Noticiário do Joãozinho

Edição da noite

8 MATÉRIAS · 12 SEÇÕES · 2 VÍDEOS VERSÃO INTERATIVA NA PÁGINA

MANCHETE

Moonshot libera os pesos open-weight do Kimi K3.1 Pro: MoE multimodal de ~2,8T (104B ativos) e contexto de 1M tokens

Em 04/11/2026 (~16:00 UTC / 13:00 BRT), a Moonshot AI (conta @Kimi_Moonshot) anunciou a liberação dos pesos open-weight do Kimi K3.1 Pro no Hugging Face (moonshotai/Kimi-K3.1-Pro). O modelo é MoE multimodal agentic: cerca de 2,8 trilhões de parâmetros totais e 104 bilhões ativos, com janela de 1 milhão de tokens. No Terminal-Bench 4.0 marca 65,7% (Claude Opus 5.5 66,4%; GPT-6 Astra 58,2%); no Kimi Code Bench 3.0 Long-horizon, 61,2% contra 55,95% do K3. No xdev-bench-hard de cyber, chega a 81,0% após filtro de dados (90,5% sem o filtro). A licença customizada "Kimi K3.1 Pro" está no repositório. O blog oficial já prometia os pesos até 4/11; o post no X da liberação somava cerca de 259 mil views e 5.184 likes na medição do vigia.

Por que isso importa: É o open-weight de fronteira do dia: pesos públicos, multimodal nativo e contexto de 1M — o marco que o vigia já tinha alertado às 13:34 e que agora fecha a edição.

Post oficial: pesos do Kimi K3.1 Pro no Hugging Face — X / @Kimi_Moonshot, 04/11/2026 https://x.com/Kimi_Moonshot/status/2118010528123773297
moonshotai/Kimi-K3.1-Pro (pesos) — Hugging Face, 04/11/2026 https://huggingface.co/moonshotai/Kimi-K3.1-Pro
Introducing Kimi K3.1 Pro (blog; prometia pesos até 4/11) — Kimi.ai, 04/11/2026 https://www.kimi.ai/blog/kimi-k3-1-pro

NESTA EDIÇÃO

Modelos	1	Regulação	0	Finanças	1	Vídeos	2
Pesquisa	1	Open Source	0	Games	0		
Produtos	0	Agentes e Código	0	Brasil	0		
Negócios e Investimento	1	Modelos Locais	0	Fofocas das redes	4		

Modelos

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA ALTA 04/11/2026 · Tencent

Tencent lança Hy4 com visão e computer-use; pesos e system card “nos próximos dias”

A Tencent anunciou o Hy4 (04/11): pela primeira vez com compreensão visual nativa e computer-use (CLI e GUI), com treino especial em Blender e CAD. No WorkBuddy, score médio 83,18 e pass rate 96,15%. Em avaliação cega com 213 profissionais de finanças (579 tarefas), Hy4 fez 3,04 — à frente de DeepSeek-V4.1-Flash-1017 (2,97), Kimi K3.1 Pro (2,68) e GLM-5.3 (2,68). A taxa de alucinação caiu de 12,95% no preview para 7,2%. Pesos e system card vêm “nos próximos dias”; o foco de open-source responsável cita cybersecurity defensiva e biosafety. Segundo o TechNode, o modelo tem 770 bilhões de parâmetros; API a US$ 0,834 / US$ 2,501 por milhão de tokens (input/output); WorkBuddy e CodeBuddy grátis de 4 a 15/11. Não há, nas fontes desta edição, número oficial de parâmetros ativos do Hy4 final.

> **Por que isso importa:** Mais um open-source chinês de ponta no mesmo dia do Kimi — agora com visão e computer-use de fábrica, e pesos ainda por sair.

Fonte original: Tencent https://www.tencent.com/tencent-launches-hy4-joining-the-ranks-of-the-worlds-leading-open-source-ai-models/
TechNode (770B; preços de API; free 4–15/11) https://technode.com/2026/11/04/tencent-launches-hy4-with-vision-and-computer-use-capabilities/

Pesquisa

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA ALTA 04/11/2026 · Anthropic

Anthropic: Claude ajuda a mapear a 1ª estrutura 3D de vault particle em vida bacteriana/fagos

O lab de biologia da Anthropic publicou (04/11) a primeira estrutura 3D de um vault particle encontrado em células procarióticas — no caso, montado a partir da proteína L05 de bacteriófago: 96 cópias, cerca de 9,5 MDa, ~54×45 nm, com cerca de 25% mais volume que vaults de mamíferos. A purificação otimizada levou cerca de 90 minutos (contra ~3 dias no método convencional); o primeiro modelo estrutural saiu em ~34 horas via cryo-EM. Claude ajudou em busca em bases, constructs de DNA, isolamento e processamento de microscopia. A função biológica dos vaults continua desconhecida; um preprint foi submetido ao bioRxiv (link a atualizar quando sair).

> **Por que isso importa:** IA de laboratório saindo do paper genérico: estrutura inédita em biologia fundamental, com Claude embutido em cada etapa.

Fonte original: Anthropic https://www.anthropic.com/research/investigating-vault-particles-with-claude
Post oficial no X (@AnthropicAI) https://x.com/AnthropicAI/status/2118073665296228863

Produtos

0 MATÉRIAS NESTA EDIÇÃO

Sem novidades relevantes nesta edição.

Negócios e Investimento

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA ALTA 04/11/2026 · FinancialMarkets.com

Meta captura US$ 15 bi em bonds em 5 tranches para financiar expansão de IA

A Meta vendeu US$ 15 bilhões em bonds em cinco partes (04/11), com yields de 5,85% a 7,53%; o book chegou a cerca de US$ 58 bilhões (~3,9× a oferta). O título de 10 anos saiu a 6,55%; o mais longo (2056) a +2,2 pontos percentuais sobre Treasuries — cerca de 0,9 pp mais largo que um comparável de abril. A S&P cortou o outlook da Meta, citando o peso do gasto em IA. Na mesma janela, Oracle captou US$ 10 bi e a Alphabet € 6 bi; a Meta prepara seu primeiro euro bond.

> **Por que isso importa:** Demanda ainda profunda, mas o preço do dinheiro da IA subiu — e o rating já piscou o alerta.

Fonte original: FinancialMarkets.com https://www.financialmarkets.com/article/meta-paid-up-to-753-to-borrow-15-billion-its-order-book-was-about-as-deep-as-in

Regulação

0 MATÉRIAS NESTA EDIÇÃO

Sem novidades relevantes nesta edição.

Open Source

0 MATÉRIAS NESTA EDIÇÃO

Sem novidades relevantes nesta edição.

Agentes e Código

0 MATÉRIAS NESTA EDIÇÃO

Sem novidades relevantes nesta edição.

Modelos Locais

0 MATÉRIAS NESTA EDIÇÃO

Sem novidades relevantes nesta edição.

Finanças

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA ALTA 03/11/2026 · The Register

Revolut corta ~90% o custo de inferência de coding com GLM-5.3 self-hosted em 56×B200

No AI Engineer NY, David Abitbol (ops/infra da Revolut) descreveu o pivot: ~2.000 engenheiros, coding time +60% e lead time +70% após o rollout de coding AI; o default passou a ser GLM self-hosted (de 5.1 para 5.3) em 56 GPUs Nvidia B200. Custo ~US$ 0,30 / milhão de tokens — cerca de 1/10 de Codex/Claude Code — com corte de ~90% na conta de inferência de coding. Stack: NVFP4, prefill/decode desagregado; TTFT sub-3s em ~256k de contexto (Abitbol: "long context is a scam"); ~200 tok/s. Um router manda visão para Gemini/GPT Luna; fronteira comercial fica como exceção.

> **Por que isso importa:** Fintech grande mostrando na prática que open-weight self-hosted vira o default econômico do coding AI — e o frontier modelo vira exceção.

Fonte original: The Register https://www.theregister.com/ai-and-ml/2026/11/03/british-fintech-revolut-slashes-ai-coding-costs-10x-by-cutting-out-claude-and-going-self-hosted/5307345

Games

0 MATÉRIAS NESTA EDIÇÃO

Sem novidades relevantes nesta edição.

Brasil

0 MATÉRIAS NESTA EDIÇÃO

Sem novidades relevantes nesta edição.

Fofocas das redes

4 MATÉRIAS NESTA EDIÇÃO

X IMPORTÂNCIA ALTA 04/11/2026 · @nebiustf

Nebius Token Factory coloca Kimi K3.1 Pro em day-0 no mesmo dia da liberação dos pesos

A conta @nebiustf (Nebius Token Factory) anunciou suporte day-0 ao Kimi K3.1 Pro no mesmo dia em que a Moonshot liberou os pesos open-weight — reforçando a corrida das plataformas de inferência para oferecer o modelo assim que sai.

> **Por que isso importa:** Day-0 em cloud de inferência acelera quem quer Kimi sem baixar os 2,8T de pesos.

Engajamento: 23,1 mil visualizações · 173 curtidas · 12 reposts · 11 citações · 12 respostas · 12 salvamentos (medido em 04/11/2026 18:10)
Post original (X): @nebiustf https://x.com/nebiustf/status/2118023422248669547

X IMPORTÂNCIA MÉDIA 04/11/2026 · @guilleflorvs

Leitura da licença Kimi K3.1 Pro: free até US$20M SaaS / 100M MAU ou US$20M/mês; depois deal ou crédito de marca

O post de @guilleflorvs resume a licença custom “Kimi K3.1 Pro”: uso free até US$20 milhões de receita SaaS / 100 milhões de MAU ou US$20 milhões/mês de receita; acima disso, exige deal ou crédito de marca. Waivers via Moonshot ou parceiros — o post cita Nebius e Vercel AI Gateway com day-0.

> **Por que isso importa:** Licença open-weight com teto comercial: quem escala de verdade precisa negociar.

Engajamento: 664 visualizações · 4 curtidas · 3 respostas · 4 salvamentos (medido em 04/11/2026 18:10)
Post original (X): @guilleflorvs https://x.com/guilleflorvs/status/2118043607781216544

X IMPORTÂNCIA MÉDIA 04/11/2026 · X News (trending)

Trending no X: “weight-space programming” — modelos frontier editando pesos de redes menores

O agregador de trending do X destaca o buzz em torno de “weight-space programming”: a ideia de modelos frontier editarem pesos de redes menores (ex.: Qwen3.5-0.8B) sem retreino clássico. Tratar como eco de redes — sem afirmar percentuais do summary do X como fato de paper.

> **Por que isso importa:** Tema quente no feed de IA do dia; ainda buzz, não paper fechado nesta edição.

Post original (X): X News (trending) https://x.com/i/trending/2117960690896118225

X IMPORTÂNCIA MÉDIA 04/11/2026 · X News (trending)

Trending no X: miHoYo e planos de IA para NPCs/agentes em games + queda de vendas na China

Segundo o agregador de trending do X, circula o gancho de investimento da miHoYo em IA para NPCs/agentes em games, ao lado de relatos de queda de vendas dos hits na China. O hook do X menciona “15 Billion”; sem press release primário confirmado nesta edição, o número não entra como fato — só como eco do agregador.

> **Por que isso importa:** Games + agentes é pauta quente; aqui ainda é buzz de redes, não confirmação de PR.

Post original (X): X News (trending) https://x.com/i/trending/2117954842907865273

Vídeos

2 VÍDEOS

Os resumos se baseiam no título verificado e nos anúncios relacionados lidos para esta edição. Clique na miniatura ou no título para abrir no YouTube.



Tencent Hy4 IS HERE! INSANE Demos Tested... But My 300-Test Benchmark Tells a Different Story !

AI WITH Rithesh · 2026

Cobertura no dia do lançamento do Hy4: demos de visão/computer-use e um benchmark próprio de ~300 testes — útil para contrastar o marketing oficial com medição independente.

https://www.youtube.com/watch?v=JfQpx2VQFK8



Kimi K3.1 Pro 5x CHEAPER than Opus 5.5: a Spider-Man game for $35

Raidar · 2026

Teste prático de custo e código com Kimi K3.1 Pro versus Claude Opus 5.5 — no mesmo dia em que a Moonshot liberou os pesos open-weight.

https://www.youtube.com/watch?v=zJTrMVH9Wbw

SOBRE ESTA EDIÇÃO
Nº 10 — Edição da noite · 04/11/2026

JANELA
Cobre o que saiu desde a edição nº 9 (fechada por volta das 8h de Brasília de 04/11) até por volta das 18h de Brasília de 04/11, incluindo as pautas do vigia (Hy4 às 08:39 e pesos do Kimi K3.1 Pro às 13:34). Fechamento: 04/11/2026, por volta das 18h de Brasília.

FONTES VARRIDAS
Blogs e changelogs oficiais (OpenAI, Anthropic, Google DeepMind, Meta, xAI, Microsoft, NVIDIA, Apple, Mistral, DeepSeek, Qwen, Moonshot), arXiv e HF Papers, trending do Hugging Face e do GitHub, X, imprensa dos EUA, Europa, China, Japão, Índia e Brasil, e YouTube.

CRITÉRIO
Só entram notícias lidas numa fonte real, de preferência primária, com todos os links checados. Páginas com bloqueio (403) foram mantidas; links quebrados (404) foram descartados.

Fontes e links

▶ VÍDEOS SUGERIDOS

1. **Tencent Hy4 IS HERE! INSANE Demos Tested... But My 300-Test Benchmark Tells a Different Story !** — AI WITH Rithesh, 2026
   https://www.youtube.com/watch?v=JfQpx2VQFK8
2. **Kimi K3.1 Pro 5x CHEAPER than Opus 5.5: a Spider-Man game for $35** — Raidar, 2026
   https://www.youtube.com/watch?v=zJTrMVH9Wbw

Manchete

Post oficial: pesos do Kimi K3.1 Pro no Hugging Face — X / @Kimi_Moonshot
https://x.com/Kimi_Moonshot/status/2118010528123773297

moonshotai/Kimi-K3.1-Pro (pesos) — Hugging Face
https://huggingface.co/moonshotai/Kimi-K3.1-Pro

Introducing Kimi K3.1 Pro (blog; prometia pesos até 4/11) — Kimi.ai
https://www.kimi.ai/blog/kimi-k3-1-pro

Modelos

Tencent lança Hy4 com visão e computer-use; pesos e system card "nos próximos dias"
Tencent: https://www.tencent.com/tencent-launches-hy4-joining-the-ranks-of-the-worlds-leading-open-source-ai-models/
TechNode (770B; preços de API; free 4–15/11): https://technode.com/2026/11/04/tencent-launches-hy4-with-vision-and-computer-use-capabilities/

Pesquisa

Anthropic: Claude ajuda a mapear a 1ª estrutura 3D de vault particle em vida bacteriana/fagos
Anthropic: https://www.anthropic.com/research/investigating-vault-particles-with-claude
Post oficial no X (@AnthropicAI): https://x.com/AnthropicAI/status/2118073665296228863

Negócios e Investimento

Meta captura US$ 15 bi em bonds em 5 tranches para financiar expansão de IA
FinancialMarkets.com: https://www.financialmarkets.com/article/meta-paid-up-to-753-to-borrow-15-billion-its-order-book-was-about-as-deep-as-in

Finanças

Revolut corta ~90% o custo de inferência de coding com GLM-5.3 self-hosted em 56×B200
The Register: https://www.theregister.com/ai-and-ml/2026/11/03/british-fintech-revolut-slashes-ai-coding-costs-10x-by-cutting-out-claude-and-going-self-hosted/5307345

Fofocas das redes

Nebius Token Factory coloca Kimi K3.1 Pro em day-0 no mesmo dia da liberação dos pesos
@nebiustf: https://x.com/nebiustf/status/2118023422248669547

Leitura da licença Kimi K3.1 Pro: free até US$20M SaaS / 100M MAU ou US$20M/mês; depois deal ou crédito de marca
@guilleflorvs: https://x.com/guilleflorvs/status/2118043607781216544

Trending no X: "weight-space programming" — modelos frontier editando pesos de redes menores
X News (trending): https://x.com/i/trending/2117960690896118225

Trending no X: miHoYo e planos de IA para NPCs/agentes em games + queda de vendas na China
X News (trending): https://x.com/i/trending/2117954842907865273

15 links únicos desta edição, sem repetir URLs.