

TÁC ĐỘNG CỦA TRÍ TUỆ NHÂN TẠO (AI) VÀ KIẾN NGHỊ GIẢI PHÁP ĐỐI VỚI VIỆT NAM

★ PGS, TS. LÊ VĂN CHIẾN

*Viện trưởng, Viện Lãnh đạo học và Hành chính công,
Học viện Chính trị quốc gia Hồ Chí Minh*

● **Tóm tắt:** Ngày nay trí tuệ nhân tạo (AI) đã len lỏi vào mọi góc ngách của cuộc sống và có tác động sâu rộng, toàn diện đến cuộc sống con người. Trí tuệ nhân tạo đã và đang mang đến rất nhiều cơ hội nhưng cũng tạo ra không ít thách thức mà con người phải tìm cách ứng phó, thậm chí một số ý kiến cảnh báo trí tuệ nhân tạo có thể hủy diệt loài người. Bài viết sẽ phân tích một số tác động cơ bản mang tính khái quát của trí tuệ nhân tạo đến đời sống con người, những cơ hội và thách thức mà nó mang lại.

● **Từ khóa:** Trí tuệ nhân tạo; AI; tác động của AI.

1. Quá trình hình thành và phát triển của trí tuệ nhân tạo

Trong những năm gần đây, nhất là từ khi trí tuệ nhân tạo tạo sinh (GenAI) xuất hiện và trở nên phổ biến với các công cụ như ChatGPT (của OpenAI) hay Gemini (của Google)... thì nhiều người mới biết đến trí tuệ nhân tạo. Tuy nhiên, thực tế là trí tuệ nhân tạo đã ra đời khá sớm từ những năm 50 của thế kỷ trước và có lịch sử phát triển tương đối lâu dài. Thuật ngữ “trí tuệ nhân tạo” xuất hiện lần đầu tiên là năm 1956 khi John McCarthy⁽¹⁾, nhà khoa học máy tính người Mỹ, đưa ra tại Hội nghị Dartmouth để mô tả ngành khoa học kỹ thuật tạo ra máy móc thông minh có thể bắt chước hành vi của con người. Kể từ đó, sự phát triển của trí tuệ nhân tạo đã trải qua nhiều thăng trầm trước khi phát triển rực rỡ như ngày nay. Nó đã từng nhận được sự kỳ vọng và đầu tư lớn của một số chính phủ, nhất là Mỹ và Nhật, nhưng cũng đã từng trải qua

ít nhất hai “mùa đông AI” vào những năm cuối thập niên 70 và cuối những năm 80, đầu những năm 90 của thế kỷ XX, khi chính phủ các nước trên giảm mạnh đầu tư vào lĩnh vực này do sự phát triển chậm chạp của nó. Bất chấp những “khoảng lặng” đó, AI vẫn âm thầm phát triển và một sự kiện có thể coi là một trang sử mới cho AI diễn ra vào năm 2016, khi robot hình người, Sophia do Hanson Robotics chế tạo đã trở thành “Công dân robot” đầu tiên trên thế giới⁽²⁾. Công dân này được Saudi Arabia công nhận là một robot hình người có khả năng nhận dạng khuôn mặt, giao tiếp bằng lời nói và biểu hiện cảm xúc trên khuôn mặt. Tuy nhiên, các thế hệ AI này được phát triển dựa trên các thuật toán logic hình thức, quy luật “Nếu - Thì” nên chủ yếu làm việc theo lệnh mà chưa có khả năng tạo ra tri thức mới, được gọi đó là AI phân tích. Phải đến năm 2022, với sự xuất hiện của trí tuệ nhân tạo tạo sinh, GenAI, một bước ngoặt lớn trong lịch

sử phát triển công nghệ AI thì trí tuệ nhân tạo mới có bước phát triển bùng nổ. Khác với các hệ thống AI thông thường vốn thiên về phân tích và nhận diện mẫu, GenAI nổi bật nhờ vào việc tạo ra dữ liệu hoàn toàn mới, không chỉ dừng lại ở việc hỗ trợ con người trong việc xử lý thông tin, mà còn trực tiếp sản xuất nội dung văn bản, hình ảnh, video, thậm chí cả âm thanh dựa trên các mẫu dữ liệu được huấn luyện trước.

2. Tác động của trí tuệ nhân tạo đến cuộc sống con người

(1) Một số tác động tích cực

Trí tuệ nhân tạo, nhất là trí tuệ nhân tạo tạo sinh đã mang lại cho con người rất nhiều lợi ích. Một số lợi ích cơ bản có thể kể đến như:

Một là, trí tuệ nhân tạo giúp tăng năng suất và chuyên đổi mô hình tăng trưởng kinh tế

Đến nay, nhiều nghiên cứu cho thấy, AI nâng cao năng suất lao động thông qua tự động hóa các công việc lặp đi lặp lại, tối ưu hóa quy trình sản xuất, quản lý chuỗi cung ứng và ra quyết định dựa trên dữ liệu lớn. Ứng dụng AI vào quy trình sản xuất không chỉ làm giảm giá thành sản phẩm mà còn thay đổi cấu trúc chuỗi giá trị, dịch chuyển nền kinh tế từ dựa nhiều vào tài nguyên thiên nhiên và lao động giá rẻ sang tăng trưởng dựa trên tri thức và năng suất cao (OECD, 2024)⁽³⁾.

Các nghiên cứu và báo cáo lớn đều ước tính AI có thể đóng góp đáng kể cho GDP toàn cầu nếu được ứng dụng hiệu quả. Năm 2017, theo một nghiên cứu của Công ty kiểm toán, PwC ước tính AI có thể đóng góp tới khoảng 14% tăng trưởng GDP toàn cầu vào năm 2030, tương đương khoảng 15,6 nghìn tỷ đô la (PwC, 2017)⁽⁴⁾. Trong lĩnh vực sản xuất, AI được ứng dụng mạnh vào tối ưu lịch trình sản xuất, tự động hóa kiểm tra chất lượng và dự báo nhu cầu trong chuỗi cung ứng nhằm nâng cao hiệu quả

sản xuất. Trí tuệ nhân tạo, nhất là trí tuệ nhân tạo tạo sinh có thể thúc đẩy tăng trưởng năng suất các yếu tố tổng hợp (TFP) khoảng 0,25-0,6 điểm phần trăm, tương đương mức tăng 0,4-0,9 điểm phần trăm năng suất lao động (OECD, 2024)⁽⁵⁾. AI cũng góp phần tái cấu trúc chuỗi giá trị của các ngành truyền thống như nông nghiệp (ví dụ sử dụng camera AI giám sát năng suất cây trồng, tưới tiêu theo nhu cầu). Đồng thời, nó cũng tạo ra các ngành mới như “kinh tế dữ liệu” và dịch vụ thông minh, thúc đẩy mô hình tăng trưởng dựa trên tài sản vô hình (cơ sở dữ liệu lớn được coi là nguồn lực đầu vào có giá trị cao). Điều này mở rộng không chỉ quy mô giá trị gia tăng mà cả tính bền vững của tăng trưởng.

Tại Việt Nam, các nghiên cứu cũng cho thấy, AI đã và đang góp phần vào tăng năng suất trong nhiều ngành như tài chính, logistics và thương mại điện tử nhờ vào khả năng xử lý dữ liệu nhanh, tối ưu hóa quy trình sản xuất. Khảo sát McKinsey năm 2019 cho thấy 63% doanh nghiệp được hỏi cho rằng AI giúp tăng doanh thu và 44% số doanh nghiệp được hỏi cho rằng AI giúp giảm chi phí vận hành (McKinsey & Company, 2019)⁽⁶⁾.

Như vậy, có thể thấy AI đã góp phần tăng năng suất lao động và chuyển đổi mô hình tăng trưởng kinh tế. Tất nhiên, lợi ích lớn mà nó tạo ra không tự động đến với tất cả các quốc gia. Chúng phụ thuộc vào năng lực hấp thụ công nghệ, đầu tư hạ tầng, kỹ năng lao động và chính sách chuyển đổi số của các quốc gia (World Bank, 2023)⁽⁷⁾.

Hai là, trí tuệ nhân tạo thúc đẩy đổi mới sáng tạo và rút ngắn chu trình nghiên cứu và phát triển (R&D)

Trí tuệ nhân tạo (AI) có tác động mạnh mẽ tới đổi mới sáng tạo, đặc biệt trong các ngành có cường độ R&D cao như: dược phẩm, hóa

học, vật liệu và công nghiệp chế tạo. Nhờ khả năng phân tích dữ liệu lớn, khám phá mẫu ẩn và mô phỏng, AI giúp rút ngắn đáng kể thời gian và chi phí nghiên cứu. Ví dụ, AI hỗ trợ phòng thí nghiệm tự động hóa, cho phép tiến hành thí nghiệm liên tục, giảm lỗi và chi phí. AI còn giúp tối ưu thử nghiệm lâm sàng thông qua phân tầng bệnh nhân và mô phỏng thiết kế. Ở lĩnh vực vật liệu, AI có thể khám phá vật liệu mới nhanh chóng và tiết kiệm hơn. Một nghiên cứu của Guy Doron và cộng sự (2025)⁽⁸⁾ cho thấy: sự phát triển nhanh chóng của Generative AI mở ra nhiều cơ hội mới trong cả giai đoạn khám phá và phát triển thuốc trong nghiên cứu và phát triển (R&D) dược phẩm. Sự hội tụ của các công nghệ tiên tiến cho phép thu thập tập dữ liệu phong phú và thể hệ AI kế tiếp sẽ cho phép tự động hóa nhanh chóng quá trình hình thành và kiểm định giả thuyết khoa học.

Tại Việt Nam, AI đã được ứng dụng trong nông nghiệp giúp dự báo sâu bệnh, tối ưu tưới tiêu và quản lý phân bón, giảm số vòng thử nghiệm thực địa (Long, L.H.N và cộng sự, 2024)⁽⁹⁾. Tại các trung tâm như VinAI, VinBrain hay FPT, AI đã đưa mô hình từ sản phẩm nghiên cứu đến sản phẩm thương mại nhanh hơn nhờ dữ liệu địa phương, nhân lực nghiên cứu, kỹ thuật tích hợp và hạ tầng đám mây (Phạm Bình Minh và cộng sự, 2023)⁽¹⁰⁾. Tuy nhiên, các nghiên cứu nhấn mạnh rằng lợi ích này chỉ bền vững khi có hạ tầng dữ liệu chuẩn hóa, nhân lực chất lượng cao và cơ chế hợp tác viện - trường - doanh nghiệp (Le & Do, 2024)⁽¹¹⁾.

Ba là, trí tuệ nhân tạo giúp cải thiện an sinh xã hội và dịch vụ xã hội

Trong an sinh xã hội, với cơ sở dữ liệu công dân dễ dàng được phân nhóm theo từng đối tượng có thể hưởng chế độ, chính sách của nhà nước, nhất là cơ sở dữ liệu các nhóm dễ bị tổn

thương trong xã hội và được cập nhật thường xuyên theo thời gian thực, AI được dùng để chi trả, trợ cấp trực tiếp tới đối tượng hưởng thụ, không qua trung gian, đồng thời dễ dàng phát hiện gian lận, tự động hóa phê duyệt và tối ưu phân bổ trợ cấp. Báo cáo của World Bank (2021)⁽¹²⁾ và Darrell M. West (2021) khẳng định ứng dụng AI giúp giảm chi phí hành chính và tăng công bằng trong quản lý phúc lợi⁽¹³⁾.

Trong y tế, AI hỗ trợ phân tích hình ảnh y khoa, sàng lọc sớm bệnh tật và theo dõi bệnh nhân từ xa, giúp giảm thời gian chẩn đoán và mở rộng năng lực chuyên môn tới các vùng xa. Ở Việt Nam, AI đã được ứng dụng trong chẩn đoán hình ảnh, dự báo điều trị và quản lý hồ sơ điện tử, góp phần rút ngắn chu trình lâm sàng và giảm chi phí (Huynh và cộng sự, 2023)⁽¹⁴⁾. Nghiên cứu quốc tế cũng cho thấy trợ giúp AI cải thiện độ chính xác của bác sĩ X-quang trong nhiều nhiệm vụ chẩn đoán (Pamela G. Anderson và cộng sự, 2024)⁽¹⁵⁾.

Trong giáo dục, AI cho phép cá nhân hóa lộ trình học, hỗ trợ giảng dạy và chấm điểm tự động. Các tổng quan hệ thống cho thấy nền tảng học tập cá nhân hóa mang lại kết quả học tập tích cực, nhất là với học sinh cần hỗ trợ đặc thù. Ở Việt Nam, thử nghiệm công nghệ số trong giáo dục đại học và y tế đã cho thấy cải thiện năng lực tự học và sự hài lòng của sinh viên⁽¹⁶⁾.

Bốn là, AI nâng cao hiệu quả quản trị nhà nước, minh bạch và xử lý khủng hoảng hiệu quả

Trong quản trị công hiện đại, trí tuệ nhân tạo (AI) đã trở thành công cụ chiến lược nhằm tăng hiệu quả điều hành, nâng cao minh bạch và cải thiện khả năng ứng phó khủng hoảng. Trong khu vực công, AI có thể được sử dụng một cách rất hiệu quả vào các lĩnh vực hoạch định chính sách dựa trên bằng chứng, tự động hóa thủ tục

hành chính và kiểm soát chi tiêu công, cùng với nâng cao năng lực cảnh báo sớm và quản trị khủng hoảng. Nhờ khả năng phân tích dữ liệu lớn, AI hỗ trợ hoạch định chính sách, giúp chính phủ dự báo nhu cầu dịch vụ công và các rủi ro chính sách có thể có. OECD (2019) nhấn mạnh rằng các hệ thống mô phỏng chính sách dựa trên AI, như ở Hàn Quốc, đã cải thiện đáng kể năng lực dự báo và chất lượng bằng chứng cho quyết định⁽¹⁷⁾. Không những vậy, AI có thể thúc đẩy tự động hóa hành chính công và giám sát chi tiêu. Ước tính khoảng 30% công việc hành chính có thể được tự động hóa bằng AI (McKinsey, 2018)⁽¹⁸⁾. Estonia là ví dụ điển hình về khả năng sử dụng hiệu quả nền tảng trợ lý ảo AI trong cung ứng dịch vụ công. Tại Việt Nam, một số địa phương đã thử nghiệm Chatbot hành chính công và hệ thống phân tích phản ánh của người dân, góp phần nâng cao mức độ hài lòng và giảm thời gian xử lý hồ sơ (Bảo Bình, 2025)⁽¹⁹⁾. Ở nhiều quốc gia đã sử dụng AI để theo dõi, dự báo và phản ứng nhanh, hiệu quả với những biến cố tự nhiên và nhân tạo mang tính khủng hoảng. Nhật Bản sử dụng AI để dự báo sóng thần và ngập lụt, giúp rút ngắn thời gian cảnh báo gần như ngay lập tức (Fumiyasu Makinoshima, 2021)⁽²⁰⁾. Trong đại dịch COVID-19, các mô hình AI tại Trung Quốc cho thấy khả năng dự đoán số ca nhiễm mới chính xác hơn so với mô hình dịch tễ truyền thống (Zixin Hu và cộng sự, 2020)⁽²¹⁾.

Năm là, trí tuệ nhân tạo hỗ trợ quản lý, bảo vệ môi trường và quản trị rủi ro khí hậu

Trí tuệ nhân tạo (AI) ngày càng được coi là công cụ trọng yếu trong quản lý môi trường và ứng phó biến đổi khí hậu, đóng góp trực tiếp vào các Mục tiêu phát triển bền vững (SDG 13 và SDG 15). AI giúp nâng cao năng lực dự báo thiên tai, quản lý tài nguyên, phát hiện vi phạm

môi trường và thúc đẩy nông nghiệp thông minh. Các mô hình Deep learning xử lý dữ liệu khí tượng và vệ tinh đã chứng minh hiệu quả trong dự báo bão, lũ và cháy rừng (Nasim Mohamadzadeh, 2024)⁽²²⁾. Ở Việt Nam, Viện Khoa học Khí tượng Thủy văn và Biến đổi khí hậu đã thử nghiệm hệ thống AI dự báo mưa lớn và ngập lụt đô thị, hỗ trợ chính quyền địa phương ứng phó kịp thời (Cục Biến đổi khí hậu, 2025)⁽²³⁾. Nhiều quốc gia đã ứng dụng AI để phát hiện phá rừng theo thời gian thực: nghiên cứu tại Việt Nam cho thấy AI nâng cao độ chính xác trong phân loại rừng ngập mặn ở Đồng bằng sông Cửu Long, góp phần xây dựng chính sách bảo tồn ven biển (Thuong V Tran và cộng sự, 2024)⁽²⁴⁾. Một số đô thị đã sử dụng AI để giám sát chất lượng không khí qua mạng lưới cảm biến, cung cấp dữ liệu thời gian thực cho hoạch định chính sách. Như vậy, AI tạo bước tiến quan trọng trong quản lý môi trường và rủi ro khí hậu, song để phát huy bền vững, Việt Nam cần đầu tư hạ tầng dữ liệu, tăng cường hợp tác quốc tế và xây dựng khung pháp lý bảo đảm minh bạch, công bằng trong ứng dụng công nghệ.

(2) Một số tác động tiêu cực của AI đối với con người

Một là, trí tuệ nhân tạo gây mất việc làm thậm chí đe dọa sự tồn tại của con người

Sự phát triển nhanh chóng của trí tuệ nhân tạo (AI) và tự động hóa đang đặt ra thách thức lớn đối với thị trường lao động. Khi nhiều công việc mang tính lặp lại hoặc kỹ năng thông thường được thay thế bởi máy móc và thuật toán, nguy cơ thất nghiệp cơ cấu, giảm thu nhập và bất ổn xã hội có thể gia tăng. Các nghiên cứu quốc tế cho thấy tỷ lệ việc làm dễ bị thay thế bởi AI là đáng kể. McKinsey (2017) ước tính khoảng 30% công việc toàn cầu có thể được tự động hóa vào năm 2030⁽²⁵⁾. Frey và Osborne (2017) cũng

dự báo 47% việc làm tại Mỹ có nguy cơ cao bị thay thế⁽²⁶⁾. Trong khi đó, OECD (2019) chỉ ra sự khác biệt giữa các ngành, với lao động phổ thông dễ bị ảnh hưởng nhất⁽²⁷⁾. Tại Việt Nam, nhóm lao động tay nghề thấp đứng trước rủi ro cao. Báo cáo của ILO (2016) dự báo 56% việc làm ở 5 nước ASEAN, trong đó có Việt Nam, có khả năng bị thay thế bởi tự động hóa trong 20 năm tới⁽²⁸⁾. Một nghiên cứu khác của World Bank (2024) chỉ ra rằng các ngành như dệt may, da giày và sản xuất điện tử, vốn sử dụng nhiều lao động giản đơn, chịu tác động lớn nhất⁽²⁹⁾.

Thậm chí, có một số ý kiến cho rằng, tương lai phát triển của trí tuệ nhân tạo, sự xuất hiện của người máy có khả năng tư duy hiệu quả hơn con người có thể hủy diệt sự sống của con người.

Hai là, trí tuệ nhân tạo có thể gây thiên kiến, bất công và tăng bất bình đẳng

Mặc dù AI hứa hẹn nâng cao hiệu quả và tính chính xác trong nhiều lĩnh vực, song một trong những lo ngại lớn nhất là khả năng hệ thống này tái sản xuất hoặc khuếch đại các thiên kiến xã hội vốn có. Nguyên nhân cốt lõi nằm ở dữ liệu huấn luyện và thiết kế thuật toán: nếu dữ liệu phản ánh bất bình đẳng lịch sử, AI sẽ “học” và nhân rộng các định kiến đó, nếu dữ liệu chứa thiên kiến lịch sử, các thuật toán có thể dẫn đến quyết định phân biệt đối xử trong tiếp cận việc làm, tín dụng, phúc lợi hoặc tư pháp. Thực tiễn quốc tế đã cho thấy các rủi ro này. Hệ thống COMPAS tại Mỹ, dùng để dự đoán tái phạm trong tư pháp hình sự, bị chỉ trích vì thiên vị đối với người da màu (Angwin và cộng sự, 2016)⁽³⁰⁾. Tại Hà Lan, hệ thống SyRI dùng AI để phát hiện gian lận phúc lợi đã bị Tòa án tuyên bố vi phạm quyền riêng tư và tạo ra sự bất công trong phân loại rủi ro (Frederik Zuiderveen Borgesius và Marvin van Bekkum, 2021)⁽³¹⁾. Nhiều

nghiên cứu cũng cảnh báo rằng AI trong tuyển dụng có thể vô tình loại trừ ứng viên nữ hoặc các nhóm thiểu số nếu dữ liệu huấn luyện phản ánh định kiến giới hoặc sắc tộc (Raghavan và cộng sự, 2020)⁽³²⁾. Ở Việt Nam, tuy ứng dụng AI trong phúc lợi và dịch vụ công còn mới, các chuyên gia đã lưu ý đến nguy cơ thiên kiến dữ liệu do thiếu hệ thống dữ liệu chuẩn hóa và minh bạch. Nếu không có cơ chế giám sát và giải thích thuật toán, các hệ thống AI có thể củng cố bất bình đẳng xã hội thay vì thu hẹp khoảng cách. Thiên kiến và bất công trong AI không chỉ là vấn đề kỹ thuật, mà còn là thách thức chính trị - xã hội.

Ba là, AI có thể dẫn tới tập trung quyền lực và phụ thuộc công nghệ

Cũng như tất cả các lĩnh vực khác, một cách rất tự nhiên, không phải ai cũng có cơ hội tiếp cận với AI giống nhau. Sự phát triển mạnh mẽ của trí tuệ nhân tạo (AI) đang tạo ra nguy cơ tập trung quyền lực vào một số ít tập đoàn công nghệ toàn cầu nắm giữ dữ liệu lớn, năng lực tính toán và thuật toán cốt lõi. Điều này có thể làm gia tăng bất cân xứng giữa các quốc gia và doanh nghiệp, đồng thời khiến nhiều nước phụ thuộc vào công nghệ nhập khẩu, làm suy giảm năng lực tự chủ số. Quốc gia nào nắm được công nghệ lõi và đi trước trong phát triển AI, quốc gia đó có cơ hội phát triển vượt trước và có cơ hội khai thác nguồn lực, nhất là nguồn lực số của các quốc gia khác. Các phần mềm ChatGPT, Gemini... là những ví dụ điển hình, khi cả thế giới sử dụng nó cũng có nghĩa là cả thế giới cung cấp dữ liệu, một yếu tố được coi là tài nguyên quan trọng nhất trong kỷ nguyên số, cho công ty sở hữu các phần mềm này. Thậm chí, một số ý kiến cho rằng đây chính là một kiểu “thực dân số”, khi các tập đoàn công nghệ ở một số quốc gia khai thác dữ liệu cá

nhân để kiểm soát hành vi và thị trường. Ronald Deibert và cộng sự (2019) cũng cảnh báo rằng AI có thể được sử dụng như công cụ quyền lực địa chính trị, làm gia tăng sự cạnh tranh và phụ thuộc giữa các quốc gia⁽³³⁾. Ở lĩnh vực kinh tế, OECD (2022) chỉ ra rằng đầu tư vào tài sản vô hình như dữ liệu số có thể làm tăng sự tập trung trong tay các công ty lớn. Các công ty này có thể tận dụng dữ liệu tốt hơn để tìm kiếm cạnh tranh và đổi mới ở các doanh nghiệp nhỏ⁽³⁴⁾.

Bốn là, AI có thể xâm phạm quyền riêng tư và gây các vấn đề về an ninh cho con người

Sự phát triển nhanh chóng của trí tuệ nhân tạo (AI) đi kèm với các rủi ro liên quan đến quyền riêng tư, giám sát xã hội và an ninh mạng. Do đặc thù dựa vào dữ liệu lớn, AI có thể xâm phạm đời tư cá nhân nếu việc thu thập và xử lý dữ liệu thiếu minh bạch hoặc vượt quá phạm vi cho phép. Các công nghệ nhận diện khuôn mặt và giám sát hành vi đã làm dấy lên lo ngại về “xã hội giám sát”, trong đó quyền tự do và riêng tư của công dân bị đe dọa (Joanna, 2019)⁽³⁵⁾. Ở cấp độ an ninh, AI có thể làm gia tăng các hình thức tấn công mạng tinh vi. Các hệ thống GenAI được sử dụng để tạo mã độc, email lừa đảo hoặc Deepfake, khiến việc phát hiện và ngăn chặn trở nên khó khăn hơn... Thậm chí, việc tích hợp AI vào hệ thống vũ khí tự động làm dấy lên lo ngại về kiểm soát chiến tranh, chu trình ra quyết định và rủi ro lỗi hệ thống. Liên hợp quốc đã nhiều lần thảo luận về việc cấm hoặc kiểm soát “vũ khí sát thương tự động”, bởi nguy cơ AI tham gia vào các quyết định sinh tử mà không có sự kiểm soát của con người (Alice Spazian, 2021)⁽³⁶⁾.

Có thể nói, AI vừa tạo ra công cụ bảo vệ an ninh, vừa tiềm ẩn nguy cơ trở thành mối đe dọa nếu thiếu cơ chế giám sát và kiểm soát. Do đó, cần có chính sách cân bằng giữa đổi mới công

nghệ và bảo vệ quyền riêng tư, an ninh mạng cũng như chuẩn mực quốc tế về sử dụng AI.

Tóm lại, trí tuệ nhân tạo mang đến cho loài người nhiều cơ hội có thể khai thác phục vụ con người như tạo ra nguồn lực mới, động lực tăng trưởng mới, thúc đẩy chuyển dịch cơ cấu kinh tế, tạo hệ sinh thái khởi nghiệp; nâng cao hiệu quả quản trị công; hỗ trợ mục tiêu phát triển bền vững. Tuy nhiên, nó cũng mang lại không ít thách thức cần quan tâm xử lý như khả năng bị bỏ lại phía sau do không làm chủ được công nghệ AI, thiếu hụt nguồn nhân lực chuyên sâu và kỹ năng số, hạ tầng số khó phát triển đòi hỏi phải đầu tư lớn cả nhân lực và tài chính. Sự thiếu hụt khung pháp lý hay khoảng trống về trách nhiệm và đạo đức, bảo mật...

3. Một số giải pháp

Một là, xây dựng và hoàn thiện khung pháp lý và tăng cường giáo dục đạo đức AI

AI đặt ra vấn đề mới về trách nhiệm pháp lý, minh bạch thuật toán, bảo vệ dữ liệu cá nhân và an toàn hệ thống. Việt Nam cần rà soát và ban hành hệ thống pháp luật nhằm tạo hành lang pháp lý cụ thể, tham chiếu kinh nghiệm quốc tế phù hợp với bối cảnh mới với sự phát triển của AI. Hành lang pháp lý đó cần nhấn mạnh các nguyên tắc AI phải dựa trên tính minh bạch, công bằng và trách nhiệm giải trình, phân loại và phòng ngừa rủi ro do AI mang lại để điều chỉnh hành vi của xã hội.

Bên cạnh xây dựng hệ thống pháp luật đủ mạnh để phòng ngừa và xử lý các hành vi vi phạm pháp luật thì việc tuyên truyền, giáo dục để mọi công dân khi sử dụng AI, tham gia hoạt động trên môi trường số phải bảo đảm chuẩn mực đạo đức, không trái với thuần phong, mỹ tục của dân tộc, không gây tổn hại đến cộng đồng và đất nước là việc làm cần duy trì thường xuyên.

Hai là, phát triển hạ tầng dữ liệu tập trung, chuẩn hóa và an toàn

Dữ liệu chất lượng là điều kiện tiên quyết để AI hoạt động hiệu quả. Tại Việt Nam, dữ liệu đang bị phân tán, thiếu chuẩn hóa và khó tiếp cận. Chính phủ cần xây dựng các kho dữ liệu công mở trong các lĩnh vực trọng yếu (y tế, giáo dục, môi trường, giao thông), đồng thời ban hành cơ chế quản lý dữ liệu bảo đảm quyền riêng tư. Kinh nghiệm từ một số nước đi trước như Phần Lan, Xin-ga-po cho thấy việc triển khai hệ thống dữ liệu mở quốc gia đã tạo nền tảng cho hệ sinh thái đổi mới số. Việc xây dựng cơ sở dữ liệu phải bảo đảm khả năng chia sẻ và tái sử dụng dữ liệu trong nghiên cứu và kinh doanh.

Ba là, đầu tư phát triển nguồn nhân lực, đào tạo lại, đào tạo kỹ năng

Tự động hóa do AI sẽ thay thế một số việc làm, đặc biệt là lao động giản đơn. Vì vậy, cần các chương trình đào tạo lại (Reskilling) và nâng cao kỹ năng (Upskilling), tập trung vào kỹ năng số, tư duy sáng tạo và quản lý là rất cần thiết. Báo cáo của ILO (2019) cảnh báo 56% việc làm tại Việt Nam và ASEAN có thể bị thay thế trong 20 năm tới, đòi hỏi chính sách đào tạo chủ động. Xin-ga-po là ví dụ điển hình với chương trình đào tạo lại và nâng cao kỹ năng, hỗ trợ tài chính cho người lao động học tập suốt đời, từ đó nâng cao khả năng thích ứng (Zhi Yong Lim và cộng sự, 2024)⁽³⁷⁾. Việt Nam có thể tham khảo để xây dựng cơ chế khuyến khích học tập liên tục, gắn với nhu cầu thị trường và chiến lược quốc gia về chuyển đổi số.

Bốn là, thúc đẩy hệ sinh thái đổi mới và doanh nghiệp AI trong nước

Để giảm phụ thuộc vào các trung tâm công nghệ số quốc tế, Việt Nam cần hỗ trợ khởi nghiệp và trung tâm nghiên cứu phát triển (R&D) trong lĩnh vực AI. Chính sách có thể bao gồm ưu đãi

thuế cho nghiên cứu, quỹ đầu tư công - tư và hỗ trợ hạ tầng tính toán hiệu năng cao. Thực tiễn từ Hàn Quốc và Trung Quốc cho thấy các trung tâm tính toán hiệu năng cao quốc gia giúp giảm chi phí cho doanh nghiệp nhỏ và đẩy nhanh quá trình thử nghiệm mô hình AI. Ở Việt Nam, các tổ chức như VinAI hay FPT AI đã chứng minh khả năng phát triển sản phẩm số có tính cạnh tranh quốc tế khi được hỗ trợ hạ tầng và hợp tác quốc tế. Điều này cho thấy chính sách khuyến khích hệ sinh thái đổi mới là yếu tố quyết định để AI đóng góp vào mô hình tăng trưởng dựa trên tri thức.

Năm là, củng cố hợp tác quốc tế và tham gia thiết lập chuẩn mực toàn cầu

AI là công nghệ xuyên biên giới, đòi hỏi hợp tác quốc tế để bảo đảm an ninh, công bằng và phát triển bền vững. Việt Nam cần tham gia sâu hơn vào các sáng kiến quốc tế như: UNESCO AI Ethics Recommendation (2021), OECD AI Principles (2019) và ASEAN Digital Masterplan 2025. Hợp tác trong chia sẻ dữ liệu cần thiết sẽ giúp Việt Nam vừa tận dụng được tri thức toàn cầu, vừa đóng góp tiếng nói vào các quy tắc quốc tế. Nghiên cứu của Ryan Calo (2017) chỉ ra rằng sự hợp tác đa phương là cách duy nhất để xử lý các rủi ro AI vượt khỏi phạm vi biên giới quốc gia⁽³⁸⁾.

Kết luận

AI mang lại cơ hội lớn cho đổi mới, tăng trưởng và quản trị công hiệu quả, nhưng đồng thời cũng đi kèm rủi ro xã hội, đạo đức và an ninh. Một chiến lược chính sách toàn diện cần hài hòa giữa thúc đẩy đổi mới và bảo đảm an sinh, quyền con người và chủ quyền công nghệ. Việt Nam, nếu triển khai đồng bộ các giải pháp về pháp lý, dữ liệu, nhân lực, đổi mới, đạo đức và hợp tác quốc tế, có thể không chỉ tận dụng lợi ích từ AI mà còn đóng góp vào quá trình định hình quản trị toàn cầu về công nghệ này ❖

- (1), (2) Văn Toàn (2023), “*Những cột mốc đánh dấu sự hình thành và phát triển của trí tuệ nhân tạo*”, trong: Báo Nhân dân điện tử, ngày 13/3/2023, tại: <https://nhandan.vn>, truy cập ngày 20/7/2025.
- (3) OECD (2024), “*AI Principles*”, tại: <https://www.oecd.org/>, truy cập ngày 20/7/2025.
- (4) Pwc (2017), “*AI to Drive GDP Gains of \$15.7 Trillion with Productivity, Personalisation Improvements*”, July 06, tại: <https://www.pwc.com>, truy cập ngày 20/7/2025.
- (5) OECD (2024), “*Miracle or Myth? Assessing the Macroeconomic Productivity Gains from Artificial Intelligence*”, OECD Artificial Intelligence Papers.
- (6) McKinsey & Company (2019), “*Global AI Survey: AI proves its worth, but few scale impact*”, November 22, tại: <https://www.mckinsey.com>, truy cập ngày 20/7/2025.
- (7) World Bank (2023), “*Digital Progress and the Productivity Paradox*”, World Bank Policy Report, tại: <https://openknowledge.worldbank.org>, truy cập ngày 20/7/2025.
- (8) Guy Doron và cộng sự (2025), “*Generative AI: Driving Productivity and Scientific Breakthroughs in Pharmaceutical R&D*”, Drug Discovery Today, Vol.30, Issue 1, January 2025, doi: 10.1016/j.drudis.2024.104272, tại: <https://pubmed.ncbi.nlm.nih.gov>, truy cập ngày 20/7/2025.
- (9) Long Le Hoang Nguyen và cộng sự (2024), “*Cooperative Performance and Lead Firm Support in Cleaner Production Adoption: SEM-fsQCA Analysis of Precision Agriculture Acceptance in Vietnam*”, Journal of Cleaner Production, October 2024.
- (10) Phạm Bình Minh và cộng sự (2023), “*Chinh phục thế giới bằng sản phẩm công nghệ AI trong y tế “made in Vietnam”*”, trong: Báo Vietnamnet, ngày 25/8/2023, tại: <https://vietnamnet.vn>, truy cập ngày 20/7/2025.
- (11) Le, V. T., & Do, M. H. (2024), “*AI Adoption and Challenges in Vietnam’s R&D Ecosystem*”, Journal of Science and Technology Policy, 12(2), pp.45-62.
- (12) World Bank (2021), “*World Development Report 2021: Data for Better Lives - Leveraging Greater Value from Data to help the Poor*”, tại: <https://blogs.worldbank.org>, truy cập ngày 20/7/2025.
- (13) Darrell M. West (2021), “*Using AI and Machine Learning to Reduce Government Fraud*”, tại: <https://www.brookings.edu>, truy cập ngày 20/7/2025.
- (14) Huynh, T. K., Nguyen, P. T., & Le, Q. H. (2023), “*Healthcare in Vietnam: Harnessing Artificial Intelligence and Robotics to Improve Patient Outcomes*”, Frontiers in Public Health, 11, 10565519, tại: <https://doi.org/10.3389/fpubh.2023.10565519>, truy cập ngày 20/7/2025.
- (15) Pamela G Anderson và cộng sự (2024), “*Deep Learning Improves Physician Accuracy in the Comprehensive Detection of Abnormalities on Chest X-rays*”, tại: <https://pubmed.ncbi.nlm.nih.gov>, truy cập ngày 20/7/2025.
- (16) Phương Thu Thuy, Van Hong Bui (2025), “*Assessing Students’ Self-Learning Competence in the Digital Learning Environment at Ho Chi Minh City University of Technology and Education*”, Journal of Technical Education Science, tại: <https://www.jte.edu.vn>, truy cập ngày 20/7/2025.
- (17), (27) OECD (2019), “*Artificial Intelligence in Society*”, 11 June 2019, tại: <https://www.oecd.org>, truy cập ngày 20/7/2025.
- (18) McKinsey & Company (2018), “*AI, Automation, and the Future of Work: ten things to solve for*”, June 1, 2018, tại: <https://www.mckinsey.com>, truy cập ngày 20/7/2025.
- (19) Bảo Bình (2025), “*Chatbot AI sẽ đóng vai trò như một “cán bộ số” trong dịch vụ công trực tuyến*”, trong: VnEconomy, ngày 20/02/2025, tại: <https://vneconomy.vn>, truy cập ngày 20/7/2025.
- (20) Fumiyasu Makinoshima (2021), “*Real-Time Tsunami Inundation Forecasting - Leveraging AI*”,

- trong: Nature Communications, June 1, 2021, tại: <https://www.fujitsu.com>, truy cập ngày 20/7/2025.
- (21) Zixin Hu và cộng sự (2020), “*Artificial Intelligence Forecasting of Covid-19 in China*”, tại: <https://arxiv.org>, truy cập ngày 20/7/2025.
- (22) Nasim MohamadiazarRolnick et al. (2024), “*Integrating Deep Learning, Satellite Image Processing, and Spatial-temporal Analysis for Urban Flood Prediction*”, trong: Journal of Hydrology, Vol. 639, tại: <https://www.sciencedirect.com>, truy cập ngày 20/7/2025.
- (23) Cục Biến đổi khí hậu (2025), “*Công nghệ AI - Giải pháp thích ứng mới trước biến đổi khí hậu*”, ngày 22/8/2025, tại: <https://dcc.mae.gov.vn>, truy cập ngày 22/8/2025.
- (24) Thuong V Tran và cộng sự (2024), “*Characterising the Distribution of Mangroves along the Southern Coast of Vietnam using Multi-spectral Indices and a Deep Learning Model*”, trong: Sci Total Environ, May 2024, doi: 10.1016/j.scitotenv.2024.171367, tại: <https://pubmed.ncbi.nlm.nih.gov>, truy cập ngày 20/7/2025.
- (25) Mc Kinsey & Company (2017), “*Jobs Lost, jobs Gained: Workforce Transitions in a Time of Automation*”, tại: <https://www.mckinsey.com>, truy cập ngày 20/7/2025.
- (26) Frey, Carl Benedikt & Osborne, Michael A (2017), “*The Future of Employment: How Susceptible are Jobs to Computerisation?*”, trong: Technological Forecasting and Social Change, Elsevier, vol. 114 (C), pp.254-280.
- (28) ILO (2016), “*ASEAN in Transformation: the Future of Jobs at Risk of Automation*”, tại: <https://www.ilo.org>, truy cập ngày 20/7/2025.
- (29) World Bank (2024), “*Báo cáo Việt Nam 2045, đột phá thể chế cho một tương lai thu nhập cao*”, tại: <https://documents1.worldbank.org>, truy cập ngày 20/7/2025.
- (30), Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016), “*Machine Bias: There’s Software used Across the Country to Predict Future Criminals and it’s Biased against Blacks*”, trong: ProPublica, tại: <https://www.propublica.org>, truy cập ngày 20/7/2025.
- (31) Frederik Zuiderveen Borgesius & Marvin van Bekkum (2021), “*Digital Welfare Fraud Detection and the Dutch SyRI Judgment*”, tại: <https://iapp.org>, truy cập ngày 20/7/2025.
- (32) Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020), “*Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices*”, trong: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp.469-481, <https://doi.org/10.1145/3351095.3372828>, tại: <https://www.researchgate.net>, truy cập ngày 20/7/2025.
- (33) Ronald Deibert et al. (2019), “*The Road to Digital Unfreedom*”, trong: Journal of Democracy, January 2019, Vol.30.
- (34) OECD (2022), “*Data Shaping Firms and Markets*”, tại: <https://www.oecd.org>, truy cập ngày 20/7/2025.
- (35) Joanna (2019), “*Shoshana Zuboff: Surveillance Capitalism is an Assault on Human Autonomy*”, trong: The Guardian Labs, tại: <https://www.theguardian.com>, truy cập ngày 20/7/2025.
- (36) Alice Spazian et al. (2021), “*UNIDIR on Lethal Autonomous Weapons Mapping our Research to the Discussions of the GGE on LAWS*”, tại: file:///C:/Users/ASUS/Downloads/UNIDIR-on-Lethal-Autonomous-Weapons-Final.pdf, truy cập ngày 20/7/2025.
- (37) Zhi Yong Lim và cộng sự (2024), “*Advancing Lifelong Learning in the Digital Age: A Narrative Review of Singapore’s SkillsFuture Programme*”, tại: <https://www.mdpi.com>, truy cập ngày 20/7/2025.
- (38) Ryan Calo (2017), “*Artificial Intelligence Policy: A Primer and Roadmap*”, 51 U.C.D. L. Rev. 399, tại: <https://digitalcommons.law.uw.edu>, truy cập ngày 20/7/2025.