AI 技术每日分析

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 5 月 23 日

摘要

今日 AI 技术主线非常集中，几乎都指向同一个关键词：Agent 开始从“会生成内容”走向“会连接系统、调用工具、治理数据”。Google 在 I/O 2026 上把 Gemini 3.5 Flash 与 Managed Agents、Antigravity 2.0 打包推出，继续把开发者生态从提示工程推向任务编排；Cohere 则以 Apache 2.0 许可开源 Command A+，把“主权 AI”和企业私有部署推到更前的位置；Oracle 围绕 Autonomous AI Database MCP Server 展示 Codex 连接企业数据库的路径，说明数据库与 Agent 的结合正在从概念变成工程实践；Confluent 则把实时数据流、托管 MCP 与 PII 治理捆到一起，表明企业 AI 下一步的竞争焦点，已经不是单一模型能力，而是“模型 + 上下文 + 权限 + 实时数据”的完整交付能力。

Contents

一、Google 把“开发者 AI”正式推进到多 Agent 工作流

5 月 19 日，Google 在 I/O 2026 开发者更新中宣布推出 Gemini 3.5 Flash、Antigravity 2.0 桌面应用以及 Gemini API 中的 Managed Agents。官方表述很直接：Google 正在加速从“prompts to action”的转变。按照 Google 给出的说明，Gemini 3.5 Flash 主打高速度与可执行性，Managed Agents 则允许开发者用单次 API 调用启动具备推理、工具使用和隔离代码执行能力的 Agent，而 Antigravity 2.0 进一步支持多 Agent 并行、动态 subagents 和后台计划任务。

这件事的重要性不在于 Google 又发了一个新模型，而在于它把“调用模型”升级为“托管执行环境”。过去开发者更多是在 API 上拼装 RAG、函数调用与外部工具；现在头部平台开始直接提供 Agent 运行时、并行子代理与可恢复环境。这意味着开发者工具市场的竞争，正在从“谁的模型更强”转向“谁能把任务执行、状态保存、工具调用和工程环境更稳地封装起来”。Google 此举也会进一步推动行业把 Agent 视为默认的软件接口，而不只是高级聊天框。

二、Cohere 开源 Command A+，企业模型竞争转向“可部署性”

5 月 20 日，Cohere 发布 Command A+，并以 Apache 2.0 许可开放模型权重。根据官方博客和模型卡，Command A+ 采用 MoE 架构，总参数 218B、激活参数 25B，支持 128K 输入上下文、64K 最大生成，覆盖文本、图像与工具使用，并支持 48 种语言；官方同时强调，该模型在 W4A4 量化下可以在 2 块 H100 上运行。

这条动态的意义很清晰：企业 AI 市场正在把“能否自主部署”置于和“模型性能”同样重要的位置。过去企业使用大模型，往往在闭源 API 和自建开源栈之间二选一。Command A+ 试图提供第三条道路：既保留企业级 Agent、推理和多语言能力，又把许可与硬件门槛压到更适合私有部署的区间。Cohere 把这件事定义为 “sovereign AI”，本质上是在争夺企业客户对于数据主权、成本可控和合规边界的需求。它也说明，2026 年的开源模型竞争，已经不是简单拼参数，而是拼“能否进入真实企业系统”。

三、Oracle 把 Codex 接进数据库，Agent 开始碰触企业核心数据层

本周 Oracle 发布博客，演示如何将 Codex 连接到 Oracle Autonomous AI Database MCP Server，让 AI Agent 通过 MCP 发现数据库工具、查看 schema、读取元数据并在权限控制下执行数据库任务。Oracle 更早发布的 MCP Server 介绍和产品文档强调，这一能力是数据库内置、托管、多租户的，并结合数据库身份、审计、ACL 和最小权限控制。

这代表 AI Agent 的一次边界前移。企业过去对 AI 最常见的做法，是让模型访问知识库、文档库和 FAQ；一旦走到数据库层，事情就完全不同，因为数据库承载的是真实业务状态。Oracle 的思路不是让 Agent “直

接连库自由发挥”，而是把 MCP、数据库权限、Select AI Agent 工具和审计体系组合起来，让 Agent 在可治理的框架里接触数据。这对整个产业都是信号：谁能率先把 Agent 安全接入企业核心系统，谁就更有机会成为下一代企业 AI 基础设施的入口。

四、Confluent 把托管 MCP、实时上下文与数据治理打包成企业 AI 底座

5 月 19 日，Confluent 宣布在 Confluent Intelligence 和 Confluent Cloud 中新增多项能力，包括托管 MCP Server、Agent Skills、PII 自动脱敏，以及通过 Azure Private Link 连接外部模型。官方给出的判断是，很多 AI 项目并不是卡在模型本身，而是卡在数据层、治理层和安全层。

这是一条很值得注意的长尾新闻。因为企业 Agent 要真正可用，不能只消费静态知识，还必须接入持续变化的实时数据流。Confluent 的动作说明，数据基础设施厂商已经不满足于做“模型上游的数据管道”，而是在主动定义 Agent 的数据工作台：既要把上下文实时送到模型面前，也要把隐私治理和私网连接内建进去。对产业而言，这类玩家的崛起意味着 AI 系统栈正在重新分层：上面是模型与 Agent，中间是 MCP 和工具编排，下面则是数据流、治理与私有网络的基础设施。

五、今日判断：Agent 时代的主战场正在从“模型”迁移到“连接层”

把今天这几条新闻放在一起看，一个趋势非常明确：2026 年的 AI 产业，正在从模型竞赛进入连接竞赛。Google 做 Agent 运行时，Cohere 做可私有部署的 Agent 模型，Oracle 做数据库 MCP 入口，Confluent 做实时数据与治理中台。它们争夺的其实是同一件事：谁能让 AI 在真实业务环境中稳定、可控地工作。

因此，接下来最值得关注的并不只是模型排行榜，而是几个更工程化的问题：Agent 如何获得上下文，如何调用工具，如何被审计，如何运行在私网，如何与数据库、流式数据、代码仓库和业务系统拼接。对企业来说，AI 的分水岭已经不是“要不要上模型”，而是“能不能把模型安全地放进流程里”。

参考文献

1. Google，**Building the agentic future: Developer highlights from I/O 2026**，2026-05-19，用于核实 Gemini 3.5 Flash、Managed Agents、Antigravity 2.0 发布内容。 https://blog.google/innovation-and-ai/technology/developers-tools/google-io-2026-developer-highlights/
2. Google，**100 things we announced at Google I/O 2026**，2026-05-20，用于补充 Google I/O 整体 AI 发布背景。 https://blog.google/innovation-and-ai/technology/ai/google-io-2026-all-our-announcements
3. Cohere，**Introducing Command A+: Making sovereign agentic capabilities available to all**，2026-05-20，用于核实 Command A+ 定位、许可和硬件门槛。https://cohere.com/blog/command-a-plus
4. Hugging Face，**Model Card for Command A+**，2026-05-21 前后可见，用于核实参数规模、上下文长度和量化部署要求。 https://huggingface.co/CohereLabs/command-a-plus-05-2026-w4a4
5. Cohere Docs，**Release Notes**，2026-05-20，用于核实 Command A+ 已开放给 Cohere API 用户。 https://docs.cohere.com/changelog
6. Oracle，**Connecting Codex to Enterprise Data with Oracle Autonomous AI Database MCP Server**，2026-05-22，用于核实 Codex 接入企业数据库的最新演示场景。https://blogs.oracle.com/machinelearning/connecting-codex-to-enterprise-data-with-oracle-autonomous-

ai-database-mcp-server

7. Oracle, **Announcing the Oracle Autonomous AI Database MCP Server**，2025-12-24，用于补充 MCP Server 产品架构与安全特性背景。https://blogs.oracle.com/machinelearning/announcing-the-oracle-autonomous-ai-database-mcp-server
8. 8. Oracle Docs，**Autonomous AI Database MCP Server**，2026 年文档版本，用于核实托管、多租户、内置 MCP 与 Select AI Agent 能力。https://docs.oracle.com/en-us/iaas/autonomous-database-serverless/doc/mcp-server.html
9. 9. OpenAI Developers，**Docs MCP**，2026 年在线文档，用于核实 Codex 与 MCP 的配置方式和官方支持情况。https://developers.openai.com/learn/mcp
10. Confluent, **Confluent Makes It Easier to Build and Secure Real-Time AI at Scale**，2026-05-19，用于核实托管 MCP Server、Agent Skills、PII 脱敏与私有连接。https://www.confluent.io/press-release/build-and-secure-real-time-ai/

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