Noticiário do Joãozinho

Edição da manhã

21 MATÉRIAS · 11 SEÇÕES · 2 VÍDEOS VERSÃO INTERATIVA NA PÁGINA

MANCHETE

Grok e Claude passam a sair com marca d'água invisível por padrão na API, por causa do AI Act europeu

A xAI publicou na documentação da API que, a partir de 07/11, o texto gerado pelos modelos grok-4.7, 4.6, 4.5, 4.3 e pela família grok-4.20 recebe uma marca d'água invisível ligada por padrão, inclusive em saídas estruturadas e argumentos de chamadas de ferramenta; a resposta traz o cabeçalho x-watermark: true, e administradores de equipe só podem desligar por modelo no Console, confirmando uma exceção legal. A Anthropic, por sua vez, estende em 03/11 a marca ao Claude Mythos 5, Opus 4.7, Opus 4.6 e Sonnet 4.6 (no Amazon Bedrock pode levar alguns dias a mais); na Anthropic a marca não pode ser desligada e não acrescenta tokens, e modelos lançados depois de 02/08/2026 já saíam marcados. Segundo o The Decoder, no Opus 4.6 e no Sonnet 4.6 a marca só vale com temperature, top_p e top_k nos valores padrão.

Por que isso importa: A mudança responde à obrigação de marcar conteúdo gerado por IA do AI Act da União Europeia e põe dois grandes laboratórios no mesmo caminho, enquanto a OpenAI só liga a marca por padrão no ChatGPT e no Codex para usuários europeus. Quem usa essas APIs em produtos, inclusive bancos, passa a entregar texto que pode ser identificado como gerado por IA, e precisa revisar políticas de transparência e ajustes de amostragem.

Watermarking (documentação da API) — xAI, 31/10/2026 https://docs.x.ai/developers/advanced-api-usage/watermarking
Anthropic's watermark rollout leaves a loophole on two older Claude models — The Decoder, 31/10/2026 https://the-decoder.com/anthropics-watermark-rollout-leaves-a-loophole-on-two-older-claude-models/
How Claude marks AI-generated content — Claude Help Center, 31/10/2026 https://support.claude.com/en/articles/16266773-how-claude-marks-ai-generated-content

NESTA EDIÇÃO

Modelos	1	Negócios e Investimento	4	Agentes e Código	3	Games	2
Pesquisa	2	Regulação	2	Modelos Locais	1	Brasil	0
Produtos	3	Open Source	1	Finanças	2	Vídeos	2

Modelos

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 30/10/2026 · Hugging Face (SenseTime)

SenseTime lança o SenseNova-Loop-9B, modelo que roda o mesmo bloco de camadas duas vezes

O SenseNova-Loop-9B-Preview, publicado em 30/10 sob licença Apache 2.0, é multimodal (texto e imagem de entrada) e foi treinado com contexto de até 262 mil tokens. Tem cerca de 9 bilhões de parâmetros, mas executa as camadas 4 a 27 duas vezes com os mesmos pesos: são 56 execuções de camada a partir de 32 camadas guardadas, o equivalente a cerca de 14,6 bilhões de parâmetros na inferência. Já há suporte via pull requests no vLLM e no SGLang.

> **Por que isso importa:** Reaproveitar camadas é um jeito de ganhar profundidade sem aumentar a memória dos pesos, o que interessa a quem roda modelos em hardware limitado. É mais um laboratório chinês abrindo arquitetura nova com licença permissiva.

Fonte original: Hugging Face (SenseTime) https://huggingface.co/sensenova/SenseNova-Loop-9B-Preview

Pesquisa

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 30/10/2026 · NVIDIA Blog

Agentes de IA passam a operar o radiotelescópio Allen, do SETI Institute, a partir de comandos em linguagem natural

O SETI Institute e o Breakthrough Listen montaram no Allen Telescope Array, na Califórnia, um sistema em que agentes apontam as 42 antenas e analisam os dados de rádio ao vivo a partir de pedidos em inglês. A plataforma usa o Holoscan da NVIDIA e processa até 2,4 terabits por segundo em GPUs RTX PRO 6000 Blackwell no próprio observatório. Como prova de conceito, o sistema fez a primeira detecção em tempo real, feita por IA, de um pulsar na Nebulosa do Caranguejo.

> **Por que isso importa:** É um exemplo concreto de agentes operando equipamento científico caro em tempo real, e não só escrevendo código. O mesmo padrão, linguagem natural controlando instrumentos com checagem humana, vale para laboratórios e indústria.

Fonte original: NVIDIA Blog https://blogs.nvidia.com/blog/ai-joins-the-search-for-extraterrestrial-signals/

IMPORTÂNCIA BAIXA 31/10/2026 · Reuters / The Information

Meta construiu com a chinesa Fourier um protótipo de robô humanoide e já trabalha na segunda versão

Segundo o The Information, citado pela Reuters neste sábado (31/10), a Meta montou em cerca de três meses, com a startup chinesa Fourier, um protótipo de humanoide chamado Max, para dar corpo ao seu agente Muse, e já trabalha numa segunda versão. O projeto é de pesquisa interna, sem plano de venda. A Reuters diz que não conseguiu confirmar a informação de forma independente.

> **Por que isso importa:** Mostra a Meta levando seus agentes de IA para o mundo físico e se apoiando em fornecedores chineses de robótica, num momento de tensão entre EUA e China em tecnologia.

Fonte original: Reuters / The Information https://www.investing.com/news/stock-market-news/meta-builds-humanoid-robot-with-chinese-startup-fourier-the-information-says-4989297

Produtos

3 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA ALTA 31/10/2026 · Apple Newsroom (China)

Apple Intelligence chega à China em dezembro, com o Qwen, da Alibaba, dentro da Siri

A Apple anunciou que os recursos de Apple Intelligence chegam aos usuários da China continental em beta a partir de 04/11, com a versão beta para desenvolvedores do iOS, e no lançamento do iOS 27.3, em dezembro. O modelo Qwen (千问), da Alibaba, vai alimentar recursos da Siri, das Ferramentas de Escrita e da Inteligência Visual. Funciona do iPhone 16 em diante, no iPhone 15 Pro e Pro Max, no iPhone Air e no iPhone Duo.

> **Por que isso importa:** A China é um dos maiores mercados do iPhone e estava sem os recursos de IA da Apple. A parceria com a Alibaba mostra como as big techs precisam de modelos locais para operar sob as regras chinesas.

Fonte original: Apple Newsroom (China) https://www.apple.com.cn/newsroom/2026/10/apple-intelligence-features-are-coming-to-users-in-china/

IMPORTÂNCIA MÉDIA 31/10/2026 · Claude Help Center

Anthropic rebatiza os planos do Claude: Max 20x vira Ultra e Max 5x passa a se chamar Max

O plano de US$ 200 por mês agora se chama Ultra: dá 10 vezes o uso semanal do Pro, tira o limite de 5 horas para contas verificadas, tem um limite semanal separado para o Claude Fable e inclui créditos mensais de API. O antigo Max 5x, de US$ 100, passa a se chamar só Max.

> **Por que isso importa:** O fim da janela de 5 horas no plano mais caro atende quem usa o Claude Code o dia todo. Para quem paga pelo time, a troca de nomes também muda o que conferir na fatura.

Fonte original: Claude Help Center https://support.claude.com/en/articles/17728952-what-is-the-ultra-plan

IMPORTÂNCIA MÉDIA 31/10/2026 · OpenAI Help Center

OpenAI vai unificar em dezembro os limites do ChatGPT e do Codex, medidos em tokens

Segundo a central de ajuda da OpenAI, a partir de dezembro os planos Plus, Pro e Business passam a ter um único limite de uso compartilhado entre ChatGPT e Codex, medido em tokens, à medida que os recursos de agente do Work entram na experiência principal do ChatGPT. Conversas no modo Instant não contam para o limite. O consumo poderá ser acompanhado em Configurações, em Uso.

> **Por que isso importa:** Quem usa muito o Codex vai disputar a mesma cota com o uso do ChatGPT. Medir em tokens deixa o custo mais transparente, mas exige que times acompanhem o consumo.

Fonte original: OpenAI Help Center https://help.openai.com/en/articles/20001625-understanding-chatgpt-and-codex-usage-limits

Negócios e Investimento

4 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 30/10/2026 · CNA / Reuters

Kling AI, de vídeo por IA da Kuaishou, fecha rodada de cerca de US$ 3 bilhões

A Kuaishou informou na sexta (30/10) que a Beijing Kling concluiu uma captação de cerca de 19,8 bilhões de yuans (US$ 2,98 bilhões). A Kuaishou fica com dois terços do capital e 52,45% dos votos. A rodada vem depois de outra de US$ 2,8 bilhões em julho, quando a Kling foi avaliada em US$ 15 bilhões antes do aporte. Segundo a Bloomberg, a empresa planeja abrir capital no ano que vem.

> **Por que isso importa:** Vídeo por IA virou uma das áreas mais disputadas, e a Kling concorre diretamente com OpenAI, Google e ByteDance. O volume mostra o apetite chinês por esse mercado mesmo com as ações de IA em queda.

Fonte original: CNA / Reuters https://www.channelnewsasia.com/business/kuaishous-kling-ai-unit-completes-near-3-billion-funding-round-6507461

IMPORTÂNCIA MÉDIA 29/10/2026 · Deepgram

Deepgram compra a Daily, criadora do Pipecat, e o framework de agentes de voz vai para uma fundação independente

A Deepgram anunciou em 29/10 a compra da Daily, sem revelar o valor. O Pipecat, framework aberto para agentes de voz em tempo real, continua com licença BSD-2 e passa a ser cuidado por uma Pipecat Foundation independente; Kwindla Hultman Kramer, da Daily, vira chefe de open source da Deepgram.

> **Por que isso importa:** O Pipecat é uma das bases mais usadas para montar atendentes de voz com IA. A fundação tenta manter a confiança de quem já depende dele, inclusive empresas que usam outros fornecedores de voz.

Fonte original: Deepgram https://deepgram.com/blog/daily-pipecat-deepgram

IMPORTÂNCIA BAIXA 30/10/2026 · 9to5Mac

Apple compra a Ascend.io, de agentes para engenharia de dados, revela lista da União Europeia

A lista de aquisições que a Apple é obrigada a publicar pela Lei de Mercados Digitais (DMA) da União Europeia revelou um acordo com a Ascension Labs, a Ascend.io. É uma contratação de equipe: a Apple pode fazer ofertas a funcionários, recebe licença não exclusiva da propriedade intelectual e pode receber patentes. A Ascend se descreve como plataforma de engenharia de dados com agentes.

> **Por que isso importa:** Mostra a Apple reforçando times de IA com aquisições pequenas. Também é um efeito prático da DMA, que tira do sigilo compras que antes passavam despercebidas.

Fonte original: 9to5Mac https://9to5mac.com/2026/10/30/apple-acquires-ascend%E2%80%A4io-an-agentic-ai-engineering-startup/

IMPORTÂNCIA MÉDIA 30/10/2026 · The Globe and Mail

Anthropic negocia alugar campus de data centers perto de Edmonton, no Canadá, com usina a gás de 920 MW

Segundo o The Globe and Mail, a Anthropic negocia ser a inquilina de um campus da Beacon Data Centers no condado de Sturgeon, em Alberta. O projeto prevê quatro salas de dados e uma usina a gás natural de 920 MW, aprovada pela agência reguladora de Alberta em 23/10, com obras a partir do fim de 2026 e operação prevista para o fim de 2027. Anthropic e Beacon não comentaram. O campus teria cerca do dobro da potência do projeto da EverWind na Nova Escócia (420 MW).

> **Por que isso importa:** Mostra a corrida da Anthropic por capacidade de computação antes do IPO e a dependência crescente de gás natural para alimentar data centers de IA.

Fonte original: The Globe and Mail https://www.theglobeandmail.com/business/economy/article-anthropic-ai-talks-to-lease-beacon-data-centre-near-edmonton-alberta/

Regulação

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 31/10/2026 · TechCrunch

Juíza nega pedido da xAI e ação da NAACP por poluição do data center Colossus vai seguir

A juíza federal Debra M. Brown, do Mississippi, negou o pedido da xAI para arquivar a ação movida pela NAACP com o Southern Environmental Law Center, que acusa a empresa de operar ao menos 35 turbinas a gás no complexo Colossus, em Southaven, perto de Memphis, sem as licenças exigidas pela Lei do Ar Limpo. A xAI alegava que as turbinas eram fontes móveis e isentas; a juíza rejeitou o argumento. Um pedido semelhante do Departamento de Justiça já tinha sido negado neste mês.

> **Por que isso importa:** É um precedente para a disputa sobre a energia dos data centers de IA, num momento em que cresce a reação pública contra eles nos EUA.

Fonte original: TechCrunch https://techcrunch.com/2026/10/31/judge-allows-naacp-lawsuit-over-xai-data-center-pollution-to-move-forward/

IMPORTÂNCIA MÉDIA 31/10/2026 · The Decoder

OpenAI vai treinar modelos com sessões do Daybreak Blue, de cibersegurança, e promete opção de recusa após reação

Em e-mail a membros do programa Trusted Access, a OpenAI avisou que a partir de 29/11 vai usar dados das sessões do Daybreak Blue, até então fora do treino, para treinar seus modelos, mesmo para quem desligou a opção de melhorar o modelo. Contas Business e Enterprise não são afetadas. Depois da reação, um segundo e-mail prometeu uma opção de recusa separada; interações sinalizadas em revisões de segurança continuam podendo ser usadas para fins de segurança.

> **Por que isso importa:** Pesquisadores de segurança temem que achados ainda não publicados acabem no treino. Para empresas, reforça a necessidade de checar os termos de cada produto, e não só a configuração geral da conta.

Fonte original: The Decoder https://the-decoder.com/openai-plans-to-train-models-on-daybreak-blue-cybersecurity-session-data-that-was-previously-off-limits/

Open Source

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 29/10/2026 · Pika

Pika abre os pesos do H3 Flex, modelo de vídeo com áudio que gera em seis passos

O Pika H3 Flex é um ajuste do MiniMax H3 que gera vídeo com áudio sincronizado em seis passos de amostragem, contra 50 do modelo original. Os pesos (adaptadores LoRA) estão no Hugging Face sob a licença comunitária do MiniMax H3; localmente gera em 480p e 768p, e a API da Pika leva a até 4K e entrega cinco trilhas de áudio separadas. A Pika diz que é o melhor modelo aberto de texto para vídeo.

> **Por que isso importa:** Coloca um gerador de vídeo com áudio de alto nível nas mãos de quem quer rodar e ajustar o próprio modelo. A licença, porém, não é totalmente livre e vale conferir antes de uso comercial.

Fonte original: Pika https://pika.art/blog/pika-h3-flex-research

Agentes e Código

3 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA ALTA 29/10/2026 · CrowdStrike

CrowdStrike: hackers ligados à China usaram agentes com DeepSeek para atacar mais de 200 organizações em 11 dias

Segundo a CrowdStrike, o grupo VAPOR PANDA usou um framework de agentes chamado ARTEX, movido principalmente pelo DeepSeek, contra mais de 200 organizações em pelo menos 14 países, em 11 dias a partir de 17/10. A invasão foi confirmada em 9 delas; entre os 19 alvos nos EUA havia empresas farmacêuticas, de semicondutores, defesa, energia nuclear e serviços financeiros. Numa única intrusão foram consumidos 1,4 bilhão de tokens de entrada, e os agentes eram instruídos a procurar máquinas sem o antivírus Falcon.

Por que isso importa: É um dos primeiros casos documentados de ataque em escala conduzido por agentes, e o setor financeiro está entre os alvos. Reforça a pressão sobre o uso do DeepSeek, que o governo americano já mandou tirar dos sistemas federais.

Fonte original: CrowdStrike https://www.crowdstrike.com/en-us/blog/vapor-panda-uses-artex-to-scale-global-campaign/

IMPORTÂNCIA MÉDIA 30/10/2026 · Gemini API Changelog

Computer Use da Gemini API sai do preview e fica disponível para produção

O Google marcou em 30/10 como disponibilidade geral o Computer Use da Gemini API, que deixa o modelo operar interfaces de computador, nas APIs Interactions e GenerateContent.

Por que isso importa: Com a versão estável, empresas podem colocar em produção agentes que clicam e digitam em sistemas sem API, um uso comum em automação de back-office.

Fonte original: Gemini API Changelog https://ai.google.dev/gemini-api/docs/changelog

IMPORTÂNCIA MÉDIA 27/10/2026 · Latent Space

Anthropic reescreveu a Messages API de Python para Rust usando o Claude

No podcast Latent Space, Katelyn Lesse e Angela Jiang, da Anthropic, contaram como a empresa reconstruiu em Rust a Messages API, antes em Python, com o Claude escrevendo a maior parte do código e testes com tráfego real em modo sombra antes da troca. Segundo o relato, que repercutiu no X neste fim de semana, o uso de CPU caiu de 40% a 50% e o tempo até o primeiro token ficou de 10% a 20% menor; a troca foi concluída em setembro.

Por que isso importa: É um caso real de migração grande de linguagem feita com agentes num sistema crítico, com um método replicável: plano humano, código do modelo e verificação com tráfego de produção.

Fonte original: Latent Space https://latent.space/p/how_anthropic_ships_the_1_llm_in_the_world_safely_separating_the_brain_from_the_hands/

Modelos Locais

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 01/11/2026 · MarkTechPost / Hugging Face

Nace.AI abre o Drex Max v1, modelo de decisão de 27B que roda em Macs com 24 GB

O Drex Max não gera texto: é um modelo de decisão que devolve probabilidades para as opções de resposta. É um modelo denso de 27 bilhões de parâmetros, com licença Open RAIL-M. A versão quantizada Q4 roda em Macs com Apple Silicon a partir de 24 GB via MLX ou Kapitsa, e a GGUF Q8_0 pede 48 GB, com forks do llama.cpp e do Ollama. Fica em 4º lugar entre 18 modelos no Decision Index da Artificial Analysis.

Por que isso importa: Classificação e triagem, como encaminhar chamados ou checar regras, são boa parte do trabalho de agentes em empresas. Um modelo que faz isso localmente ajuda onde os dados não podem sair da máquina.

Fonte original: MarkTechPost / Hugging Face https://www.marktechpost.com/2026/11/01/nace-ai-open-sources-drex-max-v1-local-inference/

Finanças

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 29/10/2026 · Zum Channel

BTG Pactual lança o Sales Trader AI, assistente que consulta o Research e envia ordens na renda variável

Lançada em 29/10, a ferramenta entende perguntas em linguagem natural, consulta dados do Research do banco e envia ordens à plataforma de renda variável: a mercado ou limitadas, cestas com vários ativos e rebalanceamento de carteira. Segundo Jerson Zanlorenzi, do BTG, a confirmação da boleta continua com o usuário.

> **Por que isso importa:** É um dos primeiros casos no Brasil de assistente de IA ligado direto à execução de ordens num grande banco, um passo além do atendimento.

Fonte original: Zum Channel https://zumchannel.com/btg-pactual-lanca-assistente-de-ia-que-consulta-dados-do-research-e-envia-ordens-na-plataforma-de-renda-variavel/

IMPORTÂNCIA BAIXA 31/10/2026 · Latent Space

Arca, gestora de patrimônio nativa em IA, conta como roda agentes em tarefas financeiras que duram até dez dias

Em entrevista ao Latent Space, Aman Singh, da Arca, explicou que os agentes da empresa tocam tarefas de até dez dias, com memória persistente por cliente e um harness próprio que escolhe o modelo por tarefa (o Claude é o principal). Os agentes só analisam e propõem: operações e movimentação de dinheiro ficam com consultores humanos, e as travas de segurança são determinísticas. A Arca captou recentemente US$ 48,5 milhões em rodada liderada pela General Catalyst.

> **Por que isso importa:** Mostra um desenho de agente para finanças com limites claros entre o que a IA faz e o que exige humano, útil para bancos que estudam o mesmo caminho.

Fonte original: Latent Space https://latent.space/p/arca/

Games

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 29/10/2026 · Atelico

Atelico abre acesso antecipado ao seu motor de IA que roda dentro do jogo, no aparelho do jogador

O Atelico AI Engine integra Godot, Unity e Unreal e roda no aparelho do jogador modelos de linguagem, geração de imagem, voz (fala e reconhecimento) e sincronia labial; o menor modelo de linguagem tem cerca de 250 MB. Já é usado nos jogos Bobium Brawlers e GARP. Segundo a empresa, um agendador que divide a GPU com o renderizador manteve uma cena em 121 FPS com um modelo de 3B gerando texto, numa RTX 5080 em 4K, contra 3 FPS sem o agendador.

> **Por que isso importa:** Custo de nuvem é o que mais trava IA generativa em jogos. Rodar localmente, sem derrubar a taxa de quadros, abre espaço para NPCs mais vivos em jogos comerciais.

Fonte original: Atelico https://atelico.studio/blog/the-atelico-ai-engine-is-here

IMPORTÂNCIA BAIXA 30/10/2026 · GamesBeat

Ramen abre o código do Aura Autopilot, agente que testa jogos sozinho no Unreal, Unity e Godot

O Aura Autopilot joga usando teclado e mouse de verdade e melhora a si mesmo enquanto testa. É open source (Apache-2.0, no GitHub), funciona com o Claude Code, o Codex ou chave de API própria e, segundo a Ramen, é três vezes mais rápido que o agente de verificação anterior da empresa.

> **Por que isso importa:** Teste de jogo é caro e repetitivo. Um agente aberto que joga e reporta falhas pode baratear o controle de qualidade, sobretudo para estúdios pequenos.

Fonte original: GamesBeat https://gamesbeat.com/ramen-launches-aura-autopilot-an-open-source-autonomous-game-tester/

Brasil

0 MATÉRIAS NESTA EDIÇÃO

Sem novidades relevantes nesta edição.

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.



Using Agents for Long-Horizon Financial Work — Aman Singh, Arca

Latent Space · 31/10/2026

Aman Singh, da Arca, mostra como agentes com memória e harness próprio cuidam de tarefas de gestão de patrimônio que duram dias, deixando operações e dinheiro sempre com humanos.

https://www.youtube.com/watch?v=KM3VQZKjpiY



How Anthropic ships the #1 LLM in the World — with Katelyn Lesse and Angela Jiang

Latent Space · 27/10/2026

Katelyn Lesse e Angela Jiang, da Anthropic, contam como a empresa lança modelos, escala a plataforma do Claude e reescreveu a Messages API em Rust com ajuda do próprio Claude.

https://www.youtube.com/watch?v=MRQTlwcW-aQ

SOBRE ESTA EDIÇÃO
Nº 3 — Edição da manhã · 01/11/2026

JANELA
Cobre o que saiu desde a edição nº 2 (fechada por volta das 18h30 de Brasília de 31/10), as pautas do vigia da madrugada e notícias de 27 a 30/10 que ainda não tinham sido publicadas. Domingo fraco em novidades: a edição é mais enxuta e a seção Brasil ficou sem notícia nova. Fechamento: 01/11/2026, por volta das 8h 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. **Using Agents for Long-Horizon Financial Work — Aman Singh, Arca** — Latent Space, 31/10/2026
   https://www.youtube.com/watch?v=KM3VQZKjpiY
2. **How Anthropic ships the #1 LLM in the World — with Katelyn Lesse and Angela Jiang** — Latent Space, 27/10/2026
   https://www.youtube.com/watch?v=MRQTlwcW-aQ

Manchete

Watermarking (documentação da API) — xAI
https://docs.x.ai/developers/advanced-api-usage/watermarking

Anthropic's watermark rollout leaves a loophole on two older Claude models — The Decoder
https://the-decoder.com/anthropics-watermark-rollout-leaves-a-loophole-on-two-older-claude-models/

How Claude marks AI-generated content — Claude Help Center
https://support.claude.com/en/articles/16266773-how-claude-marks-ai-generated-content

Modelos

SenseTime lança o SenseNova-Loop-9B, modelo que roda o mesmo bloco de camadas duas vezes
Hugging Face (SenseTime): https://huggingface.co/sensenova/SenseNova-Loop-9B-Preview

Pesquisa

Agentes de IA passam a operar o radiotelescópio Allen, do SETI Institute, a partir de comandos em linguagem natural
NVIDIA Blog: https://blogs.nvidia.com/blog/ai-joins-the-search-for-extraterrestrial-signals/

Meta construiu com a chinesa Fourier um protótipo de robô humanoide e já trabalha na segunda versão
Reuters / The Information: https://www.investing.com/news/stock-market-news/meta-builds-humanoid-robot-with-chinese-startup-fourier-the-information-says-4989297

Produtos

Apple Intelligence chega à China em dezembro, com o Qwen, da Alibaba, dentro da Siri
Apple Newsroom (China): https://www.apple.com.cn/newsroom/2026/10/apple-intelligence-features-are-coming-to-users-in-china/

Anthropic rebatiza os planos do Claude: Max 20x vira Ultra e Max 5x passa a se chamar Max
Claude Help Center: https://support.claude.com/en/articles/17728952-what-is-the-ultra-plan

OpenAI vai unificar em dezembro os limites do ChatGPT e do Codex, medidos em tokens
OpenAI Help Center: https://help.openai.com/en/articles/20001625-understanding-chatgpt-and-codex-usage-limits

Negócios e Investimento

Kling AI, de vídeo por IA da Kuaishou, fecha rodada de cerca de US$ 3 bilhões
CNA / Reuters: https://www.channelnewsasia.com/business/kuaishous-kling-ai-unit-completes-near-3-billion-funding-round-6507461

Deepgram compra a Daily, criadora do Pipecat, e o framework de agentes de voz vai para uma fundação independente
Deepgram: https://deepgram.com/blog/daily-pipecat-deepgram

Apple compra a Ascend.io, de agentes para engenharia de dados, revela lista da União Europeia
9to5Mac: https://9to5mac.com/2026/10/30/apple-acquires-ascend%E2%80%A4io-an-agentic-ai-engineering-startup/

Anthropic negocia alugar campus de data centers perto de Edmonton, no Canadá, com usina a gás de 920 MW
The Globe and Mail: https://www.theglobeandmail.com/business/economy/article-anthropic-ai-talks-to-lease-beacon-data-centre-near-edmonton-alberta/

Regulação

Juíza nega pedido da xAI e ação da NAACP por poluição do data center Colossus vai seguir
TechCrunch: https://techcrunch.com/2026/10/31/judge-allows-naacp-lawsuit-over-xai-data-center-pollution-to-move-forward/

OpenAI vai treinar modelos com sessões do Daybreak Blue, de cibersegurança, e promete opção de recusa após reação
The Decoder: https://the-decoder.com/openai-plans-to-train-models-on-daybreak-blue-cybersecurity-session-data-that-was-previously-off-limits/

Open Source

Pika abre os pesos do H3 Flex, modelo de vídeo com áudio que gera em seis passos
Pika: https://pika.art/blog/pika-h3-flex-research

Agentes e Código

CrowdStrike: hackers ligados à China usaram agentes com DeepSeek para atacar mais de 200 organizações em 11 dias
CrowdStrike: https://www.crowdstrike.com/en-us/blog/vapor-panda-uses-artex-to-scale-global-campaign/

Computer Use da Gemini API sai do preview e fica disponível para produção
Gemini API Changelog: https://ai.google.dev/gemini-api/docs/changelog

Anthropic reescreveu a Messages API de Python para Rust usando o Claude
Latent Space: https://latent.space/p/how_anthropic_ships_the_1_llm_in_the_world_safely_separating_the_brain_from_the_hands/

Modelos Locais

Nace.AI abre o Drex Max v1, modelo de decisão de 27B que roda em Macs com 24 GB
MarkTechPost / Hugging Face: https://www.marktechpost.com/2026/11/01/nace-ai-open-sources-drex-max-v1-local-inference/

Finanças

BTG Pactual lança o Sales Trader AI, assistente que consulta o Research e envia ordens na renda variável
Zum Channel: https://zumchannel.com/btg-pactual-lanca-assistente-de-ia-que-consulta-dados-do-research-e-envia-ordens-na-plataforma-de-renda-variavel/

Arca, gestora de patrimônio nativa em IA, conta como roda agentes em tarefas financeiras que duram até dez dias
Latent Space: https://latent.space/p/arca/

Games

Atelico abre acesso antecipado ao seu motor de IA que roda dentro do jogo, no aparelho do jogador
Atelico: https://atelico.studio/blog/the-atelico-ai-engine-is-here

Ramen abre o código do Aura Autopilot, agente que testa jogos sozinho no Unreal, Unity e Godot
GamesBeat: https://gamesbeat.com/ramen-launches-aura-autopilot-an-open-source-autonomous-game-tester/

26 links únicos desta edição, sem repetir URLs.