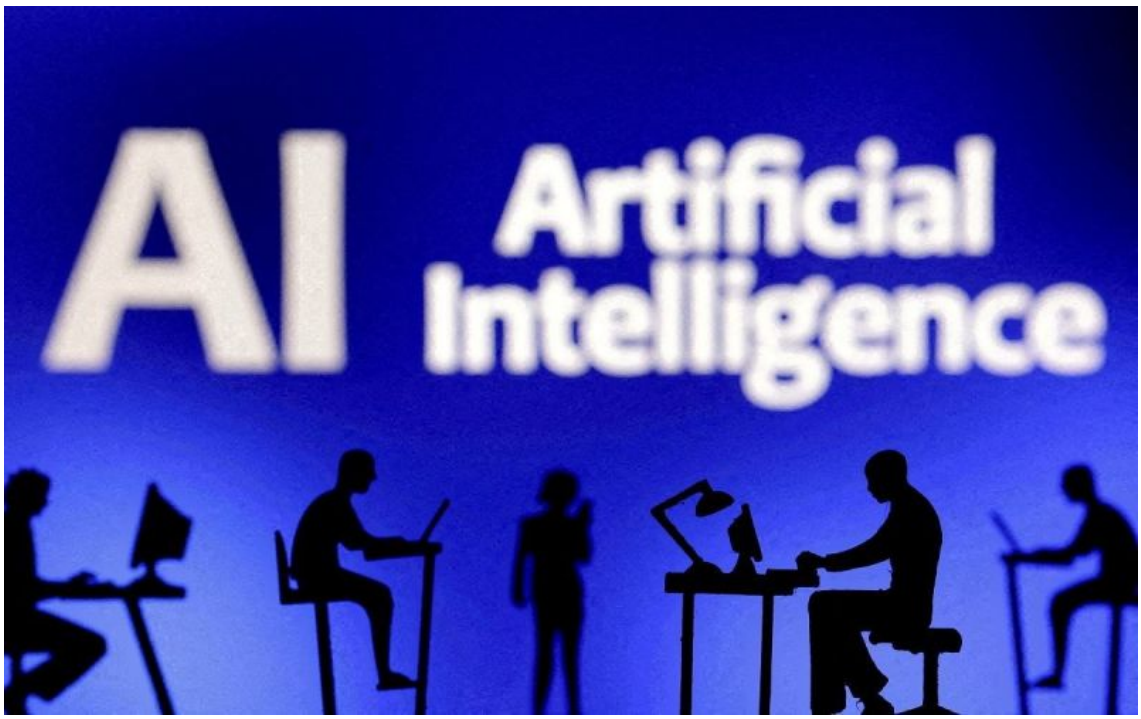
TẠP CHÍ ĐIỆN TỬ
THƯƠNG HIỆU & CÔNG LUẬN

Hôm nay, OECD xây dựng ‘hướng dẫn mới’ cho nhà phát triển trí tuệ nhân tạo

07:43 02/05/2024

Mặc dù "Nguyên tắc của OECD về AI" không có tính ràng buộc về mặt pháp lý nhưng đã được 46 quốc gia ký kết. Điều này có tác động đến việc các nước xây dựng chính sách về AI của riêng mình.

Tại Hội nghị Hội đồng Bộ trưởng ngày 2-3/5 ở Paris, Pháp, Tổ chức Hợp tác và Phát triển kinh tế (OECD) có kế hoạch xây dựng "hướng dẫn mới", yêu cầu các nhà phát triển trí tuệ nhân tạo (AI) giải quyết thông tin sai lệch do công nghệ này tạo ra.



OECD: Trí tuệ nhân tạo được ứng dụng nhiều vào đời sống nhưng cũng gây ra những mối đe dọa do thông tin sai lệch. Nguồn Reuters.

Các hướng dẫn mới sẽ được đưa vào dự thảo đề xuất thay đổi khuyến nghị “Nguyên tắc của OECD về AI” (còn gọi là "Hướng dẫn quốc tế về AI") hiện đang được tổ chức xem xét.

Trước đó, vào tháng 9/2019, tại Hội nghị Bộ trưởng thường niên của OECD ở Pháp, 42 quốc gia đã thông qua dự thảo khuyến nghị “Nguyên tắc của OECD về AI".

Văn bản trên không mang tính chất ràng buộc về mặt pháp lý mà chỉ mang tính khuyến nghị. Tuy nhiên, do tới thời điểm đó vẫn chưa có nguyên tắc quốc tế nào về AI, cho nên có thể coi đây là tiêu chuẩn quan trọng trong tương lai của công nghệ này mà bất kỳ nước nào cũng nên tham khảo khi đưa ra các chính sách liên quan.

Các thành viên OECD đều thừa nhận, AI đang cách mạng hóa lối sống và làm việc, mang lại lợi ích phi thường cho xã hội và nền kinh tế. Tuy nhiên, AI cũng đặt ra những thách thức mới và làm gia tăng những lo lắng cũng như mối quan tâm về đạo đức.

Điều này đặt trách nhiệm lên các chính phủ để đảm bảo rằng hệ thống AI được thiết kế theo cách tôn trọng luật pháp và các giá trị của con người, sự an toàn và riêng tư của con người là tối quan trọng.



Hôm nay, OECD xây dựng ‘hướng dẫn mới’ cho nhà phát triển trí tuệ nhân tạo. Ảnh internet.

"Nguyên tắc của OECD về AI" là một tham chiếu toàn cầu đáng tin cậy để các nước có thể khai thác các cơ hội mà AI mang lại theo cách tốt nhất cho tất cả.

Kể từ khi OECD áp dụng nguyên tắc trên vào năm 2019, việc sử dụng AI để tạo ra văn bản và video phức tạp đã phát triển nhanh chóng. Do đó, một điều khoản mới yêu cầu các nhà phát triển và nhà cung cấp dịch vụ AI thực hiện hành động chống lại thông tin sai lệch do AI tạo ra dự kiến sẽ được thêm vào các nguyên tắc.

Dự thảo "hướng dẫn mới" dự kiến được thông qua tại Cuộc họp Hội đồng Bộ trưởng OECD được tổ chức vào ngày 2-3/5.

Có nguồn tin cho biết, với các cuộc bầu cử quan trọng trên thế giới sắp diễn ra như bầu cử Nghị viện châu Âu (tháng 6) và tổng thống Mỹ (tháng 11), các nước thành viên OECD đã chia sẻ nhận thức về mối đe dọa từ thông tin sai lệch do AI tạo ra.

Dự thảo cũng sẽ sửa đổi các quy định về tính minh bạch và trách nhiệm giải trình của AI. Điều này phản ánh thỏa thuận đạt được trong Quy tắc dành cho AI ở Hiroshima (Nhật Bản) vào tháng 5/2023 tại Hội nghị thượng đỉnh Nhóm các nước công nghiệp phát triển hàng đầu thế giới (G7).

Bản sửa đổi yêu cầu các nhà phát triển AI công khai thông tin về khả năng và hạn chế của công nghệ này, cũng như thông tin về dữ liệu đào tạo và cách AI tạo ra thông tin.

Theo baoquocte.vn