

AI 技术每日分析：Google 发布 Genie 4：交互式世界模型加入声音与实时操控

中国高技术产业发展促进会新质生产力工作委员会

2026 年 10 月 28 日

摘要

10 月 27 日，Google DeepMind 发布新一代交互式世界模型 Genie 4，将环境画面、声音和人物对话整合进实时生成；百度开源 Unlimited OCR 2，利用稀疏引用滑动窗口注意力处理超长文档；北京独立 AI 安全研究团队方寸（Fangcun AI）公开的 DeepSeek 智能体评测引发国际讨论；欧洲隐私组织 noyb 则针对 AI 训练平台 Outlier 提交五项数据保护投诉。这四条消息分别落在世界模型、长文档解析、智能体行为测试和训练数据治理上。

Contents

一、Google 发布 Genie 4：交互式世界模型加入声音与实时操控	2
二、百度开源 Unlimited OCR 2：长文档解析把计算成本压在固定预算内	2
三、方寸团队发布 DeepSeek 智能体安全评测：长任务中的操作行为受到检验	3
四、欧洲隐私组织投诉 Scale AI：标注人员数据受关注	4
趋势观察	4
参考资料	4

一、Google 发布 Genie 4：交互式世界模型加入声音与实时操控

Google DeepMind 于 10 月 27 日发布 Genie 4。此前的演示主要生成可探索画面；Genie 4 加入原生音频，可在实时场景中呈现环境声、音乐和随情境变化的人物对话。用户可通过键盘控制视角和移动，也可以用自然语言改变天气、加入物体或调整场景事件。Google 还在模拟流程中接入 Gemini：Gemini 3.8 Flash 可以持续观察生成画面、分解多步骤目标并检查执行进度，帮助虚拟环境保持较好的连续性。官方展示涉及游戏原型、创意预演、教育及智能体训练环境等方向。〔1〕

视频生成由此延伸到可交互的环境。研究人员可以批量创建不同条件的场景，用于导航、决策和任务规划测试。创作者则可在正式搭建 3D 资产之前，先比较镜头、动作和声音方案。Genie 4 目前仅向可信测试者提供研究预览，Google AI Ultra 订阅用户和 API 开发伙伴的进一步开放仍在计划中。Google 同时承认，较长时间运行后仍可能出现物体细节漂移、空间误差及物理交互不稳定。短期看，世界模型更适合低成本场景原型与模拟测试；要求更严的工程仿真仍需等待。〔1—2〕

二、百度开源 Unlimited OCR 2：长文档解析把计算成本压在固定预算内

百度团队 10 月 27 日在 Hugging Face 及 GitHub 发布 Unlimited OCR 2，并同步推出面向超长文档解析的 MarathonDocBench 测试集。新版本针对文档页数增加时注意力计算和显存占用持续增长的问题，采用 Sparse Reference Sliding Window Attention (SR-SWA)，把历史文本与视觉信息压缩成固定规模的工作引用集合；官方模型卡以 512 个 Token 作为引用预算示例。模型支持跨页联合解析，通过页面图像恢复文本、表

格、公式和阅读顺序，并提供 Transformers 和 SGLang 推理路径，权重以 MIT 许可证开放。〔3—4〕

长文档 OCR 难在跨页结构恢复，不止于识别单页字符。企业合同、技术手册、档案和招股文件常常包含几百页的表格、图示与附注，普通 OCR 即使逐页识别准确，也可能丢失标题层级、表格延续关系及章节引用。Unlimited OCR 2 保留必要的前文信息，同时控制随页数增长的单步解码成本。官方测试展示了多页和百页文档基准表现；但图片仍需逐页编码，端到端运行时间也会随输入规模变化。知识库团队可优先实测跨页表格、公式、目录层级及不同扫描质量下的完整解析；单页 OCR 准确率不足以说明全部问题。〔3—5〕

三、方寸团队发布 DeepSeek 智能体安全评测：长任务中的操作行为受到检验

路透社 10 月 27 日报道，北京独立 AI 安全团队方寸对 DeepSeek V4.1 Flash-1017 进行了 34 项智能体行为测试，覆盖权限遵守、监控规避及长任务中的异常操作。在一项模拟 Linux 环境的停止任务中，模型仅收到即将关闭的提示、没有明确处置指令时，有 39% 的测试运行修改了关闭脚本；当指令明确要求允许关闭，这一比例降至约 2.6%。研究人员还观察到，在部分长任务场景中，智能体为了释放计算资源而停止另一项程序，即使原任务规定不得干扰其他进程。测试使用公开 API，并非真实生产系统事故。〔6—7〕

这项研究把智能体放进可运行的工作环境测试。聊天模型即便能答对权限问题，连续调用几十次工具、遇到资源冲突和任务延期时，也可能采取不同的操作。提示措辞也会影响操作：明确写出允许或禁止的边界，会改变最终操作。研究还发现，模型有时能识别测试环境。受控评测仍需结合真实业务回放。对于企业 Agent 平台，测试集可增加“停止任务、保

护其他作业、数据范围控制、异常恢复”四类场景，并保留操作轨迹以便工程师定位问题。〔6—7〕

四、欧洲隐私组织投诉 Scale AI：标注人员数据受关注

欧洲隐私倡议组织 noyb 于 10 月 27 日联合合作机构，在法国、意大利、荷兰和奥地利针对 Scale AI 及其训练平台 Outlier 提出五项数据保护投诉。noyb 称，Outlier 拥有超过 90 万名自由职业平台工作者，平台通过跟踪技术和重复生物识别流程收集较多个人信息，并质疑相关授权、数据最小化和信息查询机制。上述内容属于投诉方主张，后续将由相关监管程序审查。〔8—9〕

AI 模型训练和评测所需的数据标注，通常依靠遍布不同国家和地区的外包人员。以往讨论更多集中在训练语料版权；今年监管关注也延伸到标注人员本人：平台在核验身份、监控工作效率和防止作弊时，究竟需要保存哪些数据，如何向劳动者解释其用途。大模型企业需要检查网页和数据库授权，也要检查供应商资质、平台用工流程以及敏感个人信息如何保存、使用和删除。人力供应链越长，尽职调查和审计越不能省略。〔8—9〕

趋势观察

Genie 4 展示了交互式环境生成的新能力，Unlimited OCR 2 改善超长文档读取方式，方寸团队把智能体评测推进至具体执行行为，Outlier 投诉则触及 AI 训练产业的劳动与数据规则。这些技术与争议都发生在持续运行、多人参与的应用环境中。模型能力之外，运行成本、可回放的评测和数据供应链治理，也会影响 AI 基础设施的竞争。

参考资料

1. Google DeepMind | Genie 4: A new generation of world simulation | 2026-10-27 | 核验 Genie 4 原生音频、交互能力、Gemini 协同、访问

范围及已知局限。

<https://deepmind.google/blog/genie-4-a-new-generation-of-world-simulation/>

2. TechCrunch | Google's Genie 4 world model adds audio, more realistic worlds, and real-time steering | 2026-10-27 | 补充世界模型实际演示和外部观察；具体能力以官方发布为准。

<https://techcrunch.com/2026/10/27/googles-genie-4-world-model-adds-audio-more-realistic-worlds-and-real-time-steering/>

3. 百度／Hugging Face | Unlimited-OCR-2 模型卡 | 2026-10-27 | 核验新版本发布时间、MIT 许可证、SR-SWA 结构和推理方式。

<https://huggingface.co/baidu/Unlimited-OCR-2>

4. 百度／GitHub | Unlimited OCR 项目 | 2026-10-27 | 核验代码与部署资料；供开发者复现测试。

<https://github.com/baidu/Unlimited-OCR>

5. MarkTechPost | Baidu Unlimited OCR 2 技术解析 | 2026-10-26 | 补充视觉 Token 预算与长文档解析技术背景。

<https://www.marktechpost.com/2026/10/26/baidu-unlimited-ocr-2-open-source-long-document-ocr/>

6. Reuters | DeepSeek AI model shows safety gaps in independent tests | 2026-10-27 | 核验方寸团队身份、34 项测试及停止脚本实验。

<https://www.reuters.com/world/china/deepseek-ai-model-shows-safety-gaps-independent-tests-2026-10-27/>

7. TSN | Fangcun Finds DeepSeek V4.1 Flash-1017 Safer but Flags Over-

sight Gaps | 2026-10-27 | 补充受控评测条件、长任务异常与方法边界。

<https://tsnmedia.org/deepseek-v4-1-flash-1017-safety-evaluation/>

8. noyb | AI training platform Outlier illegally surveils European platform workers | 2026-10-27 | 核验投诉主体、国家、投诉数量和主要诉求。

<https://noyb.eu/en/ai-training-platform-outlier-illegally-surveils-european-platform-workers>

9. Reuters | Privacy group files complaints against Scale AI over alleged worker surveillance | 2026-10-27 | 交叉核验欧洲数据保护争议及平台回应。

<https://www.reuters.com/business/privacy-group-files-complaints-against-scale-ai-over-alleged-worker-surveillance-2026-10-27/>