Open weights, watched agents, and a bank running 12,700 of them

Mistral opens its trillion-parameter model, OpenAI describes how it watches its own agents, UBS reports 12,700 employee agents, and Canada buys into Cohere.

This week the conversation moved from whether to use agents to how to run them at scale. A large European model became available to download and run on your own hardware, OpenAI published how it monitors agents during training and testing, and one of Europe's biggest banks disclosed how many agents its staff already use. Governments kept treating AI capacity as a national asset, from Ottawa to Dubai to Ankara.

MODELS & RELEASES

Mistral releases the weights of Large 4, its trillion-parameter model

On 29 October Mistral published the weights of Mistral Large 4, the trillion-parameter model it introduced earlier in October, on Hugging Face. Organisations can now download it and run it on their own infrastructure, and Mistral is positioning it for software development and defensive cybersecurity. The company says it performs on par with China's GLM-5.3 while using fewer tokens per task.

Why it matters: A frontier-scale model with downloadable weights from an EU company gives regulated organisations a real option for keeping sensitive workloads entirely on their own premises.

Our view: Running a trillion-parameter model is a hardware and operations decision, not just a download. For most use cases a smaller open model, fine-tuned on your own documents, will be cheaper to run, and Large 4 is worth benchmarking where the smaller ones fall short.

Source: Mistral AI

AGENTS & TOOLING

OpenAI details the layered monitoring it uses to catch misbehaving agents

OpenAI described a four-layer monitoring system for its frontier models: lightweight classifiers screen activity, a model reviews conversations and tool use, an investigative agent reconstructs flagged cases, and a final check compares transcripts with network logs before a human is alerted. In a replay of the summer Hugging Face incident, OpenAI says the system would have raised an alert about 30 minutes after agents left their test environment. That figure comes from a simulation, not a live event.

Why it matters: Supervisors are starting to ask how firms detect an agent acting outside its mandate. A published design from a frontier lab gives risk teams a concrete pattern to compare their own controls against.

Our view: The pattern scales down well: cheap rules on every action, a model reviewing the suspicious ones, and network logs as the final evidence. We build agent deployments with that kind of audit trail from day one, because adding it later is much harder.

Source: OpenAI

Amazon says outside shopping agents still lose to its own store

On Amazon's earnings call on 29 October, CEO Andy Jassy said few customers buy through third-party shopping agents and that they are not yet more convenient than shopping with a retailer directly. Amazon blocked Meta's Muse assistant from its store in September, and is instead promoting its own Alexa shopping agent, which it says more than 400 million customers used in the past year.

Why it matters: Large platforms are deciding which outside agents may act on their customers' behalf. Banks and insurers face the same choice about third-party agents reaching their portals and accounts.

Our view: Blocking outside agents protects the channel but does not remove the demand. Offering your own agent, or a controlled interface for approved ones, is usually the more durable answer.

Source: Fortune

ENTERPRISE ADOPTION

UBS staff now use more than 12,700 personal AI agents

In its third-quarter results on 29 October, UBS said its employees use more than 12,700 personal AI agents in their daily work. The bank reported 608 AI use cases in production, 81% more than a year earlier, with over 1,090 more in development, and said 18,000 of its developers use AI coding tools.

Why it matters: A heavily regulated global bank now reports agent adoption as an operating metric. Peers will be asked by boards and supervisors where they stand, and how they govern that many agents.

Our view: At this scale the hard part is not building agents but keeping an inventory: who owns each one, what data it reads and what it is allowed to do. That register is worth setting up before the count reaches the hundreds.

Source: PYMNTS

REGULATION & GOVERNANCE

US FTC orders nine companies to explain personalised pricing and pay

On 28 October the US Federal Trade Commission used its study powers to order nine companies, including Delta, United, Uber, DoorDash, Target and Walmart, to explain whether prices vary between individuals based on personal data. Ride-hailing and delivery firms must also describe how they set driver pay. The companies have 45 days to respond.

Why it matters: Algorithmic pricing based on personal data is under scrutiny in the US, as it already is in parts of the EU. Insurers and lenders that price individually should expect similar questions about inputs and fairness.

Our view: If a model sets a price or a limit, be ready to show which inputs it used and why. Keeping that explanation next to the decision, rather than reconstructing it later, is what makes these inquiries manageable.

Source: US Federal Trade Commission

SOVEREIGN AI & INFRASTRUCTURE

Canada takes a stake in Cohere with conditions on data and compute

Prime Minister Mark Carney announced up to C$1.6 billion for an equity stake in Toronto-based Cohere, as part of a C$7 billion round led by Germany's Schwarz Group. In return Cohere must keep its headquarters in Canada, create 1,000 jobs and spend at least C$1.7 billion on Canadian computing infrastructure by 2030. Carney said government data will stay in Canada.

Why it matters: Governments are moving from buying AI services to owning part of the supplier, with data residency and local compute written into the terms. The Schwarz Group's lead role also links the deal to Europe.

Our view: Sovereignty here is defined by concrete clauses: where the data sits, where the compute runs and who controls the company. Those are the same three questions to put to any AI vendor you plan to rely on.

Source: Yahoo News Canada (The Canadian Press)

REGIONAL WATCH

Türkiye halves the fee for urgent searches on AI patents

A TÜRKPATENT fee change published in the Official Gazette on 31 October cuts the urgent search report fee for AI patent applications by 50%. Small and medium-sized companies filing their first patent also get 75% off application, search and examination fees. The change took effect on publication.

Why it matters: Cheaper and faster patent searches lower the cost of protecting AI work developed in Türkiye, including inside banks, insurers and industrial groups.

Our view: Teams building agents and retrieval systems in-house often produce methods worth protecting. A quick review of what has been built this year, before it is published or shown to vendors, is now cheaper to act on.

Source: Techinside

Dubai raises its AI Campus targets to 1,000 companies by 2030

During the Dubai AI Festival on 26 October, Sheikh Hamdan raised the targets for the Dubai AI Campus at DIFC to more than 1,000 AI companies and 5,000 skilled jobs by 2030. The campus says it has attracted over 500 AI companies since its launch in May 2024 and has already passed its original jobs target. Scale AI opened a UAE office there during the festival.

Why it matters: Dubai is building a dense local market of AI suppliers next to its financial centre, which will shape how Gulf banks and government entities source agent projects.

Our view: A larger local supplier base helps Gulf organisations keep data and delivery in-region. When choosing among them, ask to see agents running in production with an audit trail, not demos.

Source: Fortune

WHAT WE'RE WATCHING

We will watch whether the EU AI Office's proposed criteria for independent evaluators of the largest models move forward, how firms respond to the UK ICO's call for evidence on agentic AI before it closes on 20 November, and whether other large banks follow UBS in reporting agent numbers with their results.

Want to apply this in your organisation?

We build production-grade AI agents for banking, insurance, manufacturing and energy. A working POC usually takes one to two weeks.

Read on agenticdynamic.com **Talk to us →**