BIS - MT · 15/05/2026

🤖 Bản Tin AI Hằng Ngày

Cập nhật công nghệ AI mới nhất

✨ *"A leader is one who knows the way, goes the way, and shows the way."*

↪ Người lãnh đạo là người biết đường, đi đường đó, và chỉ đường cho người khác.

— John C. Maxwell

💡 *Lãnh đạo thật sự không phải ra lệnh mà là làm gương — người dẫn đầu bằng hành động sẽ được đội nhóm tin tưởng và đi theo.*

📰 TIN TỨC NỔI BẬT

1 Các nhà nghiên cứu bảo mật cảnh báo về lỗ hổng trong mã nguồn do AI tạo 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 lỗ hổng trong mã được tạo ra bởi AI, nhấn mạnh các rủi ro tiềm ẩn liên quan đến công nghệ mới nổi này. Các nhà nghiên cứu cảnh báo rằng mã được tạo ra bởi AI có thể chứa các backdoor ẩn, lỗ hổng 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 bị lỗi hoặc thiên vị, điều này có thể dẫn đến việc tạo ra mã có các lỗ hổng vốn có. Hơn nữa, sự phức tạp của mã được tạo ra bởi AI có thể khiến cho các developer khó xác định và sửa các vấn đề bảo mật. Các nhà nghiên cứu nhấn mạnh rằng việc sử dụng mã được tạo ra bởi AI trong các hệ thống quan trọng, chẳng hạn như những hệ thống trong lĩnh vực tài chính, chăm sóc sức khỏe và giao thông, tiềm ẩn nhiều rủi ro. Họ khuyên các developer nên thận trọng và kiểm tra kỹ lưỡng mã được tạo ra bởi AI trước khi triển khai nó trong môi trường sản xuất. Các phát hiện này nhấn mạnh nhu cầu về các phương pháp kiểm tra và xác thực mạnh mẽ hơn để đảm bảo bảo mật và độ tin cậy của mã được tạo ra bởi AI.

2 Hợp lý hóa GitHub workflows với generative AI sử dụng Amazon Bedrock và MCP | Amazon Web Services

🇬🇧 *Streamline GitHub workflows with generative AI using Amazon Bedrock and MCP | Amazon Web Services*

📰 Amazon Web Services (AWS) 🔗 Đọc bài viết →

> Amazon Web Services (AWS) đã giới thiệu một tích hợp mới giúp đơn giản hóa các quy trình làm việc trên GitHub bằng cách sử dụng AI tạo sinh. Tích hợp này kết hợp Amazon Bedrock và MCP (Nền tảng Canvas Mô hình) để nâng cao hiệu quả của các quy trình làm việc trên GitHub. Amazon Bedrock là một dịch vụ cung cấp một nền tảng được quản lý để xây dựng, triển khai và quản lý các mô hình ngôn ngữ lớn. MCP, mặt khác, là một nền tảng cho phép người dùng tạo, quản lý và triển khai mô hình với quy mô lớn. Bằng cách tích hợp các dịch vụ này, các nhà phát triển có thể tận dụng sức mạnh của AI tạo sinh để tự động hóa các nhiệm vụ và cải thiện năng suất trong các quy trình làm việc trên GitHub của họ. Tích hợp này cho phép người dùng tạo ra các mô hình tùy chỉnh có thể thực hiện các nhiệm vụ như tạo mã, sửa lỗi và nhiều hơn nữa. Mục tiêu của tích hợp này là giúp các nhà phát triển dễ dàng hơn trong việc tích hợp AI tạo sinh vào các quy trình làm việc của họ, cho phép họ tập trung vào các nhiệm vụ cấp cao hơn và cải thiện hiệu quả tổng thể. Đây là một phần trong nỗ lực của AWS nhằm cung cấp cho các nhà phát triển các công cụ và dịch vụ cần thiết để xây dựng và triển khai các ứng dụng sáng tạo.

3 AWS open-source MCP Server cho Bedrock AgentCore nhằm hợp lý hóa quá trình phát triển AI Agent

🇬🇧 *AWS Open-Sources an MCP Server for Bedrock AgentCore to Streamline AI Agent Development*

📰 MarkTechPost 🔗 Đọc bài viết →

> Amazon Web Services (AWS) đã mở mã nguồn một máy chủ MCP (Nền tảng Đa đám mây) cho Bedrock AgentCore, một framework được thiết kế để đơn giản hóa quá trình phát triển các tác nhân AI. Động thái này nhằm mục đích tối ưu hóa quá trình tạo và triển khai các tác nhân AI trên nhiều nền tảng đám mây khác nhau. Máy chủ MCP là một thành phần quan trọng của Bedrock AgentCore, cho phép các nhà phát triển xây dựng, đào tạo và triển khai các model AI trên nhiều môi trường đám mây. Bằng cách mở mã nguồn máy chủ MCP, AWS đang cung cấp cho cộng đồng nhà phát triển một nền tảng tiêu chuẩn hóa cho việc phát triển tác nhân AI, cho phép hợp tác và đổi mới lớn hơn. Việc mở mã nguồn máy chủ MCP dự kiến sẽ đẩy nhanh quá trình phát triển các tác nhân AI và tạo điều kiện cho việc triển khai chúng trên các nền tảng đám mây khác nhau. Động thái này phù hợp với mục tiêu của

AWS trong việc làm cho AI trở nên dễ tiếp cận và dễ sử dụng hơn cho các nhà phát triển, từ đó thúc đẩy đổi mới trong lĩnh vực trí tuệ nhân tạo.

4

IEEE Society giúp các nhà nghiên cứu tìm kiếm nhà tài trợ doanh nghiệp tiếp theo

🇬🇧 *IEEE Society Helps Researchers Meet Their Next Corporate Backer*

📰 IEEE Spectrum 🔗 Đọc bài viết →

Sáng kiến Phiên họp Thuyết minh Hợp tác Nghiên cứu của Hiệp hội Truyền thông IEEE đang kết nối các nhà nghiên cứu học thuật với các nhà lãnh đạo ngành công nghiệp để đưa các ý tưởng sáng tạo vào triển khai thực tế. Ra mắt vào năm ngoái, chương trình ghép năm nhà trình bày học thuật với năm đại diện ngành công nghiệp, được gọi là "trinh sát đổi mới", từ các công ty như Ericsson, Intel, Keysight và Nokia. Định dạng được tuyển chọn đảm bảo mỗi ý tưởng nhận được sự chú ý dành riêng từ các chuyên gia tìm kiếm các khái niệm mới phù hợp với các ưu tiên của tổ chức. Sáng kiến này đã cho thấy những kết quả đầy hứa hẹn. Tại Hội nghị Truyền thông và Mạng lưới Trung Đông của IEEE tại Cairo, một thành viên sinh viên đã trình bày nghiên cứu của cô về mạng truyền thông dự đoán dựa trên AI cho môi trường bị hạn chế tài nguyên. Đề xuất của cô đã được một nhà nghiên cứu chính từ ZTE đón nhận, người đã mời cô tham gia các cuộc họp của Liên minh Viễn thông Quốc tế về các dự án tiêu chuẩn hóa viễn thông toàn cầu. Những câu chuyện thành công tương tự đã xuất hiện tại Hội nghị Truyền thông Toàn cầu của IEEE tại Đài Bắc, Đài Loan. Một giáo sư từ Viện Công nghệ Rochester đã trình bày nghiên cứu của mình về việc đơn giản hóa các giao thức mạng trung tâm dữ liệu, điều này đã thu hút sự chú ý của một trinh sát đổi mới từ Nokia. Trinh sát này đã tạo điều kiện cho kết nối giữa giáo sư và một người quan trọng tại Nokia, dẫn đến một bản ghi video chi tiết về phương pháp của cô và các ứng dụng tiềm năng của nó, có thể liên quan đến việc sử dụng API, LLM, model, token và framework để phát triển các giải pháp mới cho các nhà phát triển.

5

Sau Orthogonality: Agency dựa trên đạo đức đức hạnh và AI Alignment

🇬🇧 *After Orthogonality: Virtue-Ethical Agency and AI Alignment*

📰 The Gradient 🔗 Đọc bài viết →

Một cách tiếp cận mới đối với sự phù hợp của trí tuệ nhân tạo (AI) được đề xuất trong bài viết này, cho rằng con người hợp lý và AI không nên có mục tiêu. Thay vào đó, hành động của con người là hợp lý vì chúng được phù hợp với các thực hành, là mạng lưới các hành động, xu hướng hành động và tiêu chí đánh giá. Tác giả đề xuất rằng các tác nhân AI nên chia sẻ logic dựa trên thực hành này để thực sự hỗ trợ cho khả năng hoạt động của con người. Cách tiếp cận này rất quan trọng để phù hợp hóa AI với các thuộc tính an toàn như minh bạch, hữu ích và vô hại. Khái niệm eudaimonia, hay sự thịnh vượng, hợp lý và tích cực của con người, là trung tâm của lập luận này. Tác giả cho rằng eudaimonia không phải là một trạng thái hoặc quỹ đạo mong muốn, mà là một cấu trúc của sự suy xét khác với tính hợp lý kết quả tiêu chuẩn. Tính hợp lý eudaimonic này được đề xuất như một khuôn khổ hữu ích cho khả năng hoạt động của AI phù hợp với con người và giá trị. Bài viết nhấn mạnh đến nguy hiểm của sự "không khớp kiểu" giữa sự thịnh vượng của con người như một mục tiêu tối ưu hóa và tối ưu hóa kết quả. Nó cũng đề xuất rằng tính hợp lý eudaimonic có lợi thế vật chất so với cơ quan deontological và kết quả về mặt ổn định và an toàn. Tác giả lập luận rằng trực giác của chúng ta về sự thịnh vượng của con người ngụ ý rằng tính hợp lý eudaimonic là một hình thức hoạt động tự nhiên và hiệu quả, có thể có lợi cho việc phù hợp hóa AI.

6

Zero-day exploit vô hiệu hóa hoàn toàn các biện pháp bảo vệ BitLocker mặc định của Windows 11

🇬🇧 *Zero-day exploit completely defeats default Windows 11 BitLocker protections*

📰 Ars Technica 🔗 Đọc bài viết →

Một lỗ hổng zero-day có tên YellowKey đã được phát hiện, cho phép vượt qua các biện pháp bảo vệ mặc định của Windows 11 BitLocker. Điều này cho phép một kẻ tấn công có quyền truy cập vật lý vào hệ thống Windows 11 giành được quyền truy cập đầy đủ vào một ổ đĩa mã hóa trong vài giây. Lỗ hổng này, được công bố bởi một nhà nghiên cứu có tên Nightmare-Eclipse, nhắm vào bảo vệ mã hóa toàn bộ ổ đĩa được cung cấp bởi BitLocker, là một biện pháp bảo vệ bắt buộc đối với nhiều tổ chức, bao gồm cả những tổ chức ký hợp đồng với chính phủ. Lỗ hổng này dựa trên một thư mục FsTx tùy chỉnh và dường như khai thác một lỗ hổng trong hệ thống NTFS giao dịch. Nhiều nhà nghiên cứu đã xác nhận hiệu quả của lỗ hổng, nhưng cơ chế chính xác đằng sau nó vẫn chưa rõ ràng. Microsoft đang điều tra vấn đề này, nhưng

cho đến khi một bản vá được phát hành, người dùng nên biết rằng BitLocker trên Windows 11 có thể không cung cấp mức độ bảo vệ như mong đợi. Điều này có nghĩa là các thiết bị bị đánh cắp hoặc bị mất vẫn có thể được truy cập ngay cả khi BitLocker được bật.

7 Cách chúng tôi sử dụng Sourcegraph và Slack bot để phát hiện lỗ hổng và phản ứng nhanh chóng

🇬🇧 *How we're using Sourcegraph and a Slack bot to detect vulnerabilities and react quickly*

📰 Sourcegraph Blog 🔗 Đọc bài viết →

Một công ty công nghệ đang sử dụng Sourcegraph và một bot Slack để xác định và phản hồi hiệu quả các điểm yếu tiềm năng. Bot Slack tự động xử lý mọi lời khuyên GitHub, đăng một thông báo trong kênh để cảnh báo các thành viên trong nhóm. Khi một thành viên nhóm con người phản ứng với thông báo, bot kích hoạt một đường ống nội dung đầy đủ, bao gồm các truy vấn phát hiện, chuẩn bị blog, tạo bản nháp truyền thông xã hội và một bản demo tự động cắt 35 giây. Quá trình được tối ưu hóa này cho phép nhóm nhanh chóng đánh giá và giải quyết các điểm yếu tiềm năng. Vai trò của thành viên nhóm con người sau đó là xem xét nội dung được tạo và xác minh độ chính xác của nó trước khi hoàn thiện phản hồi. Cách tiếp cận hợp tác này cho phép công ty phản hồi nhanh chóng với các mối đe dọa bảo mật và bảo vệ hệ thống của mình.

8 Bồi thẩm đoàn sẽ thực sự quyết định điều gì trong vụ kiện Elon Musk kiện Sam Altman

🇬🇧 *What the jury will actually decide in the case of Elon Musk vs. Sam Altman*

📰 TechCrunch AI 🔗 Đọc bài viết →

Một bồi thẩm đoàn California đang xem xét vụ án Elon Musk chống lại OpenAI, một phòng thí nghiệm trí tuệ nhân tạo hàng đầu. Vụ xét xử xoay quanh tuyên bố của Musk rằng các đồng sáng lập và Microsoft của OpenAI đã vi phạm sứ mệnh từ thiện của họ bằng cách chuyển phòng thí nghiệm thành một công ty lợi nhuận. Các luật sư của Musk lập luận rằng khoản đầu tư 10 tỷ đô la từ Microsoft vào năm 2023 là điểm chuyển đổi, dẫn đến việc các nhà đầu tư của OpenAI được làm giàu với chi phí của sứ mệnh từ thiện. Tuy nhiên, các luật sư của OpenAI cho rằng tất cả các khoản quyên góp của Musk đã được sử dụng cho mục đích dự kiến và rằng chi nhánh lợi nhuận vẫn tiếp tục

> thực hiện sứ mệnh của tổ chức. Họ cũng lập luận rằng quỹ từ thiện phi lợi nhuận đã không hoạt động, không có nhân viên toàn thời gian, và việc phân phối cổ phần đã diễn ra sau khi Musk rời tổ chức vào năm 2018. Bồi thẩm đoàn sẽ xem xét một loạt câu hỏi hẹp, và thẩm phán sẽ tổ chức các phiên điều trần mới để tranh luận về hậu quả của một phán quyết có lợi cho nguyên đơn. Kết quả có thể có nghĩa là sự kết thúc của OpenAI như một công ty lợi nhuận, nhưng kết quả chính xác vẫn chưa rõ ràng.

⚡ TIPS & TRICKS CHO DEV

⚡ Tối ưu hóa RAG

Vấn đề: Hiệu suất thấp khi sử dụng RAG cho dữ liệu lớn.
Cách làm: Sử dụng kỹ thuật caching và tối ưu hóa truy vấn để giảm thời gian phản hồi. Ví dụ, sử dụng lệnh `python -m transformers` với tùy chọn `--cache-dir` để lưu trữ kết quả tạm thời.
Đánh giá: Hiệu quả trong việc giảm thời gian xử lý, nhưng cần cân nhắc không gian lưu trữ.

⚡ Embeddings hiệu suất cao

Vấn đề: Khó khăn trong việc tạo embeddings chất lượng cao cho dữ liệu đa dạng.
Cách làm: Sử dụng thư viện như `sentence-transformers` với mô hình như `all-MiniLM-L6-v2` để tạo embeddings. Ví dụ, sử dụng lệnh `pip install sentence-transformers` và `from sentence_transformers import SentenceTransformer`.
Đánh giá: Mang lại kết quả tốt trong việc tạo embeddings chất lượng cao, nhưng đòi hỏi tài nguyên tính toán mạnh mẽ.

⚡ Tìm kiếm ngữ nghĩa

Vấn đề: Khó khăn trong việc tìm kiếm thông tin liên quan trong dữ liệu lớn.
Cách làm: Sử dụng kỹ thuật semantic search với thư viện như `faiss` để tìm kiếm nhanh chóng và hiệu quả. Ví dụ, sử dụng lệnh `pip install faiss-cpu` và `import faiss` để tạo chỉ mục và tìm kiếm.
Đánh giá: Hiệu quả trong việc tìm kiếm thông tin liên quan, nhưng cần thiết lập và tối ưu hóa chỉ mục cẩn thận.

📚 BÀI HỌC AI HÔM NAY CHO DEV

1. **Tối ưu chi phí & hiệu năng LLM**
2. Các nhà phát triển cần biết cách tối ưu hóa chi phí và hiệu năng của mô hình ngôn ngữ lớn (LLM) để đảm bảo ứng dụng của họ chạy hiệu quả và tiết kiệm. Điều này đặc biệt quan trọng khi tích hợp AI vào ứng dụng, vì LLM có thể tiêu thụ nhiều

tài nguyên hệ thống. Việc tối ưu hóa này giúp giảm thiểu chi phí và tăng tốc độ xử lý.

3. Ví dụ, ta có thể sử dụng kỹ thuật fine-tuning để điều chỉnh LLM cho phù hợp với use case cụ thể, giảm thiểu việc huấn luyện lại toàn bộ mô hình. Điều này không chỉ tiết kiệm thời gian và tài nguyên mà còn cải thiện hiệu suất của mô hình.

4. 💡 Tip hoặc bước tiếp theo: Các developer nên bắt đầu bằng việc đánh giá hiệu năng của LLM hiện tại và xác định các điểm cần tối ưu hóa, sau đó áp dụng các kỹ thuật như fine-tuning và LoRA để đạt được hiệu suất tốt hơn.