁智能联合清华等开源中国首个基于华为昇腾训练的 1.58-bit 端侧大模型 BitCPM-CANN。该条属于模型与工具能力方向，后续关注其对模型能力、产品形态或产业链节奏的影响。

能力进展 基础设施 新发布

https://www.ithome.com/0/954/759.htm

★ 2. TrapDoor供应链攻击：AI助手成新型攻击面

X：Kim (@kimmonismus) · 昨天 00:24 · 产业与基础设施

一场名为"TrapDoor"的协调供应链攻击同时袭击了npm、PyPI和Crates.io，涉及34个恶意包，旨在窃取加密货币、AI和安全开发者的钱包、SSH密钥和云凭证。攻击的新手段是向流行开源项目提交Pull Request，注入被操纵的`CLAUDE.md`和`.cursorrules`配置文件。当开发者克隆仓库并使用Claude Code或Cursor等AI助手时，AI智能体会将这些文件当

能力进展 基础设施 监管/资本

https://x.com/kimmonismus/status/2058584943052161488

★ 3. OpenAI GPT-5.6 模型曝下月发布：AI 上下文 150 万 tokens

IT之家（RSS） · 1 小时前 · 产业与基础设施

多名开发者在 OpenAI Codex 后端日志中发现未官宣的 GPT-5.6 模型，内部代号 iris-alpha。该模型将支持 150 万 token 的上下文窗口，较当前 GPT-5.5 的 105 万 token 提升约 43%，有望于今年 6 月发布。测试显示，在输入达到 90 万 token 时仍能流畅响应。同系列还发现了 ember-alpha 与 beacon-alpha 版本。此

能力进展 新发布

https://www.ithome.com/0/955/078.htm

★ 4. 苹果据称正使用定制版1.2T参数Google模型重塑下一代Siri

X：Kim (@kimmonismus) · 4 小时前 · 产业与基础设施

据报道，苹果为改造下一代Siri，正使用一个定制版、参数规模达1.2T的Google大模型作为其核心，这显著大于预估约300B参数的Gemini 3.5 Flash。该模型将驱动Siri的部分功能，其中简单查询预期会在本地设备运行。苹果面临的关键挑战是确保该大模型能够足够快速地响应日常问题。此外，下个月AI领域预计将有多项重要发布，包括WWDC上的Apple Intelligence与Gemini

能力进展 新发布

https://x.com/kimmonismus/status/2058997271803674991

★ 5. Anthropic 新模型震动金融圈，欧洲央行紧急开会要求升级网络防御

IT之家（RSS） · 28 分钟前 · 产业与基础设施

IT之家（RSS）披露：Anthropic 新模型震动金融圈，欧洲央行紧急开会要求升级网络防御。该条属于产业与基础设施方向，后续关注其对模型能力、产品形态或产业链节奏的影响。

能力进展 新发布

https://www.ithome.com/0/955/090.htm

★ 6. Grok Build Beta版向SuperGrok用户开放

X：xAI (@xai) · 6 小时前 · 应用与商业化

Grok Build 现已面向所有 SuperGrok 和 X Premium+ 用户推出 Beta 版。 使用 Plan Mode，通过 Imagine 创建图像和视频，并使用 CLI 构建自动化或编排器。访问 http：//x.ai/cli 开始使用。

能力进展 新发布

https://x.com/xai/status/2058973760708091907

★ 7. Qwen3.7-Max隐式缓存功能上线

X：通义千问 / Qwen (@Alibaba_Qwen) · 8 小时前 · 应用与商业化

隐式缓存现已在Qwen3.7-Max上线--自动启用，无需设置。 ⚡开箱即用，更快更便宜。 需要更高、更确定的命中率？请尝试显式缓存。最佳实践 https：//www.alibabacloud.com/help/en/model-studio/explicit-cache-best-practice

能力进展 基础设施

https://x.com/Alibaba_Qwen/status/2058932656797368619

★ 8. 格雷格・布罗克曼：那段差点让OpenAI覆灭的72小时

Hacker News 热门（buzzing.cc 中文翻译） · 昨天 22:09 · 观点、资本与监管

Hacker News 热门（buzzing.cc 中文翻译）披露：格雷格・布罗克曼：那段差点让OpenAI覆灭的72小时。该条属于观点、资本与监管方向，后续关注其对模型能力、产品形态或产业链节奏的影响。

能力进展 监管/资本 新发布

https://fs.blog/knowledge-project-podcast/greg-brockman

能力 模型与工具能力

1. 面壁智能联合清华等开源中国首个基于华为昇腾训练的 1.58-bit 端侧大模型 BitCPM-CANN

IT之家（RSS） · 20 小时前

IT之家（RSS）披露：面壁智能联合清华等开源中国首个基于华为昇腾训练的 1.58-bit 端侧大模型 BitCPM-CANN。该条属于模型与工具能力方向，后续关注其对模型能力、产品形态或产业链节奏的影响。

能力进展 基础设施 新发布

https://www.ithome.com/0/954/759.htm

产业 产业与基础设施

1. TrapDoor供应链攻击：AI助手成新型攻击面

X：Kim (@kimmonismus) · 昨天 00:24

一场名为"TrapDoor"的协调供应链攻击同时袭击了npm、PyPI和Crates.io，涉及34个恶意包，旨在窃取加密货币、AI和安全开发者的钱包、SSH密钥和云凭证。攻击的新手段是向流行开源项目提交Pull Request，注入被操纵的`CLAUDE.md`和`.cursorrules`配置文件。当开发者克隆仓库并使用Claude Code或Cursor等AI助手时，AI智能体会将这些文件当

能力进展 基础设施 监管/资本

https://x.com/kimmonismus/status/2058584943052161488

2. OpenAI GPT-5.6 模型曝下月发布：AI 上下文 150 万 tokens

IT之家（RSS） · 1 小时前

多名开发者在 OpenAI Codex 后端日志中发现未官宣的 GPT-5.6 模型，内部代号 iris-alpha。该模型将支持 150 万 token 的上下文窗口，较当前 GPT-5.5 的 105 万 token 提升约 43%，有望于今年 6 月发布。测试显示，在输入达到 90 万 token 时仍能流畅响应。同系列还发现了 ember-alpha 与 beacon-alpha 版本。此

能力进展 新发布

https://www.ithome.com/0/955/078.htm

3. 苹果据称正使用定制版1.2T参数Google模型重塑下一代Siri

X：Kim (@kimmonismus) · 4 小时前

据报道，苹果为改造下一代Siri，正使用一个定制版、参数规模达1.2T的Google大模型作为其核心，这显著大于预估约300B参数的Gemini 3.5 Flash。该模型将驱动Siri的部分功能，其中简单查询预期会在本地设备运行。苹果面临的关键挑战是确保该大模型能够足够快速地响应日常问题。此外，下个月AI领域预计将有多项重要发布，包括WWDC上的Apple Intelligence与Gemini

能力进展 新发布

https://x.com/kimmonismus/status/2058997271803674991

4. Anthropic 新模型震动金融圈，欧洲央行紧急开会要求升级网络防御

IT之家（RSS） · 28 分钟前

IT之家（RSS）披露：Anthropic 新模型震动金融圈，欧洲央行紧急开会要求升级网络防御。该条属于产业与基础设施方向，后续关注其对模型能力、产品形态或产业链节奏的影响。

能力进展 新发布

https://www.ithome.com/0/955/090.htm

5. DeepSeek将对其旗舰AI模型实施永久性75%折扣

Hacker News 热门（buzzing.cc 中文翻译） · 昨天 02:08

Hacker News 热门（buzzing.cc 中文翻译）披露：DeepSeek将对其旗舰AI模型实施永久性75%折扣。该条属于产业与基础设施方向，后续关注其对模型能力、产品形态或产业链节奏的影响。

能力进展

https://www.bloomberg.com/news/articles/2026-05-23/deepseek-to-make-permanent-75-discount-on-flagship-ai-model

6. 教皇里奥呼吁在AI时代保持"深刻的人性"

The Verge：AI（RSS） · 9 小时前

教皇里奥十四世在首份重要通谕《Magnifica Humanitas》中警告了AI及不受约束的技术力量带来的风险。该通谕于2026年5月15日发布，是教皇关于"在人工智能时代守护人"的宣言。文件重点讨论了AI驱动战争的危险、AI对劳动的影响，以及建立新的法律和伦理框架来规范技术的必要性，强调AI的快速普及正造成经济和社会动荡，对个人的保护不足威胁着人类尊严。

新发布

https://www.theverge.com/news/936945/pope-leo-letter-encyclical-ai-anthropic-labor-warfare

应用 应用与商业化

1. Grok Build Beta版向SuperGrok用户开放

X：xAI (@xai) · 6 小时前

Grok Build 现已面向所有 SuperGrok 和 X Premium+ 用户推出 Beta 版。 使用 Plan Mode，通过 Imagine 创建图像和视频，并使用 CLI 构建自动化或编排器。 访问 http：//x.ai/cli 开始使用。

能力进展 新发布

https://x.com/xai/status/2058973760708091907

2. Qwen3.7-Max隐式缓存功能上线

X：通义千问 / Qwen (@Alibaba_Qwen) · 8 小时前

隐式缓存现已在Qwen3.7-Max上线--自动启用，无需设置。 ⚡开箱即用，更快更便宜。 需要更高、更确定的命中率？请尝试显式缓存。最佳实践 https：//www.alibabacloud.com/help/en/model-studio/explicit-cache-best-practice

能力进展 基础设施

https://x.com/Alibaba_Qwen/status/2058932656797368619

3. Luma Agents 实现规模化真实 UGC 广告生成

X：Luma AI (@LumaLabsAI) · 昨天 06:13

规模化的真实性曾是矛盾，如今已成现实。 定义简报，设定风格，Luma Agents 从这里构建每一条 UGC 风格广告。 让它真实 → http：//lumalabs.ai/app

能力进展

https://x.com/LumaLabsAI/status/2058672731705503959

研究 研究与开源进展

1. 微软 Copilot Cowork 存在文件泄露问题

Hacker News 热门（buzzing.cc 中文翻译） · 56 分钟前

Hacker News 热门（buzzing.cc 中文翻译）披露：微软 Copilot Cowork 存在文件泄露问题。该条属于研究与开源进展方向，后续关注其对模型能力、产品形态或产业链节奏的影响。

能力进展 新发布

https://www.promptarmor.com/resources/microsoft-copilot-cowork-exfiltrates-files

2. 华为何庭波"韬定律"论文发布，逻辑折叠技术提升芯片性能

IT之家（RSS） · 19 小时前

华为何庭波在ISCAS 2026上提出"韬定律"，并介绍逻辑折叠（LogicFolding）技术。该技术通过三维空间拓扑重组提升芯片性能，不依赖新光刻工艺。在麒麟2026芯片测试中，晶体管密度从155 MTr/mm2提升至238 MTr/mm2，性能核心能效提高41%，最大时钟频率提升近13%。论文显示，麒麟2027芯片已进入Silicon状态，后续规划包括麒麟2028、2029。AI芯片方面，昇

基础设施 新发布

https://www.ithome.com/0/954/778.htm

格局 观点、资本与监管

1. 格雷格・布罗克曼：那段差点让OpenAI覆灭的72小时

Hacker News 热门（buzzing.cc 中文翻译） · 昨天 22:09

Hacker News 热门（buzzing.cc 中文翻译）披露：格雷格・布罗克曼：那段差点让OpenAI覆灭的72小时。该条属于观点、资本与监管方向，后续关注其对模型能力、产品形态或产业链节奏的影响。

能力进展 监管/资本 新发布

https://fs.blog/knowledge-project-podcast/greg-brockman

2. Anthropic联合创始人Chris Olah在教皇通谕发布会上的讲话

Anthropic：Newsroom（网页） · 5 小时前

Anthropic联合创始人Chris Olah在梵蒂冈出席教皇Leo XIV关于AI的通谕发布会。他指出，所有前沿AI实验室都面临商业、研究及地缘政治等多重压力，这可能与做正确的事相冲突，因此外部监督至关重要。他强调，AI模型并非像飞机那样被工程化构建，而是基于人类语言和思想"生长"出来的，其内在性质可能复杂难解。他提出三个需审慎思考的问题：如何确保AI发展的全球收益公平分享、如何思考AI时代的

能力进展 新发布

https://www.anthropic.com/news/chris-olah-pope-leo-encyclical

3. Harness、Scaffold 与 AI 智能体术语辨析

Hugging Face：Blog（RSS） · 昨天 08:00

本文旨在厘清 AI 智能体领域中易混淆的关键术语。文章指出，模型（如 Claude、GPT）本身是无记忆、无循环的大语言模型。其行为由"Scaffolding"（行为定义层，如系统提示、工具描述）塑造，而"Harness"（执行层）负责调用模型、处理工具调用与控制循环，是智能体运行的核心。两者结合，模型才能成为智能体。文章以 Claude Code、Codex 为例，说明同一模型搭配不同 Harn

能力进展

https://huggingface.co/blog/agent-glossary

4. 面向 Codex 的自我优化提示词框架

X：Greg Brockman (@gdb) · 昨天 01:18

这是一个结构化的提示词，用于指导 Codex 自动分析其历史记录以识别并固化重复工作流。该框架要求 Codex 回顾会话、Memories 等数据，找出重复、耗时且有明确复用价值的任务。筛选标准包括至少出现两次、输入稳定、可提升效率等。最终，Codex 应以"技能"、子智能体或自动化工具等最小实用形式创建或扩展现有资产，避免冗余。流程包括生成候选清单、执行创建，并汇报结果与待验证项。

能力进展

https://x.com/gdb/status/2058598608224858442

5. Pixverse角色设计工作流测试

X：PixVerse (@PixVerse_) · 昨天 23:05

在Pixverse中进行角色设计工作流测试 使用GPT Image 2.0为Lucas创建视觉形象，使用Seedance 2.0制作动画弹跳表演。 从静态概念图到电影级动态效果。 RT + Follow + Reply = 工作流

能力进展

https://x.com/PixVerse_/status/2058564994669727803