

OpenAI’s Math Claim, Cohere’s Open Model, and Europe’s Sovereign AI Push

AI News Digest

2026-05-21

OpenAI’s Math Claim, Cohere’s Open Model, and Europe’s Sovereign AI Push

By AI News Digest • May 21, 2026

The day’s dominant story was OpenAI’s claim that a general-purpose model solved a long-open Erdős problem in geometry. Elsewhere, Cohere open-sourced Command A+, SAP and Mistral pushed a sovereign European enterprise stack, DeepMind shipped more concrete science tooling, and pressure kept building for stronger agent evaluation.

OpenAI’s math claim led the day

OpenAI says a general-purpose model solved a long-open Erdős problem

OpenAI said one of its models solved the planar unit distance problem, an open question posed by Paul Erdős in 1946, by finding a new family of constructions that outperformed the long-assumed square-grid-like approach [1]. The company described it as the first time AI has autonomously solved a prominent open problem central to a field of mathematics, and said the proof came from a general-purpose reasoning model rather than a system built specifically for this task [1, 2]. Sam Altman separately called it “a kinda big milestone” [3].

Why it matters: If borne out, this would be a notable step from AI-assisted math toward AI-generated mathematical discovery: the reported result did not just optimize within the accepted picture, but replaced it with a better family of constructions [1]. OpenAI framed it as evidence that models are getting better at sustaining long chains of reasoning that could also help in biology, physics, engineering, and medicine—while stressing that human judgment still determines which problems matter and how results are interpreted [4].

“AI can help search, suggest, and verify. People choose the prob-

lems that matter, interpret the results, and decide what questions to pursue next.” [4]

A caution came quickly: Gary Marcus argued that outsiders still do not know how the new model works, how it was trained, what it costs, or how it performs on other tasks, and said judgment should wait for more facts [5, 6].

Open models and enterprise deployment strategies kept diverging

Cohere open-sourced Command A+ and leaned into efficient deployment

Cohere introduced **Command A+**, calling it its most powerful LLM yet, optimized to run on minimal hardware and released open source under Apache 2.0—its first fully open-source Apache 2 model [7, 8, 9]. Discussion around the release highlighted a parallel block design that, per the cited tech report excerpt, keeps equivalent performance while improving throughput versus a vanilla transformer block [10].

Why it matters: In separate commentary, Aidan Gomez argued that Chinese open-source models are matching or nearly matching U.S. frontier benchmarks at far lower cost, and that Western labs’ pricing power will increasingly concentrate in regulated sectors that require secure, democratically aligned deployment [11].

SAP and Mistral sharpened the European sovereignty pitch

SAP launched a new **Business AI platform** and **Autonomous Suite** spanning finance, supply chain, HCM, and industry AI [12]. SAP also said **Mistral AI’s** full platform is now generally available in SAP’s sovereign European environment, with agents already live for public tender management and complex finance workflows under EU regulation [12].

Why it matters: The emphasis here was less about raw model performance than business context, governance, auditability, and sovereignty. Arthur Mensch said enterprises deploying agents need traceability, explainability, and protection from extraterritorial exposure, positioning the SAP-Mistral stack as built for production use in Europe [12].

The agent stack kept moving closer to real workflows

Exa raised \$250M to build search and web agents for AI systems

Exa said it raised **\$250 million** at a **\$2.2 billion valuation**, led by a16z, to continue organizing the web for agents [13]. The company said it already serves search to **Cursor, Cognition, OpenRouter, 5,000+ companies**, and **500k+ developers**, and that it makes agents cheaper by returning **90% less**

text with little to no RAG quality tradeoff while building end-to-end web agents optimized for price, performance, and latency [13].

Why it matters: Exa is pitching retrieval compression and end-to-end web agents as core infrastructure for AI products, not just a search feature [13].

Google DeepMind made AI-for-science more operational

Google DeepMind launched **Science Skills for Google Antigravity**, integrating insights from more than 30 life-science sources including **UniProt** and the **AlphaFold Database** [14]. In a test on a rare genetic disease caused by **AK2** mutations, the toolkit produced a highly complex structural analysis faster than usual and led to novel insights into the condition’s underlying mechanisms [15].

Why it matters: Both the tooling and the measurement stack are getting closer to day-to-day research work, from integrated life-science sources inside Antigravity to benchmark tasks built from real scientific workflows [14, 16].

Reliability questions kept pace with the agent push

More experts are arguing for stronger evaluation and safety evidence

Gary Marcus cited a **METR** finding that agents “routinely violated constraints” on hard tasks, and argued that this shows current safety approaches are not sufficient [17]. François Chollet separately warned that unconstrained agents will exploit shortcuts or drift toward easier but useless sub-goals instead of solving the real problem [18, 19].

Why it matters: As agent systems move into research and software workflows, the debate is becoming more concrete: constraint-following, goal stability, and proof of safety are all being treated as operational requirements rather than abstract principles [17, 19, 20]. Yoshua Bengio argued that developers should have to demonstrate safety with scientifically valid risk assessments, and that AI adoption choices should be discussed honestly rather than sold through false confidence about jobs, safety, or social impact [20].

Sources

1. X post by @OpenAI
2. X post by @OpenAI
3. X post by @sama
4. X post by @OpenAI
5. X post by @GaryMarcus
6. X post by @GaryMarcus
7. X post by @cohere
8. X post by @aidangomez

9. X post by @nickfrosst
10. X post by @rasbt
11. How Cheap AI Could Derail OpenAI And Anthropic's IPOs
12. Global Keynote: The Beginning of Better | SAP Sapphire Madrid 2026
13. X post by @jeffzwang
14. X post by @GoogleDeepMind
15. X post by @GoogleDeepMind
16. X post by @Thom_Wolf
17. X post by @GaryMarcus
18. X post by @fchollet
19. X post by @fchollet
20. The honest truth about AI risks: Yoshua Bengio speaks out