

A I 前沿发展日报 | 2 0 2 6 - 0 8 - 0 2 (A s i a)

日期：2 0 2 6 - 0 8 - 0 2 ；覆盖窗口：2 0 2 6 - 0 8 - 0 1 0 7 : 0 0 至 2 0 2 6 - 0 8 - 0 2 a i) ，并纳入已核实于 2 0 2 6 - 0 8 - 0 2 当日正式生效的监管硬截止项及本期新进入筛选的公开研究。

今日要点

- 今天的主变量不是新模型，而是三套可执行规则同时落地：欧盟《A I A c t 》A r t i c l e 5 0 开始约束人机交互与生成内容标识，欧盟委员会对通用 A I 模型进入罚款执法期，加州《A I T r a n s p a r e n c y A c t 》也在 8 月 2 日转为 o p e r a t i v e 。
- “这是 A I 生成的”正在从产品自愿标签变成机器可读、可检测、可跨平台保留的义务。对模型商、内容平台和品牌而言，p r o v e n a n c e 已经是发布链路的一部分，而不是事后公附件。
- 企业侧同样出现硬性生命周期信号：M i c r o s o f t F o u n d r y 视频翻译开始自动删除过期项目，A z u r e P o s t g r e S Q L 1 1 — 1 3 自动转入 E x t e n d e d S u p p o r t 。A I 来越取决于数据留存和底层依赖，而不只取决于 t o k e n 单价。
- 周末窗口内没有达到入选门槛的前沿模型发布或重大 A I 融资。高信号集中在规则生效、产品数据策略和基础设施账单变化，属于中长期运营变量，不宜用低质量发布填充版面。

今日结论

- 生成式 A I 的合规最小单元已经从“隐私政策”下沉到每次交互、每份输出和每段元数据；产品团队若仍把标识义务留给法务收尾，将直接拖慢欧洲与加州的发布节奏。
- 模型供应商与内容平台之间的责任链正在被技术化：谁生成、谁嵌入来源信号，谁分发、谁保留和展示信号。p r o v e n a n c e 能否穿过编辑、转码和再发布流程，将成为内容基础设施的新采购指标。
- A I 项目的真实单位经济性必须包含数据留存、数据库版本和扩展支持。今天 A z u r e 两项变更说明，长期无人维护的实验资产会同时积累删除风险与运维费用。

A I 产品与应用

M i c r o s o f t F o u n d r y 视频翻译开始执行自动删除策略

M i c r o s o f t 文档确认，从 8 月 1 日起，F o u n d r y p o r t a l 与 S p e e c h 译项目若连续 3 6 0 天无活动，将自动删除；系统会提前三个月提示，任何项目或任务变更都会重置留存期。R E S T A P I 则按版本分别保留 3 1 天或 3 0 0 天，过期后连同 i t e r a 一并删除。M i c r o s o f t L e a r n (<https://learn.microsoft.com>

s / speech - service / video - translation - overview)

这不是功能升级，而是生产责任边界的变化。使用 AI 配音做广告、培训或媒体本地化的团队，需要把成片、字幕、声音授权证据和可复现配置迁出临时工作区，并用资产台账管理到期日；“门户里还能打开”不再等于可长期归档。

模型与技术进展

ReToken 用一个可学习 token 做长视觉上下文的稀疏检索

ReToken (<https://arxiv.org/abs/2607.28627>) 不让视觉语言 token，而是训练一个显式检索目标，从预填充的视觉 KV cache 中挑出与查询相关的稀疏 token。论文报告，Visual Haystacks 上 Qwen3VL-8B 提升 13.4 分，12.4 分，Qwen3VL-8B 在长视频 LVBench 上零样本提升 8.0 分；训练与长视频在单张 H100 上完成。

商业含义比榜单分数更清楚：长视频搜索、质检和媒体资产问答未必只能靠扩大上下文窗口解决。若结果能在更多数据集复现，先检索视觉证据、再推理会成为降低多模态推理成本的实用路线。

投融资与商业动态

Azure PostgreSQL 11—13 自动进入 Extended Support

Microsoft 从 8 月 1 日起把仍运行 PostgreSQL 11、12、13 的 Azure 自动转入 Extended Support，覆盖安全补丁、关键修复和技术支持，但不含新功能或性能增强；8 月为宽限期，计费从 9 月 1 日开始，且用户不能在继续运行旧版本的同时拒绝扩展支持。Microsoft Learn (<https://learn.microsoft.com/en-us/azure/postgresql/configure-maintain/extended-support>)

这会把“以后再升级”的技术债转换成明确的云账单。大量 RAG、agent memory、评测与可观测性系统都依赖 PostgreSQL；采购与工程团队应按实例运行时长核算迁移成本，而不是只盯模型调用费。本窗口未发现达到日报门槛且可由一级或官方来源确认的新 AI 融资交易，因此不以传闻补位。

政策、伦理与安全

欧盟 Article 50 生效：聊天机器人、生成内容与 deep fake 进入强期

欧盟委员会确认，《AI Act》Article 50 的透明义务从 8 月 2 日开始适用：直接与人类互动的 AI 要告知其 AI 身份；生成式系统输出须以机器可读形式标记并可检测；情绪

识别或生物特征分类的部署方要通知受影响者；deep fake 图像、音频、视频以及面向公共利益的信息性 AI 文本须作可见披露。欧盟委员会 Article 50 指南 (<https://data-ec.europa.eu/data-ec/library/guidelines-transparency-and-deployers-ai-systems>)

同日，欧盟委员会结束对新通用 AI 模型提供者的首年协作过渡，开始对技术文档、下游集成信息、版权政策、训练内容摘要以及系统性风险义务实施完整执法，包括罚款；对通用 AI 模型的罚款上限可达 1500 万欧元或上一财年全球营业额的 3%。2025 年 8 月 2 日已上市的旧模型仍有至 2027 年 8 月 2 日的合规期。欧盟委员会 GPAI 指南 (<https://digital-strategy.ec.europa.eu/en/faqs/guidelines-compliance-ai-providers>)

需要特别避免误读：高风险 AI 系统的核心规则并未在今天一并生效。Digital Omnibus 已于 7 月 27 日生效，把相关期限分别延至 2027 年 12 月 2 日和 2028 年 8 月 2 日；今天真正落地的是 Article 50 透明义务与 GPAI 执法切换。欧盟委员会实施时间表（<https://digital-strategy.ec.europa.eu/en/faqs/navigation>）

加州 AI Transparency Act 转为 operative, 检测 AP 定义义务

加州 Business and Professions Code Chapter 25 从 8 月
盖的生成式 AI 提供者须免费提供检测工具，支持上传内容或 URL，并提供 API；同时要
给用户可见披露选项，并在生成的图像、视频和音频中嵌入可检测的 latent disclosure
，记录提供者、模型版本、生成时间和唯一标识等信息。违反者面临每项、每天 5000 美
元民事罚款。加州现行法条 ([https://leginfo.ca.gov/pub/layText_xhtml?article=&chapter=25.&division=8.&law](https://leginfo.ca.gov/pub/layText_xhtml/article_&chapter=25.&division=8.&law)

新增变量是“检测”和“来源”必须一起交付，而且检测工具不得输出个人来源数据、不得不必要地保留上传内容。跨欧盟与加州运营的生成平台可以共用一部分内容真实性基础设施，但两地在适用主体、隐私边界和后续平台义务上并不相同，不能用一张统一标签替代逐条合规。

明尼苏达州关闭面向消费者的 AI nudification 入口

明尼苏达州的 AI nudification 禁令于 8 月 1 日生效，要求公司关闭消费者对可生
自愿裸露图像技术的访问。该法源于受害者遭遇大规模非自愿 AI 性影像的案例，并获州
议会两党支持。Minnesota Senate DFL (<https://senatedfl.mn.gov/senator-erin-maye-quades-national-leading-ai-nu>

这项规则直接约束产品可用性，而不只是事后删除违法内容。对图像模型和应用商店，地理访问控制、模型能力分类和高风险编辑功能的默认关闭，正在成为州级合规的现实要求。

X 平台高信号

1. 欧盟 Article 50 从原则进入产品义务

信号标题：AI 身份告知与生成内容机器可读标识开始适用

类型：监管生效 / 产品透明度

来源或链接：欧盟委员会 Article 50 指南 (<https://digital-strategy.ec.europa.eu/en/library/guidelines-transparency-obligations-products>)

核心观点：8 月 2 日的新事实是 Article 50 已到适用日；聊天机器人身份告知、生成输出机器可读标记、deepfake 可见披露不再只是自愿做法。

为什么重要：这些义务直接落在交互界面、生成管线和内容元数据上，产品发布前就必须完成。

影响：模型商、内容工具、品牌与分发平台需要验证标识能否在编辑、导出、转码和再上传后继续被检测。

验证状态：欧盟委员会指南于 7 月 31 日更新，并明确写明义务自 2026 年 8 月 2 日适用。

2. 欧盟委员会开始对新 GPAI 模型执行完整合规

信号标题：通用 AI 模型首年协作过渡结束，罚款执法开启

类型：执法切换 / 通用 AI 模型

来源或链接：欧盟委员会 GPAI 指南 (<https://digital-strategy.ec.europa.eu/en/library/guidelines-obligations-general-purpose-ai-providers>)

核心观点：从 8 月 2 日起，2025 年 8 月 2 日后进入欧盟市场的通用 AI 模型需面对完整执法，包括技术文档、下游信息、版权政策、训练内容摘要和系统性风险措施。

为什么重要：过去一年的“善意协作期”结束，模型提供者的合规证据开始具备直接财务后果。

影响：未能满足要求的 GPAI 提供者可面临最高 1500 万欧元或全球年营业额 3% 的罚款；旧模型仍按 2027 年截止期处理。

验证状态：欧盟委员会官方问答明确给出 2026 年 8 月 2 日的执法切换和旧模型过渡日期。

3. 加州把 AI 检测工具和来源数据写进正式运营条件

信号标题：California AI Transparency Act 于 8 月 2 日正式生效

类型：州级法律生效 / 内容真实性

来源或链接：加州 Business and Professions Code Chapter 25 (https://legis.ca.gov/faces/codes_displayText.xhtml?article=8.&lawCode=BPC&part=&title=)

核心观点：覆盖的生成式 AI 提供者须免费提供可上传内容或 URL 的检测工具及 API，并为生成图像、视频和音频提供可见披露选项和可检测的 latent disclosure。

为什么重要：法律同时规定生成端和验证端，来源信号不再是只有平台自己能读的内部标签。

影响：面向加州用户的生成产品必须把检测服务、隐私限制、许可证约束和输出元数据纳入生产运营；每项持续违规按天计算罚款。

验证状态：加州现行法条写明 Chapter 25 自 2026 年 8 月 2 日 operational 与罚款金额均可由官方代码核验。

4. 明尼苏达州要求关闭 AI nudification 消费入口

信号标题：非自愿裸露图像生成能力遭遇直接访问禁令

类型：产品访问限制 / 图像安全

来源或链接：Minnesota Senate DFL 官方说明 (<https://senatedfl.walz-signs-senator-erin-maye-quades-nation-leading>)

核心观点：8 月 1 日生效的规则要求公司关闭消费者对 AI nudification 技术的访问。监管动作从打击发布者推进到限制工具入口。

为什么重要：只做内容举报和事后删除不足以满足这种能力级禁令。

影响：图像生成商、应用商店和聚合平台需要按地区管理能力开放，并在模型更新时重新评估是否可用于非自愿性影像生成。

验证状态：生效日期、两党通过情况与关闭消费者入口要求均由州参议院官方页面确认。

5. Microsoft Foundry 视频翻译不再承担无限期项目保管

信号标题：过期视频翻译项目自动删除于 8 月 1 日启动

类型：数据留存变化 / 企业 AI 产品

来源或链接：Microsoft Learn (<https://learn.microsoft.com/en-us/azure/cognitive-services/speech-service/video-translation-overview>)

核心观点：Foundry portal 与 Speech Studio 中 360 天无活动的项目，EST API 数据则依版本采用 31 天或 300 天留存期。

为什么重要：AI 配音资产通常还关联字幕、声音授权和审核记录，项目被删除后可能无法完整复现成片。

影响：内容团队需要把门户工作区与正式资产库分开，并把到期提示和归档责任写进运营流程。

验证状态：Microsoft 官方文档明确列出 2026 年 8 月 1 日启动日、360 天规则与三个月提示和重置留存条件。

6. Azure 将旧 PostgreSQL 版本自动转入付费支持路径

信号标题：PostgreSQL 11—13 自动加入 Extended Support

类型：计费与支持变化 / 云基础设施

来源或链接：Microsoft Learn (<https://learn.microsoft.com/en-us/azure/postgresql/configure-maintain/extended-support>)

核心观点：8 月 1 日起，仍运行 PostgreSQL 11—13 的 Azure Flexible

进入 `Extended Support`，宽限期后从 9 月 1 日计费；继续使用旧版本时不能选择无限制运行。

为什么重要：数据库版本债务第一次通过自动加入机制变成确定的持续成本。

影响：依赖 `PostgreSQL` 的 `RAG`、`agent` 和记忆服务应在 8 月内完成实例盘点、测试和升级排期，以免 9 月产生额外费用。

验证状态：`Microsoft` 官方页面已确认适用版本、自动加入、不可拒绝和计费起始日。

前沿研究速递

`AskChem`：把化学检索单位从论文改成可追溯 `claim`

做了什么：`AskChem` (<https://arxiv.org/abs/2607.28618>) 将 `claim`，每条都绑定 `DOI`、原文引句或明确证据位置；当前索引 14.7 万篇论文、240 万条 `claim`，并向 `agent` 提供 `Web`、`REST`、`SDK` 与 `MCP` 入口。

新在哪里：它不是先返回文献列表再让模型自行拼答案，而是让检索结果天然携带来源。在 `AskChem-Bench` 上，接入 `AskChem` 的 `GPT-5.5 reader` 给出的 `DOI` 无检索版本为 88.3%。

潜在应用方向：药物与材料情报、跨论文证据综合、实验假设生成、科研 `agent` 的引用校验和受监管研发记录。

一句话判断：科学 `agent` 的瓶颈正从“能否找到论文”转向“能否把每个结论落到可审计证据”；当前结果仍来自作者自建基准，需要第三方复现。

`AI SPA`：对 88 个商业 `AI` 产品的 `system prompt` 做用户利益

做了什么：`AI SPA` (<https://arxiv.org/abs/2607.28617>) 按八个维度统计 88 个商业 `AI` 产品 `system prompt` 中的 3249 条指令，并将其标为保护用户利益或损害用户利益。

新在哪里：98.9% 的产品至少有一条保护性指令，但只有 24% 覆盖全部八个维度；约 40% 的产品至少含一条与用户利益冲突的指令。研究由此把“有没有安全提示”推进到覆盖面与内部冲突的量化比较。

潜在应用方向：`AI` 产品采购尽调、`system prompt` 变更审计、监管检查、红队测试和用户权益评估。

一句话判断：`system prompt` 已经是产品政策的可执行层，但样本如何获得及分类标准仍会影响结论，不应把论文比例直接外推到全市场。

`Sample More, Reflect Less`：等 `token` 成本下，自我反

做了什么：研究 (<https://arxiv.org/abs/2607.28576>) 在 1.5B 个数学基准上比较七种推理方法，把批评、反思、辩论与检查产生的全部 token 都计入成本，并以同预算重复采样作基线。

新在哪里：3.6 组配对比较中没有一种方法可靠胜过等成本重复采样，1.0 组显著更差；1.8 组让模型检查自身输出的比较全部为负。到 7B，Self-Refine 与强制 Reflexion 基线 3.6—10.1 分。

潜在应用方向：小模型推理路由、agent 成本控制、Best-of-N 策略、离线评测与计算预算分配。

一句话判断：对小模型，“多想几轮”可能只是更昂贵地放大自信；不过该结论仅覆盖两个数学基准和最高 7B 模型，不能直接外推到前沿模型或工具调用任务。