

Trí tuệ nhân tạo và khả năng "siêu lan truyền" thông tin sai lệch

07:09 05/01/2024

Theo NewsGuard, tổ chức ở New York chuyên theo dõi thông tin sai lệch, trí tuệ nhân tạo (AI) đang đưa việc sản xuất và lan truyền tin giả lên cấp độ mới, có thể tạo nội dung sai lệch về bầu cử, chiến tranh, thiên tai với tốc độ nhanh và khó phân biệt với thông tin thật.

Theo NewsGuard, kể từ tháng 05/2024, các website chứa bài viết sai sự thật do AI tạo đã tăng hơn 1.000%, từ 49 trang lên hơn 600 trang.



Trí tuệ nhân tạo có thể khiến tin giả có độ chân thực bất ngờ.

Mối nguy mới là trí tuệ nhân tạo siêu lan truyền thông tin sai lệch

Trước đây, nội dung rác hoặc sai lệch thường được tạo bởi cá nhân hoặc tổ chức cho mục đích nhất định. Tuy nhiên, chỉ cần nhập một vài yêu cầu đơn giản lên các mô hình AI tạo sinh như ChatGPT, giờ đây ai cũng có thể tạo ra các nội dung mà người dùng phổ thông sẽ khó phân biệt với tin tức thực tế.

Mới nhất, AI tạo câu chuyện bịa đặt về bác sỹ của Thủ tướng Israel Benjamin Netanyahu đã chết và để lại mảnh giấy "gợi ý có sự tham gia của ông Netanyahu". Nhân vật bác sỹ là hư cấu, nhưng câu chuyện đã được một chương trình truyền hình Iran đưa tin, sau đó một số trang truyền thông đăng tải, cũng

như được lan truyền trên TikTok, Reddit và Instagram.

“Một số website dùng trí tuệ nhân tạo để sinh ra hàng trăm, thậm chí hàng nghìn bài báo mỗi ngày. Đây là lý do chúng tôi gọi AI là kẻ siêu truyền bá thông tin sai lệch tiếp theo”, Jack Brewster chuyên gia đứng đầu nhóm nghiên cứu của NewsGuard cho hay.



Hình ảnh Giáo hoàng Francis mặc áo phao màu trắng đi trên phố do trí tuệ nhân tạo tạo ra khiến không ít người hoang mang.

Cũng theo Brewster, website đăng nội dung AI được đặt tên chung chung, như iBusiness Day hay Ireland Top News. Để tăng tính thuyết phục, chúng thường xen kẽ tin giả với tin thật lấy từ các nguồn uy tín.

“Đây là mối lo ngại lớn. Sự nguy hiểm nằm ở phạm vi và quy mô của AI, nhất là khi chúng đang trở nên thông minh, chứa thuật toán tạo nội dung phức tạp và khó phân biệt. Đó là cuộc chiến thông tin ở quy mô chúng tôi chưa từng thấy trước đây”, ông Brewster nhận định.

Theo các chuyên gia, người dùng có thể phát hiện nội dung AI thông qua cách hành văn, nhất là phần “ngữ pháp kỳ quặc” hoặc lỗi về cấu trúc câu. Tuy nhiên, công cụ hiệu quả nhất là nâng cao hiểu biết và có thói quen kiểm tra thêm thông tin từ nguồn chính thống thay vì chỉ tin vào một nguồn duy nhất.

Chống tin giả là cuộc rượt đuổi

Tiến sỹ Trần Thế Trung, Viện trưởng Viện nghiên cứu FPT, Phó Giám đốc Trung tâm sản phẩm AI-FPT Smart cho rằng: AI cũng tương tự như các công nghệ khác, nếu dùng sai đều có thể gây hại cho con người. Song, AI có khả năng gây hại lớn hơn do nó có khả năng sinh ra nội dung một cách thuyết phục khiến các

việc giả mạo và lừa đảo trở nên thật hơn từ lời nói, khuôn mặt... thậm chí là âm nhạc giả mạo.



Tiến sỹ Trần Thế Trung, Viện trưởng Viện Nghiên cứu FPT, Phó Giám đốc Trung tâm Sản phẩm AI-FPT Smart.

“Khi công nghệ ngày càng phát triển, con người cũng buộc phải nâng cấp trình độ của mình lên. Con người cần phải học cách tìm nguồn kiểm chứng thông tin. Nôm na, mỗi người cần “đóng vai” như một nhà báo, không phải là vai trò đi phát tán thông tin như thời kỳ mạng xã hội thịnh hành, mà là tư duy kiểm chứng, phản biện, tìm nguồn gốc thông tin...”, Tiến sỹ Trần Thế Trung khuyến nghị.

Theo Tiến sỹ Trung, cơ quan chức năng cũng cần xây dựng danh sách các nguồn thông tin chính thống, đáng tin cậy từ cơ quan của chính phủ, đơn vị truyền thông chính thống... để người dân có thể dựa vào kiểm chứng thông tin.

Ông Lê Quang Tự Do, Cục trưởng Cục Phát thanh – Truyền hình và Thông tin điện tử, Bộ Thông tin và Truyền thông cũng thừa nhận tin giả, tin sai sự thật xuất hiện và lan truyền với tốc độ chóng mặt. Việc kiểm soát thông tin hiện nay càng ngày càng trở nên khó khăn và phức tạp hơn.

“Chưa bao giờ chúng ta gặp tình trạng hậu kiểm như hiện nay. Tức là mọi người chia sẻ thông tin, sau đó mới có sự tham gia, kiểm tra kỹ lưỡng”, ông Lê Quang Tự Do chia sẻ.



Ông Lê Quang Tự Do, Cục trưởng Cục Phát thanh - Truyền hình và Thông tin điện tử, Bộ Thông tin và Truyền thông.

Theo ông Lê Quang Tự Do, để hạn chế được tin giả, bản thân người dùng Internet phải có ý thức kiểm duyệt thông tin trước khi đăng. Không chỉ người trẻ, người già cũng là đối tượng dễ bị tác động bởi tin giả.

“Tâm lý học nghiên cứu con người nhận bình phẩm về mình trên một lần là quá giới hạn chịu đựng. Nhưng từ khi có mạng xã hội, một người có thể nhận một triệu bình luận nói xấu của mình. Trong khi đó, có những người chưa bao giờ được trang bị kiến thức, kỹ năng đối phó với những lời bình phẩm, nhận xét như vậy”, ông Lê Quang Tự Do dẫn chứng.

“Vì thế các bạn trẻ đừng dễ dàng tin ngay những gì trên mạng. Chúng ta đọc gì cũng check hai lần, thứ nhất là xem thông tin có uy tín, đáng tin hay không; thứ hai, tập phớt lờ, chấp thuận, không nên đối xử như những lời bình phẩm trực tiếp được, nếu không khó sống, khó tồn tại”, ông Lê Quang Tự Do khuyến cáo.

Theo VOV.VN