

# AI 技术每日分析：DeepSeek 视觉模型 接入统一 Agent API

中国高技术产业发展促进会新质生产力工作委员会

博雅云创 & 中科创新驱动

2026 年 8 月 23 日星期日

## 摘要

AI 产品竞争继续向多模态、低资源语言、科研智能体和前沿安全治理扩展。DeepSeek 将 V4-Flash-Vision-Exp 接入 API，把视觉理解、工具调用和 Files API 放进同一套 Agent 接口；印度 IISc、ARTPARK 与 Google 推动 SraVaani 覆盖 65 种印度语言，试图把语音识别扩展到商业模型长期忽略的低资源语言；伦敦初创公司 Inherent 进一步披露 27B 科学 Agent Faraday 的长程强化学习路线，用论文复现检验科研执行能力；OpenAI 则公开建议加州强化 SB 53，对训练和评测中的前沿模型增加严重事件监测，并加强模型全生命周期网络安全。与此同时，OpenAI 下调 GPT-5.6 Sol 开发者价格，安全机构继续追问 Agent 是否能够被有效监控和隔离。AI 系统的竞争指标正在从单次 Benchmark 延伸到多模态工具链、语言覆盖、长程任务、运行成本和控制能力。

## 目录

|                                                               |   |
|---------------------------------------------------------------|---|
| 一、DeepSeek 把 V4-Flash-Vision-Exp 接入 API：多模态 Agent<br>进入统一开发接口 | 3 |
| 二、SraVaani 覆盖 65 种印度语言：语音 AI 开始补低资源语言空白                       | 4 |

|                                                |   |
|------------------------------------------------|---|
| 三、Inherent 继续披露 Faraday: 27B 小模型用长程强化学习做科研论文复现 | 4 |
| 四、OpenAI 转而呼吁强化加州 SB 53: 训练阶段的前沿模型也要持续监测       | 5 |
| 趋势观察                                           | 5 |
| 参考资料                                           | 6 |

AI 产品竞争继续向多模态、低资源语言、科研智能体和前沿安全治理扩展。DeepSeek 将 V4-Flash-Vision-Exp 接入 API，把视觉理解、工具调用和 Files API 放进同一套 Agent 接口；印度 IISc、ARTPARK 与 Google 推动 SraVaani 覆盖 65 种印度语言，试图把语音识别扩展到商业模型长期忽略的低资源语言；伦敦初创公司 Inherent 进一步披露 27B 科学 Agent Faraday 的长程强化学习路线，用论文复现检验科研执行能力；OpenAI 则公开建议加州强化 SB 53，对训练和评测中的前沿模型增加严重事件监测，并加强模型全生命周期网络安全。与此同时，OpenAI 下调 GPT-5.6 Sol 开发者价格，安全机构继续追问 Agent 是否能够被有效监控和隔离。AI 系统的竞争指标正在从单次 Benchmark 延伸到多模态工具链、语言覆盖、长程任务、运行成本和控制能力。

## 一、DeepSeek 把 V4-Flash-Vision-Exp 接入 API：多模态 Agent 进入统一开发接口

DeepSeek 在 8 月 21 日上线实验性多模态模型 V4-Flash-Vision-Exp，并在 8 月 22 日进一步受到开发者社区关注。官方资料显示，该模型在文本能力上保持 V4-Flash 的 Agent、推理和世界知识能力，同时增加视觉理解；DeepSeek 称其多模态 Agent Benchmark 相较 V4-Flash 有明显提升，并接近 Opus-4.8 水平。由于性能数据来自厂商，应等待更多独立评测。

更值得关注的是接口工程。V4-Flash-Vision-Exp 同时支持 Chat Completions、Messages 和 Responses API，可混合输入文字与图片；图片可通过 Base64、外部 URL 或 Files API 提供。单张图片计费最多折算 384 Token，沿用 V4-Flash 价格。新上线的 Files API 允许开发者上传一次图片后反复通过 file\_id 调用，DeepSeek Harness 0.1.1 也提供首日支持。

这意味着 DeepSeek 正在把视觉能力直接放进 Agent 工具链。网页操作、文档解析、GUI 理解和视觉检索等任务都可以沿用同一套 Agent 接口，

文件复用则有助于降低长流程任务中的重复带宽和上下文成本。

## 二、SraVaani 覆盖 65 种印度语言：语音 AI 开始补低资源语言空白

Moneycontrol 8 月 22 日报道，印度科学研究院 IISc 的 SPIRE Lab 与 ARTPARK、Google 共同推进开源语音识别模型 SraVaani，覆盖 65 种印度语言和方言。报道指出，其中相当一部分语言长期缺少主流商业 ASR 的正式支持，项目希望把语音 AI 从印度主流语言继续扩展到数据稀缺、口音复杂和混合语言频繁出现的地区。

公开模型资料显示，SraVaani 采用约 4.3 亿参数的 FastConformer 混合 TDT-CTC 架构，训练使用超过 3.1 万小时的多语言语音数据，并以 MIT 许可证开放。低资源语言之间的识别效果仍有差异，因此真正进入客服、政务、教育和本地化 Voice Agent 之前，仍需要针对具体语种逐一评测。

这条路线值得关注，因为语音 Agent 下一阶段的竞争不会只看英语和少数大语种的平均准确率。地方语言、方言、代码混合、实时流式识别以及低成本终端部署，会决定 Voice AI 能否真正覆盖更广泛的人群。

## 三、Inherent 继续披露 Faraday：27B 小模型用长程强化学习做科研论文复现

TechCrunch 8 月 22 日进一步披露伦敦 AI 实验室 Inherent 的科研 Agent Faraday。Inherent 由多名 Google DeepMind 前员工创立，近期完成 5000 万美元种子融资。公司披露的技术路线是以约 27B 参数模型为基础，通过长程强化学习训练 Agent 持续执行复杂科研任务。

Faraday 目前重点测试论文复现：Agent 读取论文后自主写代码、运行

实验、分析结果，再尝试复现论文核心结论。Inherent 构建的 Replica 评测包含约 100 篇机器学习和 AI for Science 论文衍生出的 310 个任务，公司称 Faraday 在这一特定评测中超过更大的通用模型。由于 Benchmark 和结果由公司主导，目前更适合作为一种长程 Agent 训练路线信号。

科研 Agent 需要连续调用代码、处理失败实验、保存中间状态并不断修正方案。若中等规模模型能够依靠 Harness、工具调用和长程强化学习稳定完成这类任务，AI for Science 的成本结构可能出现新的优化空间。

## 四、OpenAI 转而呼吁强化加州 SB 53：训练阶段的前沿模型也要持续监测

TechCrunch 8 月 22 日报道，OpenAI 全球事务团队公开表示，加州 SB 53 应进一步扩大安全保障。OpenAI 提出的方向包括，对训练或评测中的前沿模型监测潜在严重事件，并强化模型开发全生命周期的网络安全保护；公司表示愿与加州议会和州长办公室继续完善该法案。

这一表态说明大型模型实验室的安全治理正在从发布前评测扩展到训练过程。最近围绕 Agent 越权、测试环境逃逸和供应链攻击的事件，使训练集群、工具权限、模型隔离、凭证和运行监控逐渐成为新的安全重点。Guidelight 近期对主要前沿实验室的评估也把模型隔离、监控和第三方审查列为关键能力。

## 趋势观察

模型价格也在改变 Agent 经济性。Reuters 报道，OpenAI 将 GPT-5.6 Sol 的 API 价格在未来三个月下调 20% 以上，标准短上下文输入降至每百万 Token 4 美元、输出 20 美元。更低价格会推动更长、更复杂的 Agent 任务，但任务链越长、权限越高，对工具审计、文件边界、异常终止和运行

监控的要求也会同步提高。AI 平台竞争正在同时发生在“单位任务成本”和“系统是否可控”两条线上。

## 参考资料

1. DeepSeek API Docs | 《DeepSeek-V4-Flash-Vision-Exp Release: Multimodal API Now Live》 | 2026-08-21  
<https://api-docs.deepseek.com/news/news260821/>
2. Caixin Global | DeepSeek 多模态模型最新报道 | 2026-08-22
3. Moneycontrol | 《Inside IISc’ s SraVaani…》 | 2026-08-22  
<https://www.moneycontrol.com/artificial-intelligence/inside-iisc-s-sravaani-the-push-to-take-voice-ai-beyond-india-s-22-mainstream-languages-article-14012715.html>
4. Hugging Face | SraVaani Model Card | 2026 年
5. Hugging Face Blog | SraVaani 技术介绍 | 2026 年
6. TechCrunch | 《Inherent…AI teammate…replicating research》 | 2026-08-22  
<https://techcrunch.com/2026/08/22/inherent-founded-by-deepmind-alumni-says-its-ai-teammate-just-outperformed-anthropic-and-openai-at-replicating-research/>
7. Inherent Labs | Faraday/Replica 技术资料 | 2026-08
8. TechCrunch | 《OpenAI says California should strengthen its AI safety bill》 | 2026-08-22

<https://techcrunch.com/2026/08/22/openai-says-california-should-strengthen-its-ai-safety-bill/>

9. Reuters | 《OpenAI cuts developer pricing for frontier GPT-5.6 Sol model by more than 20%》 | 2026-08-21

<https://www.reuters.com/technology/openai-cuts-developer-pricing-frontier-gpt-56-sol-model-by-more-than-20-2026-08-21/>

10. Reuters / Guidelight | 《AI firms can't yet contain what they've built, study finds》 | 2026-08-19

<https://www.reuters.com/technology/artificial-intelligence/ai-firms-cant-yet-contain-what-theyve-built-study-finds-2026-08-19/>

# 联系我们，请扫描二维码



新质生产力工作委员会  
官方公众号



工业智能算网  
gyznsw.cn

## 新质生产力工作委员会：

中国高技术产业发展促进会新质生产力工作委员会，专注于推动工业人工智能、智能制造、数字化转型等前沿技术发展，为企业提供政策解读、技术咨询和产业对接服务。

## 工业智能算网：

专注于工业人工智能、新质生产力、工业软件 CAE、智能制造等前沿技术。提供每日动态分析、技术趋势解读、解决方案分享，推动工业智能化转型。

网站地址：<https://gyznsw.cn>