AI 技术每日分析：NVIDIA 推出 Nemotron for Science，科学智能体接入 CUDA-X 工具链

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 11 月 17 日

摘要

11 月 16 日的 AI 工程进展，明显向科研工作流和开放模型基础设施集中。NVIDIA 在 SC26 发布 Nemotron for Science，将科学领域视觉语言模型、化学数据集、后训练方法与 CUDA-X 计算工具组合起来；Hugging Face 宣布将推理服务提供商路由切换至 OpenRouter，让开放权重模型能够经由统一接口连接多个推理服务商。OpenAI 则发表基于 1819 名科研人员、58 万余条科学消息的使用研究，另展示计算机操作智能体在医疗软件中执行多步骤任务的评估结果。模型进入专业场景后，训练数据、可调用工具、计算平台和运行记录的组织方式正在成为新的竞争环节。

Contents

一、NVIDIA 发布 Nemotron for Science：科学智能体开始形成可复用工具体系 **2**

二、Hugging Face 与 OpenRouter 联手：开放模型的调用路径改由市场化路由承担 **2**

三、OpenAI 调查 1819 名科研人员：AI 已进入研究全过程 **3**

四、医疗计算机操作智能体进入跨系统工作流：可审计操作成为产品能力 **4**

五、CUDA-X 科学计算工具扩充：计算与代理模型协同进入工程设计 **4**

趋势观察 **5**
参考资料 **5**

一、NVIDIA 发布 Nemotron for Science：科学智能体开始形成可复用工具体系

11 月 16 日，NVIDIA 在芝加哥 SC26 大会正式推出 Nemotron for Science。它由 Nemotron 3.5 Super VL 开放模型、化学领域 LitmusBench v0.2 数据与评测、后训练配方以及可调用的 CUDA-X 计算工具组成，面向化学、生物和材料研究构建领域智能体。该公司介绍，Terray Therapeutics 对模型进行专门后训练后，在 RxnScribe 测试中按指定匹配标准恢复了约 76% 的标注反应；Lila Sciences 使用自动化实验反馈调整量子点制备方案，公布的任务准确率为 61%，相关对比和推理成本数字属于企业测试。模型和化学数据集已在 Hugging Face 开放，DGX Spark 为小规模本地研究提供运行路径。

这次发布把“科学知识问答”推向可执行的实验研究过程。图谱识别负责将论文里的反应物、产物和反应条件变成结构化条目，智能体负责组合实验步骤，计算工具负责调用数值模拟，实验结果再进入下一轮参数优化。在工业研发中，工程师可以采用类似结构，将材料成分、物性测试、工艺窗口和仿真报告关联，逐渐形成面向设计与中试的研发数据集。

二、Hugging Face 与 OpenRouter 联手：开放模型的调用路径改由市场化路由承担

Hugging Face 11 月 16 日发布《Hugging Face and OpenRouter join forces to make open-weight models easy to run》，宣布其 Inference Providers 语言模型请求由自建路由转接到 OpenRouter。OpenRouter 连接 80 多家推理服务提供商，并依据实时价格、速度及可靠性数据选择服务端点；首选服务不可用时，可重试其他提供商。对开发者而言，仍可用 Hugging Face 令牌与既有 API 代码调用 Hub 上由服务商实际托管的开放权重模

型。

这项合作影响的是模型分发之后的运行基础设施。过去企业选择开源模型，往往还需要分别比较托管商、GPU 集群、接口兼容性和结算方式；新路由将这些能力组合在统一入口，便于建立模型版本、服务质量和业务成本的比较表。对强调本地知识库、数据不出厂的工业智能体平台，这一模式也提供了架构参考：对普通公开任务使用外部推理服务，对专有工艺、图纸和现场数据保留本地推理节点，路由层统一管理身份、模型和工具调用。

三、OpenAI 调查 1819 名科研人员：AI 已进入研究全过程

OpenAI 经济研究团队 11 月 16 日发布《Science at the frontier》，分析 2026 年 8 月 20 日至 9 月 22 日期间 1819 位科学家的 580186 条科研相关对话，并汇总 251 份调查答卷。受访者估计，达到相同质量的相关工作若不使用 AI，平均需要 1.9 倍时间；在回答发现周期问题的人中，98% 预期 AI 会加快本领域科学发现。研究覆盖文献整理、证据分析、研究构思等不同环节，平台还同期宣布扩大科研人员使用计划，目标覆盖 20 万人。

这些数据来自特定平台用户的行为样本和自报估计，展示的是研究工作正在如何重新分工。科研机构一方面可以把数据清洗、材料检索、计算脚本和实验记录编排成连续任务，另一方面也能把智能体调用留下的代码、参数和来源映射到现有的科研管理系统。研究数据能否接入专业工具、操作记录能否回溯，比单次聊天能写多长更影响实际使用。

四、医疗计算机操作智能体进入跨系统工作流：可审计操作成为产品能力

OpenAI 11 月 16 日公布医疗计算机操作智能体的任务评测及企业应用案例：GPT-6 Astra 在 325 项经审查任务中有 99% 被评为临床安全，在 646 份医生帮助度评分中 93% 为“非常有帮助”或“极其有帮助”；其在 HealthAdminBench 行政工作流上获得 94.3% 的可得分。企业示例涉及患者转诊、保险预授权和跨应用资料传输，CVS Health 介绍了使用屏幕交互、操作记录和人工核对完成系统衔接的生产流程。

医疗的软件碎片化与制造业 PLM、ERP、MES、QMS 并存的现实十分相似：很多任务有可访问的业务界面，却没有统一开放的 API。计算机操作型智能体因此需要把屏幕状态、执行权限、人工审批点和结果证据作为一个整体设计。对于制造现场，这可迁移为工程变更审批、售后工单录入和跨系统质量追溯等任务，但运行环境与权限策略必须符合对应行业的业务规则。

五、CUDA-X 科学计算工具扩充：计算与代理模型协同进入工程设计

NVIDIA 11 月 16 日还发布 CUDA-X 科学计算工具进展，包括 cuPhoton、cuPotential 与 cuSpa，面向大型天文数据、分子模拟和单细胞分析。其介绍的 Daviator GPU 内核可加速 VASP 相关量子计算，在厂商所述工作负载上最高达到 22 倍加速、成本最多降低 75%。这类数值计算库与 Nemotron for Science 并行发布，揭示了科研智能体应用的双层结构：语言模型负责理解目标、选择工具和管理过程，计算库与仿真软件负责求解物理、化学和生物模型。

趋势观察

从工程应用角度看，CAD、CAE、材料实验和中试验证可以沿着同一模式组织。智能体根据产品目标检索历史工艺，调用网格划分、求解器和测试数据库，再将计算结果与已有实验曲线比对，形成下一轮设计候选。具备结构化输入输出、任务重试和操作留痕的软件接口，将决定科学智能体能够实际承担多少研发工作。

参考资料

1. NVIDIA | NVIDIA Introduces Nemotron for Science to Accelerate Discoveries With Agents | 2026-11-16 |原稿参考资料
https://nvidianews.nvidia.com/news/nvidia-introduces-nemotron-for-science-to-accelerate-discoveries-with-agents

2. Hugging Face | Hugging Face and OpenRouter join forces to make open-weight models easy to run | 2026-11-16 |原稿参考资料
https://huggingface.co/blog/openrouter

3. OpenAI | Science at the frontier | 2026-11-16 |原稿参考资料
https://openai.com/index/science-at-the-frontier/

4. OpenAI | Advancing computer-use agents for healthcare teams | 2026-11-16 |原稿参考资料
https://openai.com/index/advancing-computer-use-for-healthcare-teams/

5. NVIDIA | NVIDIA CUDA-X Libraries Give Agents the Tools to Accelerate Scientific Discovery | 2026-11-16 |原稿参考资料
https://blogs.nvidia.com/blog/cuda-x-libraries-agents-acc

elerate-scientific-discovery/

6. OpenAI | Broadening access to AI for scientists | 2026-11-16 | 原稿参考资料

https://openai.com/index/expanding-chatgpt-for-academic-researchers/

机构介绍与关注

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。



关注高促会新质生产力工委会公众号



关注工业智能算网平台

https://gyznsw.cn