BIS - MT · 15/06/2026

Bản Tin AI Hằng Ngày

Cập nhật công nghệ AI mới nhất

"You have brains in your head. You have feet in your shoes. You can steer yourself any direction you choose."

↪ Bạn có não trong đầu. Bạn có chân trong giày. Bạn có thể lái mình theo bất kỳ hướng nào bạn chọn.

— Dr. Seuss

Mỗi người đều có khả năng tự định hướng cuộc đời mình — đừng để hoàn cảnh hay người khác quyết định bạn sẽ đi đâu.

TIN TỨC NỔI BẬT

1

Chuyên gia bảo mật gióng lên hồi chuông cảnh báo về lỗ hổng trong code do AI sinh ra

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 điểm yếu bảo mật trong mã được tạo ra bởi AI. Mã này được tạo ra bằng cách sử dụng trí tuệ nhân tạo và các thuật toán học máy, thường được thiết kế để tự động hóa các nhiệm vụ và tăng tốc quá trình phát triển. Tuy nhiên, các nhà nghiên cứu đã phát hiện ra rằng mã được tạo ra bởi AI có thể chứa các lỗi bảo mật, điều này có thể đặt người dùng và hệ thống vào tình trạng nguy hiểm. Vấn đề này phát sinh từ việc các thuật toán AI có thể không hiểu đầy đủ ngữ cảnh và sắc thái của mã mà chúng tạo ra. Kết quả là, chúng có thể tạo ra mã dễ bị tấn công hoặc khai thác. Điều này có thể dẫn đến việc mất mát dữ liệu, hệ thống bị sập và các sự cố bảo mật khác. Các nhà nghiên cứu cảnh báo rằng việc sử dụng mã được tạo ra bởi AI trong môi trường sản xuất có thể dẫn đến hậu quả nghiêm trọng. Họ kêu gọi thực hiện các thử nghiệm và xác thực nghiêm ngặt hơn đối với mã được tạo ra bởi AI để đảm bảo tính bảo mật và độ tin cậy của nó. Điều này bao gồm việc thực hiện các kiểm tra và kiểm soát bổ sung để phát hiện và giảm thiểu các điểm yếu bảo mật tiềm ẩn.

2 Ai cũng muốn tham gia "vibe coding" — và Google cũng không ngoại lệ với Stitch, dự án kế nhiệm của 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 tham gia vào xu hướng "lập trình cảm xúc" với việc ra mắt Stitch, một dự án tiếp theo sau Jules. Thuật ngữ "lập trình cảm xúc" đề cập đến sự quan tâm ngày càng tăng trong việc tạo ra các công cụ lập trình tập trung vào trải nghiệm người dùng và kết nối cảm xúc. Sự tham gia của Google vào lĩnh vực này cho thấy sự thay đổi trong cách tiếp cận lập trình của gã khổng lồ công nghệ, ưu tiên sự dễ sử dụng và giao diện trực quan. Mặc dù chi tiết về Stitch còn khan hiếm, việc phát hành nó như một dự án tiếp theo sau Jules cho thấy Google cam kết phát triển thêm khả năng lập trình cảm xúc của mình. Jules, một dự án trước đó, nhằm mục đích đơn giản hóa quá trình xây dựng và triển khai các model học máy (machine learning model). Với Stitch, Google có thể đang mở rộng khái niệm này, tạo ra một môi trường lập trình thân thiện với người dùng và dễ tiếp cận hơn. Các tính năng và chức năng chính xác của Stitch vẫn còn chưa rõ, nhưng việc phát hành nó có khả năng được theo dõi chặt chẽ bởi cộng đồng công nghệ, đặc biệt là những người quan tâm đến lập trình cảm xúc và trải nghiệm người dùng.

3 Lớp lệnh AI đa agent cho kho hàng giúp đạt hiệu suất vận hành vượt trội và tình báo 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, một công nghệ được thiết kế để nâng cao hiệu quả hoạt động và trí tuệ chuỗi cung ứng trong các nhà kho. Lớp lệnh AI này cho phép giao tiếp và phối hợp liền mạch giữa các tác nhân khác nhau, bao gồm robot, máy bay không người lái và các hệ thống tự động khác, để tối ưu hóa hoạt động nhà kho. Lớp lệnh AI Nhà kho đa tác nhân sử dụng nền tảng Omniverse của NVIDIA, một bộ công cụ để xây dựng và mô phỏng các môi trường AI phức tạp. Công nghệ này cho phép theo dõi và kiểm soát hoạt động nhà kho theo thời gian thực, cho phép các doanh nghiệp đưa ra quyết định dựa trên dữ liệu và cải thiện khả năng chống chịu của chuỗi cung ứng. Các lợi ích chính của Lớp lệnh AI Nhà kho đa

tác nhân bao gồm: - Tăng cường hiệu quả hoạt động thông qua phân bổ nhiệm vụ và tự động hóa được tối ưu hóa - Cải thiện trí tuệ chuỗi cung ứng thông qua theo dõi và phân tích dữ liệu theo thời gian thực - Tăng cường linh hoạt và khả năng mở rộng trong hoạt động nhà kho - Ra quyết định tốt hơn thông qua thông tin chi tiết dựa trên dữ liệu Bằng cách tận dụng công nghệ của NVIDIA, các doanh nghiệp có thể tối ưu hóa hoạt động nhà kho, giảm chi phí và cải thiện hiệu suất chuỗi cung ứng tổng thể.

4

Trung Quốc có thể đã tiếp cận Mythos

China may have accessed Mythos

The Verge AI Đọc bài viết →

Tòa Bạch Ốc đã được cho là thể hiện mối quan ngại rằng một nhóm liên kết với Trung Quốc có thể đã truy cập vào mô hình AI mạnh mẽ của Anthropic, Mythos. Theo một báo cáo mới từ Semafor, quyết định của Tòa Bạch Ốc trong việc áp đặt hạn chế xuất khẩu đối với Mythos đã được thúc đẩy một phần bởi nỗi sợ về một rủi ro an ninh quốc gia tiềm năng. Nếu chính phủ Trung Quốc có quyền truy cập vào Mythos, họ có thể cố gắng đảo ngược mô hình, điều này có thể làm tổn hại thông tin nhạy cảm. Anthropic đã không bình luận về báo cáo, nhưng một người phát ngôn cho biết chính phủ không đề cập đến Trung Quốc trong các cuộc thảo luận xung quanh kiểm soát xuất khẩu. Đây không phải là lần đầu tiên mô hình mạnh mẽ của Anthropic bị xâm phạm, vì một nhóm Discord đã được cho là có quyền truy cập vào Mythos trong hai tuần trước khi sự vi phạm được phát hiện. Tòa Bạch Ốc đã không xác nhận báo cáo, và một bài đăng trên X của một cố vấn Trump đã không đề cập đến Trung Quốc, thay vào đó tập trung vào khả năng Fable và Mythos có thể bị "jailbroken".

5

Tin tức Lawmadi OS — 2026-06-05

Lawmadi OS News — 2026-06-05

Dev.to AI Đọc bài viết →

Hàn Quốc dự kiến sẽ trở thành quốc gia đầu tiên thực hiện "Đạo luật Cơ bản về Trí tuệ Nhân tạo" vào ngày 22 tháng 6. Sự phát triển này đã gây sự quan tâm về cách nó sẽ ảnh hưởng đến hệ thống pháp lý. Lawmadi OS, một hệ thống vận hành pháp lý được hỗ trợ bởi AI của Hàn Quốc, được thiết kế để phân tích câu hỏi và xác minh trích dẫn so

với cơ sở dữ liệu luật pháp chính thức của Hàn Quốc (law.go.kr) theo thời gian thực. Hệ thống này sử dụng 60 tác nhân AI chuyên dụng trên 60 lĩnh vực. Nó cũng có các mẫu để trả lời nhanh các câu hỏi thường gặp hoặc lưu trữ các đoạn mã có thể tái sử dụng. Hệ thống hỗ trợ bởi AI này nhằm mục đích tối ưu hóa quá trình pháp lý, có thể dẫn đến việc ra quyết định nhanh hơn và hiệu quả hơn.

6

Từ những dự án open source thành công đến OpenAI

From open source hits to OpenAI

Changelog Đọc bài viết →

Trong một cuộc trò chuyện gần đây, Max Stoiber, một developer tại OpenAI, đã thảo luận về các dự án mã nguồn mở và tác động của chúng đối với ngành công nghiệp công nghệ. Ông đã đề cập đến tầm quan trọng của các dự án mã nguồn mở ít được biết đến, chẳng hạn như react-boilerplate và styled-components, cùng với những dự án nổi bật hơn. Stoiber cũng chia sẻ kinh nghiệm của mình với Spectrum, đã trở thành một phần của GitHub và đóng góp vào sự phát triển của GitHub Discussions. Ngoài ra, ông cũng đề cập đến thành công của Stellate, một bộ nhớ đệm GraphQL đã được Shopify và The Guild mua lại. Công việc của Stoiber trên thư viện plugin và nền tảng ứng dụng của ChatGPT cũng được nhấn mạnh, với trọng tâm là cách các ứng dụng ChatGPT đang tạo ra một bề mặt mới cho phát triển phần mềm.

7

olmo-eval: Môi trường làm việc đánh giá cho chu trình phát triển model

olmo-eval: An evaluation workbench for the model development loop

Hugging Face Blog Đọc bài viết →

Viện Allen về Trí tuệ Nhân tạo (AI) đã phát triển một công cụ mới gọi là olmo-eval, được thiết kế để giúp quá trình đánh giá các mô hình ngôn ngữ lớn (LLMs) trong quá trình phát triển. Không giống như các công cụ hiện có, olmo-eval là một ngăn xếp đánh giá tích hợp cho phép đánh giá có thể tái tạo và là mã nguồn mở. Nó được xây dựng trên Tiêu chuẩn Đánh giá Mô hình Ngôn ngữ Mở (OLMES), được giới thiệu vào năm 2024, nhằm tiêu chuẩn hóa điểm chuẩn LLM để dễ dàng so sánh. olmo-eval mở rộng OLMES bằng cách cung cấp nhiều tính linh hoạt hơn trong việc định nghĩa và chạy đánh giá, giúp dễ dàng hơn trong việc tạo thành các thành phần riêng lẻ thành các quy

trình công việc lớn hơn. Nó hỗ trợ đánh giá agentic và đa lượt, và cung cấp các công cụ phân tích mạnh mẽ hơn để xác định xem một can thiệp có thực sự cải thiện mô hình hay sự khác biệt là do nhiễu. So với Harbor, một khuôn khổ mở để đánh giá các tác nhân AI, olmo-eval được thiết kế cho phát triển mô hình hàng ngày, cho phép đánh giá và phân tích nhanh chóng hiệu suất của mô hình. Nó cũng cung cấp một đường dẫn đánh giá nhẹ, đây là mặc định, và chỉ sử dụng một thiết lập nặng khi cần thiết. olmo-eval là mã nguồn mở và có sẵn trên GitHub.

8

Cách chọn Sandbox phù hợp cho AI Agent

How to Choose the Right Sandbox for AI Agents

LangChain Blog Đọc bài viết →

Việc chọn sandbox phù hợp cho các tác nhân AI là rất quan trọng để đảm bảo sự tự chủ của chúng không ảnh hưởng đến bảo mật dữ liệu và hệ thống. Các tác nhân AI có thể tạo ra giá trị bằng cách tạo và thực thi mã, nhưng điều này cũng tiềm ẩn nguy cơ về các mối đe dọa không lường trước. Sandboxes cách ly mã được tạo ra bởi AI, hạn chế quyền truy cập và tạo ra một ranh giới an toàn hơn cho các tác nhân hoạt động. Về cơ bản, chúng là máy tính an toàn mà các tác nhân có thể sử dụng mà không gây ra mối đe dọa. Để chọn một sandbox phù hợp, các đội cần xem xét các rủi ro bảo mật liên quan đến việc chạy mã được tạo ra bởi AI. Điều này bao gồm rủi ro của các cuộc tấn công tiêm prompt, nơi một kẻ tấn công khiến tác nhân viết mã làm tổn hại dữ liệu và hệ thống. "Tam giác tử thần" của các điều kiện, được phác thảo bởi Simon Willison, làm tăng rủi ro của các cuộc tấn công như vậy. Một giải pháp sandbox an toàn nên cung cấp các tính năng như: - Chứa chỉ dữ liệu mà tác nhân cần để thực hiện công việc của nó - Chặn tác nhân truy cập vào dữ liệu khác - Hạn chế các điểm cuối ngoài mà tác nhân có thể gửi dữ liệu đến qua internet - Kiểm soát việc sử dụng tính toán và bộ nhớ - Cho phép quyết định tái sử dụng sandbox hay không Khi chọn sandbox, các đội nên đánh giá các giải pháp tiềm năng một cách nghiêm túc để đảm bảo chúng đáp ứng các yêu cầu này.

TIPS & TRICKS CHO DEV

Tự động hóa kiểm thử đơn vị

Vấn đề: Viết test case thủ công tốn thời gian và dễ xảy ra lỗi.

Cách làm: Sử dụng AI tools như Pytest và Unittest để tự động hóa kiểm thử đơn vị.

Ví dụ, sử dụng lệnh `pytest --cov` để kiểm tra coverage.
Đánh giá: Hiệu quả cao, nên dùng khi dự án có nhiều module và chức năng.

Kiểm tra mã nguồn tự động

Vấn đề: Kiểm tra mã nguồn thủ công dễ bỏ sót lỗi và tốn thời gian.
Cách làm: Sử dụng công cụ như SonarQube và CodeCoverage để kiểm tra mã nguồn tự động. Ví dụ, sử dụng lệnh `sonar-scanner` để quét mã nguồn.
Đánh giá: Hiệu quả cao, nên dùng khi dự án có nhiều contributor.

Tự động hóa kiểm tra hiệu suất

Vấn đề: Kiểm tra hiệu suất thủ công tốn thời gian và không chính xác.
Cách làm: Sử dụng công cụ như JMeter và Locust để tự động hóa kiểm tra hiệu suất. Ví dụ, sử dụng lệnh `locust -f` để chạy test hiệu suất.
Đánh giá: Hiệu quả cao, nên dùng khi dự án yêu cầu hiệu suất cao.

BÀI HỌC AI HÔM NAY CHO DEV

1. **Fine-tuning & LoRA cho use case cụ thể**
2. Dev cần biết về fine-tuning và LoRA để tối ưu hóa mô hình AI cho ứng dụng cụ thể, giúp cải thiện hiệu suất và giảm chi phí. Điều này cho phép các mô hình AI học hỏi từ dữ liệu cụ thể và điều chỉnh để phù hợp với nhu cầu thực tế.
3. Ví dụ, sử dụng thư viện Hugging Face Transformers để fine-tune mô hình BERT cho nhiệm vụ phân loại văn bản: `from transformers import BertTokenizer, BertModel; tokenizer = BertTokenizer.from_pretrained('bert-base-uncased'); model = BertModel.from_pretrained('bert-base-uncased')`.
4. Tip hoặc bước tiếp theo: Sử dụng kỹ thuật LoRA (Low-Rank Adaptation) để giảm thiểu sự thay đổi của mô hình khi fine-tuning, giúp giữ nguyên kiến thức đã học và cải thiện hiệu suất.