



The AI Conversation



What Parents Need to Know

“Trí tuệ nhân tạo” (AI) là gì?

AI KHÔNG phải là một bộ não

Đó là một công cụ 'nhận diện khuôn mẫu' rất tinh vi.



Hãy coi AI như tính năng tự điền tối ưu trong tin nhắn văn bản của bạn. Nó đọc gần như mọi thứ trên internet và dự đoán bước tiếp theo có khả năng xảy ra nhất.



Tóm lại:

Trí tuệ nhân tạo không biết mọi thứ như chúng ta, nó tính toán câu trả lời có khả năng nhất dựa trên các mẫu. Nó không có ý thức, nó không suy nghĩ.

Nó dùng để làm gì?

Nếu sử dụng đúng cách, nó có thể là một gia sư cá nhân tuyệt vời

Giải thích các đề tài phức tạp
Vd: “Giải thích cho người 12 tuổi hiểu”

Đưa ra các ý tưởng để giúp khắc phục hội chứng “trang giấy trắng”

Luyện tập cá nhân hóa
Vd: tạo các bài kiểm tra và trắc nghiệm dựa trên sách giáo khoa

Học ngôn ngữ
Dịch văn bản sang ngôn ngữ mẹ đẻ để hỗ trợ việc học

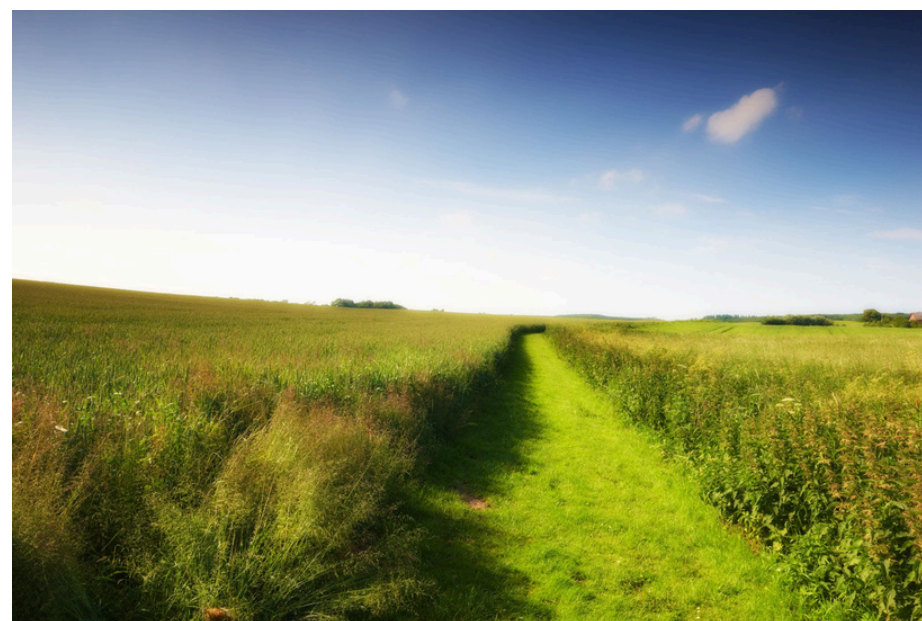
Nó KHÔNG dùng để làm gì?

Việc dựa vào AI tiềm ẩn nhiều rủi ro trong giáo dục.

AI có thể ‘tạo ảo giác’. Nó có thể dựng lên các sự kiện và số liệu.

Các thông tin từ AI cần phải được kiểm tra lại.

Giống như có một giáo viên thỉnh thoảng nói dối bạn vậy.



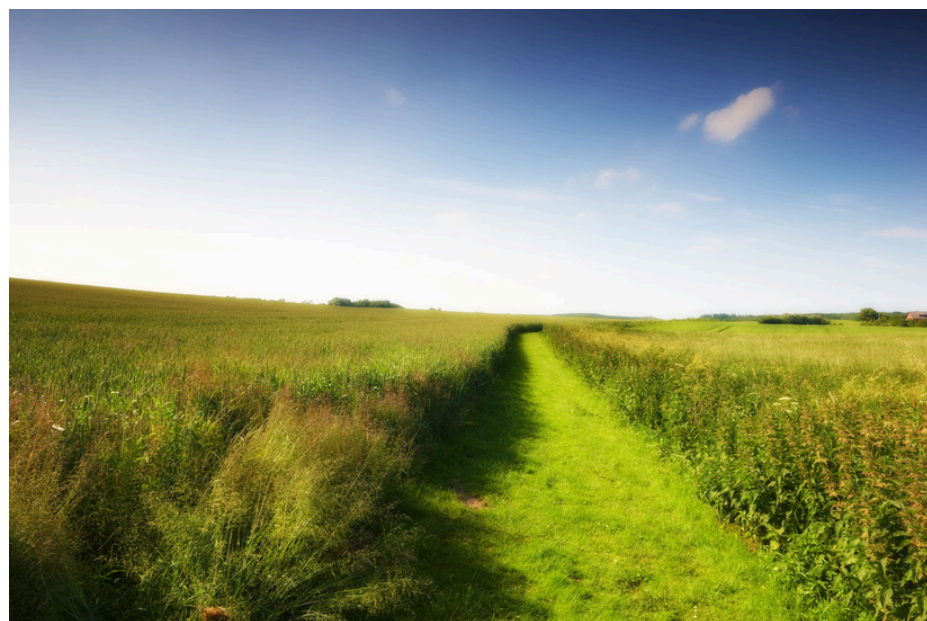
Ngăn bạn suy nghĩ:

Nếu Ai làm công việc ‘suy nghĩ’, học sinh sẽ không suy nghĩ nữa. Nó là một công cụ, không phải thay thế việc suy nghĩ.

Các đường dẫn trong não bộ sẽ không được hình thành.

Nó KHÔNG dùng để làm gì?

Các đường dẫn thần kinh và tư duy, hành vi của con người



Mọi suy nghĩ và hành vi của con người, về cơ bản, được tạo nên bởi sự hoạt động của các tế bào thần kinh trong bộ não. Chúng ta càng lặp đi lặp lại một suy nghĩ hoặc hành động, chúng ta càng củng cố sự kết nối giữa các khớp thần kinh và các tế bào thần kinh, một hoạt động được gọi là Tăng cường Tiềm năng Dài hạn. Giống như đi qua bãi cỏ cao: càng đi nhiều trên đường, con đường sẽ trở nên rõ ràng và việc di chuyển sẽ nhanh hơn. Điều này giải thích việc chúng ta có thể lái xe mà không cần nghĩ quá nhiều về thao tác đó.

Giáo viên sử dụng AI như thế nào?



Hệ thống đèn giao thông khi thiết kế các hoạt động cho học sinh

ĐỎ	Nghiêm cấm	“Đây là phần dành cho não bộ của bạn. Không sử dụng công cụ hỗ trợ.”	Đảm bảo các nhiệm vụ phải được hoàn thành trong lớp học hoặc “không thể làm bằng AI”, sử dụng các công cụ kiểm duyệt việc sử dụng AI trên Google
VÀNG	Đưa ra đánh giá về những gì bạn đã làm và hướng dẫn khi bắt đầu một nhiệm vụ.	“Sử dụng AI như la bàn, không phải tài xế của bạn. Để nó đưa cho bạn những gợi ý, nhưng đừng để nó nghĩ cho bạn. “Hãy để nó nhận xét về những gì bạn đã làm, nhưng đừng để nó trình bày những	Đặt ra giới hạn rõ ràng những mục nào trong nhiệm vụ của mình được phép sử dụng AI.

Bạn có thể làm gì?

Xem AI như một công cụ, không phải một chiếc máy tính.

Nó có thể giúp giải quyết các vấn đề về mặt kỹ thuật, nhưng bạn phải hiểu lý thuyết và tư duy phía sau công việc được hoàn thành

Chúng ta sẽ không phàn nàn việc một người sử dụng máy tính hoặc mỏ lết để làm việc. Công cụ rất quan trọng.



Bạn có thể làm gì?

Nói chuyện với con bạn.

Hãy hỏi các con chỉ cho bạn thấy cách chúng sử dụng AI

Nếu trẻ sử dụng AI để tìm ra kết quả, vậy thì đó là vấn đề.

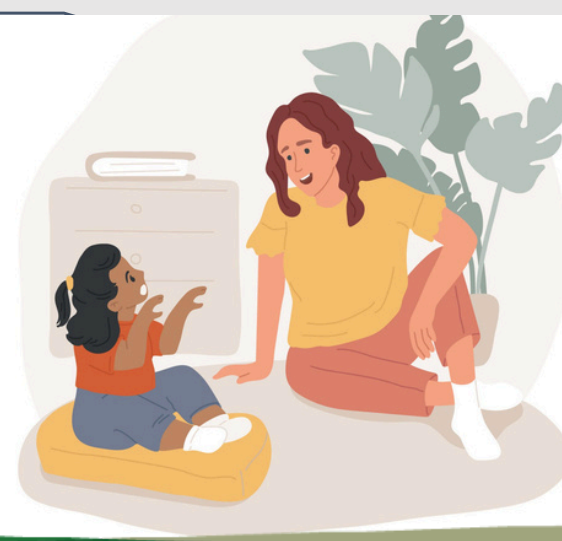
Nếu trẻ sử dụng AI để hiểu các bước thì đó là sự thành công

Ngôn ngữ chúng ta sử dụng rất quan trọng

“CON DỪNG AI ĐỂ LÀM BÀI NÀY HẢ?!” nghe giống như bạn buộc tội trẻ đang gian lận.

Hãy thử, **“Con có thể chỉ cho ba/mẹ cách con sử dụng AI để giúp con làm bài này chứ?”**

Cách hỏi này giúp chuyển hướng cuộc nói chuyện từ một cuộc tranh luận tiềm tàng sang quy trình mà trẻ sử dụng AI.



Bạn có thể làm gì?

Câu hỏi	Trả LỜI
AI là người làm bản cuối cùng	Học sinh
AI xác thực lại thông tin?	Học sinh
AI đưa ra luận điểm chính?	Học sinh
AI có giúp học sinh hiểu rõ hơn?	Có

SỬ DỤNG AI HIỆU QUẢ: GIA SƯ VS CÔNG CỤ

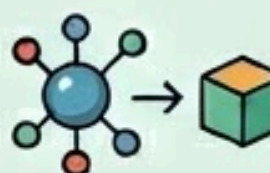


✓ AI NHƯ MỘT 'GIA SƯ'

“AI GIÚP TÔI HIỂU”



ĐỘT PHÁ Ý TƯỞNG

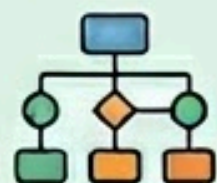


ĐƠN GIẢN HÓA KHÁI NIỆM

THỰC HÀNH & TRẮC NGHIỆM



GIẢI THÍCH CÁC BƯỚC



“Hãy dùng AI để học cách làm, không chỉ để lấy đáp án.”



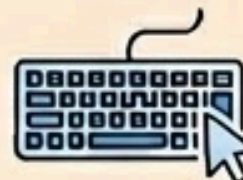
✗ AI NHƯ MỘT 'CÔNG CỤ' (THỤ ĐỘNG)



“AI LÀM THAY TÔI”



VIẾT LUẬN THUÊ



SAO CHÉP & DÁN ĐÁP ÁN



TIN TƯỞNG MÙ QUÁNG (DỮ LIỆU SAI LỆCH)



NÉ TRÁNH TƯ DUY PHẢN BIỆN

“Đó là nguồn tham khảo, không phải sự thay thế cho bộ não của bạn.”



HÃY NHỚ: BẠN CHÍNH LÀ NGƯỜI SUY NGHĨ! Hãy sử dụng AI có trách nhiệm.

Cẩm nang Phụ huynh - Biết đọc biết viết về AI

Mạng internet an toàn thế nào?

Bảo vệ chống lại thế giới trực tuyến so với thế giới thật

Trẻ em Việt Nam có nguy cơ bị người lạ quấy rối hoặc gạ gẫm trực tuyến cao hơn so với nguy cơ bị tấn công thể chất bởi người lạ ở nơi công cộng.

Trong khi các tội phạm tấn công trẻ em về thể chất vẫn ở mức thấp, trên 70% trẻ em Việt Nam đối mặt với các trải nghiệm trực tuyến không mong muốn (các tin nhắn quấy rối, các tin nhắn gợi dục, hoặc lộ thông tin cá nhân) và cứ 1 trong 5 trẻ (20%) là nạn nhân trực tiếp của bắt nạt trên mạng.

Trên các ứng dụng Zalo hoặc Facebook, trung bình trẻ không được giám sát trong 4 năm trước khi được hướng dẫn cách sử dụng ứng dụng một cách an toàn.

Trẻ em Việt Nam hiện đang có gấp 3 lần nguy cơ trở thành nạn nhân trực tuyến của lừa đảo hoặc tống tiền tình dục hơn là trở thành nạn nhân của các vụ trộm cướp hoặc cướp giật.

Bối cảnh không gian mạng đang thay đổi

Ai truy cập được các thiết bị điện tử/thông tin trực tuyến?

82% trẻ em ở độ tuổi 12-13 truy cập nền tảng trực tuyến mỗi ngày

93% trẻ em độ tuổi 14 - 15 truy cập các nền tảng trực tuyến mỗi ngày.

Độ tuổi trung bình sở hữu chiếc điện thoại đầu tiên là 9 tuổi

40% trẻ em độ tuổi 12-17 cảm thấy không an toàn và trên 70% trẻ có trải nghiệm không mong muốn khi sử dụng mạng ở Việt Nam.

Các vấn đề chính liên quan đến AI và an toàn mạng không gian mạng là gì?

Vấn nạn

Tin giả

**Deep Fakes (Công nghệ giả mạo), ảnh
khỏa thân và nội dung khiêu dâm trẻ
em do AI tạo ra (CSAM) (Nội dung
khiêu dâm trẻ em)**

**Tổng tiền bằng ảnh khiêu dâm
và AI**

Sự đồng hành của AI

AI và an toàn không gian mạng

Các mối đe dọa nào về an toàn khi liên quan đến AI và không gian mạng?

- Tin tức và nội dung giả
Các bản tin siêu thực do AI tạo ra giống như các tin tức thật (vd: VTV, BBC) nhưng lan truyền thông tin sai sự thật về sức khỏe, an toàn cộng đồng hoặc đóng cửa trường học.
- Việc sử dụng ứng dụng AI gỡ bỏ quần áo tạo ra các hình ảnh khiêu dâm không được sự đồng ý của người khác.
Khả năng AI là hình thức quấy rối trực tuyến phát triển nhanh nhất trong trường học. Trong năm 2025, phản ánh liên quan đến các nội dung rõ ràng được tạo ra từ AI gia tăng đáng kể trong các quốc gia ASEAN

Bối cảnh địa phương: Ở Malaysia và Việt Nam, dữ liệu năm 2025-2026 chỉ ra rằng cứ trong 3 thiếu niên, có 1 em là nạn nhân hoặc chứng kiến hành vi bắt nạt trực tuyến, với việc giả mạo bằng AI trở thành công cụ “trả thù” hoặc miệt thị địa vị xã hội trong trường học.

AI và An toàn trực tuyến

Các mối đe dọa đến an toàn khi liên quan đến AI và không gian mạng?

Tổng tiền bằng ảnh khiêu dâm và AI

- Tội phạm sử dụng AI để tạo ra các hình ảnh hoặc video giả mạo giống như thật của trẻ vị thành niên (kể cả khi trẻ không bao giờ gửi hình) để tổng tiền các em hoặc ép gửi thêm hình ảnh. Vào năm 2026, cứ 10 trường hợp tổng tiền bằng ảnh khiêu dâm ở trẻ vị thành niên thì có 1 trường hợp do hình ảnh giả mạo được tạo ra bằng trí tuệ nhân tạo thay vì từ ảnh thật. Các kẻ lừa đảo có thể tạo ra các hình ảnh giả mạo rất chân thật từ một hình ảnh đại diện từ Facebook hoặc Zalo.

Tháng Ba 2026, Việt Nam đã ban hành các đạo luật cụ thể để giải quyết các vấn nạn này, và thay vì chỉ là các vấn nạn trực tuyến, những vụ việc liên quan sẽ được xem xét như vấn đề hình sự và được hỗ trợ pháp lý của địa phương.

Bối cảnh khu vực: Ở Đông Nam Á, các trung tâm lừa đảo có tổ chức (thường có trụ sở đặt tại vùng biên giới các lãnh thổ trong khu vực) hiện sử dụng lớp phủ AI cho video trong thời gian thực. Kẻ lừa đảo có thể trông như một thiếu niên trong cuộc gọi trực tuyến để lừa dối tương.



Children in UK report online sextortion attempts in record numbers

Exclusive: Call for nudity-detection tech on phones as number of under-18s reporting blackmail attempts rises by 34%

● **'I felt ashamed and scared': how an online friendship became a sextortion nightmare**

which the individual harassed him with phone calls, demanding money or they would release the video.

AI và An toàn trực tuyến

Các mối đe dọa đến an toàn khi liên quan đến AI và không gian mạng?

Tình bạn AI

Ngày càng nhiều thanh thiếu niên chuyển sang những người bạn AI hoặc bạn đời ảo (như Character.ai hoặc Replika) để tương tác xã hội. Những công cụ này được thiết kế để trở thành những người bạn hoàn hảo – luôn có mặt, không bao giờ phán xét và luôn đồng thuận.

Điều này dẫn đến Rối loạn Gắn bó Kỹ thuật số, ở đó học sinh thích tương tác với AI hơn với bạn bè cùng trang lứa ngoài đời thực. Trong những trường hợp cực đoan, những công cụ này trong vài trường hợp khuyến khích việc tự gây hại hoặc đưa ra các lời khuyên nguy hiểm về sức khỏe tâm thần mà không có sự đồng cảm hoặc đào tạo chuyên nghiệp của con người.

Một tổ chức từ thiện dành cho trẻ em đã làm những đánh giá rủi ro và phát hiện rằng tất cả các nền tảng lớn đều thất bại trong việc báo cáo khi người dùng có những dấu hiệu của tư tưởng tự tử.

Những nghiên cứu ban đầu cũng chỉ ra rằng việc dựa dẫm quá mức vào tình bạn AI có thể dẫn đến việc gia tăng sự cô đơn vì giảm việc kết nối chân thực giữa người với người.

Chúng ta có thể làm gì?

Các biện pháp nhằm tăng cường an toàn cho học sinh khi sử dụng internet

8

8

8

Chúng ta có thể làm gì?

Các biện pháp nhằm tăng cường an toàn cho học sinh khi sử dụng internet

... bị. Sử dụng các công cụ cấp hệ thống như Apple Screen Time và Android Family Link, or hoặc cho các

Sự an toàn cảm xúc và các cuộc hội thoại c

không phán xét. Nếu trẻ tiết lộ rằng con đã bị 'hack' hoặc bị 'lộ' thông tin, hãy giữ bình tĩnh và không xấu hổ.

Những kẻ tấn công thường sử dụng sự xấu hổ để ép trẻ phải im lặng, phá vỡ sự xấu hổ đó sẽ làm mất đi lợi thế

Mật khẩu An toàn cho gia đình

Hãy thiết lập một "mật khẩu an toàn" riêng cho gia đình. Mật khẩu này hoạt động như một công cụ xác minh danh tính tức thì để chống lại các vụ lừa đảo bằng giọng nói hoặc các tin nhắn lừa đảo

Sự giúp đỡ từ bên ngoài và báo cáo sự việc

Hợp tác với trường, nhà trường luôn sẵn lòng giúp đỡ khi có bất kỳ lo ngại nào. Nếu con bạn là nạn nhân của một vụ lừa đảo tình dục AI và hình ảnh của trẻ đang ở trên mạng, báo cáo ngay lập tức đến takeitdown.ncmec.org.

Chúng ta có thể làm gì?

Đối phó với tin giả

~~Đặc biệt chú ý qua nhiều người hay và bạn bè như các nền tảng xã hội và kết hợp đọc theo các tin trên~~
các trang tin tức đáng tin cậy khác có đưa tin về cùng một câu chuyện hay không.

Tìm hiểu nguồn gốc và mục đích: kiểm tra mục “về chúng tôi” trên trang thông tin để xem tính minh bạch và xem tác giả có thành kiến gì hoặc mục đích ngầm nào không.

Tìm xem có các tín hiệu đáng báo động hoặc thành kiến rõ ràng: Cảnh giác trước những tiêu đề giật gân câu người xem, lỗi ngữ pháp, hoặc lỗi chính tả, vì chúng thường là những dấu hiệu của thông tin kém chất lượng hoặc tin tức giả.

Tham vấn chuyên gia hoặc dữ liệu: Sử dụng các trang kiểm chứng thông tin như Snopes và factcheck.org và tìm những mẫu khảo sát có tính đại diện lớn (trên 1,000+ phản hồi) trước khi tin vào số liệu thống kê.

Đặt câu hỏi Tại sao: Nhắc nhở trẻ cân nhắc động cơ khi một ai đó khi họ viết hoặc chia sẻ một nội dung – liệu rằng thông tin đó để thông báo hay chỉ là châm ngòi cho một phản ứng mang tính cảm xúc nào?

Chúng ta có thể làm gì?

Phát hiện Deep Fake

- 1. – Chuyển động của gương mặt và cơ thể** - Mặc dù công nghệ Deep Fake đang cải thiện nhanh chóng, công nghệ này vẫn chưa thể tạo ra biểu cảm khuôn mặt và chuyển động của cơ thể tự nhiên như người thật khi chú ý kỹ.
- 2. Phát hiện đồng bộ môi** – chuyển động môi không khớp với giọng nói có thể cho thấy đó là video deepfake, do việc chỉnh sửa âm thanh và đồng bộ hóa.
- 3. Thiếu nhất quán – hoặc thiếu – chớp mắt** – hiện tại, AI vẫn gặp khó khăn mô phỏng sự chớp mắt tự nhiên. Vì vậy, thuật toán của deepfake thường tạo ra các khuôn mẫu chớp mắt thiếu đồng nhất hoặc loại bỏ hoàn toàn việc chớp mắt.
- 4. Phản chiếu hoặc bóng đổ bất thường** – thuật toán của deepfake thường xuyên thất bại trong việc mô phỏng một cách chân thực bóng đổ và phản chiếu để phù hợp với ngữ cảnh của hình ảnh
- 5. Sự giãn nở của đồng tử** - AI thường không thay đổi được đường kính của đồng tử mắt đối tượng, do đó, thường dẫn đến đôi mắt tạo cảm giác bất thường.
- 6. Da và các đường nét khuôn mặt không phù hợp** – Các sản phẩm của ‘deep fakes’ thường có làn da giống nhau, thiếu sự đa dạng tự nhiên trong kết cấu và màu sắc đến từ các nếp nhăn, tàn nhang, nám, nốt ruồi, sẹo, và bóng đổ.
- 7. Bất thường về âm thanh** - giọng nói không tự nhiên, lặp đi lặp lại hoặc bị lỗi là điều đáng ngờ.

Chúng ta có thể làm gì?

Sử dụng các chức năng an toàn từ nhà cung cấp Dịch vụ Internet của bạn

FPT Telecom: F-Safe AnTam

Các chức năng bảo vệ bằng AI được xây dựng trong thiết bị phát sóng để tự động lọc và chặn các đường dẫn độc hại, các trang web giả mạo và các nỗ lực lừa đảo trước khi chúng tiếp cận được thiết bị của con bạn.

Viettel: Safenet

Cung cấp mạng lưới bộ lọc cấp quốc gia. Đây là dịch vụ được thiết kế và nâng cấp đặc biệt nhằm chặn các nội dung vi phạm chuẩn mực văn hóa tại Việt Nam cũng như vi phạm đạo luật bảo vệ trẻ em.

VNPT: GreenNet

Hoạt động như một dịch vụ làm sạch internet có thể tùy chỉnh. Hệ thống này trao quyền cho phụ huynh bằng việc cho phép phụ huynh tự động bật tắt và tự điều chỉnh các thông tin chi tiết cần được lọc hoặc các nội dung rõ ràng được tạo ra bởi AI.

Chúng ta có thể làm gì?

ĐẠO LUẬT VỀ TRÍ TUỆ NHÂN TẠO Nghị quyết số 203/2025/QH15

Mục đích: điều luật này quy định việc nghiên cứu, phát triển, việc cung cấp, triển khai, và sử dụng các hệ thống trí tuệ nhân tạo (từ bây giờ trở đi sẽ được ghi là các hoạt động sử dụng AI); các quyền và nghĩa vụ của các tổ chức và cá nhân liên quan

Điều 7: Đặc biệt nghiêm cấm việc lạm dụng AI để chuộc lợi từ các nhóm đối tượng dễ bị tổn thương, với nhóm trẻ em được đưa vào đầu danh sách. Việc sử dụng AI để “thao túng tư tưởng và hành vi” của trẻ là phạm pháp.

Điều luật này đánh vào các đối tượng sẵn mồi cố dụ dỗ trẻ em hoặc các ứng dụng cố xúi các thử thách nguy hiểm. Nếu một ứng dụng bị phát hiện tạo ảnh khỏa thân hoặc sử dụng thuật toán có thể dẫn đến các nguy hại về tinh thần cho giới trẻ, chính phủ Việt Nam từ giờ có quyền đình chỉ hoặc thu hồi ngay lập tức hệ thống AI đó khỏi thị trường Việt Nam.

Chúng ta có thể làm gì?

ĐẠO LUẬT VỀ TRÍ TUỆ NHÂN TẠO Nghị quyết số. 203/2025/QH15

Bất kỳ âm thanh, video, hoặc hình ảnh được tạo ra bởi AI mà giống một người thật hoặc sự kiện phải được dán nhãn rõ ràng là được tạo ra bởi AI.

Nếu con của bạn xem video của một người bạn hoặc một tin tức không được gắn nhãn, việc đó là hành vi vi phạm luật. Việc này giúp phụ huynh giáo dục trẻ dễ hơn: “Nếu không được gắn nhãn AI nhưng nhìn thấy quá kỳ lạ để có thể là sự thật thì đó là giả mạo bất hợp pháp.

Nếu con bạn là nạn nhân của deepfake hoặc lừa đảo bằng AI, bạn có thể báo cáo việc đó qua cổng trung tâm này (sẽ được công bố sau). Luật pháp yêu cầu các nhà phát triển và nhà cung cấp phải “khẩn trương triển khai các bản sửa chữa kỹ thuật” hoặc