



TreScout

RELATÓRIO DIÁRIO DE TECNOLOGIA

As tendências de tecnologia de hoje

4 de junho de 2026 (quinta-feira)

EM POUCOS GOLES

A agenda tecnológica atual é moldada em torno de ferramentas que automatizam processos de processamento de documentos e arquiteturas de agentes autônomos desenvolvidas para grandes modelos de linguagem. Os scanners de segurança de código aberto e as melhorias de interface concentram-se em aumentar a capacidade de processamento de dados e a segurança operacional dos sistemas.

24 EM DESTAQUE · ~6 MIN DE LEITURA

14 [GitHub Trending](#)

5 [Hacker News](#)

5 [HuggingFace](#) · modelos do dia

01 chopratejas/headroom

O headroom reduz o uso de tokens em 60 a 95 por cento ao compactar arquivos de log, saída de código e fragmentos de recuperação de informações (RAG) enviados para grandes modelos de linguagem (LLM). A biblioteca visa reduzir custos de processamento de dados com opções de servidor proxy e Model Context Protocol (MCP).

Python · ★ 10.887 · +3.530 hoje <https://github.com/chopratejas/headroom>

02 affaan-m/ECC

O ECC fornece um sistema de desempenho de agente que otimiza as camadas de capacidade, memória e segurança para ferramentas de desenvolvedor como Claude Code e Cursor. O sistema foi projetado para tornar mais eficientes os fluxos de trabalho dos agentes de inteligência artificial utilizados nos processos de desenvolvimento de software.

JavaScript · ★ 206.286 · +2.141 hoje <https://github.com/affaan-m/ECC>

03 aquasecurity/trivy

Trivy é uma ferramenta abrangente de verificação de segurança que detecta vulnerabilidades, configurações incorretas e segredos em contêineres, clusters Kubernetes e infraestruturas em nuvem. Ele automatiza auditorias de segurança em processos de desenvolvimento com a capacidade de criar listas de materiais de software (SBOM).

Go · ★ 35.511 · +24 hoje <https://github.com/aquasecurity/trivy>

04 NousResearch/hermes-agent

Desenvolvido pela NousResearch, o Hermes Agent oferece uma infraestrutura de agente autônomo que aprende com as interações do usuário e se desenvolve ao longo do tempo. Este sistema baseado em Python usa aprendizado contínuo para criar fluxos de trabalho personalizados.

Python · ★ 179.808 · +1.735 hoje <https://github.com/NousResearch/hermes-agent>

05 microsoft/markitdown

Desenvolvido pela Microsoft, MarkItDown é uma ferramenta Python que converte vários formatos de arquivo e documentos de escritório em formato de linguagem markdown (Markdown). Essa ferramenta facilita a tradução de dados não estruturados em formatos de texto mais processáveis para modelos de linguagem grandes.

Python · ★ 143.384 · +1.984 hoje <https://github.com/microsoft/markitdown>

06 nesquena/hermes-webui

Hermes WebUI é uma interface desenvolvida para gerenciar agentes Hermes através de navegadores web ou dispositivos móveis. Esta ferramenta baseada em Python fornece uma plataforma que facilita a interação dos usuários com agentes de inteligência artificial.

Python · ★ 13.256 · +719 hoje <https://github.com/nesquena/hermes-webui>

07 D4Vinci/Scrapling

Scrapling é uma estrutura adaptativa de web scraping que gerencia todo o processo, desde solicitações únicas até operações abrangentes de rastreamento. Esta ferramenta baseada em Python foi projetada para simplificar tarefas complexas de coleta de dados.

Python · ★ 60.629 · +1.067 hoje <https://github.com/D4Vinci/Scrapling>

08 opendataloader-project/opendataloader-pdf

OpenDataLoader PDF é um analisador de PDF de código aberto que disponibiliza dados para modelos de inteligência artificial. Este projeto baseado em Java acelera os processos de processamento de dados, automatizando a acessibilidade de documentos PDF.

Java · ★ 23.530 · +570 hoje <https://github.com/opendataloader-project/opendataloader-pdf>

09 odoo/odoo

Odoo é uma plataforma de planejamento de recursos empresariais de código aberto que permite às empresas gerenciar todos os seus processos operacionais sob o mesmo teto. Desenvolvido em linguagem Python, este sistema oferece uma ampla gama de aplicações modulares de negócios, desde vendas até contabilidade.

Python · ★ 52.082 · +29 hoje <https://github.com/odoo/odoo>

10 Open-LLM-VTuber/Open-LLM-VTuber

Open-LLM-VTuber suporta interação de voz sem usar as mãos com grandes modelos de linguagem e suporte para animação de personagens ao vivo (Live2D) que permite a execução de interrupção de voz em sistemas locais. Este projeto, que possui compatibilidade multiplataforma, automatiza processos de publicação virtual apoiados por inteligência artificial.

Python · ★ 9.185 · +693 hoje <https://github.com/Open-LLM-VTuber/Open-LLM-VTuber>

11 jwasham/coding-interview-university

A Universidade de Entrevista de Codificação oferece um plano de estudo abrangente que cobre os fundamentos da ciência da computação para aqueles que desejam se tornar engenheiros de software. Este recurso combina o conhecimento teórico e prático necessário na preparação para entrevistas técnicas com um currículo estruturado.

★ 349.290 · +330 hoje <https://github.com/jwasham/coding-interview-university>

12 **lyogavin/airllm**

AirLLM permite que grandes modelos de linguagem com 70 bilhões de parâmetros sejam executados em unidades de processamento gráfico com apenas 4 GB de memória de vídeo (VRAM). Esta biblioteca utiliza técnicas de otimização de memória para permitir o uso de modelos de alta capacidade com baixos requisitos de hardware.

[Jupyter Notebook](#) · ★ 19.113 · +208 hoje <https://github.com/lyogavin/airllm>

13 **supermemoryai/supermemory**

Supermemory oferece um mecanismo de memória escalonável e uma interface de aplicativo para a era da inteligência artificial. Essa estrutura, desenvolvida em TypeScript, fornece uma interface de programação de memória especializada (API Memory) para rápido processamento e gerenciamento de dados.

[TypeScript](#) · ★ 25.363 · +600 hoje <https://github.com/supermemoryai/supermemory>

14 **HKUDS/Vibe-Trading**

A Vibe-Trading oferece um agente comercial pessoal desenvolvido para negociação nos mercados financeiros. O projeto permite aos usuários gerenciar estratégias de negociação automáticas com sua estrutura baseada em Python.

[Python](#) · ★ 10.343 · +197 hoje <https://github.com/HKUDS/Vibe-Trading>

01 Gemma 4 12B: A unified, encoder-free multimodal model

O Google anunciou o Gemma 2 12B, um modelo multimodal de inteligência artificial que pode processar texto e dados visuais em uma única arquitetura. Graças à sua estrutura sem codificador, o modelo visa criar respostas mais consistentes processando diretamente entradas visuais e textuais.

814 pontos · **321 comentários**

<https://blog.google/innovation-and-ai/technology/developers-tools/introducing-gemma-4-12b/>

02 Elixir v1.20: Now a gradually typed language

A versão Elixir v1.20 adiciona suporte de digitação graduada ao núcleo da linguagem. Esta atualização visa reduzir erros de tempo de execução, permitindo que os desenvolvedores executem verificação de tipo estático em suas bases de código.

730 pontos · **264 comentários**

<https://elixir-lang.org/blog/2026/06/03/elixir-v1-20-0-released/>

03 I was recently diagnosed with anti-NMDA receptor encephalitis

O desenvolvedor de software Andrew Gallant compartilha sua recuperação de saúde após ser diagnosticado com encefalite anti-receptor NMDA e os efeitos desta rara condição neurológica na vida diária. Esta narrativa, centrada na experiência pessoal, chama a atenção para a complexidade dos processos de diagnóstico médico e para as repercussões dos problemas de saúde na vida profissional.

596 pontos · **181 comentários**

<https://burntsushi.net/encephalitis/>

04 "They're made out of weights"

Este estudo, que examina o princípio de funcionamento de grandes modelos de linguagem por meio de pesos, explica que os complexos processos de raciocínio dos sistemas de inteligência artificial consistem, na verdade, na interação de coeficientes matemáticos. O artigo fornece uma visão técnica dos mecanismos de processamento de dados subjacentes às capacidades de aprendizagem e previsão dos modelos.

562 pontos · **199 comentários**

<https://maxleiter.com/blog/weights>

05 Uber's \$1,500/month AI limit is a useful signal for AI tool pricing

A Uber está estabelecendo uma nova referência em estratégias de preços de ferramentas de software ao impor um limite de gastos mensais de US\$ 1.500 no uso empresarial de IA. Esta prática mostra que os modelos de preços baseados na utilização, que as empresas utilizam para manter os seus custos operacionais sob controle, estão a tornar-se generalizados.

467 pontos · **580 comentários**

<https://simonwillison.net/2026/Jun/3/uber-caps-usage/>

01 nvidia/LocateAnything-3B

Desenvolvido pela NVIDIA, LocateAnything-3D é um modelo de imagem texto para texto que executa tarefas de localização de objetos processando entradas visuais e textuais. O modelo foi projetado para identificar objetos em cenas complexas e apresentar as coordenadas desses objetos como saída baseada em texto.

♥ 1.210 · 91.834 downloads <https://huggingface.co/nvidia/LocateAnything-3B>

02 LiquidAI/LFM2.5-8B-A1B

Desenvolvido pela LiquidAI, o LFM2.5-8B-A1B foi apresentado como um modelo focado na geração de texto. O modelo visa proporcionar eficiência nos processos de processamento de linguagem com sua arquitetura computacional de alto desempenho.

♥ 486 · 72.114 downloads <https://huggingface.co/LiquidAI/LFM2.5-8B-A1B>

03 HauhauCS/Qwen3.6-35B-A3B-Uncensored-HauhauCS-Aggressive

O modelo Qwen3.6-35B-A3B-Uncensored-HauhauCS-Aggressive compartilhado no HuggingFace é um modelo de inteligência artificial que produz saída de texto processando dados visuais e de texto (imagem-texto-para-texto). O modelo oferece a capacidade de analisar diferentes tipos de dados com sua estrutura sem censura.

♥ 1.368 · 2.646.756 downloads <https://huggingface.co/HauhauCS/Qwen3.6-35B-A3B-Uncensored-HauhauCS-Aggressive>

04 openbmb/MiniCPM5-1B

Desenvolvido pela OpenBMB, MiniCPM5-1B é um modelo de linguagem de baixo parâmetro otimizado para tarefas de geração de texto. Esse modelo, que visa proporcionar alto desempenho em recursos computacionais limitados, está entre as soluções eficientes de inteligência artificial (IA).

♥ 762 · 79.427 downloads <https://huggingface.co/openbmb/MiniCPM5-1B>

05 google/gemma-4-12B-it

Desenvolvido pelo Google, o Gemma 2 12B oferece uma estrutura de modelo multimodal (qualquer para qualquer) que pode converter entre diferentes tipos de dados. Este modelo fornece uma experiência flexível de IA ao processar vários tipos de entrada, como texto e imagens.

♥ 274 · 14.866 downloads <https://huggingface.co/google/gemma-4-12B-it>
