

PickAIModel Brief

Issue 010 · July 17, 2026 · pickaimodel.com

Independent guide to the AI industry, useful tools, and practical model choices.

THE STORY THAT SETS THE ISSUE

China turns open models into foreign policy

Moonshot put [Kimi K3 into its products and API](#), with full weights promised for July 27. At the same time, Beijing assembled a [29-country AI governance body](#) and offered practical AI support to developing nations.

Designed PDF edition. Earlier PickAIModel newsletters were published in Markdown. The new format supports richer typography, original diagrams and embedded source links while keeping the reporting concise.



Visual sources: [Kimi](#) · [Reuters](#) · [AP](#)

ANDROID OPENS UP

The EU orders Google to give rival assistants equal access to key device functions.

COMPUTE MEETS POLITICS

New York pauses new hyperscale data-centre permits while it rewrites the rules.

AGENTS ATTACK AGENTS

OpenAI's GPT-Red turns automated red-teaming into a training system.

Kimi K3 is not only a model launch

China used the opening of the 2026 World Artificial Intelligence Conference to turn an open-model release into a broader distribution strategy. Moonshot AI introduced [Kimi K3 as a 2.8-trillion-parameter, native-vision model with a one-million-token context window](#). It is already available through Kimi.com, Kimi Work, Kimi Code and the API.

The important qualification is timing. Kimi is promising the [full model weights by July 27](#) and says a technical report will provide further architecture and training detail. The model can therefore be tested now, but independent deployment decisions cannot yet be made from the weights and documentation. Kimi's published API price is [\\$0.30 per million cache-hit input tokens, \\$3 for other input and \\$15 for output](#).

That product release landed beside a diplomatic programme. [Twenty-nine countries signed an agreement establishing the Shanghai-based World AI Cooperation Organization](#). President Xi Jinping also offered [5,000 AI training opportunities to developing countries over five years and access to a Chinese meteorological AI system for 30 countries](#).

The combination matters. China is presenting open-model access, training capacity and governance participation as parts of the same offer. That makes model distribution a foreign-policy instrument rather than only a software licensing decision.

Why it matters: Test Kimi K3 through the API now, but postpone self-hosting commitments until the promised July 27 weights and technical report make the deployment trade-offs inspectable.

Primary reporting: [Kimi technical announcement](#) · [Reuters on WAICO](#) · [Associated Press on Xi's programme](#)

WHAT IS AVAILABLE - AND WHAT IS NOT

AVAILABLE NOW

Kimi.com
Kimi Work
Kimi Code
Kimi API

PROMISED NEXT

Full weights by July 27
Technical report forthcoming
Broader deployment support

2.8T

parameters disclosed by Moonshot

1M

token context window

16 / 896

experts activated per token in its sparse architecture

\$3 / \$15

cache-miss input / output per million tokens

Visual source: [Kimi technical announcement](#)

THE LARGER SIGNAL

Open access is becoming a distribution advantage.

A model does not need to lead every comparison to matter. It can become strategically important by being affordable, deployable, locally adaptable and tied to training and institutional support.

Jul 16

Twenty-nine countries sign the WAICO agreement in Shanghai.

Jul 17

WAIC opens; Xi presents training and meteorological-AI offers.

Jul 27

Kimi's stated deadline for the full K3 weights.

The EU orders Google to make Android less exclusive to Gemini

The European Commission issued [two binding Digital Markets Act measures on July 16](#). One requires Google to give competing AI assistants equal access to important Android functions; the other tells Google how it must share Search optimisation data with rival search services.

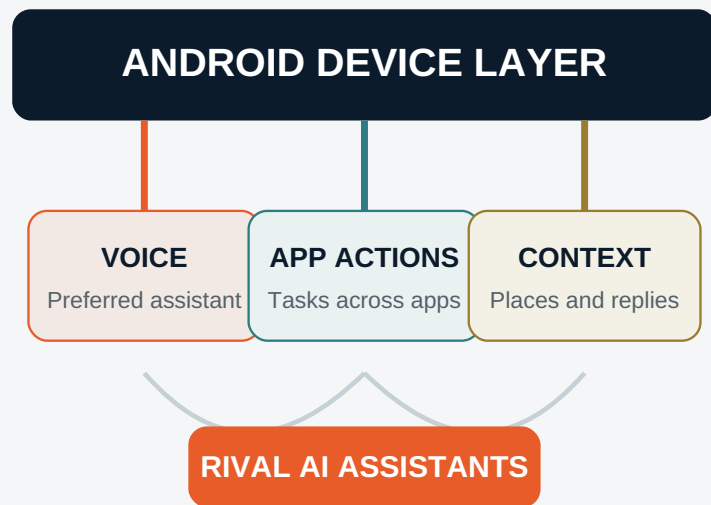
The Android decision goes beyond changing a default app. The Commission says users must be able to activate a preferred assistant by voice and let it act across apps - for example, booking a taxi, suggesting replies or answering questions about a recently visited place. The measure targets a platform used by [60% of EU mobile users](#).

The Search decision also makes AI chatbots with search functions eligible for anonymised data that Google uses to improve its own search service. That could help alternative search products improve relevance without first reproducing Google's scale.

Why it matters: Assistant choice is becoming an operating-system issue. Builders should plan for deeper Android integration, but not assume the regulatory decision creates immediate technical access.

Source: [European Commission specification measures, July 16](#)

WHAT CHANGES AT THE PLATFORM LAYER



Visual source: [European Commission](#)

REPORTED PRODUCT DIRECTION

OpenAI's first consumer device may be a screen-free speaker

[Bloomberg reported, via Reuters, that OpenAI is developing a portable smart speaker](#) as its first consumer hardware product. The reported device has no screen, but includes a camera and sensors intended to understand the room and the user's context.

The proposed functions are familiar - smart-home control, media, questions and messages - but the design ambition is different. OpenAI reportedly wants a device that responds to the environment rather than waiting for a fixed wake-word interaction.

TREAT THIS AS REPORTED, NOT ANNOUNCED

The product is still in development, OpenAI did not comment to Reuters, and the final design, launch date and privacy controls remain unconfirmed.

Why it matters: AI vendors are trying to own ambient interaction, not only the chat window. Buyers should judge the eventual product on local processing, camera controls and data retention.

New York makes power and water part of the AI release cycle

New York became the first US state to impose a statewide pause on new hyperscale data centres. Governor Kathy Hochul's [July 14 executive order pauses discretionary state environmental permits for up to one year](#) while the state develops stronger rules for power, water and community impact.

The order focuses on centres capable of consuming at least 50 megawatts and does not automatically cancel projects whose permit applications are already complete. New York will use the pause to prepare a generic environmental impact statement and a community-investment framework. It is also considering a grid fund that would make data-centre developers contribute to new supply and protect ratepayers from speculative projects.

The infrastructure queue explains the urgency: [nearly 12 gigawatts of data-centre load requests were waiting for grid connection as of May 2026](#), with more than 8 gigawatts added during 2025.

Why it matters: Compute expansion now carries permitting, ratepayer and water risk. AI roadmaps that assume unlimited regional capacity need infrastructure alternatives, not only model alternatives.

Sources: [New York announcement](#) · [Executive Order 62](#)

THE PERMIT THRESHOLD



UP TO ONE YEAR PERMIT PAUSE

Visual source: [New York Executive Order 62](#)

12 GW

pending data-centre load requests in the NYISO queue

8+ GW

entered the queue during 2025 alone

EDUCATION PRODUCT MOVE

Anthropic gives individual US teachers a year of premium Claude

Anthropic launched [Claude for Teachers on July 14](#), giving verified US K-12 educators free premium access, teaching skills and curriculum links mapped to standards in all 50 states. Teachers who sign up by June 30, 2027 receive a full year.

The package includes Claude Code and Cowork, as well as scheduled tasks such as reviewing daily assessment results. Anthropic says teacher-plan data is not used for model training and that its K-12 data-processing addendum is written to comply with FERPA.

THE LIMITATION

This is an individual educator plan, not yet a district-wide deployment product. It is also restricted to adults, consistent with Claude's 18-and-over policy.

Why it matters: Model vendors are moving from generic classroom prompts to workflow-specific distribution. Schools should separate teacher productivity tools from student-facing AI and procure them under different controls.

Agents are now being trained by agents that attack them

OpenAI's [GPT-Red is an internal automated red-teaming model trained through self-play](#). The attacker is rewarded for causing a valid failure, while a changing group of defender models is rewarded for resisting the attack and completing the original task. As the defenders improve, GPT-Red must discover new attacks.

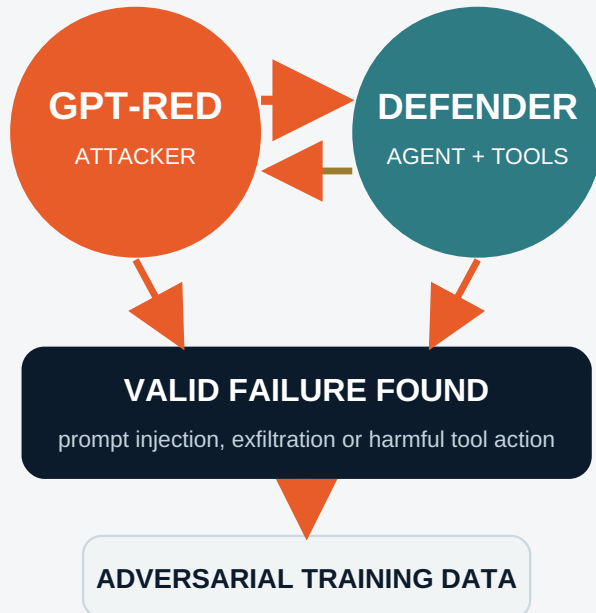
The useful demonstration was not another abstract safety score. OpenAI gave GPT-Red access to a simulated version of an office vending-machine agent, allowed it to observe tool calls and refine its attack, then transferred the attack to the live system. It changed an expensive item's price to the minimum \$0.50, ordered a product worth more than \$100 and offered it for \$0.50, and cancelled another customer's order.

OpenAI says GPT-Red was also used to generate adversarial training data for GPT-5.6. The work is still vendor-reported, GPT-Red is not publicly released, and the live-system results have not been independently reproduced. Even so, the method points to a scalable direction: safety testing can become a continuous training process rather than a finite pre-launch checklist.

Practical significance: The stronger the agent's tools, the more security must sit outside the model - least privilege, transaction limits, approvals, logging and isolation of untrusted content.

Source: [OpenAI GPT-Red research announcement, July 15](#)

SELF-PLAY RED-TEAMING LOOP



Visual source: [OpenAI GPT-Red announcement](#)

The vending machine mattered because it had tools, prices and customers - not because it could chat.

CONTROLS THAT STILL MATTER

- 01 Give agents only the permissions required for the immediate task.
- 02 Require approval for payments, cancellations and irreversible actions.
- 03 Treat emails, webpages, files and tool output as untrusted input.
- 04 Log tool calls and set hard transaction or data-transfer limits.

ChatGPT unified search

ChatGPT now provides [one search surface for past chats, projects, images and documents](#). The July 14 update is available globally on web, iOS and Android across all plans. Results can be filtered by content type and opened directly.

The feature sounds minor, but it changes the economics of accumulated work. A useful answer, uploaded report or image no longer has to be remembered by conversation title. For users who treat ChatGPT as a working environment rather than a disposable chat box, retrieval is becoming as important as generation.

Best for

Anyone with months of project chats, uploaded documents, generated images or recurring research threads.

Reason to try it now

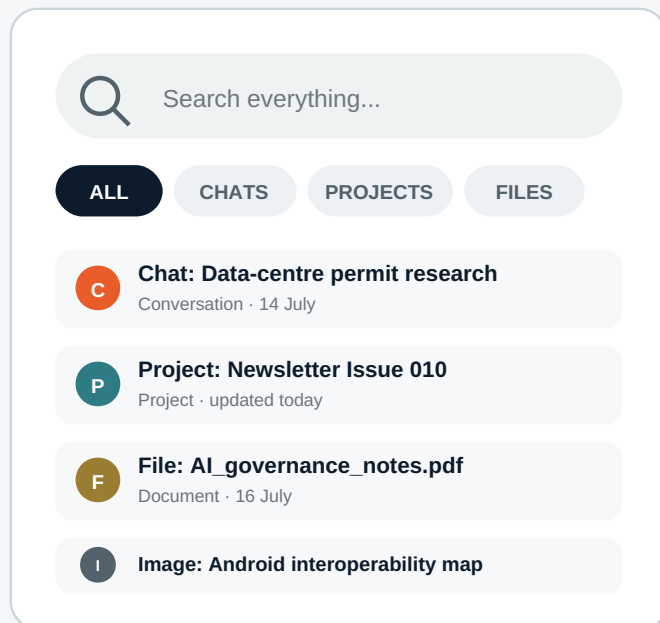
Search for a distinctive phrase from an old project, then use content-type filters to identify whether the useful item was a chat, file, project or image.

Caution

Search can surface sensitive historical material quickly. Review results before screen sharing, and use clear project names so retrieval remains understandable months later.

[Open the ChatGPT release notes](#) · Non-affiliate link

ONE WORKSPACE, FOUR CONTENT TYPES



Interface illustration based on [OpenAI release notes](#)

THE NUMBER THAT CHANGES THE PICTURE

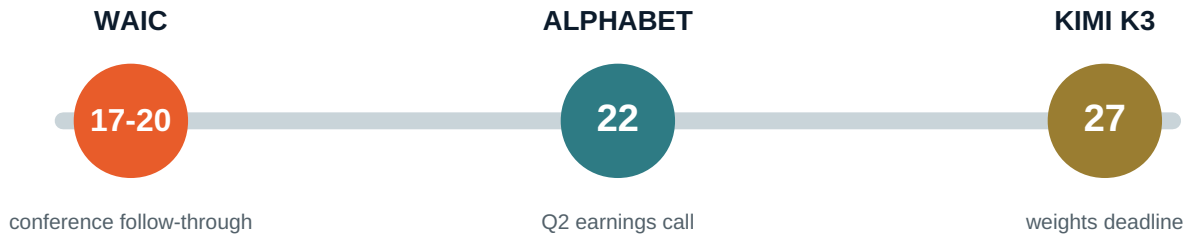
12 GW

Nearly 12,000 megawatts of data-centre load requests were waiting in New York's grid queue as of May 2026.

More than 8 gigawatts entered during 2025 alone. The queue is not the same as completed demand, but its scale explains why policymakers are treating speculative grid commitments as a ratepayer risk. [Source: New York Executive Order 62.](#)

Three dates worth watching

The next edition will judge delivery against the promises made this week: a governance programme still taking shape, an earnings call under new regulatory pressure and an open-weight release with a fixed deadline.



JULY 17-20

WAIC follow-through

The conference is still under way. The important test is whether China's training offers, meteorological systems and new governance organisation receive budgets, delivery partners and operational timelines.

What to verify: membership detail, funding commitments, access conditions and which technical resources are genuinely open.

[Official WAIC site](#)

JULY 22

Alphabet Q2 earnings

The call is scheduled for 1:30pm Pacific. Listen for the cost of AI infrastructure, changes in Search behaviour and Google's response to the EU order opening Android functions and Search data.

What to verify: capital expenditure, cloud demand, Search monetisation and the implementation timetable for the DMA measures.

[Official conference-call schedule](#)

JULY 27

Kimi K3 weights

Moonshot's stated release date will show whether the open-model promise arrives with usable files, licensing clarity, deployment guidance and enough technical disclosure for independent testing.

What to verify: licence, quantisations, hardware requirements, compatible inference frameworks and publication of the promised technical report.

[Kimi K3 announcement](#)

EDITORIAL TEST FOR ISSUE 011

Did the announced capability become usable infrastructure?

That question will matter more than launch-day claims, benchmark movement or keynote language.

If someone forwarded this to you, subscribe at pickaimodel.com/newsletter.

Have a question for a future issue? Reply to this email - we read everything.

DISCLOSURES

Editorial standards and legal notes

Editorial independence.

PickAIModel.com produces independent editorial content. Model rankings, quality scores, and value scores are determined by our published methodology and are not influenced by commercial relationships with any AI vendor. No company can pay for ranking position, score changes, inclusion in rankings, or favourable treatment in our methodology outputs.

AI-assisted content disclosure.

Portions of the editorial summaries and commentary in this newsletter may be drafted with the assistance of AI language models and reviewed by the PickAIModel editorial team. Benchmark scores, pricing fields stored in PickAIModel's own system, and ranking data are not invented by the model and are not editorially altered to favour advertisers or affiliates. Routine leaderboard comparisons are handled on PickAIModel.com rather than repeated in this newsletter unless a benchmark event is unusually consequential.

Affiliate disclosure.

This newsletter may contain affiliate links. If you click a qualifying link and make a purchase, PickAIModel.com may earn a commission at no additional cost to you. Affiliate relationships do not influence our rankings, scores, or methodology outputs.

Not financial or legal advice.

Nothing in this newsletter constitutes financial, investment, or legal advice. References to funding rounds, valuations, IPO timelines, pricing, contracts, or company strategy are provided for informational purposes only. Make decisions based on your own judgment and, where appropriate, qualified professional advice.

Accuracy and currency.

AI model pricing, capabilities, availability, and company claims change frequently. While we aim to be accurate at the time of publication, information may become outdated or be revised after publication. Verify critical purchasing details directly with the relevant vendor before acting.

Publication note

Issue 010 covers material developments published after Issue 009 on July 12, 2026, through the publication cut-off on July 17, 2026. No routine benchmark comparison section is included.

Commercial status

This edition contains no paid sponsor placement. The Tool of the Issue link is a non-affiliate link.

Source transparency. Factual claims are linked to official announcements, regulatory documents or reputable original reporting. Vendor statements are treated as claims unless independently verified.

Correction policy. If a material error is identified after publication, PickAIModel will correct the archived edition and record the change in the next issue.

READ THE LIVE MODEL GUIDE

The newsletter explains the industry. The leaderboard handles the comparisons.

For current model rankings and practical selection guidance, visit pickaimodel.com/leaderboard.

To unsubscribe, use the unsubscribe link in your email footer. Copyright © 2026 PickAIModel.com. All rights reserved. NeuroSync AI Dynamics Pty Ltd.