

COVER STORY · CONSUMER AGENTS

Meta Goes All-In on Muse: The Agent Gets a Face, an Inbox, a Wallet — and a Business Model

At Connect, Zuckerberg made tokens nearly free and bet the company on skimming transactions instead. That quietly rewrites who pays for AI.

THE PRICE WAR

Opus 5.5 at \$4/\$20, GPT-6 Luna at \$0.10 — and incumbents discounting to keep you.

TURTLE BAY

Altman and Amodei ask the Security Council for rules; Washington says no, thanks.

MOVE 37

“Cache gravity”: why the price cuts are a lock-in instrument, not a discount.

EDITOR'S NOTE

Tonight's edition draws on four desks — X, Semafor, The Information and TechCrunch — and ranks each story by corroboration: how many of those desks independently carried it, with significance breaking ties. The day's theme is **state as the new lock-in**: agents are acquiring email addresses, payment rights and cached context inside vendor platforms, just as the labs cut sticker prices and governments argue over who should hold the off-switch.

Free Tokens, Paid Transactions: Meta's Muse Becomes a Platform

Carried by 2 desks · TechCrunch (4 pieces) · The Information

Meta used Wednesday's Connect keynote in Menlo Park to declare Muse — its personal agent, only weeks old — the centerpiece of the company. Zuckerberg framed it as the future “personal superintelligence” for billions, and announced a slate of new surfaces: a real-time video avatar (powered by a new Muse Realtime Avatar model), wake-word integration on Meta's smart glasses “in the coming months,” full computer-use on the Mac, and — tellingly — Muse getting its own email address so it can be cc'd on threads and work in the background.

The commercial architecture is the real news. Meta says it will make Muse “free for a huge number of tokens,” expecting to profit “by taking a small fee from transactions.” It has wired in Shopify's full catalog and Shop Pay, PayPal and Stripe, plus retail connectors for Best Buy, Gap, Sephora, Walmart and Wayfair, travel via Expedia, and Instacart soon. On the work side, GitHub, Granola and Notion connectors ship now, and Meta says more than 1,500 developers applied to build connectors in under a week.

“We're standing behind this by making Muse free for a huge number of tokens, with the expectation that over time we will profit by taking a small fee from transactions.” — Mark Zuckerberg, via TechCrunch

The competitive ripple is already visible. The Information reports OpenAI is developing features to counter xAI's Grok bot and weighing a response to Muse (headline only; paywalled). The same day, OpenAI pushed voice-driven agentic tasks into the ChatGPT mobile Work tab. Earlier in the week TechCrunch reported Meta conceded Muse's design was deliberately modeled on the open-source OpenClaw agent.

BY THE NUMBERS

1,500+

developer applications to build Muse connectors in under a week

Alexandr Wang at Connect, via TechCrunch

4

new Muse surfaces: live avatar, AI glasses, Mac computer-use, own email address

TechCrunch

5

named retail partners (Best Buy, Gap, Sephora, Walmart, Wayfair) plus Expedia; Instacart “soon”

TechCrunch

≈\$0

consumer token price for a “huge number of tokens”; revenue via transaction fee

Zuckerberg, via TechCrunch

Weeks

Muse's age — it launched in early September

TechCrunch

Why it matters to a CTO: an ad-funded giant pricing agent inference near zero changes the reference price your employees carry into work. More importantly, a transaction-fee model means Muse’s incentives are aligned with *completing purchases*, not with your procurement policy. If you sell online, agent-originated traffic arriving through Shopify, PayPal and Stripe rails becomes a channel with a platform tax — plan for “agent-legible” catalogs and APIs now. If you run IT, a consumer agent with Mac control and its own inbox is a shadow-IT vector; your DLP and identity policy should name it explicitly.

Sources: [TechCrunch — Everything new coming to Muse](#) · [TechCrunch — Muse wearable](#) · [TechCrunch — OpenClaw likeness](#) · [The Information — OpenAI mulls response to Muse \(headline only\)](#)

MOVE 37 · THE NON-OBVIOUS READ

Cache Gravity: The Price Cuts Are a Lock-In Instrument, Not a Discount

Everyone is reading this week’s pricing through the headline input/output rates. Look instead at the line most reports skip: Opus 5.5 cuts *cache reads* from \$0.50 to \$0.20 per million tokens. In agentic workloads, the same system prompt, tool schemas and working context are resent on every step, so the cached share of tokens routinely dwarfs fresh input. The effective bill is set by the cache rate — and a cache lives on one vendor’s servers.

That makes cheap cache reads a gravitational field. The more your agents run, the more warm state you accumulate with a single provider, and the more expensive it becomes to route a given workflow elsewhere (you pay full cold-start price to rebuild context). Meta’s Muse is the consumer version of the same move: free tokens, but the agent’s inbox, connectors and purchase history accrue inside Meta. And the UN’s CIOs, per Semafor, are drafting “exit plans” precisely because Mythos access was cut off once already.

The counter-intuitive conclusion: in a price war, the cheapest vendor is not the one with the lowest rate card — it is the one whose state you can walk away from. Measure it.

Try this Monday: for each production agent, log *cached-token share* and compute a “switching premium” = (tokens to re-warm context on vendor B × B’s uncached rate) per day. Any workflow where that premium exceeds a week of savings is already locked in — and your routing layer should know it.

1

CARRIED BY 2 DESKS · TECHCRUNCH · THE INFORMATION

Meta makes Muse the company's centerpiece at Connect NEW

Avatar, glasses, Mac control, its own email address, and a commerce stack (Shopify, Shop Pay, PayPal, Stripe, major retailers). Tokens near-free; Meta takes a cut of transactions. OpenAI is reportedly weighing a response.

CTO READ Treat consumer agents with desktop control and inboxes as a new endpoint class in policy; treat agent-originated shoppers as a channel with its own take-rate.

Sources: [TechCrunch](#) · [The Information](#) (headline only)

2

CARRIED BY 2 DESKS · TECHCRUNCH · X

Frontier price war: Claude Opus 5.5 vs GPT-6 Sol and LunaNEW

Anthropic's Opus 5.5 lands at \$4/\$20 per million tokens with what TechCrunch calls "Fable-level performance"; OpenAI's GPT-6 Sol (\$2/\$10) and Luna (\$0.10/\$0.50) roughly halve the prior tier. On X, even xAI's Grok account named Opus 5.5 the top coding model. Adjacent: The Information reports software incumbents are discounting their own AI to keep customers from defecting to the labs.

CTO READ Don't pick a winner — build the router. Route long-horizon coding/computer-use to the strongest agentic model, bulk extraction to the cheapest tier, and re-bench on *your* evals before migrating; secondary reporting notes Sol regresses on some agentic-coding benchmarks versus its predecessor. Use the incumbents' discounting as leverage at renewal.

Sources: [TechCrunch](#) — Opus 5.5 · [TechCrunch](#) — GPT-6 Sol & Luna · [X live search](#) (Grok reply re: Opus 5.5) · [The Information](#) — incumbents discount AI (headline only) · Pricing detail (Build Fast with AI)

3

CARRIED BY 1 DESK (3 PIECES) · SEMAFOR · + CNN, AL JAZEERA/AP/REUTERS, BLOOMBERG

Altman and Amodei ask the UN Security Council for global AI rules; the US rejects “centralised control” **NEW**

Amodei told the council AI “could be a risk to humanity as a whole” and proposed narrow global bans (e.g., AI for bioweapons), verification regimes and a shared incident-notification system. Altman said decisions “cannot be made by labs in San Francisco alone.” US OSTP chief Michael Kratsios answered that Washington “totally rejects” international efforts at global AI governance. Trump hosts Xi today; analysts expect at most shared safety definitions and an incident hotline.

CTO READ Expect incident *notification* to arrive before capability limits. Stand up an internal AI-incident log with severity taxonomy now — it will map onto whatever reporting regime emerges, domestic or bilateral.

Sources: [Semafor](#) — bosses to address UN · [Semafor](#) — OpenAI’s US-led coalition · [Al Jazeera/AP/Reuters](#) · [CNN](#) · [PIIE](#) on Trump–Xi

4

CARRIED BY 1 DESK · SEMAFOR (EXCLUSIVE)

UN agencies draft “exit plans” from US cloud and AI vendors **NEW**

The ITU’s CIO Khuloud Odeh says last year’s ICC sanctions and this year’s Mythos model cutoff were a “wake-up call.” UN agencies are exploring shared compute and data infrastructure, sorting systems by sensitivity and asking “what would be an exit plan?” — with “sanctions-driven service suspension” now a named cloud risk.

CTO READ If a UN agency with 1,000+ AI use cases is writing vendor-exit runbooks, a multinational should too. Tier workloads by sensitivity and keep at least one open-weight fallback warm for tier-1 flows.

Source: [Semafor](#)

5

CARRIED BY 1 DESK · TECHCRUNCH · (AMODEI POST ON X CITED)

Anthropic's wet lab says Claude found a CRISPR-like enzyme system **NEW**

Anthropic says its Bay Area lab surfaced a previously unknown phage enzyme system that can cut, copy and paste DNA — found “mostly, though not entirely, by Claude” in about 21 hours using ~950 agents and 210M tokens. Humans performed all bench work, at BSL-1/2 only; Amodei acknowledged a similar Stanford finding. Validation by the field is pending.

CTO READ The reusable pattern is the architecture: massive parallel agent search over data, with humans owning the physical, irreversible step. That is the right template for any R&D or ops domain where actions can't be undone.

Source: [TechCrunch](#)

6

CARRIED BY 1 DESK · THE INFORMATION (HEADLINE ONLY)

Same flaw found in Claude Code, Codex, Gemini CLI and GitHub Copilot **NEW**

The Information's Applied AI desk reports a single vulnerability class spanning all four major coding agents. Details are paywalled; we could not verify the specifics. Context: Cisco Talos this week disclosed CLOSEDQUORUM, an implant that lets a four-model LLM panel vote on its next post-exploitation step, with no human operator.

CTO READ A cross-vendor flaw means “we use a different agent” is not a mitigation. Sandbox coding agents (no ambient creds, egress allow-lists), and add AI-API call patterns to EDR hunting.

Sources: [The Information](#) (headline only) · [Cisco Talos](#) — [CLOSEDQUORUM](#)

7

CARRIED BY 1 DESK · TECHCRUNCH

ChatGPT brings voice-driven agentic “Work” tasks to mobile **NEW**

Plus and Pro users can now draft documents, summarize Slack, build sites and use the cloud browser by voice from their phones, with handoff to desktop. Free and Go users get plugins and connected apps.

CTO READ Voice-to-agent on personal phones bypasses keyboard-centric DLP. Revisit which enterprise connectors (Slack, mail) are OAuth-approvable from consumer ChatGPT accounts.

Source: [TechCrunch](#)

DeepSeek bets big on Huawei chips to route around US export controls **NEW**

An exclusive reports DeepSeek is committing heavily to Huawei silicon. Paywalled; headline only. It lands as Trump and Xi meet with chips and AI on the agenda.

CTO READ A credible non-Nvidia training stack in China would strengthen the open-weight tier you may use as a fallback — and complicate its provenance review. Track model origin in your AI inventory.

Source: [The Information](#) (headline only)

ALSO NEW TODAY

- **Superintelligence ban bill:** Sanders and Casar propose a pause on advanced AI and a cabinet-level Department of AI. [AI Jazeera](#)
- **Meta camera-free AI glasses** debut at Connect. [TechCrunch](#)
- **Enveda raises \$311M** to push nature-derived AI drugs into trials. [TechCrunch](#)
- **Ema raises \$77M** as AI eats into enterprise software and services. [TechCrunch](#)
- **Snorkel AI triples valuation to \$3.5B** on training-data demand. [TechCrunch](#)
- **YouTube lets users build their own algorithm** with AI. [TechCrunch](#)
- **Heavy AI users are still worried about it**, new survey finds. [TechCrunch](#)
- **OpenAI “close” on another Millennium Prize problem** (teaser). [The Information](#)
- **Anthropic IPO waiting game** puts Wall Street on edge (headline only). [The Information](#)
- **Jane Street-linked data center debt sours** (headline only). [The Information](#)
- **Tsinghua professor’s stealth LLM startup** hits \$1.4B (headline only). [The Information](#)
- **Nvidia exec:** “fear of the unknown” drives data-center pushback. [Semafor](#)

CONTRARIAN WATCH · WHERE THE DESKS DIVERGE

1 · Anthropic’s biology discovery: breakthrough, or governance story?

Anthropic and Amodei lead with the find and the speed (21 hours, ~950 agents). TechCrunch frames the bigger reveal as the opposite — that humans, not Claude, still run the bench — and places it against the same CEO’s bioterror warnings and his call to “pace the frontier.” Read it as both: a capability signal *and* a template for human-gated autonomy.

2 · “Global rules” vs “US-led coalition”

At the UN, both CEOs urged international standards. Semafor’s earlier reporting shows OpenAI simultaneously pitching a *US-led* safety coalition, while the administration rejects international

governance outright. The public “call for global oversight” framing overstates convergence: the live debate is over *who leads*, not whether to coordinate.

BACK PAGE · COVERAGE GAPS

- **X (thin):** Browser session reached X live search, but page extraction returned only the top post per query. One credible AI-relevant item was used (Grok’s reply naming Opus 5.5 top for coding); the general “AI OR LLM” stream returned off-topic noise and was discarded. X home feed not sampled.
- **The Information (paywall):** All items are headline/teaser only — never bypassed. Specifics of the cross-agent coding flaw, DeepSeek–Huawei, and OpenAI’s Muse response are unverified.
- **Semafor newsletter page** served a stale June 13 edition from cache; we relied on the Semafor tech index and article pages (Sept 21–23) instead.
- **Freshness:** The Opus 5.5 / GPT-6 price moves are dated Sept 22 (~48h) — included because no prior-edition memory file was readable, so all stories are marked NEW. Wire coverage (Al Jazeera/AP/Reuters, CNN, PIIE, Cisco Talos) supplemented the four desks for the UN and security items; secondary pricing detail from an aggregator blog is flagged as such.
- **Not yet reported at press time:** outcomes of the Trump–Xi meeting on AI.

Desks: X / Twitter · Semafor Tech · The Information · TechCrunch AI

Method: stories deduplicated by normalized ID across desks, ranked by number of desks carrying them, ties broken by significance to technology leadership; NEW = not in the prior 14-day memory.

Disclaimer: Generated content — verify market-moving items against primary sources.

THE SIGNAL · Vol. 26 · No. 267 · Thursday, September 24, 2026