

Phát triển, quản trị trí tuệ nhân tạo: Kinh nghiệm quốc tế và gợi mở đối với Việt Nam

Ngày nhận bài: 28/08/2026 | Ngày phản biện: 28/08/2026 | Ngày xuất bản: 22/09/2026

Tác giả: Hoàng Xuân Phương (Trường Đại học Văn Lang; NCS ngành Quan hệ Công chúng, Học viện Báo chí và Tuyên truyền)

Tóm tắt: Trí tuệ nhân tạo (AI) đang trở thành công nghệ nền tảng, định hình lại phương thức phát triển kinh tế - xã hội và năng lực quản trị của các quốc gia. Trên cơ sở phân tích bốn mô hình quản trị AI đại diện - quản trị dựa trên rủi ro của Liên minh châu Âu (EU), quản trị mềm linh hoạt của Singapore, quản trị thúc đẩy tối giản của Nhật Bản và quản trị theo ngành, dịch vụ của Trung Quốc, bài viết làm rõ cơ sở lựa chọn, tiêu chí phân loại, cũng như thành tựu và hạn chế đã được kiểm chứng của từng mô hình. Trên nền tảng đó, nghiên cứu phân tích thực trạng thể chế, chính sách và pháp luật về AI của Việt Nam giai đoạn 2020-2026, đặc biệt là dấu mốc ban hành Luật Trí tuệ nhân tạo năm 2025, đồng thời phân biệt định hướng chính trị của Đảng với chính sách, pháp luật của Nhà nước; từ đó đề xuất, gợi mở một số hàm ý chính sách đối với Việt Nam.

1. Đặt vấn đề

Sự phát triển nhanh chóng của trí tuệ nhân tạo (AI) đang tác động sâu rộng đến kinh tế - xã hội và năng lực quản trị của mỗi quốc gia, đặt ra yêu cầu mang tính chiến lược trong xây dựng khung thể chế và mô hình quản trị AI phù hợp. Theo ước tính của McKinsey Global Institute, AI tạo sinh "có tiềm năng tạo ra từ 2,6 đến 4,4 nghìn tỷ USD giá trị kinh tế hằng năm" trên phạm vi toàn cầu (Chui, Hazan, Roberts và cộng sự, 2023). Cùng với cơ hội đó, AI cũng đặt ra thách thức quản trị từ nguy cơ xâm phạm quyền riêng tư, sai lệch thuật toán, thao túng thông tin đến rủi ro đối với quốc phòng - an ninh.

Đối với mỗi quốc gia, bài toán đặt ra là dung hòa giữa thúc đẩy đổi mới AI với kiểm soát rủi ro xã hội và bảo vệ lợi ích quốc gia. Trên cơ sở khảo sát hệ thống văn bản pháp lý và chính sách AI hiện hành, nghiên cứu nhận thấy các quốc gia đã lựa chọn những hướng tiếp cận thể chế khác biệt, có thể phân loại theo bốn tiêu chí: (i) mức độ ràng buộc pháp lý (luật cứng hay hướng dẫn mềm); (ii) thời điểm can thiệp quản lý (tiền kiểm hay hậu kiểm); (iii) cấu trúc điều phối thể chế (tập trung theo một đạo luật khung hay phân tán theo ngành, dịch vụ); (iv) mục tiêu chính sách ưu tiên (kiểm soát rủi ro hay thúc đẩy đổi mới). Trên cơ sở bốn tiêu chí này, bài viết lựa chọn phân tích bốn trường hợp đại diện cho bốn hướng tiếp cận điển hình: EU AI Act (2024) đại diện cho quản trị dựa trên rủi ro, luật hóa toàn diện; khung quản trị AI mẫu của Singapore đại diện cho quản trị mềm, phi ràng buộc; đạo luật Thúc đẩy AI của Nhật Bản (2025) đại diện cho quản trị khuyến khích, hậu kiểm tối thiểu; hệ thống quy định theo ngành của Trung Quốc đại diện cho quản trị phân tán, tiếp cận theo dịch vụ cụ thể. Đây cũng là bốn nền kinh tế có ảnh hưởng đáng kể đến chuẩn mực quản trị AI khu vực và toàn cầu, với mức độ phát triển công nghệ và thể chế khác nhau, phù hợp để đối chiếu với điều kiện của Việt Nam.

Ở Việt Nam, từ năm 2020 đến nay, hệ thống chính sách và pháp luật về AI từng bước được hình thành, đặc biệt sau dấu mốc ban hành Luật Trí tuệ nhân tạo năm 2025. Tuy nhiên, quá trình triển khai trên thực tế vẫn bộc lộ không ít khoảng trống giữa cam kết chính sách và năng lực thực thi. Trong bối cảnh đó, nghiên cứu kinh nghiệm quốc tế về quản trị AI với trọng tâm là thể chế, chính sách và pháp luật hơn là mô hình phát triển công nghệ có ý nghĩa gợi mở cho Việt Nam trong tiếp tục hoàn thiện thể chế, nâng cao hiệu quả quản trị AI thời gian tới.

2. Phương pháp nghiên cứu

Để làm rõ cơ sở lý luận và thực tiễn cho việc hoàn thiện thể chế quản trị AI ở Việt Nam, nghiên cứu sử dụng kết hợp hai phương pháp.

Phương pháp nghiên cứu tài liệu được sử dụng để rà soát, phân tích hệ thống văn bản quy phạm pháp luật và văn kiện chiến lược của Việt Nam như: Nghị quyết 57-NQ/TW, Văn kiện Đại hội XIV của Đảng, Chiến lược AI theo Quyết định 127/QĐ-TTg, Luật Dữ liệu 2024, Luật Công nghiệp công nghệ số 2025, Luật Trí tuệ nhân tạo 2025; văn bản

pháp lý quốc tế như: EU AI Act 2024, Khung quản trị AI mẫu của Singapore, Đạo luật Thúc đẩy AI của Nhật Bản 2025, các quy định quản lý AI tạo sinh của Trung Quốc, cùng dữ liệu thứ cấp từ các báo cáo quốc tế và trong nước giai đoạn 2023-2026 như: McKinsey & Company, Oxford Insights, World AI Index, IMARC Group, Bộ Khoa học và Công nghệ, làm cơ sở hệ thống hóa nội dung và xác định mốc pháp lý phục vụ so sánh.

Phương pháp so sánh thể chế được sử dụng để đối chiếu bốn mô hình quản trị AI nêu tại phần Đặt vấn đề, theo bốn tiêu chí: mức độ ràng buộc pháp lý, thời điểm can thiệp quản lý, cấu trúc điều phối thể chế và mục tiêu chính sách ưu tiên. Tiêu chí đánh giá “hiệu quả” của từng mô hình được xác định dựa trên ba nhóm chỉ báo có thể kiểm chứng qua dữ liệu công bố: (i) mức độ tuân thủ, triển khai trên thực tế (số chủ thể tuân thủ, tốc độ đăng ký/cấp phép); (ii) mức độ lan tỏa chuẩn mực ra ngoài phạm vi quốc gia/khu vực ban hành; (iii) các chỉ số xếp hạng quốc tế độc lập (Oxford Insights, World AI Index). Trên cơ sở so sánh, nghiên cứu bóc tách những yếu tố có khả năng chuyển giao và những yếu tố phụ thuộc điều kiện thể chế đặc thù của từng quốc gia, làm căn cứ luận giải khả năng vận dụng vào thực tiễn Việt Nam.

3. Kết quả và thảo luận

3.1. Kết quả nghiên cứu

3.1.1. Kinh nghiệm quốc tế trong quản trị trí tuệ nhân tạo

Liên minh châu Âu - quản trị dựa trên rủi ro. Đạo luật AI của Liên minh châu Âu (EU Artificial Intelligence Act - Regulation (EU), 2024/1689), có hiệu lực từ ngày 1/8/2024, là khung pháp lý toàn diện đầu tiên về AI được ban hành ở cấp độ khu vực (European Parliament and Council of the European Union, 2024). Đạo luật áp dụng phương pháp phân loại rủi ro bốn cấp: Điều 5 cấm hoàn toàn các thực hành AI có rủi ro không thể chấp nhận (thao túng nhận thức - hành vi, nhận diện sinh trắc học từ xa theo thời gian thực tại không gian công cộng, trừ ngoại lệ cho cơ quan thực thi pháp luật); Điều 6 và Phụ lục III quy định các hệ thống AI rủi ro cao (y tế, tuyển dụng, tín dụng) phải tuân thủ nghiêm ngặt trước khi đưa ra thị trường; Điều 50 áp đặt nghĩa vụ minh bạch với AI rủi ro giới hạn; hệ thống rủi ro tối thiểu không bị điều chỉnh trực tiếp. Quy định về mô hình AI đa mục đích (GPAI) có hiệu lực từ tháng 8/2025.

Về tác động ngoài lãnh thổ, giới nghiên cứu chính sách còn nhiều tranh luận. Theo cách tiếp cận “Hiệu ứng Brussels”, quy định khắt khe của EU có thể lan tỏa ra ngoài phạm vi lãnh thổ do doanh nghiệp toàn cầu muốn tiếp cận thị trường EU phải tuân thủ, từ đó gián tiếp định hình chuẩn mực toàn cầu mà không cần cơ chế cưỡng chế quốc tế. Tuy nhiên, phân tích gần đây của Crum (2025) cho rằng do tính bất định nội tại của công nghệ AI và cách tiếp cận mang tính thủ tục, mở của Đạo luật, tác động thực tế của EU AI Act phù hợp với mô hình “quản trị thử nghiệm” hơn là một hiệu ứng Brussels áp đặt chuẩn mực đơn phương, trong đó các khu vực pháp lý độc lập cùng giám sát, học hỏi và điều chỉnh chính sách lẫn nhau, thay vì đơn thuần sao chép tiêu chuẩn của EU.

Dù vậy, mô hình này cũng bộc lộ hạn chế đã được chính EU thừa nhận, gánh nặng tuân thủ đối với doanh nghiệp vừa và nhỏ khiến Ủy ban châu Âu phải đề xuất điều chỉnh thông qua Digital Omnibus (tháng 11/2025) (European Commission, 2025); chu kỳ lập pháp kéo dài (từ đề xuất năm 2021 đến hiệu lực năm 2024) tạo nguy cơ quy định lạc hậu trước tốc độ phát triển công nghệ. Với các quốc gia đang phát triển, việc áp dụng trực tiếp mô hình tiền kiểm toàn diện như EU đòi hỏi năng lực thực thi hành chính và hạ tầng kỹ thuật số ở mức cao là điều kiện chưa phổ biến ở nhiều nước, trong đó có Việt Nam.

Singapore- quản trị mềm và thực dụng. Singapore ban hành Khung quản trị AI mẫu năm 2019, được xem là một trong những khung quản trị AI mềm đầu tiên trên thế giới, xây dựng trên hai nguyên tắc nền tảng: quyết định do AI hỗ trợ phải giải thích được, minh bạch và công bằng; hệ thống AI phải lấy con người làm trung tâm. Khung này được cập nhật thành phiên bản 2 vào tháng 1-2020 (Personal Data Protection Commission Singapore, 2020). Đến năm 2024, Singapore ban hành hướng dẫn quản lý rủi ro mô hình AI cho lĩnh vực tài chính và giới thiệu Project Moonshot - nền tảng mã nguồn mở phục vụ kiểm thử an toàn cho các mô hình ngôn ngữ lớn (Infocomm Media Development Authority, 2024).

Về hiệu quả thực tiễn, theo Chỉ số Sẵn sàng AI Chính phủ 2024 của Oxford Insights, Singapore xếp hạng 2/188 quốc gia toàn cầu với điểm số 84,25 - cao nhất khu vực Đông Á (Oxford Insights, 2024). Đây là minh chứng cho hiệu quả của cách tiếp cận quản trị mềm, gắn với cơ chế cập nhật hướng dẫn linh hoạt, không phụ thuộc chu kỳ

lập pháp.

Tuy nhiên, khả năng nhân rộng mô hình này ra các quốc gia khác còn hạn chế bởi điều kiện đặc thù của Singapore với quy mô dân số và diện tích nhỏ, bộ máy quản lý tập trung cao độ, hạ tầng dữ liệu tương đối đồng nhất, là những yếu tố khó tái lập ở các quốc gia có quy mô và mức độ phức tạp quản trị lớn hơn. Hiệu quả của quản trị mềm còn phụ thuộc nhiều vào năng lực tự giác tuân thủ và niềm tin thể chế giữa nhà nước - doanh nghiệp, điều kiện chưa hình thành đầy đủ ở nhiều quốc gia đang phát triển.

Nhật Bản - quản trị thúc đẩy tối giản. Năm 2025, Nhật Bản ban hành Đạo luật Thúc đẩy AI (Act on Promotion of Research and Development and Utilization of AI-related Technologies) - khung pháp lý can thiệp tối thiểu, khuyến khích doanh nghiệp tự nguyện phối hợp với biện pháp an toàn của Chính phủ, đồng thời trao quyền cho cơ quan nhà nước công bố danh tính tổ chức sử dụng AI vi phạm quyền con người, thay vì áp dụng chế tài xử phạt trực tiếp.

Ở cấp độ quốc tế, đóng góp nổi bật của Nhật Bản là khởi xướng Tiến trình AI Hiroshima (Hiroshima AI Process) trong khuôn khổ G7 từ tháng 5/2023, khung quốc tế đầu tiên kết hợp nguyên tắc hướng dẫn và bộ quy tắc ứng xử cho hệ thống AI tiên tiến, được các nhà lãnh đạo G7 chính thức thông qua trong tuyên bố tháng 10/2023. Điều này cho thấy một quốc gia không thuộc nhóm dẫn đầu công nghệ AI vẫn có thể định hình chuẩn mực quốc tế thông qua chiến lược ngoại giao công nghệ chủ động.

Hạn chế lớn nhất của Đạo luật Thúc đẩy AI là không thiết lập cơ chế xử phạt, khiến việc tuân thủ phụ thuộc hoàn toàn vào thiện chí của doanh nghiệp. Cách tiếp cận này phù hợp trong giai đoạn ứng dụng AI còn tập trung ở các lĩnh vực rủi ro thấp như sản xuất, song tiềm ẩn nguy cơ thiếu bảo đảm thực chất khi AI mở rộng sang các ứng dụng có mức độ rủi ro cao hơn.

Trung Quốc - quản trị theo ngành và dịch vụ. Trung Quốc tiếp cận quản trị AI theo phương thức ban hành quy định riêng cho từng loại hình dịch vụ, kết hợp hệ thống tiêu chuẩn kỹ thuật quốc gia. Các văn bản nổi bật gồm: Quy định về Khuyến nghị thuật toán (2022), Quy định về Quản lý dịch vụ tổng hợp nội dung chuyên sâu trên Internet (2022), Biện pháp tạm thời về dịch vụ AI tạo sinh (2023) và Quy tắc dán nhãn nội dung do AI tạo ra (2025). Đây là cách tiếp cận theo chiều dọc, được cập nhật liên tục

theo từng loại hình công nghệ, cho phép phản ứng chính sách nhanh mà không cần chờ hoàn thiện một đạo luật tổng thể (Sheehan, M, 2023).

Về hiệu quả triển khai, theo Cơ quan Quản lý không gian mạng Trung Quốc (CAC), tính đến ngày 31/3/2025, đã có 346 dịch vụ AI tạo sinh hoàn tất đăng ký hợp pháp trên cả nước, tăng đáng kể so với thời điểm Biện pháp tạm thời có hiệu lực năm 2023 - minh chứng định lượng cho khả năng hấp thụ chính sách nhanh của thị trường (CGTN, 2025).

Hạn chế của mô hình này nằm ở nguy cơ phân tán pháp lý, thiếu nhất quán khi số lượng quy định ngành gia tăng theo thời gian. Các yêu cầu về nội địa hóa dữ liệu và kiểm duyệt nội dung, dù phù hợp mục tiêu quản lý thông tin trong nước, cũng tạo rào cản cho doanh nghiệp nước ngoài và hạn chế hợp tác nghiên cứu quốc tế, không phải quốc gia nào cũng phù hợp áp dụng.

3.1.2. Chính sách AI của Việt Nam giai đoạn 2020 đến nay

Về định hướng chính trị của Đảng. Nền tảng cho phát triển AI ở Việt Nam được xác lập từ Đại hội XIII (2021), trong chủ trương đẩy mạnh ứng dụng thành tựu Cách mạng công nghiệp lần thứ tư gắn với chuyển đổi số quốc gia. Nghị quyết số 57-NQ/TW, ngày 22/12/2024, của Bộ Chính trị, "Về đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia" nhấn mạnh vai trò nền tảng của thể chế, nhân lực, hạ tầng, dữ liệu và công nghệ chiến lược, trong đó thể chế phải đi trước một bước. Định hướng này tiếp tục được Đại hội XIV củng cố. Trong các định hướng phát triển đất nước giai đoạn 2026-2030, Đại hội XIV xác lập mô hình tăng trưởng mới "lấy khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số làm động lực chính" (Đảng Cộng sản Việt Nam, 2026), đồng thời xác định "đột phá phát triển khoa học, công nghệ, đổi mới sáng tạo và chuyển đổi số quốc gia" là một trong các nhiệm vụ then chốt của nhiệm kỳ, yêu cầu phát triển trí tuệ nhân tạo và các ngành công nghệ mới (Đảng Cộng sản Việt Nam, 2026). Đây là văn kiện chính trị định hướng, làm căn cứ để Nhà nước thể chế hóa thành chiến lược và pháp luật cụ thể.

Về chiến lược và pháp luật. Trên cơ sở định hướng chính trị nêu trên, hệ thống chiến lược và pháp luật về AI từng bước được thể chế hóa. Quyết định số 127/QĐ-TTg, ngày 26/1/2021, của Thủ tướng Chính phủ ban hành Chiến lược quốc gia về nghiên

cứu, phát triển và ứng dụng AI đến năm 2030, xác định AI là công nghệ nền tảng của Cách mạng công nghiệp lần thứ tư, đặt hai mốc mục tiêu, đến năm 2025, Việt Nam thuộc nhóm 5 nước dẫn đầu ASEAN và nhóm 60 nước dẫn đầu thế giới về nghiên cứu AI; đến năm 2030, trở thành trung tâm đổi mới sáng tạo AI trong khu vực (Thủ tướng Chính phủ, 2021). Trước đó, Quyết định số 749/QĐ-TTg, ngày 3/6/2020, phê duyệt Chương trình Chuyển đổi số quốc gia, xác định mục tiêu phát triển ba trụ cột chính phủ số - kinh tế số - xã hội số.

Giai đoạn 2024-2025 chứng kiến sự ra đời của bộ ba luật tạo khung pháp lý tương hỗ cho AI. Luật Dữ liệu (Luật số 60/2024/QH15), hiệu lực từ 1/7/2025, thiết lập khuôn khổ pháp lý cho quản lý, khai thác, chia sẻ dữ liệu số quốc gia - nền tảng hạ tầng thiết yếu cho AI. Luật Công nghiệp công nghệ số (Luật số 71/2025/QH15), hiệu lực từ 1/1/2026, quy định Nhà nước đẩy mạnh phát triển, ứng dụng AI và có chính sách ưu đãi thúc đẩy nghiên cứu, triển khai AI.

Cột mốc quan trọng nhất là Luật Trí tuệ nhân tạo (Luật số 134/2025/QH15), Quốc hội thông qua ngày 10/12/2025, có hiệu lực từ 1/3/2026, gồm 8 chương, 35 điều. Điều 4 của Luật quy định bốn nguyên tắc cốt lõi: lấy con người làm trung tâm; AI phục vụ con người, không thay thế thẩm quyền con người; bảo đảm công bằng, minh bạch, không thiên lệch, không gây hại; thúc đẩy AI xanh, giảm thiểu tác động tiêu cực đến môi trường. Khác với khung phân loại bốn cấp của EU, Điều 9 của Luật phân loại hệ thống AI theo ba mức độ rủi ro (cao, trung bình, thấp) làm căn cứ áp dụng nghĩa vụ tuân thủ tương ứng, đây là lựa chọn thể chế đơn giản hóa, phù hợp năng lực quản lý hiện tại của Việt Nam. Đồng thời, Luật AI 2025 áp dụng cơ chế hậu kiểm thay vì tiền kiểm, doanh nghiệp không phải xin cấp phép trước khi triển khai AI, nhưng phải tự đánh giá mức độ an toàn và chịu trách nhiệm khi cơ quan quản lý thực hiện hậu kiểm.

Trên nền tảng thể chế đó, hệ sinh thái AI và kinh tế số bước đầu ghi nhận kết quả tích cực, dù chưa thể quy hoàn toàn cho tác động của khung pháp luật AI vừa mới có hiệu lực. Kinh tế số Việt Nam đạt khoảng 18,3% GDP, đưa Việt Nam vào nhóm các nền kinh tế số tăng trưởng nhanh tại Đông Nam Á - một con số cần được nhìn nhận thận trọng, vì mức độ đóng góp cụ thể của khung pháp luật AI (vừa mới có hiệu lực từ năm 2026) vào kết quả này còn cần thêm thời gian và dữ liệu để kiểm chứng (VietnamNet, 2025). Thị trường AI nội địa đạt 865,1 triệu USD năm 2024, dự báo tăng lên 2,99 tỷ

USD vào năm 2034, tốc độ tăng trưởng kép hằng năm 14,37% giai đoạn 2025-2034 (IMARC Group, 2026). Về năng lực, Việt Nam xếp hạng 73/188 quốc gia trong Chỉ số Sẵn sàng AI Chính phủ 2024 của Oxford Insights, thuộc nhóm 5 nước dẫn đầu ASEAN (Oxford Insights, 2024).

Tuy nhiên, thực tiễn triển khai còn một số khoảng trống cần khắc phục, khoảng cách giữa mục tiêu Chiến lược AI 2030 với chỉ số đo lường cụ thể và cơ chế giám sát độc lập; nguồn nhân lực AI chất lượng cao còn thiếu, phân bố không đồng đều giữa các cơ sở đào tạo; năng lực điều phối chính sách AI còn phân tán giữa nhiều bộ, ngành do chưa có đầu mối điều phối liên ngành đủ thẩm quyền.

3.2. Thảo luận

Từ việc so sánh bốn mô hình quản trị AI quốc tế với thực trạng thể chế Việt Nam sau cột mốc Luật Trí tuệ nhân tạo 2025, nghiên cứu tập trung trả lời câu hỏi: Việt Nam nên học gì, không nên học gì từ mỗi mô hình, và yếu tố nào có thể chuyển giao, yếu tố nào phụ thuộc điều kiện thể chế đặc thù.

Từ mô hình EU, Việt Nam có thể tiếp thu nguyên tắc phân loại rủi ro làm căn cứ áp dụng nghĩa vụ tuân thủ tương ứng - nguyên tắc mà Luật AI 2025 đã vận dụng dưới dạng đơn giản hóa còn ba cấp độ. Tuy nhiên, áp dụng nguyên trạng cơ chế tiền kiểm nghiêm ngặt và gánh nặng tuân thủ hành chính như EU AI Act là chưa phù hợp với năng lực thực thi và quy mô doanh nghiệp công nghệ hiện tại của Việt Nam, phần lớn là doanh nghiệp vừa và nhỏ; việc lựa chọn cơ chế hậu kiểm, do đó, là lựa chọn thể chế phù hợp hơn với điều kiện trong nước.

Từ mô hình Singapore, yếu tố có thể chuyển giao là cơ chế cập nhật hướng dẫn kỹ thuật linh hoạt, không phụ thuộc chu kỳ lập pháp, cùng công cụ kiểm thử an toàn mã nguồn mở tương tự Project Moonshot. Ngược lại, mức độ ràng buộc pháp lý thấp của Singapore khó áp dụng nguyên trạng cho Việt Nam do chênh lệch về quy mô quản trị, mức độ đồng nhất hạ tầng dữ liệu và niềm tin thể chế giữa nhà nước - doanh nghiệp.

Từ mô hình Nhật Bản, bài học đáng chú ý là khả năng định hình chuẩn mực quốc tế thông qua ngoại giao công nghệ chủ động dù không thuộc nhóm cường quốc AI hàng đầu. Đây là kinh nghiệm phù hợp để Việt Nam tham khảo khi tham gia các diễn đàn

quản trị AI khu vực và toàn cầu. Tuy nhiên, cơ chế tự nguyện thiếu chế tài của Nhật Bản là yếu tố cần thận trọng, bởi năng lực tự giác tuân thủ của doanh nghiệp công nghệ số Việt Nam vẫn đang trong giai đoạn hình thành.

Từ mô hình Trung Quốc, bài học tham khảo là cách tiếp cận theo chiều dọc, cập nhật quy định theo từng loại hình dịch vụ để phản ứng linh hoạt với thay đổi công nghệ. Tuy nhiên, nguy cơ phân tán, chồng chéo quy định giữa các ngành là bài học cảnh báo trực tiếp cho Việt Nam, trong bối cảnh trách nhiệm quản lý AI hiện đang phân bổ cho nhiều bộ, ngành mà chưa có cơ chế điều phối liên ngành đủ thẩm quyền.

Nhìn chung, kinh nghiệm quốc tế đặt ra bài toán đánh đổi giữa kiểm soát rủi ro và thúc đẩy đổi mới trong lập pháp AI. Quản trị quá chặt chẽ theo hướng tiền kiểm toàn diện như EU tạo gánh nặng tuân thủ, kìm hãm doanh nghiệp khởi nghiệp; ngược lại, quản trị mềm, tự nguyện như Nhật Bản thiếu chế tài xử lý khi AI xâm phạm nghiêm trọng quyền riêng tư và an ninh thông tin. Việc Việt Nam lựa chọn cơ chế hậu kiểm là bước đi cân bằng, phù hợp xu thế linh hoạt được nhiều nước áp dụng, nhưng đặt ra câu hỏi cần tiếp tục nghiên cứu: năng lực kỹ thuật của bộ máy quản lý nhà nước hiện nay có đủ sức thực thi hậu kiểm hiệu quả hay không, khi cơ chế này chỉ phát huy tác dụng nếu đi kèm hệ thống tiêu chuẩn kỹ thuật minh bạch và công cụ giám sát tự động. Trên cơ sở đó, nghiên cứu đề xuất một số hàm ý chính sách đối với Việt Nam.

Một là, khẩn trương ban hành nghị định, thông tư hướng dẫn thi hành Luật AI 2025, cụ thể hóa danh mục hệ thống AI theo ba cấp độ rủi ro, trách nhiệm giải trình của chủ thể phát triển và sử dụng AI trong y tế, giáo dục, truyền thông, hành chính công, bảo đảm tính tương thích với Luật Dữ liệu và Luật An ninh mạng.

Hai là, sớm thiết lập cơ chế điều phối chiến lược AI cấp quốc gia, có thể dưới hình thức Hội đồng quốc gia về AI trực thuộc Chính phủ đảm nhận chức năng điều phối liên ngành, theo dõi tác động chính sách dựa trên dữ liệu, đồng thời là đầu mối tiếp nhận, xử lý hồ sơ hậu kiểm theo Luật AI 2025; đây là giải pháp trực tiếp cho tình trạng phân tán trách nhiệm quản lý hiện nay.

Ba là, đẩy nhanh chính sách dữ liệu mở trong khu vực công, đầu tư hạ tầng tính toán và hoàn thiện cơ chế bảo vệ dữ liệu cá nhân theo Chiến lược dữ liệu của Bộ Khoa học và Công nghệ đến năm 2030, gắn với yêu cầu bảo đảm chủ quyền số.

Bốn là, chuyển từ đào tạo đơn lẻ sang chiến lược nhân lực AI toàn diện, tích hợp AI và đạo đức công nghệ vào chương trình đào tạo, thu hẹp khoảng cách giới và khoảng cách vùng, miền trong tiếp cận giáo dục AI.

Năm là, thúc đẩy ứng dụng AI có trọng tâm trong hành chính công, y tế, giáo dục và nông nghiệp, ưu tiên phát triển mô hình ngôn ngữ tiếng Việt “Make in Vietnam” vừa vì giá trị kinh tế, vừa vì lý do chủ quyền dữ liệu.

Sáu là, chủ động tham gia các diễn đàn quốc tế về quản trị AI để vừa học hỏi kinh nghiệm, vừa đóng góp xây dựng chuẩn mực AI toàn cầu phù hợp lợi ích các quốc gia đang phát triển, theo kinh nghiệm ngoại giao công nghệ của Nhật Bản.

Nghiên cứu cũng cần được nhìn nhận trong giới hạn phạm vi, đây là nghiên cứu tiếp cận ở góc độ phân tích chính sách công và pháp lý vĩ mô, dựa trên dữ liệu thứ cấp và văn bản pháp luật. Tác động thực tế của Luật AI 2025 đối với chi phí tuân thủ của doanh nghiệp công nghệ vừa và nhỏ tại Việt Nam vẫn cần thời gian kiểm chứng khi luật có hiệu lực thi hành đầy đủ. Các nghiên cứu tiếp theo cần bổ sung điều tra thực chứng đối với cộng đồng doanh nghiệp và người dùng, cũng như đánh giá tác động của AI đến thị trường lao động Việt Nam giai đoạn 2026-2030.

4. Kết luận

Việc ban hành Luật Trí tuệ nhân tạo năm 2025 đánh dấu bước tiến quan trọng, đưa Việt Nam vào nhóm các quốc gia sớm xây dựng khung pháp lý chuyên biệt cho AI. So sánh bốn mô hình quản trị AI quốc tế cho thấy không có mô hình nào có thể áp dụng nguyên trạng cho Việt Nam: EU cung cấp bài học về phân loại rủi ro nhưng đặt gánh nặng tuân thủ cao; Singapore cho thấy giá trị của quản trị linh hoạt nhưng phụ thuộc điều kiện thể chế đặc thù; Nhật Bản khẳng định khả năng định hình chuẩn mực quốc tế của quốc gia không thuộc nhóm dẫn đầu công nghệ, song thiếu chế tài ràng buộc; Trung Quốc minh chứng tốc độ phản ứng chính sách theo ngành nhưng tiềm ẩn phân tán pháp lý. Việc Việt Nam lựa chọn cơ chế hậu kiểm là hướng đi cân bằng, phù hợp xu thế linh hoạt được nhiều nước áp dụng, song hiệu quả thực thi phụ thuộc trực tiếp vào năng lực kỹ thuật của bộ máy quản lý nhà nước và mức độ hoàn thiện của hệ thống tiêu chuẩn kỹ thuật đi kèm.

Để những cam kết chính sách được chuyển hóa thành năng lực quản trị thực chất, Việt Nam cần sớm hoàn thiện văn bản hướng dẫn thi hành Luật AI 2025, thiết lập cơ chế điều phối liên ngành đủ thẩm quyền, đầu tư đồng bộ cho hạ tầng dữ liệu và nguồn nhân lực, đồng thời chủ động tham gia định hình chuẩn mực quản trị AI khu vực và toàn cầu. Đây là những điều kiện nền tảng để AI thực sự trở thành động lực cho tăng trưởng bền vững và nâng cao năng lực quản trị quốc gia trong giai đoạn tới./.

Tài liệu tham khảo

- CGTN. (2025). "346 generative AI services registered with China's cyber authority", ngày 9/4/2025, <https://news.cgtn.com/news/2025/04/09/346-generative-AI-services-registered-with-China-s-cyber-authority-1CqpnxsMee4/p.html>
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Yee, L., Zimmel, R., Sukharevsky, A. (2023). "The economic potential of generative AI: The next productivity frontier". McKinsey & Company (McKinsey Digital), ngày 14/6/2023, <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>
- Crum, B. (2025). "Brussels effect or experimentalism? The EU AI Act and global standard-setting". Internet Policy Review, 14(3), tr. 1-21, DOI: 10.14763/2025.3.2032, <https://policyreview.info/articles/analysis/brussels-effect-or-experimentalism>
- Đảng Cộng sản Việt Nam. (2026). Văn kiện Đại hội đại biểu toàn quốc lần thứ XIV, Nxb Chính trị quốc gia Sự thật, Hà Nội, t. I.
- European Commission. (2025). "Digital Omnibus Regulation Proposal", ngày 19/11/2025, <https://digital-strategy.ec.europa.eu/en/library/digital-omnibus-regulation-proposal>
- European Parliament and Council of the European Union. (2024). "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence ... (Artificial Intelligence Act). Official Journal of the European Union, L, 2024/1689, 1–144. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
- IMARC Group. (2026). "Vietnam artificial intelligence market size, share, trends and forecast 2026 – 2034". <https://www.imarcgroup.com/vietnam-artificial-intelligence-market>
- Infocomm Media Development Authority (IMDA). (2024). "Singapore launches Project Moonshot – a generative artificial intelligence testing toolkit to address LLM safety and security challenges". <https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2024/sg-launches-project-moonshot>

Monetary Authority of Singapore. (2024). "Artificial intelligence model risk management: Observations from a thematic review". Information Paper. <https://www.mas.gov.sg/publications/monographs-or-information-paper/2024/artificial-intelligence-model-risk-management>

Oxford Insights. (2024). "Government AI Readiness Index 2024". <https://oxfordinsights.com/wp-content/uploads/2024/12/2024-Government-AI-Readiness-Index-2.pdf>

Personal Data Protection Commission Singapore (PDPC). (2020). "Model Artificial Intelligence Governance Framework (2nd ed.)". <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf>

Sheehan, M. (2023). "China's AI regulations and how they get made", Carnegie Endowment for International Peace. <https://carnegieendowment.org/research/2023/07/chinas-ai-regulations-and-how-they-get-made>

Thủ tướng Chính phủ. (2021). Quyết định số 127/QĐ-TTg, ngày 26/1/2021, ban hành Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng trí tuệ nhân tạo đến năm 2030.

VietnamNet. (2025). "Vietnam's digital economy expands at record pace, now 18.3% of GDP". <https://vietnamnet.vn/en/vietnam-s-digital-economy-expands-at-record-pace-now-18-3-of-gdp-2369146.html>