BIS - MT · 11/06/2026

Bản Tin AI Hằng Ngày

Cập nhật công nghệ AI mới nhất

“Three things cannot be long hidden: the sun, the moon, and the truth.”

↪ Ba thứ không thể che giấu lâu: mặt trời, mặt trăng và sự thật.

— Phật Thích Ca

Sự thật và chính trực luôn là nền tảng bền vững nhất — dù có thể che giấu tạm thời, sự thật cuối cùng luôn được bộc lộ.

TIN TỨC NỔI BẬT

1

Các nhà nghiên cứu bảo mật phát tín hiệu cảnh báo về lỗ hổng trong mã được tạo bởi AI

Security Researchers Sound the Alarm on Vulnerabilities in AI-Generated Code

Infosecurity Magazine Đọc bài viết →

> Các nhà nghiên cứu bảo mật đã bày tỏ mối quan ngại về các lỗ hổng trong mã được tạo ra bởi AI, đang ngày càng được sử dụng trong phát triển phần mềm. Các nhà nghiên cứu cảnh báo rằng mã được tạo ra bởi AI có thể chứa các cửa hậu ẩn, lỗi bảo mật và các điểm yếu khác có thể làm tổn hại đến bảo mật hệ thống. Vấn đề này phát sinh khi các model AI được đào tạo trên dữ liệu không đầy đủ hoặc bị thiên vị, dẫn đến việc tạo ra mã không chỉ kém hiệu quả mà còn dễ bị tấn công. Hơn nữa, sự thiếu minh bạch và giải thích trong mã được tạo ra bởi AI khiến cho các developer khó xác định và sửa chữa các lỗ hổng này. Các nhà nghiên cứu nhấn mạnh rằng việc sử dụng mã được tạo ra bởi AI không phải là xấu, mà là do thiếu các quy trình kiểm tra và xác thực phù hợp. Họ khuyến nghị các developer nên triển khai các thủ tục kiểm tra và xác thực mạnh mẽ để đảm bảo bảo mật và độ tin cậy của mã được tạo ra bởi AI. Các phát hiện này nhấn mạnh nhu cầu về một cách tiếp cận tinh tế hơn đối với mã được tạo ra bởi AI, một cách tiếp cận cân bằng giữa lợi ích của tự động hóa với nhu cầu về bảo mật và độ tin cậy.

2

Mọi người đều muốn tham gia vào việc mã hóa theo cảm hứng - và Google cũng không ngoại lệ với Stitch, sản phẩm tiếp theo sau Jules

Everyone's looking to get in on vibe coding — and Google is no different with Stitch, its follow-up to Jules

Venturebeat Đọc bài viết →

Google đang bước vào không gian mã hóa vibe với Stitch, phát triển mới nhất của họ. Động thái này diễn ra khi nhiều công ty đang khám phá khái niệm mã hóa vibe, liên quan đến việc sử dụng AI để tạo mã. Stitch là một phần tiếp theo của dự án trước đó của Google, Jules, cho thấy sự tập trung liên tục vào công nghệ này. Chi tiết về Stitch hiện đang bị giới hạn, nhưng việc phát hành nó cho thấy cam kết của Google trong việc phát triển mã hóa vibe. Sự tham gia của gã khổng lồ công nghệ này vào không gian này có thể báo hiệu sự cạnh tranh và đổi mới tăng lên trong lĩnh vực này. Khi các công ty như Google đầu tư vào mã hóa vibe, có khả năng chúng ta sẽ thấy nhiều mã được tạo bởi AI hơn trong các ứng dụng khác nhau. Việc phát hành Stitch cũng đặt ra câu hỏi về tác động tiềm năng đối với ngành công nghiệp mã hóa và vai trò của các nhà phát triển con người. Mặc dù sự tham gia của Google có thể mang lại cơ hội và tiến bộ mới, nhưng nó cũng đặt ra lo ngại về việc thay thế công việc và tương lai của công việc mã hóa.

3

Lớp lệnh AI nhà kho đa agent cho phép đạt được sự xuất sắc trong hoạt động và thông minh chuỗi cung ứng

Multi-Agent Warehouse AI Command Layer Enables Operational Excellence and Supply Chain Intelligence

NVIDIA Developer Đọc bài viết →

NVIDIA đã giới thiệu Lớp lệnh AI Nhà kho Đa tác nhân, được thiết kế để nâng cao hiệu quả hoạt động và trí tuệ chuỗi cung ứng trong các nhà kho. Lớp lệnh này được hỗ trợ bởi AI cho phép giao tiếp liền mạch giữa các hệ thống nhà kho khác nhau, bao gồm robot, xe nâng và phần mềm quản lý hàng tồn kho. Lớp lệnh AI Nhà kho Đa tác nhân này sử dụng nền tảng Omniverse của NVIDIA để tạo ra một môi trường thống nhất cho việc chia sẻ dữ liệu và ra quyết định. Điều này cho phép các nhà điều hành nhà kho tối ưu hóa các quy trình làm việc, giảm thiểu sai sót và cải thiện năng suất tổng thể. Hệ thống cũng cung cấp khả năng hiển thị thời gian thực về mức hàng tồn kho, cho phép dự báo nhu cầu và lập kế hoạch chuỗi cung ứng chính xác hơn. Bằng cách tích hợp các khả năng AI và học máy, Lớp lệnh AI Nhà kho Đa tác nhân có thể phân tích dữ liệu từ các nguồn khác nhau để xác định các lĩnh vực

cần cải thiện và cung cấp thông tin chi tiết có thể hành động. Điều này cho phép các nhà điều hành nhà kho đưa ra quyết định dựa trên dữ liệu, tối ưu hóa hoạt động và nâng cao sự hài lòng của khách hàng. Giải pháp của NVIDIA nhằm mục đích cách mạng hóa quản lý nhà kho bằng cách cung cấp một môi trường hiệu quả, thông minh và kết nối hơn.

4

Tại sao tôi ngừng sử dụng Cursor cho mã sản xuất (Và tôi sử dụng gì bây giờ)

Why I Stopped Using Cursor for Production Code (And What I Use Now)

Dev.to AI Đọc bài viết →

Một nhà phát triển công nghệ đã chia sẻ kinh nghiệm của mình khi chuyển từ trình chỉnh sửa Cursor sang Claude Code cho các nhiệm vụ mã hóa sản xuất. Ban đầu, Cursor hiệu quả cho việc tạo mẫu, cung cấp các tính năng như hoàn thành tab sắc nét và chỉnh sửa trực tuyến. Tuy nhiên, khi các dự án của nhà phát triển ngày càng phức tạp, những hạn chế của Cursor đã trở nên rõ ràng. Nó không hiểu mối quan hệ giữa các tệp, làm hỏng các phần khác của cơ sở mã và không chạy các thử nghiệm hoặc tuân theo các quy ước đã thiết lập. Để bù đắp, nhà phát triển phải duy trì một quy trình làm việc thủ công, dán một tài liệu 600 từ về các quy tắc vào trình chỉnh sửa mỗi phiên. Điểm bùng phát xảy ra khi Cursor cập nhật không chính xác một lược đồ sản phẩm, gây ra thời gian ngừng hoạt động 20 phút. Nhà phát triển sau đó chuyển sang Claude Code, cung cấp một tính năng gọi là hooks. Những hooks này cho phép chạy các lệnh tự động tại các thời điểm cụ thể, chẳng hạn như trước khi cam kết hoặc sau khi chỉnh sửa tệp. Nhà phát triển đã thiết lập hooks để chạy trình kiểm tra cú pháp, kiểm tra từ và bộ thử nghiệm của họ, đảm bảo rằng các lỗi được phát hiện và sửa ngay lập tức. Điều này đã cải thiện đáng kể quy trình làm việc của họ, cho phép họ tập trung vào việc viết mã thay vì duy trì thủ công các quy tắc.

5

AGI không phải là đa phương thức

AGI Is Not Multimodal

The Gradient Đọc bài viết →

Những tiến bộ gần đây trong các mô hình AI tạo sinh đã khiến một số người tin rằng Trí tuệ Nhân tạo Tổng quát (AGI) đang sắp xảy ra. Tuy

nhiên, giả định này có thể là quá sớm. Sự thành công của những mô hình này có thể được quy cho khả năng của chúng trong việc mở rộng hiệu quả trên phần cứng hiện có, chứ không phải là những giải pháp sâu sắc cho vấn đề về trí tuệ. Phương pháp đa mô thức, bao gồm tối ưu hóa các mạng mô-đun lớn cho các mô thức khác nhau, là một chiến lược không thể dẫn đến AGI ở mức độ con người trong thời gian tới. Phương pháp này tập trung vào việc kết hợp các mô thức khác nhau, nhưng bỏ qua tầm quan trọng của việc thể hiện và tương tác với môi trường. Một cách tiếp cận hiệu quả hơn đến AGI sẽ là ưu tiên hiểu biết vật lý và trí tuệ được đặt trong tình huống, có thể giải quyết các vấn đề xuất phát từ thực tế vật lý, chẳng hạn như sửa chữa ô tô hoặc chuẩn bị thức ăn. Các Mô hình Ngôn ngữ Lớn (LLM) hiện tại có thể không học được một mô hình thực sự của thế giới, mà chỉ là các quy tắc tạm thời để dự đoán token, dẫn đến những ấn tượng sai lầm về trí tuệ của chúng.

6 Đang bị khóa trong sự cạnh tranh gay gắt với nhà nghiên cứu, Microsoft sửa lỗi 0 ngày mà họ đã tiết lộ

Locked in heated rivalry with researcher, Microsoft fixes 0-day they disclosed

Ars Technica Đọc bài viết →

Microsoft đã phát hành bản vá cho hai lỗ hổng zero-day mức độ nghiêm trọng được tiết lộ bởi một nhà nghiên cứu có tên Nightmare Eclipse. Nhà nghiên cứu này, người đã tham gia vào một cuộc tranh cãi công khai với Microsoft, trước đó đã phát hành mã concept chứng minh cho các lỗ hổng này, những lỗ hổng có khả năng bị khai thác trong tự nhiên. Một trong những lỗ hổng đã được vá, CVE-2026-45586, là một lỗi leo thang đặc quyền cục bộ có thể được kết hợp với một lỗ hổng riêng biệt để đạt được quyền SYSTEM đầy đủ. Microsoft tuyên bố rằng lỗ hổng này yêu cầu độ phức tạp tối thiểu để khai thác và có khả năng cao bị khai thác tích cực. Lỗ hổng đã được vá khác, MiniPlasma, ban đầu đã được sửa chữa sáu năm trước nhưng lại xuất hiện do một sự hồi quy hoặc bản vá không hoàn chỉnh. Microsoft vẫn chưa phát hành bản vá cho các lỗ hổng khác được tiết lộ bởi Nightmare Eclipse, bao gồm YellowKey, ảnh hưởng đến Windows Defender RedSun, và BlueHammer. Công ty đã cung cấp hướng dẫn thủ công để giảm thiểu YellowKey nhưng chưa sửa chữa nguyên nhân cơ bản của lỗ hổng.

7 Sự bất ngờ: GPT-5.5 đánh bại Claude Fable 5 trên điểm chuẩn mới gay gắt Agents' Last Exam

Surprise upset: GPT-5.5 beats Claude Fable 5 on brutal new Agents' Last Exam benchmark

VentureBeat Đọc bài viết →

Các nhà nghiên cứu từ Trung tâm Trí tuệ Phân quyền Có trách nhiệm của Đại học California, Berkeley đã ra mắt một điểm chuẩn mới gọi là Agents' Last Exam (ALE) để đo lường khả năng của trí tuệ nhân tạo (AI) trong việc thực hiện các quy trình làm việc chuyên nghiệp có giá trị kinh tế, tầm nhìn dài hạn. Điểm chuẩn này bao gồm 1.490 trường hợp nhiệm vụ, được lấy từ các chuyên gia trong ngành, và bao phủ 55 phân ngành công nghiệp phi vật lý. Trong một sự bất ngờ, GPT-5.5 của OpenAI đã chiếm vị trí đầu tiên trên Bảng xếp hạng ALE với tỷ lệ đậu 24,0%, vượt qua mô hình Claude Fable 5 của Anthropic với điểm số 22,0%. Kiến trúc đánh giá của ALE buộc các mô hình phải điều hướng các quy trình làm việc phức tạp, sử dụng nhận thức thị giác, gọi công cụ và nền tảng thời gian chạy, và từ chối mô hình "LLM-as-a-judge" để chấm điểm. Tính xác thực và đường cong chấm điểm nghiêm ngặt của điểm chuẩn này làm nổi bật những hạn chế của AI hiện tại, với hầu hết các mô hình không thể vượt qua kỳ thi, đặc biệt là ở cấp độ "Last-Exam" khó nhất. Dự án hoạt động như một sáng kiến nghiên cứu mã nguồn mở, nhưng bảo vệ nghiêm ngặt dữ liệu đánh giá của mình để ngăn chặn "ô nhiễm điểm chuẩn".

8 Các agent giọng nói có thể xử lý khách hàng song ngữ không? Đánh giá Frontier ASR trên giọng nói chuyển đổi mã

Can Voice Agents Handle Bilingual Customers? Benchmarking Frontier ASR on Code-Switched Speech

Hugging Face Blog Đọc bài viết →

Các nhà nghiên cứu đã tạo ra một điểm chuẩn để đánh giá cách các trợ lý giọng nói xử lý ngôn ngữ chuyển đổi, một hiện tượng phổ biến ở những người nói song ngữ. Chuyển đổi ngôn ngữ liên quan đến việc chuyển đổi mượt mà giữa các ngôn ngữ, thậm chí giữa câu, và phổ biến trong các cuộc trò chuyện hàng ngày, trung tâm liên lạc và bàn hỗ trợ CNTT. Điểm chuẩn tập trung vào nhận dạng giọng nói tự động (ASR), bước đầu tiên trong đường ống trợ lý giọng nói, vì lỗi chuyển đổi có thể có hậu quả đáng kể trong môi trường doanh nghiệp. Điểm chuẩn bao gồm bốn cặp ngôn ngữ: Tây Ban Nha-Anh, Pháp-Anh, Pháp Canada-Anh và Đức-Anh. Nó sử dụng một tập dữ liệu gồm 1.018 bản

> ghi chuyển đổi ngôn ngữ, bao gồm các kịch bản quản lý nguồn nhân lực (HR) và quản lý dịch vụ CNTT (ITSM) khác nhau. Để đo hiệu suất của model, các nhà nghiên cứu báo cáo ba chỉ số: Tỷ lệ lỗi từ (WER), Tỷ lệ lỗi từ ngữ nghĩa (SWER) và Tỷ lệ lỗi câu trả lời (AER). Kết quả cho thấy chi phí chuyển đổi ngôn ngữ khác nhau tùy thuộc vào cặp ngôn ngữ và model được thử nghiệm. Các model hoạt động tốt nhất trên các chỉ số là ElevenLabs Scribe V2, Gemini 3 Flash và Assembly AI Universal 3-Pro. Các nhà nghiên cứu phát hành điểm chuẩn và dữ liệu của họ thông qua AU-Harness, cung cấp một tài nguyên quý giá để đánh giá các model giọng nói.

TIPS & TRICKS CHO DEV

Tối ưu hóa truy vấn

Vấn đề: Truy vấn dữ liệu từ database chậm và không hiệu quả.
Cách làm: Sử dụng MCP để kết nối với database và tối ưu hóa truy vấn bằng cách sử dụng lệnh `Claude Code` với prompt "Optimize SQL query".
Đánh giá: Hiệu quả khi cần truy vấn dữ liệu lớn, tiết kiệm thời gian và tài nguyên.

Tích hợp Git

Vấn đề: Quản lý phiên bản code không hiệu quả.
Cách làm: Sử dụng MCP để kết nối với Git và quản lý phiên bản code bằng lệnh `git commit` với prompt "Commit changes".
Đánh giá: Hiệu quả khi cần quản lý phiên bản code, giúp theo dõi thay đổi và hợp tác.

Xử lý file system

Vấn đề: Xử lý file system không hiệu quả.
Cách làm: Sử dụng MCP để kết nối với file system và xử lý file bằng lệnh `Claude Desktop` với prompt "Create new file".
Đánh giá: Hiệu quả khi cần xử lý file lớn, tiết kiệm thời gian và tài nguyên.

BÀI HỌC AI HÔM NAY CHO DEV

1. **Tối ưu chi phí & hiệu năng LLM**
Dev cần biết cách tối ưu chi phí và hiệu năng LLM để đảm bảo ứng dụng AI hoạt động hiệu quả và tiết kiệm chi phí. Điều này đặc biệt quan trọng khi triển khai mô hình AI trên quy mô lớn.

2. Việc tối ưu hóa LLM giúp giảm thiểu chi phí tính toán và lưu trữ, đồng thời cải thiện tốc độ xử lý và độ chính xác của mô hình.

3. Ví dụ, có thể sử dụng kỹ thuật fine-tuning và LoRA để tối ưu hóa mô hình LLM cho use case cụ thể, giúp giảm kích thước mô hình và cải thiện hiệu năng.

4. Tip hoặc bước tiếp theo: Sử dụng các công cụ như Hugging Face Transformers để thực hiện fine-tuning và LoRA cho mô hình LLM, và đánh giá hiệu năng của mô hình trên các thiết bị khác nhau để đảm bảo tối ưu hóa chi phí và hiệu năng.