

# Bản Tin AI Hằng Ngày

Cập nhật công nghệ AI mới nhất

*"The mind is everything. What you think you become."*

↳ Tâm trí là tất cả. Bạn nghĩ gì thì trở thành người đó.

— Phật Thích Ca

*Tư duy định hình thực tại — những suy nghĩ ta nuôi dưỡng hàng ngày dần dần hình thành nên con người ta và những gì ta thu hút vào cuộc sống.*

## TIN TỨC NỔI BẬT

### Google cũng tham gia làn sóng "vibe coding" với Stitch, dự án tiếp nối Jules

1

*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 "vibe coding" với việc ra mắt Stitch, một dự án tiếp nối từ dự án trước đó là Jules. Thuật ngữ "vibe coding" đề cập đến sự quan tâm ngày càng tăng trong việc phát triển các công cụ coding có thể hiểu và tái tạo cảm xúc cũng như hành vi của con người. Mặc dù các chi tiết về Stitch chưa được tiết lộ đầy đủ, nhưng rõ ràng là Google đang đầu tư vào lĩnh vực nghiên cứu này. Jules, một dự án trước đây của Google, nhằm mục đích tạo ra một conversational AI có thể hiểu và phản hồi cảm xúc của con người. Stitch dường như là sự tiếp nối của công việc này, với việc Google có thể đang tìm cách phát triển hơn nữa các capability của mình trong lĩnh vực này. Sự quan tâm của công ty đối với vibe coding cho thấy sự công nhận ngày càng tăng về tầm quan trọng của emotional intelligence trong phát triển AI. Khi ngành công nghệ tiếp tục phát triển, khoản đầu tư của Google vào Stitch và Jules làm nổi bật tiềm năng của AI để trở nên đồng cảm và giống con người hơn. Tuy nhiên, cần thêm thông tin để hiểu đầy đủ về phạm vi và goal của dự án này.

### AWS Open-Source MCP Server cho Bedrock AgentCore để tinh gọn phát triển AI Agent

2

Amazon Web Services (AWS) đã open-source một MCP (Multi-Cloud Platform) server cho Bedrock AgentCore, một development framework dành cho các AI agent. Động thái này nhằm mục đích tinh gọn quá trình phát triển AI agent bằng cách cung cấp một platform tiêu chuẩn để xây dựng và deploy các AI model trên nhiều cloud environment. MCP server được thiết kế để hoạt động với Bedrock AgentCore, cho phép các developer tạo và quản lý các AI agent có thể tương tác với nhiều application và service khác nhau. Bằng cách open-source MCP server, AWS đang cung cấp một platform do cộng đồng định hướng để các developer xây dựng, test và deploy các AI agent hiệu quả hơn. MCP server được kỳ vọng sẽ thúc đẩy sự hợp tác và đổi mới trong cộng đồng phát triển AI bằng cách cho phép các developer chia sẻ và phát triển dựa trên công việc của nhau. Động thái này phù hợp với nỗ lực của AWS nhằm thúc đẩy phát triển open-source và cung cấp nhiều sự linh hoạt cũng như lựa chọn hơn cho các developer làm việc với các công nghệ AI và machine learning.

### **Xây dựng intelligent financial analysis agent với LangGraph và Strands Agents | Amazon Web Services**

3

*Build an intelligent financial analysis agent with LangGraph and Strands Agents | Amazon Web Services*

Amazon Web Services (AWS) [Đọc bài viết →](#)

Amazon Web Services (AWS) đã giới thiệu một giải pháp mới để xây dựng các intelligent financial analysis agent sử dụng LangGraph và Strands Agents. Cách tiếp cận đổi mới này cho phép người dùng tạo ra các công cụ financial analysis tinh vi có thể xử lý và phân tích lượng lớn financial data. LangGraph là một graph-based language model cho phép người dùng biểu diễn các mối quan hệ và cấu trúc tài chính phức tạp dưới định dạng có cấu trúc. Strands Agents, mặt khác, là một framework để xây dựng các intelligent agent có thể tương tác với các LangGraph model để thực hiện các tác vụ như data analysis và visualization. Bằng cách kết hợp LangGraph và Strands Agents, người dùng có thể tạo ra các intelligent financial analysis agent có khả năng xác định các pattern, trend và anomaly trong financial data. Các agent này cũng có thể cung cấp insights và recommendation cho người dùng, giúp họ đưa ra các investment decision sáng suốt. Giải pháp này đặc biệt hữu ích cho các financial institution, investment firm và các tổ

chức khác yêu cầu khả năng financial analysis nâng cao. Với LangGraph và Strands Agents, người dùng có thể xây dựng các công cụ financial analysis tùy chỉnh đáp ứng các nhu cầu và yêu cầu cụ thể của họ.

4

## Tự động hóa Security Triage với HackerOne và Deep Search

*Automating Security Triage with HackerOne and Deep Search*

Sourcegraph Blog [Đọc bài viết →](#)

Sourcegraph đã triển khai một hệ thống automated security triage sử dụng HackerOne webhooks và Deep Search. Hệ thống này cho phép công ty validate, triage và investigate các vulnerability report một cách hiệu quả trước khi các engineer bắt đầu ngày làm việc của họ. Bằng cách tận dụng webhooks của HackerOne, Sourcegraph có thể tự động nhận và xử lý các vulnerability report từ HackerOne platform. Deep Search, một feature của code search và intelligence platform của Sourcegraph, sau đó được sử dụng để phân tích các report và cung cấp thêm context. Quy trình automated này cho phép các engineer của Sourcegraph bắt đầu ngày làm việc với sự hiểu biết rõ ràng về security posture của codebase của họ, mà không cần phải manually review và investigate từng report. Việc tích hợp HackerOne và Deep Search tinh gọn quy trình security triage, giúp Sourcegraph phản ứng nhanh chóng với các mối đe dọa bảo mật tiềm ẩn và duy trì tính toàn vẹn của codebase của họ.

5

## Tự động hóa với tốc độ của Swamp

*Automation at the speed of Swamp*

Changelog [Đọc bài viết →](#)

Trọng một cuộc trò chuyện

6

## Tôi đã thử Siri AI, và cho đến nay nó thực sự hoạt động

*I tried Siri AI, and so far it actually works*

The Verge AI [Đọc bài viết →](#)

Trợ lý ảo Siri AI mới được nâng cấp của Apple cuối cùng đã đạt được mục tiêu cung cấp trải nghiệm hữu ích cho người dùng iPhone. Trợ lý

AI này hiện có thể thực hiện các nhiệm vụ như tham khảo thông tin trong email và lịch, đặt nhắc nhở và thêm sự kiện vào lịch từ email hoặc tờ rơi. Trong một thử nghiệm thực tế, tác giả đã phát hiện ra rằng Siri AI có thể hoàn thành thành công các nhiệm vụ này, bao gồm thêm danh sách các trận đấu bóng đá vào lịch và đặt nhắc nhở cho một dự án làm vườn. Khả năng hiểu ngữ cảnh và cung cấp câu trả lời hữu ích của AI là một bước cải tiến đáng kể so với các phiên bản trước của Siri. Tuy nhiên, đáng chú ý là khả năng của Siri AI tương đối cơ bản so với các trợ lý AI khác, như Gemini của Google, đã có thể thực hiện các nhiệm vụ tương tự trong nhiều năm. Mặc dù vậy, Siri AI mới là một bước tiến cho Apple, và khả năng hoạt động hiệu quả của nó là minh chứng cho nỗ lực của công ty trong việc cải thiện công nghệ AI của mình.

## 7 Cá mập trắng lớn đã có mặt ở Địa Trung Hải hàng triệu năm — nhưng việc nhìn thấy chúng cực kỳ hiếm

*Great White Sharks Have Been in the Mediterranean Sea for Millions of Years—but Sightings Are Incredibly Rare*

Wired [Đọc bài viết →](#)

Một sự xuất hiện hiếm của cá mập trắng lớn trong biển Địa Trung Hải đã được ghi lại bởi một đội lặn từ Ghost Diving và Healthy Seas. Cuộc gặp gỡ, được quay lại bằng máy ảnh, cho thấy một con cá mập trắng lớn trưởng thành đực trong môi trường sống tự nhiên của nó. Loài này, được biết đến với tên *Carcharodon carcharias*, được coi là có nguy cơ tuyệt chủng cao trong biển Địa Trung Hải do các hoạt động của con người như bắt cá tình cờ trong các hoạt động đánh bắt cá, mất môi trường sống và sự suy giảm của con mồi tự nhiên của nó. Cá mập trắng đã sống trong biển Địa Trung Hải hàng triệu năm, nhưng các lần nhìn thấy đã trở nên ngày càng hiếm. Sự nhìn thấy được ghi lại cung cấp thông tin quý giá cho nghiên cứu khoa học và chiến lược bảo tồn loài. Theo các nhà nghiên cứu, các quan sát như thế này rất quan trọng để cải thiện sự hiểu biết về phân bố, thói quen và hành vi của loài có nguy cơ tuyệt chủng cao. Sự nhìn thấy này cũng được coi là một dấu hiệu khuyến khích cho sự đa dạng sinh học biển và một lời nhắc nhở về tầm quan trọng của việc tiếp tục nỗ lực bảo tồn trong các hệ sinh thái biển.

## 8 Sau Orthogonality: Virtue-Ethical Agency và AI Alignment

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

Trong bài viết này, tác giả cho rằng con người hợp lý và các tác nhân trí tuệ nhân tạo (AI) không nên có mục tiêu, mà thay vào đó nên căn chỉnh hành động của họ với các thực hành, là mạng lưới các hành động, khuynh hướng, tiêu chí đánh giá và tài nguyên. Cách tiếp cận này là cần thiết để các AI thực sự hỗ trợ và cộng tác với khả năng hành động của con người. Tác giả đề xuất rằng các khái niệm như vô hại và khả năng sửa đổi, thường được coi là mục tiêu hoặc quy tắc, được giải thích một cách tự nhiên hơn như các động lực trong mạng lưới hành động này. Tác giả khám phá khái niệm eudaimonia, hay sự thịnh vượng hợp lý và tích cực của con người, và cho rằng nó chỉ ra một cấu trúc suy nghĩ khác với tính hợp lý kết quả tiêu chuẩn. Hình thức hoạt động hợp lý này, được gọi là tính hợp lý eudaimonic, được coi là một khuôn khổ hữu ích cho khả năng hành động và giá trị của các AI được căn chỉnh với con người. Tác giả tuyên bố rằng tính hợp lý eudaimonic ổn định và an toàn hơn các hình thức khả năng hành động khác, và nó có thể mạnh mẽ về mặt chức năng theo cách quan trọng đối với việc căn chỉnh AI. Cuối cùng, tác giả cho rằng nếu mục tiêu của chúng ta là căn chỉnh AI để hỗ trợ sự thịnh vượng của con người, thì việc truyền tải cho các AI tính hợp lý eudaimonic có thể là một cách tiếp cận hiệu quả hơn so với các phương pháp truyền thống. Cách tiếp cận này có thể giúp giải quyết nhiều vấn đề an toàn AI cổ điển và nghịch cảnh của việc căn chỉnh AI.

## TIPS & TRICKS CHO DEV

### Tạo Code Cơ Bản

**Vấn đề:** Khi cần tạo code cơ bản cho một dự án mới.

**Cách làm:** Sử dụng Claude Code với lệnh `claude create` và cung cấp thông tin về dự án. Ví dụ: `claude create --lang python --proj_name my_project`.

**Đánh giá:** Hiệu quả trong việc tạo code nhanh chóng, nhưng nên chỉnh sửa lại code sau khi tạo.

### Debug Code

**Vấn đề:** Khi gặp lỗi trong code và cần hỗ trợ debug.

**Cách làm:** Sử dụng Claude Code với lệnh `claude debug` và cung cấp thông tin về lỗi. Ví dụ: `claude debug --code "print('Hello World')" --error "SyntaxError"`.

**Đánh giá:** Hiệu quả trong việc tìm ra nguyên nhân lỗi, nhưng cần kiểm tra lại kết quả.

### Refactor Code

**Vấn đề:** Khi cần tối ưu hóa code để cải thiện hiệu suất.

**Cách làm:** Sử dụng Claude Code với lệnh `claude refactor` và cung cấp thông tin về code cần tối ưu hóa. Ví dụ: `claude refactor --code "for i in range(10): print(i)" --optimize`.

**Đánh giá:** Hiệu quả trong việc cải thiện hiệu suất code, nhưng cần xem xét lại logic sau khi refactor.

## BÀI HỌC AI HÔM NAY CHO DEV

### 1. Tối ưu chi phí & hiệu năng LLM

2. Để tối ưu hóa chi phí và hiệu năng của mô hình ngôn ngữ lớn (LLM), các nhà phát triển cần hiểu cách giảm thiểu việc sử dụng tài nguyên và tăng tốc độ xử lý. Điều này giúp giảm chi phí vận hành và cải thiện trải nghiệm người dùng.

3. Ví dụ, sử dụng kỹ thuật fine-tuning và LoRA (Low-Rank Adaptation) có thể giúp giảm kích thước mô hình và tăng tốc độ huấn luyện. Ví dụ code: `model = transformers.AutoModelForSequenceClassification.from_pretrained('facebook/llama-13b')` sau đó áp dụng fine-tuning và LoRA.

4. Tip hoặc bước tiếp theo: Áp dụng kỹ thuật quantization và pruning để giảm kích thước mô hình và tăng tốc độ xử lý, từ đó tối ưu hóa hiệu năng và chi phí của LLM.

Luôn đi đầu trong thế giới AI! · Stay ahead in AI!

Nguồn: Google News · Groq AI