Noticiário do Joãozinho

Edição da noite

21 MATÉRIAS · 12 SEÇÕES · 4 VÍDEOS VERSÃO INTERATIVA NA PÁGINA

MANCHETE

Atualização: dois grandes acionistas pedem que o conselho da NVIDIA considere a volta de Jensen Huang, e petição pela reintegração passa de 100 mil assinaturas

Brad Gerstner, da Altimeter Capital, e Gavin Baker, da Atreides Management, publicaram no X, na noite de sexta (13/11), uma carta enviada ao conselho da NVIDIA que pede um processo "lento, medido e deliberado" para definir o futuro comando da empresa e que avalie todas as opções, inclusive "o possível retorno de Jensen Huang". Segundo a CNBC, as duas gestoras declaravam juntas cerca de US$ 2,13 bilhões em ações da NVIDIA em 30 de junho. A carta diz que a empresa será "bem conduzida" pela CEO interina, Colette Kress, e não pede a reintegração imediata. Ao mesmo tempo, a petição "Stand with Jensen" passava de 108 mil assinaturas na manhã de sábado (14/11), segundo o Business Insider. Os organizadores dizem que a lista tem mais de 58 mil acionistas, mais de 5.600 funcionários atuais e mais de mil ex-funcionários, e que trabalham para tirar assinaturas falsas: apareceram nomes como "Tim Cruise" e "Xi Jingping". Elon Musk e Amjad Masad, CEO da Replit, divulgaram a petição. O Business Insider lembra que o procurador-geral dos EUA, Todd Blanche, disse na quinta que o Departamento de Justiça ainda não concluiu se Alice Huang é espiã.

Por que isso importa: A pressão de acionistas com bilhões em ações e de milhares de funcionários dificulta uma troca rápida de comando, e a NVIDIA segue sem CEO definitivo no meio da crise dos chips desviados para a China. A volta de Huang ainda é só hipótese: a carta não a exige e o conselho mantém a investigação independente.

Nvidia board urged to consider 'potential return of Jensen Huang' by two major investors — CNBC, 14/11/2026 https://www.cnbc.com/2026/11/14/nvidia-investors-jensen-huang-return-board-letter.html

A petition to reinstate Jensen Huang as Nvidia CEO has over 100,000 signatures, including some big names — Business Insider, 14/11/2026 https://www.businessinsider.com/petition-reinstate-jensen-huang-nvidia-elon-musk-2026-11

Brad Gerstner publica a carta enviada ao conselho da NVIDIA — X (@altcap), 13/11/2026 https://x.com/altcap/status/2121389561414623679

NESTA EDIÇÃO

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

Modelos

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 14/11/2026 · The Decoder

Atualização: Gemini 4 Carbon assume o 1º lugar no MathArena e tira o GPT-6.1 Sol do topo, mas custou cerca de 15 vezes mais nas configurações do teste

O pesquisador Jasper Dekoninck anunciou no X, na sexta (13/11), que, "pela primeira vez em muito tempo", um modelo que não é da OpenAI lidera o ranking de matemática: o Gemini 4 Carbon, do Google, passou o GPT-6.1 Sol, e o Claude Fable 5.2, da Anthropic, ficou em segundo. Segundo o The Decoder, a falha mais comum do Carbon foi estourar o tempo limite, e o modelo quase nunca desiste de uma tarefa; num caso, rodou um script por mais de 11 horas sem conferir o progresso. Com as configurações de raciocínio usadas no teste, Carbon e Fable 5.2 custaram cerca de 15 e 40 vezes mais que o GPT-6.1 Sol, respectivamente.

> **Por que isso importa:** A liderança em matemática mudou de mãos, mas a um preço bem maior por resposta. Para quem paga por uso, a escolha do modelo passa a depender de quanto vale cada ponto a mais no teste.

Fonte original: The Decoder https://the-decoder.com/gemini-4-carbon-takes-the-matharena-crown-but-costs-15-times-more-than-gpt-6-1-sol/
Post de Jasper Dekoninck (MathArena) https://x.com/j_dekoninck/status/2121263753006374997

Pesquisa

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA ALTA 13/11/2026 · Google Research

Google diz que agentes do /teamwork, no Antigravity, ajudaram a provar a conjectura de Courtade–Kumar, aberta desde 2013, numa prova de 50 milhões de linhas em Lean

Num post assinado por Vahab Mirrokni, o Google Research reuniu resultados do /teamwork, sistema multiagente do Antigravity que combina tentativas em paralelo com verificadores automáticos (como o verificador de provas Lean 4 e suítes de teste) e divide o contexto entre agentes especializados, com críticos e auditores revisando o trabalho. Pesquisadores do Google, da CUHK, da USC e da CMU usaram o padrão Long Proof para provar, em generalidade total, a conjectura de Courtade–Kumar, da teoria da informação, de 2013; o Gemini 4 Argon e modelos anteriores desenvolveram a maior parte dos novos limites analíticos, e a prova formalizada em Lean 4 tem 50 milhões de linhas. O post cita ainda resultados em análise adaptativa de dados (uma lacuna aberta desde 2014), computação quântica e multiplicação de matrizes booleanas, e diz que as ferramentas já estão no Antigravity, com mais recursos para usuários externos nas próximas semanas.

> **Por que isso importa:** É mais um caso de IA produzindo matemática nova conferida por máquina, agora com uma ferramenta aberta a outros pesquisadores. A checagem em Lean reduz o risco de erro escondido, mas os resultados são relatados pelo próprio Google e pelos coautores.

Fonte original: Google Research https://research.google/blog/advancing-science-engineering-with-antigravitys-teamwork-feature/

Produtos

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 12/11/2026 · Anthropic (Claude)

Anthropic recomenda tirar o “pense bem” dos prompts do Claude Opus 5.5, agora o modelo padrão dos planos pagos, e começar no esforço médio

Num guia publicado no site do Claude, a Anthropic diz que o Opus 5.5 virou o modelo padrão dos planos Pro, Max, Ultra e Team e o recomenda como modelo do dia a dia para pesquisa, análise e escrita. Segundo o The Decoder, a empresa pede para cortar expressões como “think hard” e “think carefully”: o modelo já ajusta sozinho quanto raciocina, e nos testes internos as respostas sem essas frases começaram mais rápido, sem perda de qualidade medida. A Anthropic também recomenda começar no esforço “Medium”, o padrão, que nos testes internos igualou o Opus 5 no “High”. Para quem usava o Sonnet no dia a dia e guardava o Opus para tarefas pesadas, a empresa diz que vale repensar, porque um modelo mais capaz pode terminar a tarefa com menos trabalho e compensar boa parte do preço maior por token.

> **Por que isso importa:** Hábitos antigos de prompt deixam o modelo mais lento e caro sem ganho. Para quem paga por uso, escolher o nível de esforço certo pesa mais do que frases de efeito.

Fonte original: Anthropic (Claude) https://claude.com/resources/articles/using-claude-opus-5-5-as-your-daily-driver-for-knowledge-work

The Decoder: os hábitos que a Anthropic manda abandonar https://the-decoder.com/anthropic-says-outdated-prompting-habits-are-slowing-down-claude-opus-5-5/

IMPORTÂNCIA BAIXA 14/11/2026 · heise online

Canonical apresenta o Inference Manager, ferramenta do Ubuntu que instala e serve modelos de IA locais por uma interface compatível com a API da OpenAI

No Ubuntu Summit, a Canonical mostrou o Inference Manager, que traz o comando “inference” para instalar modelos no próprio computador e deixá-los disponíveis para outros programas por uma interface HTTP compatível com a API da OpenAI, segundo o heise. A ferramenta organiza os “inference snaps”, pacotes da Snap Store com modelos que rodam sem internet, e atende na porta TCP 8400 assistentes de chat, agentes de código no editor e programas de ditado. O Ubuntu 26.10 já trouxe o Myna, de fala para texto, baseado nesses pacotes.

> **Por que isso importa:** Uma grande distribuição Linux passa a tratar modelos locais como parte do sistema, com um ponto único de acesso, o que facilita usar IA sem nuvem no desktop.

Fonte original: heise online https://www.heise.de/news/Inference-Manager-neues-Ubuntu-Tool-verwaltet-lokale-KI-Modelle-11525275.html

Negócios e Investimento

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 14/11/2026 · CNBC

Foundation, startup de robôs humanoides militares apoiada por Eric Trump, não paga salários desde 30/10, e o CEO tomou US$ 4,6 milhões emprestados da empresa para comprar uma casa, diz a CNBC

Segundo pessoas ouvidas pela CNBC, a Foundation Future Industries, de São Francisco, que faz robôs humanoides para uso militar e tem Eric Trump como conselheiro e investidor, não paga os funcionários desde que perdeu a folha de 30 de outubro, com o caixa no fim. O CEO e cofundador, Sankaet Pathak, tomou US$ 4,6 milhões emprestados da empresa no ano passado para comprar uma casa de luxo e é o único membro do conselho, o que preocupa possíveis investidores. Ainda segundo as fontes, os contratos de US$ 24 milhões com o Pentágono, muito divulgados, vêm da compra da Boardwalk Robotics, em 2024. A empresa anterior de Pathak, a Synapse, quebrou em 2024 e deixou milhares de clientes sem acesso aos depósitos.

> **Por que isso importa:** Mostra o risco de governança em startups de robótica e defesa que captam com base em contratos públicos e padrinhos políticos, num setor que atrai muito dinheiro.

Fonte original: CNBC https://www.cnbc.com/2026/11/14/foundation-payroll-ceo-loan.html

Regulação

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 14/11/2026 · Tom's Hardware (com o Malay Mail)

Atualização: Malásia prepara lei que obriga data centers a declarar GPUs importadas e consumo de energia para pegar desvio de chips para a China

Segundo o Malay Mail, citado pelo Tom's Hardware, a Malásia vai propor uma lei que obriga operadores de data centers a informar quantas GPUs importam, o consumo de energia, onde ficam as instalações e quem são os operadores e os grandes clientes, com acompanhamento durante toda a vida do projeto, e não só na aprovação. O ministro de Investimento, Comércio e Indústria, Johari Abdul Ghani, admitiu na quinta (12/11) que empresas alugaram espaço em data centers do país para importar GPUs legalmente e depois mandá-las para fora. Os operadores também terão de pagar por toda a capacidade elétrica reservada, mesmo a que não usarem. O Tom's Hardware avalia que cruzar consumo e importações pode ajudar, mas que a variação real do consumo das GPUs e a falta de auditoria física podem limitar a medida.

> **Por que isso importa:** A Malásia é um dos maiores polos de data centers da Ásia e aparece nas investigações sobre chips da NVIDIA desviados para a China; uma lei assim encarece usar o país como ponte.

Fonte original: Tom's Hardware (com o Malay Mail) https://www.tomshardware.com/tech-industry/artificial-intelligence/malaysia-plans-new-law-using-data-center-power-consumption-to-catch-ai-gpu-smugglers-but-real-world-tdp-variance-and-lack-of-physical-audits-could-undermine-it-new-law-also-forces-operators-to-pay-for-unused-electricity-capacity

Open Source

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 14/11/2026 · MarkTechPost

Stanford abre o Memorilla, módulo de memória que resume uma coleção de documentos em 16 tokens para um LLM sem retreino e supera o contexto completo no FactKG

O grupo SNAP, de Jure Leskovec, em Stanford, com coautores de Toronto, Penn e MIT, publicou sob licença Apache 2.0 o código, as receitas de treino e os checkpoints do Memorilla. O módulo, de 194 milhões de parâmetros, lê a coleção inteira de documentos e entrega ao modelo sempre 16 “tokens de memória”, sem mexer no modelo principal (o Qwen3-8B, com o Qwen3-Embedding-4B como codificador). No FactKG, o Qwen3-8B acertou 84,33% com o Memorilla, contra 80,43% com todos os documentos no contexto, usando 108 vezes menos tokens para representar os documentos; no PersonaMem-v2, foram 44,60% contra 27,92%. Os números são dos autores, e o início rápido pede uma GPU de 80 GB.

> **Por que isso importa:** Memória longa é um gargalo de agentes e assistentes pessoais. Uma memória compacta, aberta e com licença comercial pode baratear o uso de históricos grandes sem encher a janela de contexto.

Fonte original: MarkTechPost https://www.marktechpost.com/2026/11/14/stanford-memorilla-memory-system-16-memory-tokens/
Artigo no arXiv https://arxiv.org/abs/2610.21375
Código no GitHub https://github.com/snap-stanford/memorilla

IMPORTÂNCIA BAIXA 14/11/2026 · GitHub / r/LocalLLaMA

tiny-voca: voz sintética totalmente aberta, com 4,5 milhões de parâmetros e 18 MB, que roda 70 vezes mais rápido que o tempo real numa CPU

O desenvolvedor n4ze3m publicou no r/LocalLLaMA o tiny-voca, modelo de texto para fala em inglês com 4.537.291 parâmetros (18 MB, já incluindo a parte que gera o som), duas vozes e áudio em 24 kHz. Segundo o autor, ele roda cerca de 70 vezes mais rápido que o tempo real em 4 threads de CPU, funciona em PyTorch, ONNX e no navegador e foi treinado do zero em 7,7 horas numa RTX 4090. Código, pesos, receita de treino e dados estão abertos. O próprio autor lista os limites: só inglês e duas vozes, sem clonagem, erros com nomes estrangeiros e som um pouco sintético, já que o áudio de treino também foi gerado, com o Qwen3-TTS.

> **Por que isso importa:** Mostra que dá para ter voz sintética útil num modelo minúsculo, que roda em qualquer máquina, com tudo aberto, inclusive os dados, o que é raro.

Fonte original: GitHub / r/LocalLLaMA https://github.com/n4ze3m/tiny-voca
Pesos no Hugging Face https://huggingface.co/n4ze3m/tiny-voca
Post no r/LocalLLaMA https://www.reddit.com/r/LocalLLaMA/comments/1xw7zrg/tinyvoca_a_fully_open_45mparameter_tts_code/

Agentes e Código

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA ALTA 13/11/2026 · Google Project Zero

Project Zero, do Google, encurta o prazo para divulgar falhas por causa da IA: 90 dias corridos agora, 60 em fevereiro de 2027 e 30 em janeiro de 2028

Num post de Natalie Silvanovich, o Project Zero diz que a descoberta e a exploração de falhas com IA tornaram insuficiente a regra "90 + 30", usada desde 2021, que dava 90 dias para corrigir e mais 30 de folga para distribuir o patch. Segundo a equipe, os LLMs aceleraram os atacantes, aumentaram a chance de várias pessoas acharem o mesmo bug e reduziram quase a zero o tempo entre a publicação de um patch e a exploração; o próprio grupo passou a achar várias falhas por semana em alvos onde achava poucas por ano. A nova regra, "Full 90", divulga a falha 90 dias depois do aviso ao fabricante, sem período de graça; o prazo cai para 60 dias em 2 de fevereiro de 2027 e para 30 dias em 3 de janeiro de 2028. As exceções são falhas já exploradas, falhas já públicas e pedido do próprio fabricante.

> **Por que isso importa:** Uma das equipes de segurança mais influentes do mundo admite que a IA mudou o ritmo dos ataques. Fabricantes de software terão menos tempo para corrigir, e prazos menores tendem a virar padrão no setor.

Fonte original: Google Project Zero https://projectzero.google/2026/11/disclosure-policy.html

IMPORTÂNCIA BAIXA 10/11/2026 · Lexifina

Agentes de IA e o verificador Lean acham 138 erros de redação no código tributário dos EUA, diz a Lexifina

A Lexifina, que mantém um repositório aberto (CC0) de leis versionadas, traduziu para Lean o Título 26 do Código dos EUA, o da lei do imposto de renda, na versão de 3 de novembro de 2026, e conferiu cada remissão entre artigos. O teste apontou 257 casos, cada um revisado por um agente de IA contra o texto oficial: 98 eram intencionais, 21 eram alarmes falsos nos links dos próprios editores e 138 eram erros reais, como referências a dispositivos que não existem mais, cada um com um certificado que o Lean reconfere em cerca de dois minutos e meio. A empresa também usou IA para escrever em Lean o cálculo do imposto de 2025 de uma família, testado contra tabelas e exemplos da Receita americana e contra três programas de imposto em 5.000 domicílios, com um segundo agente revisando cada parte para achar erros e trapaças. O texto chegou ao Hacker News neste sábado.

> **Por que isso importa:** Mostra agentes de código saindo do software para o direito: com um verificador formal por trás, a IA vira uma ferramenta de auditoria de leis, e não só de redação.

Fonte original: Lexifina https://lexifina.com/blog/formalizing-the-us-tax-code

Modelos Locais

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 14/11/2026 · GitHub / r/LocalLLaMA

tandem.cpp, patches para o llama.cpp, roda o DeepSeek-V4.1 Flash num PC com 64 GB de RAM e placa de 16 GB a 14,2 tokens por segundo, quase o dobro do llama.cpp puro

O autor (u/smflx) atualizou no r/LocalLLaMA o tandem.cpp, que deixa na GPU os especialistas mais usados de modelos MoE grandes demais para a memória de vídeo e põe a CPU para ajudar com o resto. Com o DeepSeek-V4.1-Flash em IQ2_XS (139 GiB de especialistas), uma RTX 5070 Ti de 16 GB com 64 GB de RAM gerou 14,2 tokens por segundo, contra 7,4 no llama.cpp; como o modelo tem cerca do dobro da RAM, parte dele é lida do SSD. Numa RTX 5090 com 128 GB de RAM, foram 33,1 contra 11,5. O autor relata ganhos de 1,4 a 2,1 vezes também no gpt-oss-120b, no GLM-4.5-Air e no Qwen3.6, e a versão nova compila para placas AMD (ROCm).

> **Por que isso importa:** Os melhores modelos abertos ficaram grandes demais para placas comuns; técnicas que juntam GPU, RAM e SSD trazem esses modelos para PCs domésticos, ainda que devagar. Os números são do próprio autor.

Fonte original: GitHub / r/LocalLLaMA https://github.com/ewhacc/tandem.cpp
Post no r/LocalLLaMA https://www.reddit.com/r/LocalLLaMA/comments/1xwan2j/update_tandemcpp_deepseekv41_on_a_64_gb_pc_with_a/

IMPORTÂNCIA BAIXA 14/11/2026 · Tom's Hardware

Advogado que montou servidor de IA com oito GPUs V100 usadas passa 80% do DeepSeek V4.1 Flash para 4 bits (NVFP4) e acrescenta 1 TB de RAM e uma RTX 3090

Segundo o Tom's Hardware, o advogado Nick Mermiges, de Columbia, na Carolina do Sul, fundador da PLI Labs, atualizou o servidor privado que usa no escritório com o DeepSeek V4.1 Flash e uma camada própria de fluxos e checagens jurídicas: cerca de 80% do trabalho passou de 2 bits para a precisão NVFP4, de 4 bits, e ele acrescentou 1 TB de RAM e uma RTX 3090. A geração de código caiu de 160–215 para 150–180 tokens por segundo, e os travamentos, antes atribuídos ao PCIe, provavelmente vinham de uma fonte com defeito. Ele diz que o Claude Code, com o Opus 5.5 e o Fable 5.2, escreveu os kernels e rodou os testes sob sua direção.

> **Por que isso importa:** Um escritório pequeno roda um modelo de ponta em casa, com hardware usado, para não mandar dados de clientes para a nuvem, e usa a própria IA para montar a infraestrutura.

Fonte original: Tom's Hardware https://www.tomshardware.com/tech-industry/artificial-intelligence/lawyer-upgrades-eight-v100-private-ai-server-to-run-80-percent-of-deepseek-v4-1-flash-at-4-bit-nvfp4-precision-adds-1tb-of-ram-an-nvidia-rtx-3090-and-fixes-crashes-caused-by-a-bad-psu
Discussão no r/LocalLLaMA https://www.reddit.com/r/LocalLLaMA/comments/1xw2ogq/the_lawyer_who_posts_here_about_his_v100s_got_his/

Finanças

1 MATÉRIA NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 13/11/2026 · Databento

Databento, fornecedora de dados de mercado com contratos com cinco dos seis maiores laboratórios de IA, capta US$ 100 milhões e passa a valer US$ 1,1 bilhão

A Databento anunciou uma rodada Série C de US$ 100 milhões, a uma avaliação de US$ 1,1 bilhão, liderada pela Fifth Down Capital, com a Fitch Ventures como investidora estratégica e novos aportes de NEA, DRW Venture Capital, Redpoint e outros investidores antigos. Segundo a empresa, a receita mais que dobrou desde a Série B, no começo deste ano, e ela virou fornecedora de dados de mercado para IA e fluxos com agentes, com contratos corporativos com cinco dos seis maiores laboratórios de IA de fronteira. A Databento diz ter chegado a essa avaliação com menos de 40 funcionários.

> **Por que isso importa:** Mostra os laboratórios de IA comprando dados financeiros de alta qualidade para treinar e alimentar agentes, e empresas enxutas de infraestrutura de dados ganhando valor com essa demanda.

Fonte original: Databento https://databento.com/blog/series-c
FinTech Global sobre a rodada https://fintech.global/2026/11/13/databento-hits-1-1bn-valuation-with-100m-series-c/

Games

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 14/11/2026 · Insider Gaming (com o IGN)

A cinco dias do GTA 6, CEO da Take-Two diz que jogos clonados por IA não ameaçam o negócio e cobra que Anthropic e OpenAI protejam a propriedade intelectual dos outros

Perguntado pelo IGN sobre o número crescente de jogos desmontados com IA e de versões não oficiais para PC, Strauss Zelnick, CEO da Take-Two, dona da Rockstar, disse que não vê ameaça financeira, mas uma questão criativa: para ele, quase todo mundo quer "a experiência da Rockstar como a Rockstar pensou". Sobre as medidas da Anthropic para impedir que o Claude faça engenharia reversa de jogos, disse que as empresas de IA devem proteger a propriedade intelectual de todos: "se você está no negócio de propriedade intelectual, como a Anthropic e a OpenAI, parece irônico pisar na de outra pessoa e esperar que a sua seja protegida". O GTA 6 sai em 19 de novembro para PlayStation 5 e Xbox Series.

> **Por que isso importa:** O maior lançamento do ano chega num momento em que a IA facilita copiar e portar jogos; a fala mostra que as grandes editoras vão pressionar as empresas de IA, mais do que os jogadores.

Fonte original: Insider Gaming (com o IGN) https://insider-gaming.com/take-two-ai-game-cloning-creative-concern/

IMPORTÂNCIA BAIXA 14/11/2026 · Clubic

"Não": divulgador da GeForce na NVIDIA responde a quem pergunta se a saída de Huang é o fim das placas de vídeo para jogos

Segundo o Clubic, Jacob Freeman, ex-EVGA e um dos "GeForce evangelists" da NVIDIA, que divulgam a linha de placas para jogos, escreveu no X que o trabalho continua e que a equipe estava na PDXLAN, uma das maiores LAN parties, onde NVIDIA e GeForce marcam presença desde 2003. Perguntado por um seguidor se tinha opinião sobre o que muitos chamam de "fim da GeForce", respondeu só "Não 😁". O site lembra que Freeman não é executivo nem decide os rumos da empresa, e que os jogos hoje pesam pouco perto do negócio de IA.

> **Por que isso importa:** Com a NVIDIA sob comando interino da diretora financeira, jogadores temem cortes na área que deu origem à empresa; por enquanto, só há uma resposta informal, sem posição oficial.

Fonte original: Clubic https://www.clubic.com/actualite-639313-jacob-freeman-assure-que-nvidia-va-poursuivre-le-business-des-gpu-gaming.html

Brasil

2 MATÉRIAS NESTA EDIÇÃO

IMPORTÂNCIA MÉDIA 14/11/2026 · Agência Brasil

Presidente do Porto Digital diz que a IA abre chance de o Brasil virar “celeiro de aplicações”, mas que o país precisa decidir se quer ser fornecedor ou comprador

Em entrevista coletiva durante o Rec’n’Play, no Recife, Pierre Lucena, presidente do Porto Digital, disse que o país está numa “encruzilhada” e defendeu que ele seja fornecedor de aplicações, apostando na indústria nacional de TI; citou o Pix, a Embraer e o agro como exemplos de soluções brasileiras que viraram referência. Para ele, o caminho passa pelos ambientes de inovação que já existem e por mais apoio das universidades: “você não desenvolve nenhum lugar em economia do conhecimento sem a universidade junto”. Segundo a Agência Brasil, o Porto Digital reúne mais de 24 mil pessoas em mais de 541 empresas e faturou R$ 7,4 bilhões no ano passado, 19% a mais que no anterior.

> **Por que isso importa:** Coloca a discussão sobre IA no Brasil em termos de política industrial: desenvolver aplicações próprias ou só comprar as das big techs.

Fonte original: Agência Brasil https://agenciabrasil.ebc.com.br/geral/noticia/2026-11/porto-digital-com-ia-brasil-pode-se-tornar-celeiro-de-aplicacoes

IMPORTÂNCIA MÉDIA 14/11/2026 · Showmetech

Antes da Black Friday, estudo da Podqi acha 351 sites falsos entre as lojas que ChatGPT, Gemini e Perplexity indicaram em buscas de compras

A Podqi, empresa de proteção de propriedade intelectual, testou 151 marcas com cinco perguntas sobre compras e avaliou 6.795 respostas e 4.437 sites indicados pelos assistentes, segundo o Showmetech, que cita levantamento divulgado pelo Tecnoblog. Apareceram 351 sites falsos e outros 367 suspensos, inativos ou em construção, 718 domínios no total. O Perplexity teve a maior proporção de respostas com algum site de risco, seguido do ChatGPT e do Gemini, e os sites de risco eram mostrados 6,3 vezes mais do que vinham acompanhados de alerta. Na pergunta sobre sites regionais e envio internacional, 16% das respostas exibiram um site de risco. O estudo não mediu lojas brasileiras especificamente, e as empresas contestaram ou contextualizaram o resultado.

> **Por que isso importa:** Com mais gente pedindo indicação de loja à IA, o golpe pode vir dentro da própria resposta. A orientação, que inclui a do CERT.br, é entrar pelo site oficial da marca e conferir o destinatário do Pix ou do boleto.

Fonte original: Showmetech https://www.showmetech.com.br/chatgpt-gemini-lojas-falsas-risco/

Fofocas das redes

4 MATÉRIAS NESTA EDIÇÃO

X 13/11/2026 · @hilbertspaess

Ex-pesquisador da Anthropic diz que um slide do roadshow do IPO pesou na decisão de sair: “eles já poderiam retreinar o Mythos com uma fração do cômputo”

Jacob Coxon, que passou pela OpenAI e pela Anthropic, postou um slide do roadshow do IPO da Anthropic sobre a eficiência de cômputo do pré-treino ao longo do tempo e disse que ele foi “um grande fator” para deixar a empresa. Segundo o post, o eixo é logarítmico e mostra progresso algorítmico “superexponencial”, e o Mythos “não foi um desvio isolado”. Ele próprio ressalva que isso não implica, por si, uma explosão de inteligência só por software, e numa resposta defendeu pausar o desenvolvimento de fronteira no ano que vem. É a leitura dele sobre o gráfico, sem comentário da Anthropic.

> **Por que isso importa:** O gráfico saiu do material de venda do IPO e virou argumento de quem pede pausa na IA de fronteira, a poucos dias da precificação.

Engajamento: 1,3 mi visualizações · 5,4 mil curtidas · 471 reposts · 240 citações · 340 respostas · 3 mil salvamentos (medido em 14/11/2026 17:58)
Post original (X): @hilbertspaess https://x.com/hilbertspaess/status/2121331333654086124
The Decoder sobre o post https://the-decoder.com/anthropic-can-already-retrain-mythos-with-a-fraction-of-the-compute-originally-needed-says-former-researcher-coxon/
Resposta de Coxon pedindo a pausa https://x.com/hilbertspaess/status/2121425977876341092

X 14/11/2026 · @elonmusk

Musk diz que muitos acordos de Huang na NVIDIA eram “um aperto de mão, não um contrato escrito”, e que alguns vão se perder sem ele

Ao citar um trecho de entrevista de Dylan Patel, da SemiAnalysis, Elon Musk escreveu que Jensen Huang conduziu a cadeia de suprimentos “brilhantemente”, mas que muitos acordos eram fechados no aperto de mão, e que “pelo menos alguns desses acordos críticos serão perdidos”. Numa resposta, disse que, quando Huang lhe dava a palavra sobre entrega de computação, sempre cumpria, e que a ordem de compra vinha depois. O post passou de 3,3 milhões de visualizações até o fim da tarde de sábado.

> **Por que isso importa:** Musk é um dos maiores compradores de GPUs do mundo; a fala alimenta a dúvida sobre o que acontece com os contratos de fornecimento da NVIDIA sem o fundador.

Engajamento: 3,4 mi visualizações · 52,3 mil curtidas · 4,9 mil reposts · 519 citações · 2,2 mil respostas · 3,4 mil salvamentos (medido em 14/11/2026 17:58)
Post original (X): @elonmusk https://x.com/elonmusk/status/2121613402846204253
Resposta de Musk sobre a palavra de Huang https://x.com/elonmusk/status/2121624209675665519
Valor Econômico sobre o post https://valor.globo.com/empresas/noticia/2026/11/14/musk-afirma-que-nvidia-perder-negcios-crticos-sem-jensen-huang.ghtml

HACKER NEWS 14/11/2026 · armcat

Estudo sobre apps feitos no Lovable que deixam passaportes e mais de 200 mil mensagens privadas abertos passa de 390 pontos no Hacker News

O post leva ao estudo da CrissyField, feito com o jornal Handelsblatt, que analisou 159.041 projetos do Lovable, plataforma sueca que cria apps a partir de descrições em texto. Dos 34.545 que usavam o banco Supabase, 24.664 enviavam dados ao abrir a página, e em 7.403 deles (30%) os autores acharam tabelas com e-mails sem exigir login. Entre os casos: um fórum alemão com mais de 80 mil perfis e mais de 200 mil mensagens, e um prestador com mais de mil registros de documentos, incluindo números de passaporte. Os autores avisam que os 30% não são uma taxa confirmada de vazamento, e lembram que os termos do Lovable põem a segurança sob responsabilidade do cliente.

> **Por que isso importa:** Apps criados por IA por quem não programa chegam ao ar com dados sensíveis desprotegidos; a conta da segurança continua com quem publica.

Engajamento: 392 pontos · 132 comentários (medido em 14/11/2026 17:59)
Post original (Hacker News): armcat https://news.ycombinator.com/item?id=50499017
O estudo da CrissyField https://www.crissyfield.de/research/lovable-en/

REDDIT 14/11/2026 · r/LocalLLaMA

"Aquecedor perfeito para o inverno": usuário mostra servidor caseiro com oito RTX 3090 rodando agentes em paralelo

O usuário u/dazzou5ouh mostrou no r/LocalLLaMA o servidor que terminou de montar: oito RTX 3090 compradas no começo de 2026, segundo ele por 600 a 700 libras cada, "antes de os preços enlouquecerem", com duas fontes, refrigeração líquida na CPU e limite de 270 W por placa. Com um Qwen quantizado no vLLM e no opencode, ele diz ter tido cerca de 60 tokens por segundo com um agente e pico de 370 tokens por segundo somados com dez agentes em paralelo.

> **Por que isso importa:** Mostra até onde vai quem prefere rodar IA em casa, num momento em que GPUs ficaram mais caras e difíceis de achar.

Post original (Reddit): r/LocalLLaMA https://www.reddit.com/r/LocalLLaMA/comments/1xwc3m7/ai_rig_finally_done_and_stable_perfect_winter/

Vídeos

4 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.



Petition To 'Bring Jensen Back' Signed By Musk, Top CEOs & Employees; Nvidia Board Under Fire | N18G

Firstpost · 14/11/2026

Em cerca de cinco minutos, o Firstpost explica a campanha "Stand With Jensen", divulgada por Elon Musk, que pede ao conselho da NVIDIA devido processo, uma audiência justa e a volta de Jensen Huang enquanto corre a investigação independente.

https://www.youtube.com/watch?v=7Xm72n1uppI

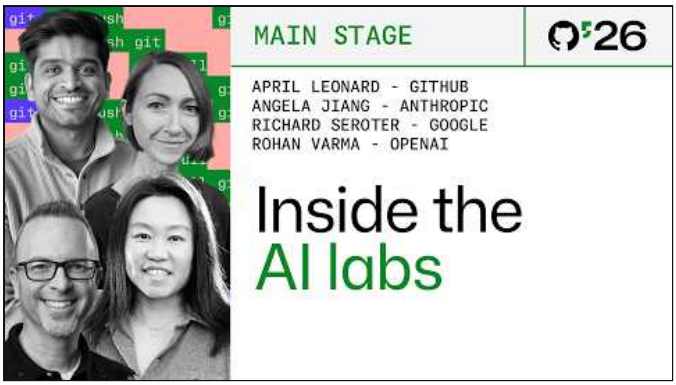


Special Report: The Fall Of Jensen Huang & What Comes Next

Alex Kantrowitz · 13/11/2026

Alex Kantrowitz e Ranjan Roy, do Margins, discutem por cerca de uma hora a saída de Huang, os chips restritos enviados à China, os números da NVIDIA em Singapura, a concorrência e quem pode suceder o fundador.

https://www.youtube.com/watch?v=NPLDSyxxt-0



The view from the labs: What's next for AI-assisted development

GitHub · 14/11/2026

Painel gravado no GitHub Universe 2026 com líderes da Anthropic, da OpenAI, do Google Cloud e do GitHub Copilot sobre os próximos dois anos do desenvolvimento com IA: modelos de fronteira, agentes autônomos, integrações MCP e segurança de código.

https://www.youtube.com/watch?v=eYPxL3pTylk



Watch me build a full AI app (no code)

Google Cloud Tech · 14/11/2026

Episódio do Build It em que o plano de implementação é entregue ao Google Antigravity, que monta passo a passo um app de personal trainer com IA; o vídeo mostra os erros no caminho e como investigá-los pelo console do navegador.

https://www.youtube.com/watch?v=-nRo5q63b2Y

SOBRE ESTA EDIÇÃO
Nº 30 — Edição da noite · 14/11/2026

JANELA
Cobre o que saiu desde a edição nº 29 (fechada por volta das 8h de Brasília de 14/11) até por volta das 18h de Brasília de 14/11, mais pautas de 12 e 13/11 que ainda não tinham saído (a carta de Gerstner e Baker ao conselho da NVIDIA, a nova regra do Project Zero, o /teamwork do Google, o guia do Opus 5.5 e a Série C da Databento).

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. **Petition To 'Bring Jensen Back' Signed By Musk, Top CEOs & Employees; Nvidia Board Under Fire | N18G** – Firstpost, 14/11/2026
   https://www.youtube.com/watch?v=7Xm72n1uppI
2. **Special Report: The Fall Of Jensen Huang & What Comes Next** – Alex Kantrowitz, 13/11/2026
   https://www.youtube.com/watch?v=NPLDSyxxt-0
3. **The view from the labs: What's next for AI-assisted development** – GitHub, 14/11/2026
   https://www.youtube.com/watch?v=eYPxL3pTylk
4. **Watch me build a full AI app (no code)** – Google Cloud Tech, 14/11/2026
   https://www.youtube.com/watch?v=-nRo5q63b2Y

Manchete

Nvidia board urged to consider 'potential return of Jensen Huang' by two major investors – CNBC
https://www.cnbc.com/2026/11/14/nvidia-investors-jensen-huang-return-board-letter.html

A petition to reinstate Jensen Huang as Nvidia CEO has over 100,000 signatures, including some big names – Business Insider
https://www.businessinsider.com/petition-reinstate-jensen-huang-nvidia-elon-musk-2026-11

Brad Gerstner publica a carta enviada ao conselho da NVIDIA – X (@altcap)
https://x.com/altcap/status/2121389561414623679

Modelos

Atualização: Gemini 4 Carbon assume o 1º lugar no MathArena e tira o GPT-6.1 Sol do topo, mas custou cerca de 15 vezes mais nas configurações do teste
The Decoder: https://the-decoder.com/gemini-4-carbon-takes-the-matharena-crown-but-costs-15-times-more-than-gpt-6-1-sol/
Post de Jasper Dekoninck (MathArena): https://x.com/j_dekoninck/status/2121263753006374997

Pesquisa

Google diz que agentes do /teamwork, no Antigravity, ajudaram a provar a conjectura de Courtade–Kumar, aberta desde 2013, numa prova de 50 milhões de linhas em Lean
Google Research: https://research.google/blog/advancing-science-engineering-with-antigravitys-teamwork-feature/

Produtos

Anthropic recomenda tirar o "pense bem" dos prompts do Claude Opus 5.5, agora o modelo padrão dos planos pagos, e começar no esforço médio
Anthropic (Claude): https://claude.com/resources/articles/using-claude-opus-5-5-as-your-daily-driver-for-knowledge-work
The Decoder: os hábitos que a Anthropic manda abandonar: https://the-decoder.com/anthropic-says-outdated-prompting-habits-are-slowing-down-claude-opus-5-5/

Canonical apresenta o Inference Manager, ferramenta do Ubuntu que instala e serve modelos de IA locais por uma interface compatível com a API da OpenAI
heise online: https://www.heise.de/news/Inference-Manager-neues-Ubuntu-Tool-verwaltet-lokale-KI-Modelle-11525275.html

Negócios e Investimento

Foundation, startup de robôs humanoides militares apoiada por Eric Trump, não paga salários desde 30/10, e o CEO tomou US$ 4,6 milhões emprestados da empresa para comprar uma casa, diz a CNBC
CNBC: https://www.cnbc.com/2026/11/14/foundation-payroll-ceo-loan.html

Regulação

Atualização: Malásia prepara lei que obriga data centers a declarar GPUs importadas e consumo de energia para pegar desvio de chips para a China
Tom's Hardware (com o Malay Mail): https://www.tomshardware.com/tech-industry/artificial-intelligence/malaysia-plans-new-law-using-data-center-power-consumption-to-catch-ai-gpu-smugglers-but-real-world-tdp-variance-and-lack-of-physical-audits-could-undermine-it-new-law-also-forces-operators-to-pay-for-unused-electricity-capacity

Open Source

Stanford abre o Memorilla, módulo de memória que resume uma coleção de documentos em 16 tokens para um LLM sem retreino e supera o contexto completo no FactKG
MarkTechPost: https://www.marktechpost.com/2026/11/14/stanford-memorilla-memory-system-16-memory-tokens/
Artigo no arXiv: https://arxiv.org/abs/2610.21375
Código no GitHub: https://github.com/snap-stanford/memorilla

tiny-voca: voz sintética totalmente aberta, com 4,5 milhões de parâmetros e 18 MB, que roda 70 vezes mais rápido que o tempo real numa CPU
GitHub / r/LocalLLaMA: https://github.com/n4ze3m/tiny-voca
Pesos no Hugging Face: https://huggingface.co/n4ze3m/tiny-voca
Post no r/LocalLLaMA: https://www.reddit.com/r/LocalLLaMA/comments/1xw7zrg/tinyvoca_a_fully_open_45mparameter_tts_code/

Agentes e Código

Project Zero, do Google, encurta o prazo para divulgar falhas por causa da IA: 90 dias corridos agora, 60 em fevereiro de 2027 e 30 em janeiro de 2028
Google Project Zero: https://projectzero.google/2026/11/disclosure-policy.html

Agentes de IA e o verificador Lean acham 138 erros de redação no código tributário dos EUA, diz a Lexifina
Lexifina: https://lexifina.com/blog/formalizing-the-us-tax-code

Modelos Locais

tandem.cpp, patches para o llama.cpp, roda o DeepSeek-V4.1 Flash num PC com 64 GB de RAM e placa de 16 GB a 14,2 tokens por segundo, quase o dobro do llama.cpp puro

GitHub / r/LocalLLaMA: https://github.com/ewhacc/tandem.cpp

Post no r/LocalLLaMA: https://www.reddit.com/r/LocalLLaMA/comments/1xwan2j/update_tandemcpp_deepseekv41_on_a_64_gb_pc_with_a/

Advogado que montou servidor de IA com oito GPUs V100 usadas passa 80% do DeepSeek V4.1 Flash para 4 bits (NVFP4) e acrescenta 1 TB de RAM e uma RTX 3090

Tom's Hardware: https://www.tomshardware.com/tech-industry/artificial-intelligence/lawyer-upgrades-eight-v100-private-ai-server-to-run-80-percent-of-deepseek-v4-1-flash-at-4-bit-nvfp4-precision-adds-1tb-of-ram-an-nvidia-rtx-3090-and-fixes-crashes-caused-by-a-bad-psu

Discussão no r/LocalLLaMA: https://www.reddit.com/r/LocalLLaMA/comments/1xw2ogq/the_lawyer_who_posts_here_about_his_v100s_got_his/

Finanças

Databento, fornecedora de dados de mercado com contratos com cinco dos seis maiores laboratórios de IA, capta US$ 100 milhões e passa a valer US$ 1,1 bilhão

Databento: https://databento.com/blog/series-c

FinTech Global sobre a rodada: https://fintech.global/2026/11/13/databento-hits-1-1bn-valuation-with-100m-series-c/

Games

A cinco dias do GTA 6, CEO da Take-Two diz que jogos clonados por IA não ameaçam o negócio e cobra que Anthropic e OpenAI protejam a propriedade intelectual dos outros

Insider Gaming (com o IGN): https://insider-gaming.com/take-two-ai-game-cloning-creative-concern/

“Não”: divulgador da GeForce na NVIDIA responde a quem pergunta se a saída de Huang é o fim das placas de vídeo para jogos

Clubic: https://www.clubic.com/actualite-639313-jacob-freeman-assure-que-nvidia-va-poursuivre-le-business-des-gpu-gaming.html

Brasil

Presidente do Porto Digital diz que a IA abre chance de o Brasil virar “celeiro de aplicações”, mas que o país precisa decidir se quer ser fornecedor ou comprador

Agência Brasil: https://agenciabrasil.ebc.com.br/geral/noticia/2026-11/porto-digital-com-ia-brasil-pode-se-tornar-celeiro-de-aplicacoes

Antes da Black Friday, estudo da Podqi acha 351 sites falsos entre as lojas que ChatGPT, Gemini e Perplexity indicaram em buscas de compras

Showmetech: https://www.showmetech.com.br/chatgpt-gemini-lojas-falsas-risco/

Fofocas das redes

Ex-pesquisador da Anthropic diz que um slide do roadshow do IPO pesou na decisão de sair: “eles já poderiam retreinar o Mythos com uma fração do cômputo”

@hilbertspaess: https://x.com/hilbertspaess/status/2121331333654086124

The Decoder sobre o post: https://the-decoder.com/anthropic-can-already-retrain-mythos-with-a-fraction-of-the-compute-originally-needed-says-former-researcher-coxon/

Resposta de Coxon pedindo a pausa: https://x.com/hilbertspaess/status/2121425977876341092

Musk diz que muitos acordos de Huang na NVIDIA eram “um aperto de mão, não um contrato escrito”, e que alguns vão se perder sem ele

@elonmusk: https://x.com/elonmusk/status/2121613402846204253

Resposta de Musk sobre a palavra de Huang: https://x.com/elonmusk/status/2121624209675665519

Valor Econômico sobre o post: https://valor.globo.com/empresas/noticia/2026/11/14/musk-afirma-que-nvidia-perder-negcios-crticos-sem-jensen-huang.ghtml

Estudo sobre apps feitos no Lovable que deixam passaportes e mais de 200 mil mensagens privadas abertos passa de 390 pontos no Hacker News

armcat: https://news.ycombinator.com/item?id=50499017

O estudo da CrissyField: https://www.crissyfield.de/research/lovable-en/

“Aquecedor perfeito para o inverno”: usuário mostra servidor caseiro com oito RTX 3090 rodando agentes em paralelo

r/LocalLLaMA: https://www.reddit.com/r/LocalLLaMA/comments/1xwc3m7/ai_rig_finally_done_and_stable_perfect_winter/

42 links únicos desta edição, sem repetir URLs.