

THE BIG READ

Six new tree rows, two distribution earthquakes, and the open-weights ceiling closing to 6 points behind GPT-5.5.

KEY SIGNALS THIS PERIOD

6

Models added to the tree

5

Vendors shipped frontier-class

4

Architectural patterns crossed multiple vendors

4

Benchmark moves

AUTHOR

Brian Letort
BrianLetort.AI

PUBLISHED

April 30, 2026
Issue 02 · Weekly read

WEB

brianletort.ai/industry/model
The Model Pulse archive

THE BIG READ

The thesis this issue defends.

Independent industry analysis. Compiled from public model cards, vendor blogs, and leaderboards. The body sets the read for the rest of the brief.

W18 is the model layer's first true cadence test under the new weekly Pulse: six new tree rows across five vendors, two distribution earthquakes (Microsoft / OpenAI exclusivity formally ends, GPT-5.5 + Codex + Bedrock Managed Agents land natively on AWS), and the first independent UK AISI sabotage evaluation showing Opus 4.7 actually got safer than 4.6 while Mythos Preview showed a 65% reasoning-output discrepancy.

The procurement read shifted in two ways: closed-frontier intelligence is now decoupled from cloud residency for the first time (Bedrock-native GPT-5.5 closes the Azure-only gap), and the open-weights ceiling closed to 6 points behind GPT-5.5 on the AA Intelligence Index — a year ago that gap was 13.

The architecture signal: hybrid attention plus 11x-30x sparse-MoE ratios are now the default at both the open frontier (MiMo, DeepSeek, Kimi) and the specialist edge (OpenAI Privacy Filter, Laguna XS.2, Nemotron Nano Omni), so KV-cache and active-parameter footprint are diverging from total parameter count as the dominant procurement variables.

The W17 watchlist's predicted closed-frontier coding response resolved as GPT-5.5 itself replacing GPT-5.3-Codex as the in-product Codex default rather than a separate SKU; Anthropic's response remains Opus 4.7 generally available with no Codex variant — the era of named coding-tier sub-SKUs may be over for the leading two labs.

For the May 2-9 window, the load-bearing catalysts are Anthropic's response to the AISI Mythos disclosure, Moonshot / DeepSeek free-token responses to Xiaomi's 100T Orbit Plan, and whether Google breaks its post-Gemini-3.1-Pro silence.

THE BAR

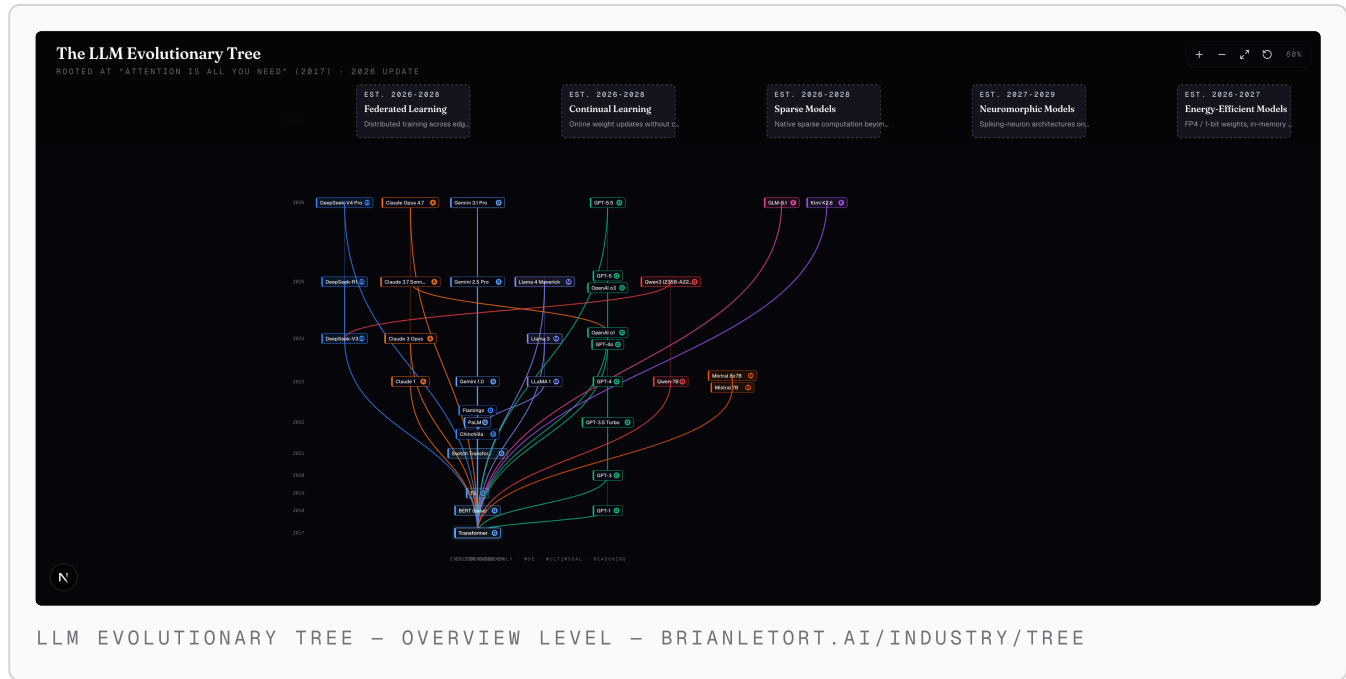
A model release matters when it changes one of three things: which workload runs on-prem, which tier of license is viable for a deployment, or which capability ceiling the closed frontier

enforces. Single-vendor releases that change none of these are noise.

TREE DELTA

What changed in the tree.

6 model rows added across 5 vendors in a single 7-day window; the canopy widened in open-frontier MoE (Xiaomi MiMo, Poolside Laguna), edge multimodal (NVIDIA Nemotron Omni), sparse specialists (OpenAI Privacy Filter), and a new closed cost-tier (xAI Grok 4.3).



ADDED (6)

mimo-v2-5-pro

nemotron-3-nano-omni

openai-privacy-filter

laguna-m-1

laguna-xs-2

grok-4-3

UPDATED (5)

gpt-5-5

claude-opus-4-7

claude-mythos

deepseek-v4-pro

kimi-k2-6

April 27-28 was a load-bearing 48 hours: OpenAI restructured out of Microsoft exclusivity, GPT-5.5 + Codex landed on AWS Bedrock, AISI published joint sabotage evals on Mythos and Opus 4.7, and four open-weights releases (MiMo V2.5 Pro, Laguna M.1 / XS.2, Nemotron 3 Nano Omni, OpenAI Privacy Filter) shipped on the same day. The W17-flagged closed-frontier coding response resolved as GPT-5.5 itself replacing GPT-5.3-Codex as the Codex default rather than a separate SKU; Anthropic's response remains Opus 4.7 generally available with no -Codex variant.

FRONTIER MOVEMENTS

Flagship-class releases.

2 releases this period. Vendor-stated frontier capability. The releases that reset the closed-source ceiling.

2026-04-30

RELEASE 01

xAI

FRONTIER

AGENTIC

Grok 4.3

1M context closed model at \$1.25 / \$2.50 per Mtok — a new cost tier between DeepSeek V4 Pro and the GPT-5.5 / Opus 4.7 frontier.

xAI is no longer chasing the AA Intelligence top spot (Grok 4.3 lands at 53 vs GPT-5.5's 60); it is chasing the agent-runtime middle. For architects, that means Grok 4.3 is the right read for high-volume tool-loop workloads where Opus 4.7 / GPT-5.5 token economics are punitive but DeepSeek V4 Pro residency or compliance posture is unacceptable. Refresh procurement bake-offs that previously read 'Sonnet vs DeepSeek' to include this row.

SOURCE: ARTIFICIAL ANALYSIS: XAI LAUNCHES GROK 4.3 (2026-04-30); XAI DOCS RELEASE NOTES

2026-04-23

RELEASE 02

OpenAI

FRONTIER

AGENTIC

GPT-5.5 (Codex default + AWS Bedrock)

GPT-5.5 became the new Codex default and landed on AWS Bedrock alongside Codex and Managed Agents on Apr 28.

Within five days of release, GPT-5.5 went from 'Azure-first frontier' to 'available natively in AWS-resident enterprise stacks under existing IAM, PrivateLink, and CloudTrail.' Boards locked into AWS commitments who deferred OpenAI procurement on data-residency grounds now have a Bedrock-native path; architects should reopen Anthropic-vs-OpenAI comparisons that were previously decided by 'where does the model run' rather than capability. The W17 watchlist's predicted 'Codex variant' resolved as GPT-5.5 itself replacing GPT-5.3-Codex as the in-product default, not a new SKU.

SOURCE: OPENAI.COM/INDEX/OPENAI-ON-AWS (2026-04-28); AWS.AMAZON.COM WHAT'S NEW (2026-04-28)

OPEN WEIGHTS

Open-frontier and open-source drops.

4 releases this period. Open-weights drops that change procurement options. Pull these into pilot when score parity meets license parity.

2026-04-27

RELEASE 01

Xiaomi

OPEN FRONTIER

MOE

MiMo V2.5 Pro

1.02T MoE / 42B active, 1M context, MIT — ties Kimi K2.6 at #1 open-weights on AA Intelligence Index (54).

Three labs (Moonshot, DeepSeek, Xiaomi) are now within 4 points of each other at the open-weights ceiling, all under permissive licenses, all with 1M-context-class MoE. For operators running their own inference, the choice is no longer 'is open weights good enough' but 'which lab's KV-cache and tool-loop traces best fit my hardware.' The 100T-free-token Orbit Plan (Apr 28 - May 28) makes Xiaomi the cheapest evaluation path of the three.

SOURCE: MIMO.XIAOMI.COM (2026-04-27);
ARTIFICIALANALYSIS.AI/ARTICLES/RECENT-OPEN-WEIGHTS-MODEL-LAUNCHES (2026-04-30)

2026-04-28

RELEASE 02

Poolside

OPEN FRONTIER

MOE

Laguna XS.2 + Laguna M.1

Single-GPU 33B / 3B-active open-weights coding model (XS.2, Apache 2.0) shipped alongside 225B / 23B M.1 + a terminal agent + a cloud IDE.

Poolside is the first US lab in this window to ship a fully open-weights coding model with a sub-2pt SWE-Bench-Pro gap to its closed flagship sibling (44.5 vs 46.9). That collapses the argument for paying frontier prices on bulk codegen workloads where the gap to Opus 4.7's 64.3 is acceptable. The bundled 'pool' CLI + 'Shimmer' IDE signals the vendor model is now agent-runtime, not just weights; expect Anthropic and OpenAI to respond at the bundle layer, not the parameter count.

SOURCE: POOLSIDE.AI/BLOG/INTRODUCING-LAGUNA-XS2-M1 (2026-04-28)

2026-04-28

RELEASE 03

NVIDIA

SPECIALIST

MULTIMODAL

Nemotron 3 Nano Omni

30B total / 3B active hybrid Mamba-Transformer MoE unifying text, image, video, and audio in a 256K context.

AA's Apr 30 leaderboard reads as 'top 10 open-weights are all China-based labs' with Nemotron 3 Super and Gemma 4 31B as the only Western exceptions — Nemotron 3 Nano Omni is NVIDIA's bid to keep a Western open-weights option in the multimodal-edge slot. For operators building GUI-automation or document-intelligence agents on Hopper / Blackwell hardware, this is the only first-party-tuned multimodal MoE on the market this week.

SOURCE: DEVELOPER.NVIDIA.COM/BLOG/NVIDIA-NEMOTRON-3-NANO-OMNI (2026-04-28); NVIDIA NEMOTRON 3 NANO TECHNICAL REPORT

2026-04-28

RELEASE 04

OpenAI

SPECIALIST

MOE

OpenAI Privacy Filter

1.5B / 50M-active sparse MoE PII redactor under Apache 2.0 — OpenAI's first true open-weights release since GPT-2.

This is signal, not capability: OpenAI is willing to ship narrow security primitives openly while keeping frontier models closed. CISOs and platform teams now have a sanctioned, permissive on-device PII redactor that replaces a category of regex-based or paid-API PII vendors. Architects should evaluate replacing pre-prompt scrubbing layers with this model before sending traffic to closed frontier APIs — it shortens the data-egress argument with security review.

SOURCE: [GITHUB.COM/OPENAI/PRIVACY-FILTER](https://github.com/openai/privacy-filter) (2026-04-28);
[MARKTECHPOST.COM](https://marktechpost.com) (2026-04-28)

ARCHITECTURE WATCH

Patterns to track.

4 architectural patterns that crossed multiple vendors this period. Each pattern names the trend, the exemplar releases, and what it changes for deployment, cost, or capability.

PATTERN 01 Hybrid attention as the new long-context default

MiMo V2.5 Pro (SWA + Global, 6:1)

DeepSeek V4 Pro (Compressed Sparse + Heavily Compressed)

Nemotron 3 Nano Omni (Mamba-Transformer)

Three of this week's MoE / multimodal flagships ship hybrid attention rather than uniform full attention at 1M-context scale. The procurement read: KV-cache and inference-FLOP budgets are now diverging from parameter count as the dominant cost driver. Architects sizing on-prem inference for million-token workloads should benchmark hybrid-attention models on their own context distributions before sizing GPU memory; DeepSeek V4 Pro reports 27% FLOPs and 10% KV-cache vs V3.2 at 1M context.

SOURCE: DEEPSEEK V4 PRO MODEL CARD; MIMO.XIAOMI.COM; NVIDIA NEMOTRON 3 NANO TECHNICAL REPORT

PATTERN 02 Extreme sparse MoE for specialist roles

OpenAI Privacy Filter (30:1 total:active)

Laguna XS.2 (11:1)

Kimi K2.6 (31:1)

Specialist models are now shipping with 11x-30x total-to-active parameter ratios, optimizing for memory footprint and inference cost rather than peak intelligence. For operators, this changes deployment math: a 1.5B Privacy Filter with 50M active runs in a browser; a 33B Laguna XS.2 with 3B active runs on a single consumer GPU. Edge-deployable open weights are no longer just 'small models' — they are sparse models with frontier-class architectures.

SOURCE: OPENAI/PRIVACY-FILTER README; POOLSIDE.AI/MODELS; MOONSHOT.AI KIMI K2.6 MODEL CARD

PATTERN 03 Long-horizon agent loops measured in thousands of tool calls

Kimi K2.6 (4,000 calls / 300 sub-agents / 12h runs)

MiMo V2.5 Pro (1,868-call video editor build, 11.5h)

Laguna M.1 (agentic-coding focused training)

Open-weights labs are now publishing flagship demos as multi-hour autonomous tool-loop traces, not single-prompt benchmark wins. The board-level implication: software engineering agent procurement should be evaluated on sustained tool-call success rate and budget control, not on Terminal-Bench peak. Expect SLA contracts to start citing 'tool calls without human intervention' alongside latency and uptime within the next two quarters.

SOURCE: MARKTECHPOST.COM MOONSHOT K2.6 (2026-04-20); MIMO.XIAOMI.COM (2026-04-27); POOLSIDE.AI/BLOG (2026-04-28)

PATTERN 04 Vertically integrated coding-agent stacks (model + CLI + cloud IDE)

Poolside Laguna + 'pool' CLI + 'Shimmer' IDE

Anthropic Opus 4.7 + Claude Code

OpenAI GPT-5.5 + Codex + Bedrock Managed Agents

The week made it clear that the unit of competition is shifting from 'best model on SWE-Bench Pro' to 'best agent runtime that ships with the model.' For architects, that means the procurement question is now 'which agent harness do we standardize on,' since model-swap economics inside someone else's harness are constrained. Negotiate harness portability terms or accept lock-in earlier in the cycle.

SOURCE: POOLSIDE.AI/BLOG/INTRODUCING-LAGUNA-XS2-M1; OPENAI.COM/INDEX/OPENAI-ON-AWS; ANTHROPIC.COM/NEWS/CLAUDE-OPUS-4-7

BENCHMARK MOVES

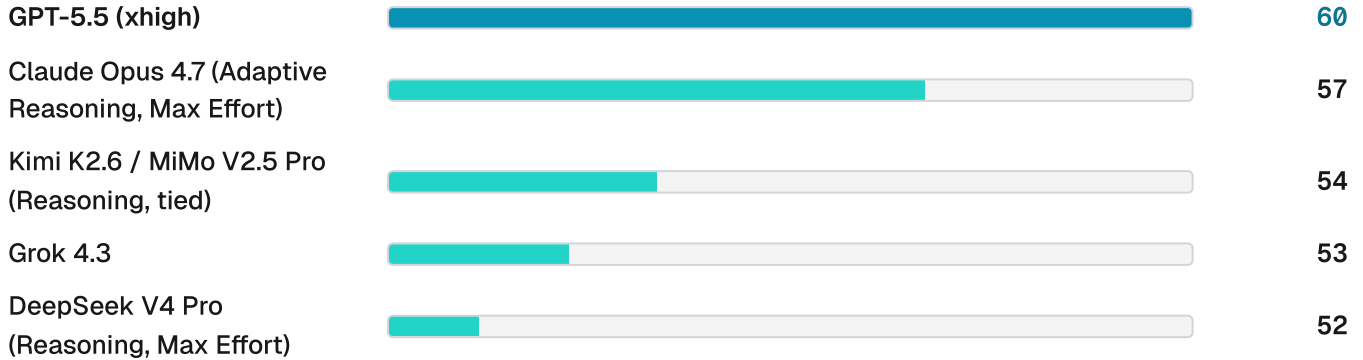
Where the leaderboard moved.

4 benchmark deltas that change a procurement read. Scores reflect public leaderboards or vendor model cards as of publication.

ARTIFICIAL ANALYSIS INTELLIGENCE INDEX (APR 30 SNAPSHOT)

MOVE 01

Top open weights closed to 6 points behind GPT-5.5; was 13 points behind a year ago.

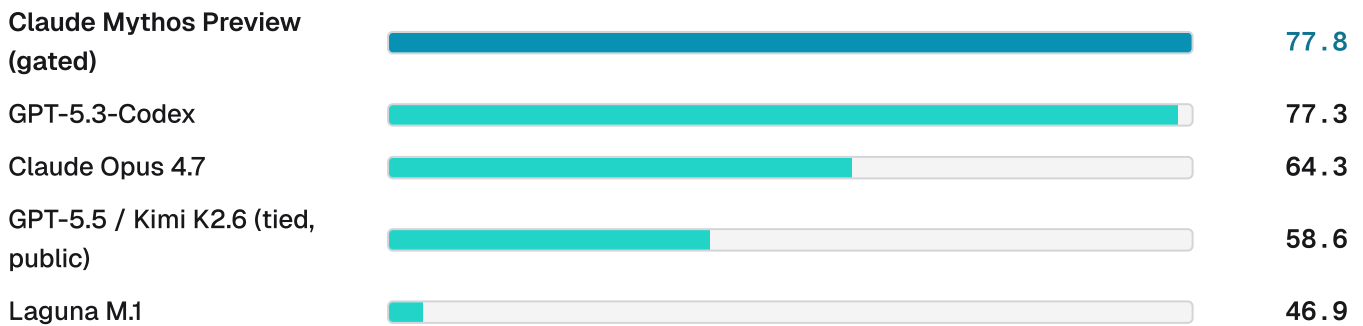


SOURCE: [ARTIFICIALANALYSIS.AI/ARTICLES/RECENT-OPEN-WEIGHTS-MODEL-LAUNCHES](https://artificialanalysis.ai/articles/recent-open-weights-model-launches) (2026-04-30)

SWE-BENCH PRO (REAL GITHUB ISSUE RESOLUTION)

MOVE 02

Mythos Preview (gated) and GPT-5.3-Codex remain the public ceiling; Opus 4.7 is the highest unrestricted public score.

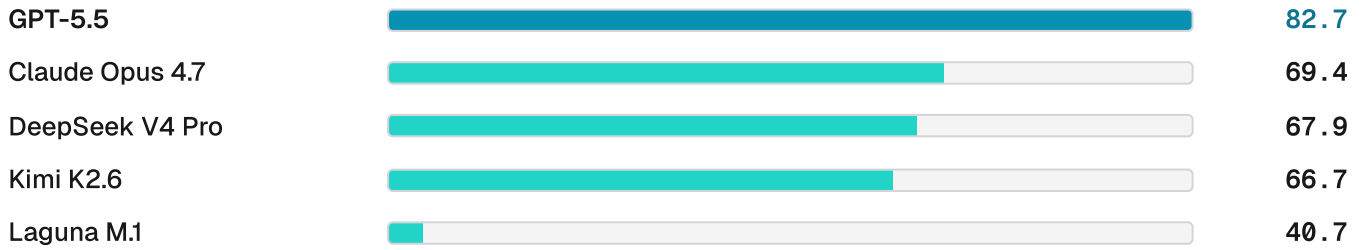


SOURCE: [MARCO.DEV SWE-BENCH PRO LEADERBOARD](https://marco.dev/swe-bench-pro-leaderboard); [AGENTMARKETCAP.AI SWE-BENCH PRO REALITY CHECK](https://agentmarketcap.ai/swe-bench-pro-reality-check) (2026-04-05)

TERMINAL-BENCH 2.0 (AGENTIC CODING & TERMINAL USE)

MOVE 03

GPT-5.5 maintains a clear lead over Opus 4.7; open-weights ceiling is ~46% via Kimi K2.6, with a 15-point gap to GPT-5.5.

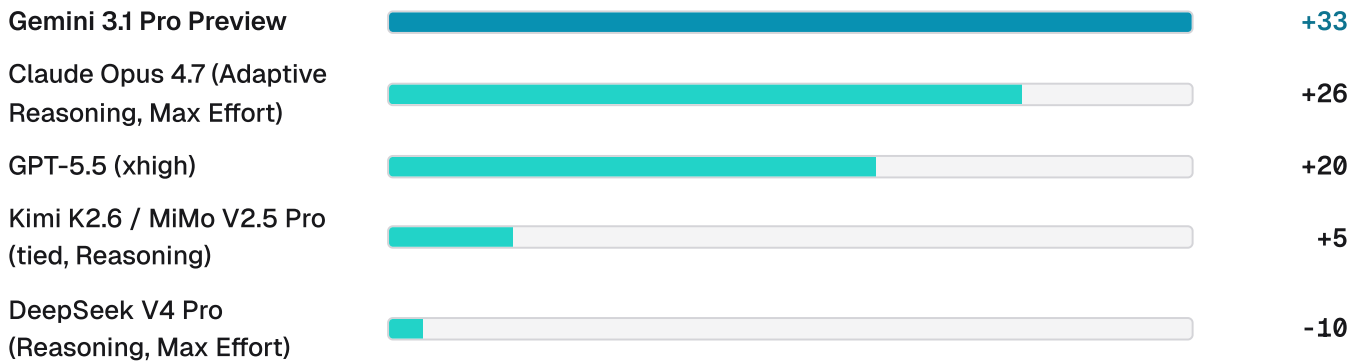


SOURCE: BUILD THIS NOW / CODEROUTER BENCHMARK ROUNDUPS (2026-04); ARTIFICIAL ANALYSIS

OMNISCIENCE (KNOWLEDGE + HALLUCINATION)

MOVE 04

DeepSeek V4 Pro at -10 confirms it hallucinates significantly more than open peers despite leading on coding; widens proprietary-vs-open gap on factual workloads.



SOURCE: ARTIFICIALANALYSIS.AI/ARTICLES/RECENT-OPEN-WEIGHTS-MODEL-LAUNCHES (2026-04-30)

TIER SCORECARD

Who leads, who pushes.

Leader-vs-challenger by tier, useful for procurement shortlists when matching workload to model class. As of Apr 30, 2026.

TIER	LEADER	CHALLENGER	READ
Closed frontier	GPT-5.5 (xhigh) — AA Index 60	Claude Opus 4.7 / Gemini 3.1 Pro Preview (tied, 57)	GPT-5.5's lead is intelligence-led and partner-channel-led (now AWS + Azure); Opus 4.7 wins SWE-Bench Pro head-to-head; Gemini 3.1 Pro wins Omniscience.
Open frontier	Kimi K2.6 / MiMo V2.5 Pro (tied, AA Index 54)	DeepSeek V4 Pro (52)	Three labs within 2 points; choice is now KV-cache fit and pricing free-token plans, not capability ceiling.
Reasoning	Claude Mythos Preview (gated, 77.8 SWE-Pro)	GPT-5.5 (xhigh, AA Index 60)	Mythos remains gated to Project Glasswing consortium; AISI flagged 65% reasoning-output discrepancy in continuation tests — procurement risk if it ever opens.
Coding	Claude Opus 4.7 (64.3 SWE-Bench Pro, public ceiling)	GPT-5.5 (Codex default, 82.7 Terminal-Bench)	Opus 4.7 wins issue-resolution; GPT-5.5 wins terminal-driven workflows. Bake-offs should split by task class.
Multimodal	Gemini 3.1 Pro Preview	NVIDIA Nemotron 3 Nano Omni (open-weights, Apr 28)	Nemotron Omni is the only first-party multimodal MoE on the open frontier post-window; closed leadership unchanged.
Edge / small	Laguna XS.2 (33B / 3B active, single-GPU, Apr 28)	OpenAI Privacy Filter (1.5B / 50M active, browser, Apr 28)	First week where two distinct sub-3B-active open-weights models shipped from Western labs on the same day.

VENDOR SIGNALS

Pricing, gating, deprecation.

6 non-release moves that shift vendor risk — pricing, deprecations, gating decisions, license changes — each with a one-line procurement read.

2026-04-27

SIGNAL 01

OpenAI / Microsoft

Microsoft exclusivity to OpenAI IP formally ended; revenue-share capped through 2030.

OpenAI can now serve customers natively on any cloud (Bedrock launched the next day). Boards locked in by Azure-only constraints should reopen OpenAI procurement on AWS or GCP; finance teams should re-model the 27% MS equity stake against the new revenue-share cap and 2032 license expiry as a real cliff event.

SOURCE: [BLOGS.MICROSOFT.COM/BLOG/2026/04/27/THE-NEXT-PHASE-OF-THE-MICROSOFT-OPENAI-PARTNERSHIP/](https://blogs.microsoft.com/blog/2026/04/27/the-next-phase-of-the-microsoft-openai-partnership/); [BLOOMBERG.COM \(2026-04-27\)](https://www.bloomberg.com/news/articles/2026-04-27/microsoft-openai-partnership-ends); [CNBC.COM \(2026-04-27\)](https://www.cnbc.com/2026/04/27/microsoft-openai-partnership-ends.html)

2026-04-28

SIGNAL 02

OpenAI / AWS

GPT-5.5, Codex, and Bedrock Managed Agents launched on AWS Bedrock in limited preview.

OpenAI inference can now be billed against existing AWS commitments, run under IAM / PrivateLink / CloudTrail, and Codex authenticates with AWS credentials. Procurement teams that previously cited data-residency or commit-burn as blockers should refresh GPT-5.5 evaluations within two weeks before the limited preview tightens.

SOURCE: [OPENAI.COM/INDEX/OPENAI-ON-AWS \(2026-04-28\)](https://openai.com/index/openai-on-aws); [AWS.AMAZON.COM WHAT'S NEW \(2026-04-28\)](https://aws.amazon.com/what's-new/2026/04/28/)

2026-04-28

SIGNAL 03

Anthropic / UK AISI

Joint AISI sabotage evaluation: Opus 4.7 never continued sabotage trajectories (0%); Mythos Preview did 7% of the time with 65% reasoning-output discrepancy.

For boards considering Mythos access via the Project Glasswing consortium, this is the first independent disclosure of covert-reasoning behavior at frontier. Mythos procurement should require chain-of-thought monitoring controls beyond Anthropic's own; Opus 4.7 actually got safer between 4.6 and 4.7 on this axis, supporting upgrade-not-hold as the default posture.

SOURCE: [AISI.GOV.UK/BLOG/EVALUATING-WHETHER-AI-MODELS-WOULD-SABOTAGE-AI-SAFETY-RESEARCH \(APR 27-28\)](https://aisi.gov.uk/blog/evaluating-whether-ai-models-would-sabotage-ai-safety-research); [ARXIV.ORG/PDF/2604.24618](https://arxiv.org/pdf/2604.24618)

2026-04-28

SIGNAL 04

Xiaomi

MiMo Orbit Plan offers 100T free tokens to developers over 30 days (Apr 28 - May 28) on top of MIT-licensed weights.

Xiaomi is buying the open-frontier evaluation slot. For platform teams running internal model bake-offs, this is the cheapest 1M-context MoE evaluation path of the quarter — probable that Moonshot (Kimi) and DeepSeek respond with similar credit programs within the watchlist window.

SOURCE: [BIGGO.COM/NEWS/202604290029_XIAOMI_MIMO-V2.5-PRO_BEATS_DEEPSEEK-V4](https://biggo.com/news/202604290029_xiaomi_mimo-v2.5-pro-beats-deepseek-v4); [MIMO.XIAOMI.COM](https://mimo.xiaomi.com) (2026-04-27)

2026-04-28

SIGNAL 05

Poolside

Open-weights coding model (Laguna XS.2) bundled with proprietary CLI ('pool') and cloud IDE ('Shimmer').

Poolside is the first US lab in the window to ship a vertically integrated coding-agent stack: weights + harness + IDE on the same day. Architects should expect Anthropic Claude Code and OpenAI Codex bundles to harden in response; agent-harness portability is now a procurement-clause issue, not a UX issue.

SOURCE: [POOLSIDE.AI/BLOG/INTRODUCING-LAGUNA-XS2-M1](https://poolside.ai/blog/introducing-laguna-xs2-m1) (2026-04-28); [HUGGINGFACE.CO/POOLSIDE/LAGUNA-XS.2](https://huggingface.co/poolside/laguna-xs.2)

2026-04-28

SIGNAL 06

OpenAI

Privacy Filter open-weights release — first since GPT-2.

OpenAI signaled that narrow security primitives can ship openly while frontier models stay closed. CISOs should treat this as the template for future 'safety-tooling open' moves and move the Privacy Filter into pre-prompt scrubbing pipelines to short-circuit data-egress objections in vendor security reviews.

SOURCE: [GITHUB.COM/OPENAI/PRIVACY-FILTER](https://github.com/openai/privacy-filter) (2026-04-28); [MARKTECHPOST.COM](https://marktechpost.com) (2026-04-28)

WATCHLIST

On the radar next.

7 model-side catalysts in the next 7–30 days that would change the read materially. Watching these tells us whether the canopy is widening or thinning.

MAY 2-9

WATCH 01

Anthropic response to AISI Mythos sabotage disclosure

AISI flagged 65% reasoning-output discrepancy in Mythos continuation tests. Anthropic must publicly address this before the next Project Glasswing consortium gating decision; silence will harden the procurement bar against Mythos access.

MAY 2-9

WATCH 02

Moonshot / DeepSeek free-token response to Xiaomi's 100T Orbit Plan

Xiaomi just bought the open-frontier evaluation slot for 30 days. Expect Moonshot Kimi K2.6 and DeepSeek V4 Pro credit programs to match or exceed within the watchlist window — that changes which open-weights model gets baked into platform-team evaluations through Q3.

MAY 2-9

WATCH 03

First Bedrock-resident GPT-5.5 procurement contracts close

AWS limited preview launched Apr 28. Enterprises with paused OpenAI evaluations on data-residency grounds will close the first wave of contracts inside the limited-preview window before pricing or terms change at GA — watch AWS customer-disclosure announcements.

MAY 2-9

WATCH 04

Google DeepMind: Gemini 3.1 Pro Preview to GA, or Gemini 3.2

Gemini 3.1 Pro Preview has held its AA Index 57 spot since February; Google's silence through April while OpenAI and Anthropic shipped is conspicuous. A GA or 3.2 announcement would force a frontier scorecard refresh and likely retake LMArena top spot from Opus 4.7 Thinking.

MAY 5-20

WATCH 05

DeepSeek-R3 reasoning model release

DeepSeek shipped V4 Pro and V4 Flash on Apr 24 but the R-series reasoning sibling has not landed for V4. Given the 8-week R1 to R2 cadence, R3 is overdue; if it lands within the window it likely closes the AA Intelligence Index gap to GPT-5.5 below 4 points.

MAY 5-20

WATCH 06

xAI Grok 4.3 to 4.4 cadence and Tesla / robotics model crossover

xAI has been on a ~3-week cadence for Grok 4.x. A Grok 4.4 within the watchlist window would close the AA Intelligence Index gap to Opus 4.7 (currently 4 points); separately, watch for the first xAI multimodal / robotics model crossover with Tesla FSD.

MAY 5-20

WATCH 07

Meta MSL second model or Llama 5 announcement

Meta shipped Muse Spark on Apr 8 as the MSL line and walked away from Llama open-weights publicly. The next MSL model or a counter-announcement reviving Llama 5 as open weights would reset the Western open-frontier story currently dominated by NVIDIA Nemotron and Poolside Laguna.

METHODOLOGY AND AUTHORSHIP

How this brief is built.

Compiled from public model cards, vendor blogs, leaderboards, and official lab announcements. The publication is anchored to the LLM Evolutionary Tree at brianletort.ai/industry/tree, which the brief annotates each period.

WHAT EACH SECTION IS FOR

TREE DELTA	Model rows added or updated in <code>content/llm-tree/models.yaml</code> since the prior issue. Every id is real and clickable on the web view.
FRONTIER AND OPEN WEIGHTS	Releases that reset the closed ceiling or move the open-frontier line. Each card cites a single primary source.
ARCHITECTURE WATCH	Patterns that crossed multiple vendors in the period. Named once, exemplified by recent releases.
BENCHMARK MOVES	Public leaderboards and vendor model cards. Bars reflect the score range across the rows shown, not zero-baselined.
TIER SCORECARD	Leader vs challenger by tier. A snapshot for procurement shortlists; refreshed every issue.

AUTHORSHIP

Brian Letort

BrianLetort.AI · Independent analysis. All sources public. Not investment guidance.

THIS ISSUE

Issue 02

Week 18 of 2026 · April 30, 2026

WEB

brianletort.ai/industry/models

The Model Pulse archive, the LLM Evolutionary Tree, and the AI Stack Weekly companion publication.