



TreScout

RELATÓRIO DIÁRIO DE TECNOLOGIA

As tendências de tecnologia de hoje

13 de agosto de 2026 (quinta-feira)

EM POUCOS GOLES

No mundo do desenvolvimento de software, destacam-se novos estudos acadêmicos com foco em problemas de infraestrutura profundamente enraizados, como erros de banco de dados SQLite e evolução autônoma de software. No que diz respeito aos grandes modelos de linguagem, os dados de desempenho das arquiteturas de nova geração, como DeepSeek e Qwen, estão amplamente disponíveis na comunidade de desenvolvedores.

33 EM DESTAQUE · ~10 MIN DE LEITURA

17 GitHub
Trending

4 Hacker
News

5 HuggingFace ·
modelos do dia

5 HuggingFace ·
artigos do dia

2 Lobsters

01 cathrynlavery/diagram-design

O design de diagramas, desenvolvido por Cathryn Lavery, inclui 29 tipos diferentes de diagramas editoriais preparados para a ferramenta de codificação apoiada por inteligência artificial Claude Code. Esta biblioteca apresenta diagramas complexos diretamente usando HTML e gráficos vetoriais escaláveis (SVG) em vez de ferramentas de diagramação padrão.

HTML · ★ 11.456 · +2.855 hoje <https://github.com/cathrynlavery/diagram-design>

02 macro-inc/macro

Macro é um espaço de trabalho que combina ferramentas de negócios como e-mail, chat, documentos e gerenciamento de tarefas com uma memória comum de inteligência artificial. Desenvolvido com a linguagem Rust, esta plataforma conecta diferentes fluxos de trabalho em uma única interface.

Rust · ★ 2.066 · +227 hoje <https://github.com/macro-inc/macro>

03 semantica-agi/semantica

Semantica é uma infraestrutura nativa de gráficos que fornece gerenciamento contextual de dados e responsabilidade em sistemas de inteligência artificial. Desenvolvida em linguagem Python, essa ferramenta ajuda modelos de inteligência artificial a produzir resultados mais consistentes e auditáveis por meio do gerenciamento de redes de informações complexas.

Python · ★ 5.886 · +845 hoje <https://github.com/semantica-agi/semantica>

04 stablyai/orca

Desenvolvido pela Stably AI, o Orca é um ambiente de desenvolvimento de agentes (ADE) que permite gerenciar vários agentes de inteligência artificial simultaneamente. Esta ferramenta, que funciona em desktop, celular e servidor virtual privado (VPS), permite utilizar diferentes agentes de codificação em paralelo com suas próprias assinaturas.

TypeScript · ★ 44.177 · +1.235 hoje <https://github.com/stablyai/orca>

05 msitarzewski/agency-agents

Agency-agents é uma estrutura de automação que reúne agentes de inteligência artificial com diferentes áreas de atuação. Esses agentes especializados, que realizam diversas tarefas como design, software e gerenciamento de comunidade, visam gerenciar os processos de negócios de ponta a ponta.

Shell · ★ 144.741 · +1.873 hoje <https://github.com/msitarzewski/agency-agents>

06 shiyu-coder/Kronos

Kronos é um modelo básico projetado para compreender e analisar a linguagem complexa dos mercados financeiros. Oferece uma infraestrutura de inteligência artificial especializada no processamento de dados financeiros e na interpretação dos movimentos do mercado.

Python · ★ 36.993 · +266 hoje <https://github.com/shiyu-coder/Kronos>

07 NanmiCoder/MediaCrawler

MediaCrawler é um web scraper baseado em Python para extrair conteúdo e dados de comentários de plataformas populares de mídia social (como Xiaohongshu, TikTok, Bilibili, Weibo). Oferece aos usuários a oportunidade de coletar informações de postagens e interações em diferentes plataformas por meio de uma estrutura central.

Python · ★ 62.109 · +215 hoje <https://github.com/NanmiCoder/MediaCrawler>

08 hugohe3/ppt-master

O ppt-master baseado em Python é uma ferramenta de inteligência artificial que converte documentos ou tópicos em arquivos de apresentação editáveis (PowerPoint). O software permite que os usuários usem seus próprios modelos, transformando texto em apresentações profissionais com formas nativas, efeitos de transição, animações e gráficos baseados em dados.

Python · ★ 45.916 · +476 hoje <https://github.com/hugohe3/ppt-master>

09 infiniflow/ragflow

RAGFlow é um mecanismo de geração baseada em busca (RAG) de código aberto que cria uma camada de contexto para grandes modelos de linguagem (LLM). O objetivo é melhorar o processamento de dados e a precisão das respostas, combinando técnicas avançadas de RAG com capacidades de agente.

Go · ★ 87.648 · +139 hoje <https://github.com/infiniflow/ragflow>

10 paperclipai/paperclip

Paperclip é um aplicativo de código aberto que permite gerenciar agentes de IA no local de trabalho por meio de uma única interface. Desenvolvido em TypeScript, esta ferramenta permite controlar os fluxos de trabalho de diferentes agentes a partir de um ponto central.

TypeScript · ★ 77.843 · +571 hoje <https://github.com/paperclipai/paperclip>

11 NVIDIA-NeMo/Switchyard

Desenvolvido pela NVIDIA, Switchyard é um mecanismo de inferência de inteligência artificial de alto desempenho escrito em linguagem Rust. Ele oferece um ambiente de tempo de execução otimizado para executar modelos de linguagem grandes (LLM) com eficiência em diferentes infraestruturas de hardware.

Rust · ★ 923 · +421 hoje <https://github.com/NVIDIA-NeMo/Switchyard>

12 ZuodaoTech/everyone-can-use-english

Everyone-can-use-english é um projeto educacional de código aberto que apoia o processo de aprendizagem de inglês com um currículo estruturado. Oferece métodos práticos e recursos de estudo para usuários que desejam melhorar suas habilidades no idioma.

TypeScript · ★ 36.233 · +86 hoje <https://github.com/ZuodaoTech/everyone-can-use-english>

13 smicallef/spiderfoot

SpiderFoot é uma ferramenta que automatiza a coleta de inteligência de código aberto (OSINT) para inteligência de ameaças e processos de mapeamento de superfície de ataque. Este software baseado em Python acelera a análise de segurança coletando dados sobre os sistemas alvo.

Python · ★ 20.413 · +74 hoje <https://github.com/smicallef/spiderfoot>

14 localsend/localsend

LocalSend é um aplicativo de código aberto que permite o compartilhamento de arquivos entre diferentes sistemas operacionais. Ele cria uma alternativa independente de plataforma ao recurso AirDrop oferecido pelos dispositivos Apple.

Dart · ★ 87.961 · +213 hoje <https://github.com/localsend/localsend>

15 Lightricks/LTX-2

LTX-2, desenvolvido pela Lightricks, é um modelo de inteligência artificial (modelo generativo) que produz áudio e vídeo. Este pacote fornece as ferramentas Python necessárias para executar os processos de inferência do modelo e ajustá-lo usando o método de adaptação de baixa ordem (LoRA).

Python · ★ 8.759 · +65 hoje <https://github.com/Lightricks/LTX-2>

16 embabel/embabel-agent

Embabel é um framework projetado para desenvolver agentes de inteligência artificial rodando na Java Virtual Machine (JVM). Esta estrutura, escrita em Kotlin, facilita a integração de capacidades generativas de IA em ecossistemas de software empresarial.

Kotlin · ★ 4.255 · +40 hoje <https://github.com/embabel/embabel-agent>

17 cactus-compute/needle

Desenvolvido pela Cactus Compute, o Needle oferece um modelo básico de 14 MB que pode ser executado em pequenos hardwares, como telefones, dispositivos vestíveis e robôs. Essa estrutura leve visa executar aplicativos de inteligência artificial localmente em dispositivos de ponta com poder de processamento limitado.

Python · ★ 4.412 · +315 hoje <https://github.com/cactus-compute/needle>

01 Tailscale Traces Database Corruption to 16y/o SQLite WAL-Reset Bug

A ferramenta de segurança de rede Tailscale descobriu que um erro de redefinição de log de gravação antecipada do SQLite de 16 anos estava na raiz da corrupção do banco de dados. A empresa eliminou o risco de integridade dos dados ao lançar a atualização que corrigiu esse problema técnico.

914 pontos · 177 comentários <https://tailscale.com/blog/sqlite-wal-reset-bug>

Discussão: news.ycombinator.com/item?id=49272832 (914 votes · 177 commentaires) · lobste.rs/s/e0lkmi/how_tailscale_helped_find_sqlite_wal (91 votes · 3 commentaires)

02 DeepSeek V4 Pro 0813

DeepSeek V4 Pro 0813 é um novo modelo de linguagem disponível através da plataforma OpenRouter que fornece acesso a modelos de IA multimodais. Este modelo visa demonstrar alto desempenho em tarefas complexas de inferência lógica e codificação.

831 pontos · 330 comentários <https://openrouter.ai/deepseek/deepseek-v4-pro-0813>

Discussão: news.ycombinator.com/item?id=49274600 (831 votes · 330 commentaires)

03 AI is removing the middle class of software engineering?

As ferramentas de inteligência artificial reduzem o espaço de trabalho dos desenvolvedores de software de nível médio, automatizando tarefas básicas nos processos de desenvolvimento de software. Esta mudança traz consigo uma nova estrutura de força de trabalho que permite às empresas alcançar elevada eficiência com menos engenheiros seniores.

812 pontos · 740 comentários <https://blog.florianherrengt.com/ai-removing-middle-class-software-engineering.html>

Discussão: news.ycombinator.com/item?id=49271994 (812 votes · 740 commentaires)

04 License plate reader searches should require a warrant

Está sendo discutido que buscas de rastreamento de veículos feitas através de leitores automáticos de placas deveriam estar sujeitas a um mandado de busca legal com base na privacidade pessoal. A tônica está nos riscos colocados pela utilização descontrolada dos dados recolhidos por estes sistemas sobre a liberdade de circulação dos indivíduos.

568 pontos · 336 comentários <https://andrewpwheeler.com/2026/08/12/license-plate-reader-searches-should-require-a-warrant/>

Discussão: news.ycombinator.com/item?id=49273165 (568 votes · 336 commentaires)

01 Qwen/Qwen3.8-2.4T-A95B

Desenvolvido pela Alibaba, Qwen3.8-2.4T-A95B é um grande modelo de linguagem focado na geração de texto. Este modelo, compartilhado na plataforma Hugging Face, é utilizado em diversas tarefas de processamento de linguagem natural com sua estrutura treinada em grandes conjuntos de dados.

♥ 587 · 978 downloads <https://huggingface.co/Qwen/Qwen3.8-2.4T-A95B>

Discussão: news.ycombinator.com/item?id=49273478 (558 votes · 128 commentaires)

02 meta-models/Muse-Glimmer-30B

Muse-Glimmer-30B, publicado no HuggingFace, é um modelo multimodal de IA que processa dados visuais e de texto e produz resultados baseados em texto. O modelo analisa e dá sentido a conteúdos visuais complexos e gera respostas baseadas em texto nas interações do usuário.

♥ 1.324 · 0 downloads <https://huggingface.co/meta-models/Muse-Glimmer-30B>

03 MiniMaxAI/MiniMax-H3

MiniMax-H3, desenvolvido pela MiniMax, é um modelo de inteligência artificial que produz vídeos usando entradas visuais e de texto. Esta tecnologia oferece a capacidade de converter conteúdo estático ou comandos escritos fornecidos pelos usuários em imagens em movimento.

♥ 3.747 · 1.575.808 downloads <https://huggingface.co/MiniMaxAI/MiniMax-H3>

04 larryvrh/MiniMax-H3-Turbo-Lora

MiniMax-H3-Turbo-Lora é um modelo de adaptação de baixa classificação usado para melhorar os recursos de texto para vídeo. Este arquivo compartilhado no HuggingFace é um peso de ajuste fino para otimizar o desempenho de produção de vídeo de modelos de inteligência artificial existentes.

♥ 705 · 0 downloads <https://huggingface.co/larryvrh/MiniMax-H3-Turbo-Lora>

05 deepseek-ai/DeepSeek-V4-Flash-0731

DeepSeek-V4-Flash-0731 é um modelo de geração de texto desenvolvido pela DeepSeek. Este modelo, hospedado no Hugging Face, oferece uma estrutura otimizada para tarefas de processamento de linguagem de alto desempenho.

♥ 3.255 · 1.048.685 downloads <https://huggingface.co/deepseek-ai/DeepSeek-V4-Flash-0731>

01 Persistent Recursive Worlds Enable Autonomous Software Evolution

EvoX Genesis define projetos de software como mundos persistentes e recursivos, garantindo um processo de desenvolvimento contínuo apesar da vida útil limitada dos agentes de inteligência artificial. Esta abordagem mantém o próprio software como uma estrutura centralizada, permitindo que agentes de codificação locais avancem no projeto com tarefas ad-hoc.

1 curtidas · 1 comentários <https://huggingface.co/papers/2608.10450>

02 OpenART: Scaling Agent Red Teaming via Open-Ended Environment Evolution

OpenART é um ambiente de teste aberto desenvolvido para avaliar os riscos cumulativos de agentes de inteligência artificial em fluxos de trabalho de longo prazo. Ao contrário dos tradicionais testes estáticos de benchmark, visa detectar vulnerabilidades (red teaming) de forma escalável, simulando as interações dos agentes em uma situação compartilhada.

62 curtidas · 1 comentários <https://huggingface.co/papers/2608.00677>

03 AtlasVLA: Persistent World-Ego State Modeling for Vision-Language-Action Models

AtlasVLA é uma estrutura de modelo Visão-Linguagem-Ação que permite aos robôs lembrar objetos fora de vista e o progresso da tarefa. Este método permite que sistemas robóticos superem a limitação da resposta instantânea e gerenciem tarefas complexas e de longo prazo de forma mais consistente.

1 curtidas · 1 comentários <https://huggingface.co/papers/2608.06729>

04 Self-Geometry: GT-Free and Plug-and-Play Test-Time Adaptation for Geometrically Consistent 3D Vision Foundation Models

A autogeometria é um método que elimina inconsistências geométricas em dados como profundidade e posição da câmera produzidos por modelos de base de visão sem exigir treinamento adicional durante a fase de teste. Esta abordagem torna os resultados tridimensionais dos modelos mais consistentes, sem a necessidade de métodos tradicionais que criam custos computacionais complexos.

8 curtidas · 1 comentários <https://huggingface.co/papers/2608.10708>

05 Can LLM Agents Stick to the Script? A Benchmark for Long-Horizon Consistency in Interactive Narratives

Grandes modelos de linguagem têm dificuldade em manter a consistência lógica, apesar das intervenções do usuário na narrativa interativa. Desenvolvido por pesquisadores, o NCP-Bench oferece uma estrutura de avaliação padrão para medir a integridade narrativa de longo prazo desses modelos.

20 curtidas · 2 comentários <https://huggingface.co/papers/2608.08160>

01 I hate packaging my software for Linux

Os desenvolvedores de software enfrentam sérios problemas de compatibilidade no processo de empacotamento de aplicativos devido à diversidade das distribuições Linux. Diferentes requisitos de sistema e formatos de pacotes complicam os processos de implantação dos desenvolvedores, levando à perda de tempo.

71 pontos · 57 comentários <https://getfresh.dev/docs/blog/packaging-for-linux/>

Discussão: lobste.rs/s/ixbrmc/i_hate_packaging_my_software_for_linux (71 votes · 57 commentaires)

02 Guide to (not) fucking up QR codes

A utilização de códigos QR envolve detalhes técnicos que requerem atenção em termos de segurança e acessibilidade. Para reduzir os riscos de projeto defeituoso e de direcionamento incorreto, a clareza visual dos códigos e a segurança da conexão precisam ser verificadas.

83 pontos · 8 comentários <https://infosec.exchange/@rebane2001/117078420917152774>

Discussão: lobste.rs/s/omcyaz/guide_not_fucking_up_qr_codes (83 votes · 8 commentaires)

SVG

Arquivo de imagem digital desenhado com fórmulas matemáticas que pode ser ampliado ou reduzido conforme desejado sem perda de qualidade.

graph-native

Uma estrutura de sistema que funciona com base em relações complexas entre dados e coloca as conexões diretamente no seu centro.

ADE

Um ambiente de trabalho digital onde todas as ferramentas são utilizadas para acelerar os processos de desenvolvimento de software.

VPS

Um servidor virtual acessado pela internet, onde os recursos de um poderoso computador host são divididos e reservados especificamente para você.

framework

Um conjunto de bibliotecas e ferramentas que fornecem a estrutura básica e facilitam o desenvolvimento de software.

foundation model

Um modelo básico de inteligência artificial projetado para executar muitas tarefas diferentes, sendo treinado com uma ampla gama de dados.

web scraper

Uma ferramenta de software que copia automaticamente dados de sites e os organiza em uma lista organizada.

PowerPoint

Um programa de preparação de apresentações que permite apresentar informações em slides visuais.

LLM

Modelos de inteligência artificial treinados com enormes dados que podem compreender a linguagem humana e produzir texto.

RAG

Um método que permite que a inteligência artificial extraia informações de seus documentos atuais e privados enquanto gera respostas.

AI agents

Software autônomo de inteligência artificial que pode tomar decisões por conta própria e executar tarefas específicas.

inference

O processo de uma inteligência artificial treinada produzindo uma resposta a uma nova pergunta usando as informações que aprendeu.

runtime

O período durante o qual um software é executado no computador e executa ativamente sua tarefa.

OSINT

Um método de coleta e análise de informações de fontes disponíveis publicamente na Internet.

generative model

Inteligência artificial que pode criar novos conteúdos nunca vistos antes, do zero, usando os padrões que aprendeu.

LoRA

É uma técnica para especializar um grande modelo de inteligência artificial em um assunto específico, atualizando apenas uma pequena parte dele, sem alterar todo o modelo.
