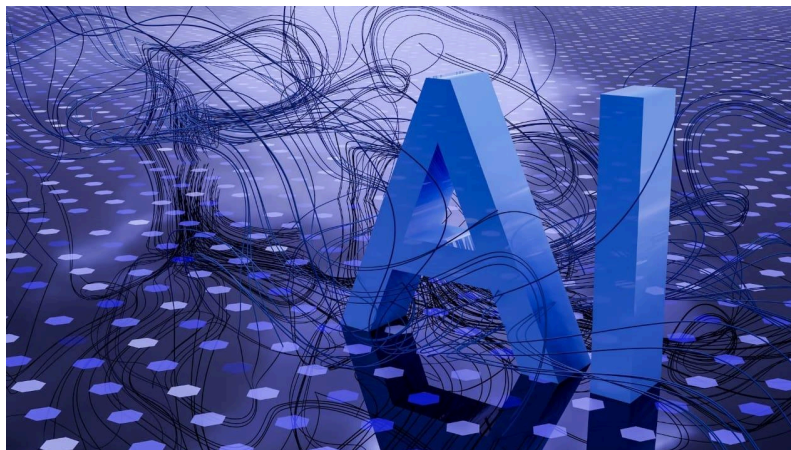


Đi tìm sự an toàn trong công nghệ, coi chừng phản tác dụng!

Lê Tây Sơn/SGN

14 tháng 11, 2023



(ảnh: Steve Johnson/Unsplash)

Kiện tụng không phải là cách tốt nhất để giữ an toàn cho ngành công nghệ. Đó là kết luận bài phân tích về vấn đề này trên tờ [The Washington Post](#) của hai chuyên viên truyền thông Louisa Bartolo và Dominique Carlon tại Đại học Công nghệ Queensland và Noah Giansiracusa, phó giáo sư toán học và khoa học dữ liệu tại Đại học Bentley, tác giả cuốn sách “How Algorithms Create and Prevent Fake News” (cách thuật toán tạo và ngăn chặn tin tức giả).

Trí khôn nhân tạo (AI) làm cho công nghệ nguy hiểm hơn với trẻ em

Hai năm trước, một cựu nhân viên trở mặt tố cáo Facebook đã hé lộ cho thế giới thấy, công ty (hiện có tên mới là Meta) biết rất rõ Instagram có hại cho sức khỏe tâm thần của người dùng tuổi teen nhưng vẫn không có biện pháp điều chỉnh lại nền tảng để tránh gây nghiện cho người dùng.

Hai tuần trước, 41 tiểu bang của Mỹ đã khởi kiện Meta. Vụ kiện có vẻ là một bước tiến lớn hướng tới công lý và một môi trường trực tuyến an toàn hơn cho con em chúng ta, nhưng nếu không cẩn thận, nó có thể gây tác dụng ngược và khiến ngành công nghệ từ bỏ mọi nỗ lực tìm hiểu mặt tiêu cực của sản phẩm để giảm những tác hại của nó.

Điểm chính trong vụ kiện không chỉ là Meta không chỉ gây ra thiệt hại mà còn “cố tình gây thiệt hại” sau khi biết được các nguy cơ. Đơn kiện nêu rõ, Meta từng nghiên cứu và phát hiện ra

những rủi ro đáng kể về sức khỏe của sản phẩm (Instagram) nhưng vẫn che giấu bất chấp ảnh hưởng nghiêm trọng đến trẻ em.

Đây là sự thật “không thể chối cãi”. Vấn đề còn lại là những tác động ngược phía sau đơn kiện. Nếu việc một công ty tự truy tìm những rủi ro trong sản phẩm của mình lại gây ra thiệt hại về tài chính, pháp lý và danh tiếng cho chính nó thì cách an toàn nhất nhìn từ góc độ kinh doanh là không truy tìm chúng nữa!

Việc các công ty công nghệ đang chạy đua để đưa chatbot và các sản phẩm trí tuệ nhân tạo tổng quát (generative artificial intelligence) đến hàng triệu trẻ em, định hình lại đời sống xã hội và sự phát triển nhận thức của chúng mà chưa biết có lợi hay hại nên được xem “liều lĩnh không thể chấp nhận được”.

Dẫn chứng, vào đầu năm nay, Snap, công ty điều hành nền tảng nhắn tin phổ biến Snapchat đã phát hành chatbot “My AI” cho phép người dùng trao đổi tin nhắn với bạn đồng hành kỹ thuật số ở khắp mọi nơi. Chỉ hai tháng sau, Snap khoe “thu hút được trên 150 triệu người gửi hơn 10 tỷ tin nhắn tới My AI”! Tín đồ của Snapchat chủ yếu là giới trẻ, chiếm 1/5 số người dùng (hơn 125 triệu) dưới 18 tuổi.

Chúng háo hức thử nghiệm cái mới. Đây là một con số đáng kinh ngạc. Hai câu hỏi đặt ra là: liệu trẻ em có bị nghiện chatbot 24 giờ một ngày không? Liệu chúng sẽ giảm giao tiếp bạn bè ngoài đời thật không? Sẽ cần nghiên cứu sâu để trả lời những câu hỏi này, nhưng thật ngây thơ khi hy vọng Snap sẽ tự nguyện bù đắp vào những hậu quả tiêu cực của “My AI”, đặc biệt khi xuất hiện những nguy cơ không được dự báo trước.

Không chịu thua Snap, Meta đang phát hành các nhân vật nổi tiếng được hỗ trợ bởi AI để giới trẻ có thể trò chuyện bất cứ lúc nào. Liệu các bản sao AI của những người nổi tiếng như Tom Brady và Kendall Jenner có làm trầm trọng thêm xu hướng nghiện Instagram của thanh thiếu niên?



Thế giới của người mẫu thời trang đang xuất hiện sự “cạnh tranh” của người mẫu A.I. (ảnh: Oleksii Chumachenko/SOPA Images/LightRocket via Getty Images)

Đừng trông chờ Meta trả lời câu hỏi này ngay bây giờ. Hiểu biết về tác động của chatbot và AI không chỉ để ngăn chặn thiệt hại cho cả một thế hệ mà còn tránh gây ra nỗi sợ hãi vô căn cứ và đưa ra “những quy định thiếu hiểu biết” để giải quyết “những vấn nạn tưởng tượng”.

An toàn trực tuyến của trẻ em đã được các chính trị gia và công chúng quan tâm rất nhiều, và sắp tới sẽ có thêm nhiều hành động pháp lý về lĩnh vực này hơn bất kỳ lĩnh vực quản lý internet nào khác.

Tránh các quy kết cực đoan

Cách giải quyết tác hại của thế giới mạng là đừng quá cực đoan, một chiều, mà nên có cái nhìn bao quát hơn. Ví dụ như: bảo vệ quyền riêng tư dữ liệu của trẻ em tốt hơn (một phần quan trọng trong vụ kiện Meta), không cho các công ty công nghệ sử dụng thông tin người dùng trẻ, không thu hút sự tò mò của trẻ bằng các tính năng gây nghiện và quảng cáo lôi cuốn; tăng cường sự giám sát của cha mẹ ở lứa tuổi nhạy cảm.

Sắc lệnh hành pháp của Tòa Bạch Ốc về AI ban hành tháng trước có nhiều điều khoản quan trọng liên quan đến an toàn và minh bạch trên không gian mạng, nhưng sức khỏe tâm thần của trẻ lại bị xem nhẹ.

Ngoài ra, cũng cần có nhiều công ty công nghệ dám tự điều tra sự an toàn sản phẩm của mình và mời các nhóm chuyên viên tham gia để có hành động kịp thời dựa trên kết quả nghiên cứu, trước hết là cảnh báo người dùng về những tác hại tiềm ẩn của sản phẩm và đưa ra các lựa

chọn điều chỉnh thiết kế nhằm giảm thiểu tác hại ngay cả khi động thái này sẽ làm mất bớt doanh thu trong ngắn hạn.

Vậy làm thế nào thực hiện giải pháp này? Đầu tiên, hãy điều tiết ngành công nghệ theo cách đang làm với ngành y tế và dược phẩm. Nếu các công ty dược phẩm phải thử nghiệm thuốc của họ, các công ty công nghệ cũng phải kiểm tra kỹ lưỡng và thiết lập độ an toàn cho sản phẩm của mình trước khi tung nó ra thị trường (không nhất thiết phải áp dụng cho tất cả các sản phẩm, nhưng áp dụng trước cho những sản phẩm trẻ em dễ tiếp cận).

Ngành công nghệ lập luận sự thận trọng này sẽ kéo chậm đáng kể quá trình phát triển sản phẩm, nhưng các cuộc thăm dò gần đây về AI cho thấy đây chính xác là điều công chúng mong muốn.



GPT-4, phiên bản A.I. mới nhất, mạnh nhất nâng cấp từ ChatGPT của công ty OpenAI, được cho là một bộ não tinh tế không thua kém bộ óc con người. (ảnh: Jaap Arriens/NurPhoto via Getty Images)

Nếu người Mỹ có thể chờ đợi vaccine trong một trận đại dịch chết người thì chúng ta có thể chờ đợi những chatbot an toàn dành cho trẻ em. Vấn đề không phải là buộc các công ty công nghệ phải loại bỏ “sản phẩm trong tầm ngắm” mà là thúc đẩy họ thiết kế sản phẩm cẩn thận hơn, loại bỏ các tính năng có hại cho giới trẻ và thêm nhãn cảnh báo khi chưa chắc chắn về độ an toàn.

Thực tế cho thấy, các sản phẩm công nghệ không đứng yên tại chỗ mà được điều chỉnh liên tục theo thời gian. Cách sử dụng và tác động của chúng cũng thay đổi, từ vô hại trở thành có

hại. Ví dụ TikTok xuất thân là một nền tảng khiêu vũ đã trở thành một nguồn tin tức, giải trí, kiến thức và lời khuyên về đủ thứ trên đời. Vì vậy, việc theo dõi và nghiên cứu liên tục về sản phẩm là điều bắt buộc phải làm.

Hoa Kỳ có thể rút ra một số bài học từ Đạo luật Dịch vụ Kỹ thuật số (Digital Services Act) được Liên minh Châu u (EU) thông qua gần đây, trong đó yêu cầu các công ty như Meta tự đánh giá rủi ro và công bố kết quả, đặc biệt là rủi ro đối với trẻ em.

Đạo luật này cũng yêu cầu các công ty công nghệ cho phép các nhà nghiên cứu độc lập bên ngoài truy cập vào dữ liệu nền tảng để thẩm tra, đánh giá, tránh việc các công ty tự chấm điểm tốt cho mình.

Mục đích đằng sau các vụ kiện chống lại Meta là chính đáng, nhưng trong khi trừng phạt các công ty có hành vi sai trái, pháp luật cũng tránh làm trầm trọng thêm vấn đề bằng cách đẩy các công ty vào thế thụ động và thiếu minh bạch với sản phẩm của mình.

<https://saigonnhonews.com>