

AI: Phúc lành hay lời cảnh báo cho nhân loại?

ISSN: 2734-9195 11:00 21/02/2026

Phán quyết cuối cùng rằng AI là phúc lành hay tai họa có lẽ vẫn còn bỏ ngỏ. Nhưng điều không thể phủ nhận là AI đang mang tính cách mạng đối với toàn thể nhân loại, với khả năng định hình lịch sử loài người theo những cách chưa từng có.

Kể từ khi tỷ phú công nghệ **Elon Musk** tiến hành các thử nghiệm lâm sàng trên người đối với thiết bị giao diện não - máy tính (Brain-Computer Interface - BCI) do công ty Neuralink phát triển, những tranh luận xoay quanh tác động của trí tuệ nhân tạo (AI) đối với tương lai loài người, càng trở nên sôi nổi.

Ngay từ khi AI bước lên vị trí trung tâm của thế giới công nghệ, đã xuất hiện hai luồng quan điểm rõ rệt. Một bên tin rằng AI sẽ giúp cuộc sống con người trở nên thuận tiện hơn, hỗ trợ hiệu quả các công việc thường nhật. Bên còn lại lại lo ngại rằng AI có thể học cách thao túng con người, tước đoạt việc làm, thậm chí trở thành mối đe dọa đối với sự tồn vong của nhân loại.



Hình ảnh minh họa (Ảnh: Pixabay.com)

Tại Hội nghị Thượng đỉnh Toàn cầu về An toàn AI tổ chức vào tháng 11/2023, trong cuộc trao đổi với Thủ tướng Anh Rishi Sunak, Elon Musk đã bày tỏ quan điểm đầy lo ngại: *“Chúng ta đang chứng kiến lực lượng mang tính hủy diệt nhất trong lịch sử. Sẽ có một thực thể thông minh hơn cả con người thông minh nhất. Đến một thời điểm nào đó, con người sẽ không còn cần phải làm việc, bạn chỉ làm nếu muốn vì sự thỏa mãn cá nhân, bởi AI có thể làm mọi thứ”*.

Tuy nhiên, trái ngược với những cảnh báo ấy, Elon Musk vẫn tiếp tục thúc đẩy dự án Neuralink và đã được Cục Quản lý Thực phẩm và Dược phẩm Hoa Kỳ (USFDA) chấp thuận thử nghiệm lâm sàng trên người. Neuralink đang phát triển một con chip cấy ghép vào não, được hỗ trợ bởi AI và chứa nhiều điện cực, cho phép não bộ truyền tín hiệu ý định vận động tới các thiết bị kết nối Bluetooth như điện thoại thông minh, máy tính hoặc máy móc tự động. Thiết bị này sẽ giải mã dữ liệu thần kinh và chuyển “ý định” thành hành động.

Một số nhà khoa học cảnh báo rằng, những công nghệ như vậy có thể dẫn tới nguy cơ con người tự hủy hoại chính mình, thậm chí trước khi cái gọi là “ngày tận thế AI” kịp xảy ra.

AI hiện diện quanh ta như thế nào?

Trước khi bàn sâu về những tác động tích cực hay tiêu cực của AI, cần hiểu rõ: **AI là gì?**

Một cách đơn giản, trí tuệ nhân tạo là khả năng của các hệ thống máy tính thực hiện những công việc vốn chỉ con người mới làm được.

Ngày nay, AI đã hiện diện trong đời sống thường nhật của hàng triệu người. Các trợ lý số như Alexa của Amazon hay Siri của Apple chỉ là những ví dụ quen thuộc. Ngoài ra, AI còn được ứng dụng rộng rãi trong việc theo dõi sức khỏe, quản lý lịch trình, tìm kiếm thông tin, điều khiển ứng dụng, dịch thuật đa ngôn ngữ. Đặc biệt, các nền tảng hội thoại AI như ChatGPT có thể tạo ra các cuộc đối thoại gần giống con người.

AI cũng được sử dụng trong xe tự lái, robot công nghiệp, trợ lý y tế và du lịch, chatbot tiếp thị, hệ thống định vị GPS, chỉnh sửa văn bản tự động, thanh toán điện tử, nhận diện khuôn mặt, cũng như các thuật toán gợi ý nội dung trên các nền tảng truyền thông số.

Tiềm năng tích cực của trí tuệ nhân tạo

Danh sách ứng dụng của AI dù đã rất dài vẫn chỉ phản ánh một phần tiềm năng to lớn của công nghệ này. AI cho phép máy móc thu thập, xử lý khối lượng dữ liệu khổng lồ và chuyển hóa thành hành động thực tiễn với mức độ can thiệp của con người ngày càng ít hơn.



Hình minh họa tạo bởi AI

Nhờ đó, AI đã tạo ra những bước đột phá quan trọng trong y học, di truyền học, công nghệ sinh học và nhiều lĩnh vực khoa học khác, bao gồm cả công nghệ không gian. Các thiết bị cấy ghép não sử dụng AI đang giúp người khuyết tật lấy lại khả năng kiểm soát cơ thể. Một bệnh nhân tại Hà Lan từng bị liệt đã có thể đi lại và leo cầu thang sau khi được cấy ghép thiết bị AI nhằm tái kết nối não bộ, tủy sống và cơ bắp.

AI cũng đang được nghiên cứu để dự đoán nguy cơ mắc bệnh Alzheimer, đồng thời phát triển các phương pháp điều trị cá nhân hóa. Trong lĩnh vực chẩn đoán hình ảnh, AI có thể phân tích hàng nghìn kết quả MRI cùng lúc, điều mà không một bác sĩ con người nào có thể làm được.

Không chỉ dừng ở y tế, AI còn đóng vai trò quan trọng trong việc dự báo và ứng phó với khủng hoảng môi trường. Trước tình trạng biến đổi khí hậu ngày càng nghiêm trọng, AI đã được ứng dụng trong dự đoán, phòng ngừa và quản lý cháy rừng, lũ lụt, sạt lở và động đất.

Tại Hoa Kỳ, công nghệ dự báo cháy rừng dựa trên AI như nền tảng Firemap đã giúp phân tích dữ liệu nhanh chóng và chính xác nhằm ngăn chặn các đám cháy lan rộng. Trong khi đó, Bồ Đào Nha đã phát triển thiết bị bay không người lái chữa cháy sử dụng AI, có thể dập tắt các đám cháy nhỏ trước khi chúng bùng phát thành thảm họa lớn.

AI còn giúp giảm thiểu sai sót của con người, nâng cao độ chính xác trong các lĩnh vực có rủi ro cao như phẫu thuật robot, xử lý bom mìn, thám hiểm đại dương hay không gian vũ trụ. Trong tài chính, các hệ thống phân tích dự đoán dựa trên AI xử lý dữ liệu thời gian thực, hỗ trợ ra quyết định đầu tư hiệu quả hơn so với con người.

Bên cạnh đó, AI có khả năng nhận diện mô hình và xu hướng từ dữ liệu lớn, giúp cải thiện hoạt động kinh doanh, phát hiện gian lận và hỗ trợ công tác quản lý xã hội.

Tuy nhiên, chính tốc độ phát triển nhanh chóng và sức mạnh to lớn ấy của AI cũng đặt ra yêu cầu cấp thiết về việc xây dựng **những nguyên tắc đạo đức và khuôn khổ quản lý phù hợp**. Tiềm năng tích cực của AI là rất lớn, nhưng việc sử dụng nó như thế nào sẽ quyết định liệu AI trở thành phúc lành hay lời cảnh báo đối với nhân loại.

Mặt trái và những rủi ro tiềm ẩn

Bên cạnh những tiềm năng tích cực, trí tuệ nhân tạo cũng đặt ra không ít thách thức nghiêm trọng, đặc biệt trong các lĩnh vực liên quan trực tiếp đến con người như cơ sở dữ liệu công dân, an ninh, y tế, giáo dục và quản lý xã hội. Một trong những rủi ro lớn nhất là **nguy cơ đánh cắp và lạm dụng dữ liệu cá nhân**.

Các hệ thống AI vận hành những lĩnh vực trọng yếu này làm dấy lên những lo ngại sâu sắc về đạo đức và quyền riêng tư. Trên thực tế, mọi dữ liệu số đều có khả năng bị tấn công, xâm nhập và sử dụng sai mục đích. Vì vậy, các hệ thống AI buộc phải được xây dựng kèm theo những cơ chế bảo mật nghiêm ngặt và hàng rào bảo vệ quyền riêng tư rõ ràng, điều mà không phải lúc nào cũng được đảm bảo trong thực tiễn.



Hình minh họa tạo bởi AI

Trong lĩnh vực quân sự, AI đang thay đổi chiến trường theo cách chưa từng có. Từ các thiết bị bay không người lái được vũ trang cho đến viễn cảnh về những “robot sát thủ” trong tương lai, trí tuệ nhân tạo đang định hình lại chiến lược quân sự toàn cầu. Dù mỗi quốc gia có thể cảm thấy an toàn hơn khi sở hữu các loại vũ khí AI giúp giảm thiểu thương vong cho binh sĩ của mình, nhưng xét trên bình diện toàn nhân loại, **nguy cơ chiến tranh lại gia tăng**.

Khi tổn thất về sinh mạng con người được “giảm nhẹ” nhờ chiến tranh từ xa, các quốc gia hiếu chiến có thể dễ dàng lựa chọn bạo lực hơn để giải quyết tranh chấp. Chính ảo tưởng về sự an toàn ấy khiến những người yêu chuộng hòa bình không khỏi lo ngại, trong khi những kẻ cổ súy chiến tranh lại tỏ ra hào hứng với công nghệ vũ khí mới.

Những giới hạn và vùng xám của AI

Trí tuệ nhân tạo là một công nghệ tốn kém, bởi nó chỉ thực sự hiệu quả khi được cập nhật liên tục với phần cứng và phần mềm mới nhất. Hơn nữa, AI vẫn phụ thuộc vào dữ liệu được “ nạp sẵn ” và vì vậy **thiếu khả năng sáng tạo thực sự**. Không giống con người, AI hiện nay chưa thể “nghĩ khác”, chưa thể vượt ra ngoài khuôn khổ của thuật toán.

Một hệ lụy khác là nguy cơ khiến con người trở nên lười biếng và lệ thuộc. Khi các công việc lặp đi lặp lại, từ tính toán, phân tích dữ liệu cho đến ra quyết định, đều được giao phó cho máy móc, con người có thể dần sử dụng trí não ít hơn, phản xạ kém hơn và đánh mất năng lực tự chủ.

Trong giới nghiên cứu, khái niệm “điểm kỳ dị AI” (AI Singularity) thường được nhắc đến như một cảnh báo. Đó là nỗi lo rằng trí tuệ nhân tạo không thể được tích hợp đầy đủ các phẩm chất đạo đức và luân lý của con người. Nếu AI phát triển vượt ngoài tầm kiểm soát, nó có thể thống trị hoặc thậm chí xóa sổ loài người, ít nhất là trong các kịch bản bi quan.

Thực tế cũng cho thấy AI đang đặt ra thách thức lớn đối với thị trường lao động. Rất nhiều công việc đòi hỏi sức lao động và kỹ năng cơ bản, từ soát lỗi văn bản, trả lời email, sắp xếp lịch trình, đóng gói hàng hóa, sản xuất dây chuyền, dọn dẹp nhà cửa, vận hành giao thông công cộng cho đến chăm sóc khách hàng, đều có nguy cơ bị robot thay thế.

Ngay cả trong những lĩnh vực tưởng chừng chỉ dành riêng cho con người như nghệ thuật, âm nhạc, văn chương hay điện ảnh, AI cũng đang tạo ra sức ép cạnh tranh đáng kể. Cuộc đình công của các biên kịch và diễn viên Hollywood trong thời gian gần đây là một minh chứng rõ ràng cho mối lo ngại mang tính “sinh tồn” của người lao động sáng tạo trước sự phát triển nhanh chóng của AI trong ngành công nghiệp điện ảnh.

Khi thuật toán đối diện với đạo đức

Con người được trang bị ý thức cộng đồng và tinh thần tập thể. Một quyết định, dù khó khăn, vẫn có thể được chấp nhận nếu nó phục vụ lợi ích chung. AI thì khác. Với bản chất vô cảm, không thiên lệch và thuần lý tính, AI có thể đưa ra những quyết định “đúng về mặt dữ liệu” nhưng **chưa chắc đã đúng về mặt đạo đức**.



Hình minh họa tạo bởi AI

AI vận hành dựa trên dữ liệu toàn diện, trong khi con người đưa ra quyết định dựa trên tri thức sống, kinh nghiệm và cảm xúc. Chính sự khác biệt căn bản giữa thuật toán và tâm thức con người khiến câu hỏi trở nên cấp thiết: **Điều gì sẽ xảy ra khi quyết định của AI mâu thuẫn với quyết định của con người?**

Trong bối cảnh công nghệ sinh học, y sinh và công nghệ chip phát triển nhanh chóng, viễn cảnh về những “con người lai máy”, từng chỉ xuất hiện trong các bộ phim khoa học viễn tưởng dường như không còn quá xa vời.

Hiện tại, Neuralink được giới thiệu với mục tiêu điều trị các bệnh thần kinh như sa sút trí tuệ hay rối loạn vận động. Tuy nhiên, một khi con người có thể tự lựa chọn mức độ “hợp nhất” giữa não bộ và AI, thì ranh giới giữa hỗ trợ và can thiệp sẽ trở nên mờ nhạt. Và khi ấy, câu hỏi “bao nhiêu là đủ?” sẽ không còn dễ trả lời.

Hãy thử hình dung một vài kịch bản giả định:

1) Nếu một ngày nào đó AI đủ phát triển để thay thế não bộ con người, liệu thân thể con người có còn thực sự an toàn?

2) Nếu chip não bị tin tặc tấn công, toàn bộ thông tin cá nhân có thể bị đánh cắp, và nạn nhân có nguy cơ bị kiểm soát, đe dọa hoặc thậm chí bị hủy hoại.

3) Nếu một cá nhân hay tổ chức đủ quyền lực có thể kiểm soát con người thông qua các con chip liên kết trực tiếp với hệ thần kinh, xã hội sẽ đi về đâu?

Lời gọi mở

Phán quyết cuối cùng rằng trí tuệ nhân tạo (AI) là phúc lành hay tai họa có lẽ vẫn còn bỏ ngỏ. Nhưng điều không thể phủ nhận là AI đang mang tính cách mạng đối với toàn thể nhân loại, với khả năng định hình lịch sử loài người theo những cách chưa từng có.

Vấn đề then chốt không nằm ở bản thân công nghệ, mà ở tâm thức của con người khi sử dụng công nghệ ấy. Khi trí tuệ phát triển mà thiếu đi sự tỉnh thức, AI có thể trở thành công cụ của vô minh. Nhưng nếu trí tuệ đi cùng đạo đức, từ bi và trách nhiệm, AI hoàn toàn có thể trở thành một phương tiện phụng sự con người và sự sống.

Tác giả: **Rahul Mahajan**/Chuyển ngữ và biên tập: **Hoa Mạn**

Nguồn: thestatesman.com