



# Bản Tin AI Hằng Ngày

Cập nhật công nghệ AI mới nhất

✨ *“Creativity is intelligence having fun.”*

↳ Sáng tạo là trí tuệ đang vui đùa.

— Albert Einstein

💡 *Tư duy sáng tạo nảy sinh khi ta tiếp cận vấn đề với sự tò mò và niềm vui — đừng để áp lực giết chết khả năng sáng tạo.*

## TIN TỨC NỔI BẬT

### 1 Công cụ AI mới của Anthropic có ảnh hưởng đến tất cả chúng ta – dù chúng ta có thể sử dụng nó hay không | Shakeel Hashim

🇬🇧 *Anthropic's new AI tool has implications for us all – whether we can use it or not | Shakeel Hashim*

The Guardian [🔗 Đọc bài viết →](#)

Anthropic, một công ty chuyên về nghiên cứu AI, đã gần đây phát triển một công cụ AI mới. Những ý nghĩa của công cụ này rất sâu rộng và có thể có ý nghĩa quan trọng đối với mọi người, bất kể họ có thể truy cập vào nó hay không. Các khả năng và ứng dụng tiềm năng của công cụ này không được nêu rõ trong thông tin cung cấp. Tuy nhiên, việc đề cập đến Claude Mythos, một AI có vẻ có khả năng hack siêu phàm, gợi ý rằng công cụ này có thể có các khả năng tiên tiến trong các lĩnh vực như trí tuệ nhân tạo, học máy hoặc an ninh mạng. Các chuyên gia được cho là đã cảnh báo về khả năng của Claude Mythos, ngụ ý rằng công cụ này có thể gây ra những rủi ro hoặc thách thức đáng kể. Trong khi đó, chính quyền Trump được đề cập là "mù quáng bởi sự thù địch", điều này có thể cho thấy sự thiếu tham gia hoặc hiểu biết về tác động tiềm năng của công cụ. Bài viết không cung cấp thêm chi tiết về sự phát triển, khả năng hoặc hậu quả tiềm năng của công cụ, để lại nhiều câu hỏi chưa được trả lời.

### 2 Mỹ triệu tập các ông chủ ngân hàng về rủi ro mạng từ mô hình AI mới nhất của Anthropic

🇬🇧 *US summons bank bosses over cyber risks from Anthropic's latest AI model*

The Guardian [🔗 Đọc bài viết →](#)

Các cơ quan quản lý của Mỹ đã tổ chức một cuộc họp tại Washington để thảo luận về các rủi ro mạng tìn ẩn liên quan đến mô hình AI mới nhất của Anthropic. Cuộc họp, diễn ra gần đây, đã tập hợp các giám đốc điều hành hàng đầu của ngân hàng và các quan chức từ Cục Dự trữ Liên bang, bao gồm cả chủ tịch Jerome Powell. Trung tâm của cuộc thảo luận là về các điểm yếu tìn ẩn của hệ thống tài chính Mỹ đối với các mối đe dọa mạng do các mô hình AI tiên tiến như của Anthropic gây ra. Mục tiêu của cuộc họp là đánh giá rủi ro và tìm cách giảm thiểu chúng. Chi tiết cụ thể của cuộc thảo luận và kết quả của cuộc họp không được công bố công khai. Tuy nhiên, việc các bên liên quan quan trọng, bao gồm cả người đứng đầu Cục Dự trữ Liên bang, tham dự cho thấy mức độ nghiêm trọng của vấn đề và nhu cầu hành động phối hợp để giải quyết các mối đe dọa tìn ẩn. Cuộc họp nhấn mạnh mối quan ngại ngày càng tăng về sự giao thoa giữa AI và an ninh mạng trong lĩnh vực tài chính.

3

### Các sản phẩm AI đang thâm nhập sâu hơn vào cuộc sống của chúng ta. Liệu điều đó có quan trọng khi ai kiểm soát các công ty đứng sau chúng? | Van Badham

 *AI products are reaching further into our lives. Does it matter who controls the companies behind them? | Van Badham*

 The Guardian [Đọc bài viết →](#)

Trong một bài viết đầy suy ngẫm, Van Badham bày tỏ mối quan ngại về sự hiện diện ngày càng tăng của các sản phẩm AI trong cuộc sống hàng ngày của chúng ta. Khi những công nghệ này trở nên tích hợp hơn, điều quan trọng là phải xem xét ai kiểm soát các công ty đứng sau chúng. Tác giả nhấn mạnh nhu cầu về "lối chặn" để ngăn chặn thiệt hại có thể gây ra bởi sự dễ sai lầm của con người. Những lối chặn này sẽ hoạt động như một khuôn khổ quản lý để giảm thiểu các hậu quả tiêu cực của các quyết định và hành động được thúc đẩy bởi AI. Bài viết nhấn mạnh tầm quan trọng của trách nhiệm và giám sát trong việc phát triển và triển khai các sản phẩm AI, đề xuất rằng các công ty đứng sau chúng nên chịu trách nhiệm về tác động của chúng. Bằng cách kiểm tra quyền kiểm soát và trách nhiệm của các công ty AI, chúng ta có thể hiểu rõ hơn về tác động của các sản phẩm của họ đối với cuộc sống và làm việc để giảm thiểu thiệt hại tìn ẩn.

4

### Tôi đã do dự trước ý tưởng về 'tối ưu hóa ma sát'. Nhưng có nhiều điều hơn meets mắt | Gaby Hinsliff

 *I balked at the idea of 'friction-maxxing'. But there's more to it than meets the eye | Gaby Hinsliff*

 The Guardian [Đọc bài viết →](#)

Nhật báo viên Gaby Hinsliff khám phá khái niệm "friction-maxxing", một xu hướng tự giúp bản thân có vẻ vô lý khi nhìn thoáng qua. Ý tưởng này liên quan đến việc tạo ra các chướng ngại vật hoặc thách thức trong cuộc sống của một người, chẳng hạn như nấu ăn từ đầu hoặc gặp gỡ bạn bè trực tiếp, thay vì dựa vào sự tiện lợi hoặc công nghệ. Ban đầu, Hinsliff hoài nghi về cách tiếp cận này, nhưng sau khi xem xét kỹ lưỡng, cô tin rằng có thể có nhiều hơn những gì nhìn thấy ở bề mặt. Cô đề xuất rằng xu hướng này khai thác một mong muốn sâu sắc hơn của con người về kết nối, cộng đồng và những trải nghiệm có ý nghĩa. Bằng cách chấp nhận ma sát và không hoàn hảo, các cá nhân có thể có thể nuôi dưỡng một cuộc sống chân thực và đầy đủ hơn. Bài viết của Hinsliff cung cấp một quan điểm sâu sắc về giá trị của việc giảm tốc độ và tham gia với thế giới xung quanh chúng ta, thay vì dựa vào các đường tắt và công nghệ.

5

## 'Không có sự thiếu hụt về công nghệ đáng sợ': làm thế nào AI trở thành nhân vật phản diện mới của phim truyền hình

 *'There's no shortage of terrifying technology': how AI became TV drama's new go-to villain*

 The Guardian [Đọc bài viết →](#)

Trí tuệ nhân tạo đã trở thành nhân vật phản diện phổ biến trong các bộ phim truyền hình hiện đại. Các nhà viết kịch hiện đang mô tả AI như một lực lượng có thể cứu hoặc phá hủy nhân loại. Xu hướng này được phản ánh trong số lượng ngày càng tăng các bộ phim kinh dị có AI là nhân vật phản diện chính. Việc mô tả AI như một thực thể đáng sợ là một sự thay đổi so với những miêu tả trước đây về công nghệ như một lực lượng nhân từ. Thay vào đó, các chương trình truyền hình hiện đang khám phá mặt tối của AI, nhấn mạnh tiềm năng của nó trong việc phá vỡ và hủy hoại xã hội loài người. Sự thay đổi này trong câu chuyện phản ánh mối quan ngại ngày càng tăng về tác động của AI đối với cuộc sống của chúng ta và thế giới nói chung. Khi AI tiếp tục phát triển và trở nên tích hợp nhiều hơn vào cuộc sống hàng ngày, các nhà viết kịch truyền hình đang sử dụng nó như một câu chuyện cảnh báo, cảnh báo người xem về những nguy hiểm tiềm ẩn của việc phụ thuộc quá nhiều vào công nghệ.

6

## Amazon sẽ cuối cùng ra mắt internet vệ tinh Leo vào 'giữa năm 2026', theo CEO

 *Amazon to finally launch Leo satellite internet in 'mid-2026', says CEO*

 The Guardian [Đọc bài viết →](#)

Amazon dự kiến sẽ ra mắt dịch vụ internet vệ tinh Leo vào giữa năm 2026, theo lời CEO Andy Jassy. Sự phát triển này đánh dấu một cột mốc quan trọng cho đối thủ cạnh tranh lâu đời của dịch vụ internet vệ tinh Starlink của Elon Musk. Khi nói chuyện với các cổ đông, Jassy cho biết Leo đang "trên bờ vực" của việc ra mắt, cho thấy dịch vụ đang gần đến ngày ra mắt. Thời gian và chi tiết chính xác của việc ra mắt vẫn chưa được tiết lộ. Dịch vụ internet vệ tinh Leo của Amazon nhằm cung cấp kết nối internet tốc độ cao cho các khu vực thiếu thốn và hẻo lánh trên toàn thế giới, cạnh tranh trực tiếp với Starlink. Việc ra mắt Leo dự kiến sẽ đưa Amazon vào thị trường internet vệ tinh đang phát triển, thị trường này đã chứng kiến sự tăng trưởng đáng kể trong những năm gần đây do nhu cầu ngày càng tăng về kết nối internet đáng tin cậy và nhanh chóng. Với việc ra mắt Leo, Amazon sẵn sàng thách thức sự thống trị của Starlink trên thị trường.

7

## xAI của Elon Musk kiện Colorado về các quy tắc mới cho trí tuệ nhân tạo

 *Elon Musk's xAI sues Colorado over new rules for artificial intelligence*

 The Guardian [Đọc bài viết →](#)

Công ty trí tuệ nhân tạo của Elon Musk, xAI, đã đệ đơn kiện lên tiểu bang Colorado. Lý do đằng sau vụ kiện là các quy tắc mới của tiểu bang về việc quản lý các hệ thống trí tuệ nhân tạo, dự kiến sẽ có hiệu lực vào tháng Sáu. xAI cho rằng những quy định này xâm phạm quyền của họ theo Tu chính án thứ Nhất. Công ty chưa cung cấp thêm chi tiết về các khía cạnh cụ thể của quy định mà họ phản đối hoặc cách chúng vi phạm quyền hiến pháp của họ. Vụ kiện này là diễn biến mới nhất trong cuộc tranh luận đang diễn ra về việc quản lý các hệ thống AI. Các quy tắc mới của Colorado nhằm đảm bảo tính minh bạch và trách nhiệm giải trình trong các quy trình ra quyết định của AI. Vụ kiện này có khả năng được ngành công nghệ và các nhà nghiên cứu AI theo dõi chặt chẽ, vì nó có thể có ảnh hưởng đến việc quản lý

các hệ thống AI trên toàn quốc. Kết quả của vụ kiện sẽ phụ thuộc vào cách tòa án giải thích Tu chính án thứ Nhất và việc áp dụng nó vào quản lý AI.

8

## The Crunch: Đối tượng không gian được trực quan hóa, nhiều bản đồ của Hormuz và các mùa của Melbourne

 *The Crunch: Space objects visualised, the many maps of Hormuz and Melbourne's seasons*

 The Guardian [Đọc bài viết →](#)

Trong những tin tức công nghệ mới nhất, một dự án trực quan hóa mới đã được ra mắt để lập bản đồ các vật thể không gian quay quanh Trái Đất. Dự án này nhằm cung cấp một cái nhìn tổng quan về các vật thể trong quỹ đạo của hành tinh của chúng ta, đặt ra câu hỏi về việc định nghĩa "không gian" trong một môi trường đông đúc và đông dân. Ngoài ra, cuộc khủng hoảng nhiên liệu đang diễn ra đã được nhấn mạnh trong một bản cập nhật gần đây. Bài viết cũng đề cập đến chủ đề lập bản đồ, đặc biệt tập trung vào Eo Hormuz, một tuyến đường thủy quan trọng kết nối Vịnh Ba Tư với Vịnh Oman. Eo Hormuz đã được lập bản đồ bởi các thực thể khác nhau, bao gồm Hải quân Mỹ và chính phủ Iran, mỗi thực thể có quan điểm và mục tiêu riêng. Cuối cùng, bài viết đề cập đến sự thay đổi theo mùa của Melbourne, nhưng không cung cấp thêm chi tiết về chủ đề này.

### ⚡ TIPS & TRICKS CHO DEV

#### ⚡ Sử dụng GitHub Copilot để tạo code nhanh chóng

GitHub Copilot là một công cụ AI mạnh mẽ giúp tạo code nhanh chóng. Bạn có thể sử dụng nó để tạo code mới hoặc hoàn thiện code đã có. Ví dụ, bạn có thể nhập prompt "Tạo một hàm tính diện tích của hình chữ nhật với chiều rộng là 5 và chiều cao là 3" vào GitHub Copilot, và nó sẽ tự động tạo code hoàn chỉnh cho bạn.

#### ⚡ Sử dụng OpenAI ChatGPT để tạo tài liệu hướng dẫn

ChatGPT là một công cụ AI giúp tạo tài liệu hướng dẫn và thông tin. Bạn có thể nhập prompt "Viết một tài liệu hướng dẫn về cách sử dụng GitHub Copilot" vào ChatGPT, và nó sẽ tự động tạo một tài liệu hướng dẫn chi tiết và dễ hiểu cho bạn.

### BÀI HỌC AI HÔM NAY CHO DEV

#### Tối ưu chi phí & hiệu năng LLM (Large Language Model)

Dev cần biết cách tối ưu chi phí và hiệu năng LLM vì nó ảnh hưởng trực tiếp đến hiệu suất của ứng dụng AI. Việc sử dụng LLM có thể tốn kém và đòi hỏi nguồn lực mạnh mẽ, nếu không được tối ưu hóa, có thể dẫn đến giảm hiệu suất ứng dụng. Vì vậy, dev cần biết cách sử dụng LLM một cách hiệu quả và tiết kiệm.

Ví dụ: Để tối ưu hóa hiệu năng LLM trong ứng dụng Chatbot, bạn có thể sử dụng technique "model pruning" để loại bỏ các phần không cần thiết của mô hình, giúp giảm kích thước và đòi hỏi tính toán của LLM.

```
import torch
from transformers import AutoModelForSequenceClassification, AutoTokenizer
```

#### Tính toán mô hình LLM

```
model = AutoModelForSequenceClassification.from_pretrained("distilbert-base-uncased")
tokenizer = AutoTokenizer.from_pretrained("distilbert-base-uncased")
```

### Loại bỏ các phần không cần thiết của mô hình

```
pruned_model = torch.prune.l1_unstructured(model, "weight", amount=0.5)
```

💡 Tip: Hãy xem xét các tính năng "scaling" và "parallelization" để tăng hiệu năng LLM và giảm chi phí tính toán.

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

Nguồn: Google News · Groq AI