

Quản trị rủi ro trong phát triển trí tuệ nhân tạo, kinh nghiệm quốc tế và bài học cho Việt Nam

Hoàng Ngọc Quang¹

Ngày nhận bài: 07/5/2026 | Ngày gửi phản biện: 11/5/2026 | Ngày duyệt đăng: 28/5/2026

Tóm tắt: Bài viết phân tích kinh nghiệm quản trị rủi ro trong phát triển trí tuệ nhân tạo (AI) của ba quốc gia: Hoa Kỳ, Vương quốc Anh và Trung Quốc, qua đó rút ra các bài học định hướng cho Việt Nam. Hoa Kỳ áp dụng khung quản trị rủi ro đa tầng, chú trọng an toàn kỹ thuật, công bằng, minh bạch và quyền kiểm soát của con người. Vương quốc Anh triển khai mô hình linh hoạt, thích ứng theo mức độ rủi ro và gắn với cơ chế giải trình, bảo vệ quyền lợi xã hội. Trung Quốc xây dựng hệ thống quản trị tập trung, coi rủi ro AI là vấn đề mang tính hệ thống liên quan đến an ninh, đạo đức và ổn định thể chế. Từ những kinh nghiệm đó, Việt Nam cần hướng tới mô hình quản trị AI chủ động, toàn diện và cân bằng giữa đổi mới công nghệ với kiểm soát rủi ro, bảo đảm sự phát triển bền vững và an toàn cho xã hội.

Từ khóa: Quản trị rủi ro, trí tuệ nhân tạo, chính sách.

Risk Governance in Artificial Intelligence Development:

International Experience and Lessons for Vietnam

Abstract: This article analyzes the risk management experiences in artificial intelligence (AI) development of three countries: the United States, the United Kingdom, and China, thereby drawing orienting lessons for Vietnam. The United States applies a multi-layered risk management framework, emphasizing technical safety, fairness, transparency, and human control. The United Kingdom implements a flexible model that adapts to risk levels and is tied to accountability mechanisms and the protection of social interests. China builds a centralized governance system, viewing AI risks as systemic issues related to security, ethics, and institutional stability. From these experiences, Vietnam needs to aim for a proactive, comprehensive AI governance model that balances technological innovation with risk control, ensuring sustainable development and safety for society.

Keywords: Risk management, artificial intelligence, policy.

1. Đặt vấn đề

Trong thập kỷ qua, sự phát triển đột phá của công nghệ trí tuệ nhân tạo (AI), đặc biệt là trí tuệ nhân tạo tạo sinh (Generative AI), đã đưa thế giới bước vào một kỷ nguyên chuyển đổi số sâu rộng. Không chỉ dừng lại ở một công cụ tối ưu hóa, AI đang định hình lại cấu trúc kinh tế toàn cầu. Theo báo cáo của Viện Toàn cầu McKinsey (2023), AI tạo sinh có khả năng đóng góp thêm từ 2,6 đến 4,4 nghìn tỷ USD hàng năm cho nền kinh tế thế giới. Tuy nhiên, sự phát triển theo cấp số nhân của AI cũng mang theo nhiều vấn đề rủi ro đối với xã hội loài người. Trong Báo cáo Rủi ro Toàn cầu năm 2024, Diễn đàn Kinh tế Thế giới (WEF) đã xếp hạng "Thông tin sai lệch và xuyên tạc do AI tạo ra" là rủi ro toàn cầu lớn nhất trong ngắn hạn. Đặc biệt là các mối đe dọa đến an ninh quốc gia (như tội phạm mạng, lừa đảo deepfake) đang gây ra nhiều tác động tiêu cực tới các thiết chế xã hội truyền thống.

¹ TS., Trường Đại học Kinh tế, Đại học Quốc gia Hà Nội; Email: Quanghn1903@vnu.edu.vn

Trước bối cảnh đó, các siêu cường công nghệ đã bắt đầu thiết lập các chốt chặn pháp lý nhằm tối thiểu hóa rủi ro trong khi vẫn duy trì không gian cho đổi mới sáng tạo. Dù chung một mục tiêu, nhưng do sự khác biệt về thể chế chính trị và ưu tiên kinh tế, các quốc gia đang hình thành những trường phái quản trị rủi ro hoàn toàn khác biệt: Từ cách tiếp cận bảo vệ quyền dân sự của Hoa Kỳ, mô hình quản lý linh hoạt, tỷ lệ thuận với rủi ro của Vương quốc Anh, cho đến khung pháp lý can thiệp sâu, đặt nặng trật tự xã hội của Trung Quốc.

Tại Việt Nam, AI đang được định vị là một trong những công nghệ cốt lõi thúc đẩy chuyển đổi số quốc gia. Theo Quyết định số 127/QĐ-TTg (2021) của Thủ tướng Chính phủ về Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng AI đến năm 2030, Việt Nam đặt mục tiêu đưa AI trở thành lĩnh vực công nghệ quan trọng, đưa Việt Nam vào nhóm bốn nước dẫn đầu ASEAN. Trên thực tế, việc ứng dụng AI tại Việt Nam đang diễn ra mạnh mẽ. Vì thế, Quốc hội đã ban hành Luật số 134/2025/QH15 ngày 10/12/2025 về trí tuệ nhân tạo, trong đó đưa ra các định hướng quản trị rủi ro cho công nghệ này. Tuy nhiên, hệ thống pháp luật hiện hành vẫn chưa đưa ra khung pháp lý chuyên biệt nào để phân loại, đánh giá và kiểm soát các rủi ro do AI gây ra; dẫn đến tình trạng lúng túng trong việc xử lý các hệ lụy như lừa đảo bằng hình ảnh giả mạo hay vi phạm quyền sở hữu trí tuệ thời gian qua.

Xuất phát từ thực tiễn trên, bài viết này đi sâu phân tích phương án quản trị rủi ro trong phát triển AI tại ba quốc gia tiêu biểu, bao gồm Hoa Kỳ, Vương quốc Anh và Trung Quốc. Từ đó, nghiên cứu rút ra những bài học kinh nghiệm và đề xuất hàm ý chính sách nhằm giúp Việt Nam thiết kế một khung pháp lý về AI hài hòa: Vừa thiết lập được các ranh giới để bảo vệ an ninh quốc gia và người dân, vừa đảm bảo sự cởi mở để thúc đẩy hệ sinh thái đổi mới sáng tạo trong nước.

2. Kinh nghiệm quốc tế về quản trị rủi ro trong phát triển AI

2.1. Hoa Kỳ

Là một quốc gia đi đầu trong phát triển trí tuệ nhân tạo, Hoa Kỳ nhận thức rõ rằng, tốc độ ứng dụng nhanh chóng của công nghệ này kéo theo nhiều rủi ro tiềm ẩn cần được quản trị thận trọng. Theo văn bản “Blueprint for an AI Bill of Rights” do Văn phòng Chính sách Khoa học và Công nghệ Nhà Trắng ban hành tháng 10/2022, sẽ có năm nhóm rủi ro chính có thể phát sinh gồm: (i) *Thứ nhất*, rủi ro về an toàn và hiệu quả hệ thống bắt nguồn từ khả năng các mô hình AI gặp lỗi kỹ thuật hoặc thể hiện hành vi không mong muốn nếu thiếu khâu kiểm thử và giám sát đầy đủ, đặc biệt nguy hiểm trong các ứng dụng liên quan trực tiếp đến tính mạng và quyền lợi của con người như y tế, giao thông hay tư pháp; (ii) *Thứ hai*, rủi ro bất bình đẳng thuật toán thể hiện ở chỗ các mô hình học máy có thể khuếch đại định kiến xã hội sẵn có, gây phân biệt đối xử với các nhóm yếu thế liên quan đến chủng tộc, giới tính, tuổi tác hoặc khuyết tật; (iii) *Thứ ba*, rủi ro xâm phạm quyền riêng tư và mất kiểm soát dữ liệu cá nhân phát sinh từ việc thu thập, xử lý dữ liệu quy mô lớn, tạo khoảng cách thông tin giữa các tổ chức phát triển công nghệ và người dùng; điều này đặc biệt đáng lo trong những lĩnh vực nhạy cảm như y tế, giáo dục, tài chính hay lao động; (iv) *Thứ tư*, rủi ro thiếu minh bạch và hiện tượng “hộp đen” cũng được lưu ý, vì khi người dân không hiểu rõ cơ chế vận hành và tiêu chí ra quyết định của hệ thống tự động, niềm tin xã hội cùng khả năng giám sát sẽ suy giảm; (v) *Thứ năm*, Hoa Kỳ nhận diện rủi ro mất quyền tự quyết của con người, khi việc giao quá nhiều quyền cho hệ thống tự động có thể khiến con người bị loại khỏi quá trình tham gia, can thiệp hoặc phản hồi trong những quyết định hệ trọng.

Để quản lý và kiểm soát những rủi ro nêu trên, Hoa Kỳ thiết lập một khung chính sách gồm năm nguyên tắc cốt lõi bao gồm: *Nguyên tắc thứ nhất* về quản trị rủi ro và đảm bảo chất lượng hệ thống gồm kiểm soát hai giai đoạn: trước triển khai (ex-ante) – yêu cầu thử nghiệm, thẩm định độc lập và tham vấn

đa bên nhằm phát hiện rủi ro từ sớm; sau triển khai (ex-post) – duy trì giám sát, đánh giá kết quả và cho phép áp dụng quyền ngừng sử dụng với hệ thống không bảo đảm an toàn. *Nguyên tắc thứ hai* là ngăn ngừa bất bình đẳng thuật toán, trong đó, Hoa Kỳ làm rõ định nghĩa pháp lý về hành vi phân biệt đối xử do AI gây ra, tiến hành đánh giá công bằng của thuật toán dựa trên dữ liệu đại diện, thử nghiệm chênh lệch ở cả giai đoạn phát triển và vận hành, đồng thời công khai báo cáo tác động để tăng cường trách nhiệm giải trình trước xã hội. *Nguyên tắc thứ ba* liên quan đến bảo vệ quyền riêng tư dữ liệu, sử dụng cách tiếp cận tối thiểu hóa dữ liệu – chỉ thu thập thông tin thật sự cần thiết, thiết lập cơ chế đồng ý minh bạch để người dùng có quyền lựa chọn cách dữ liệu được sử dụng, và phân loại rủi ro dữ liệu theo từng lĩnh vực, áp dụng mức độ bảo vệ nghiêm ngặt hơn ở các ngành có ảnh hưởng lớn đến đời sống. *Nguyên tắc thứ tư* là tăng cường minh bạch và quyền được biết, yêu cầu tài liệu hóa và truyền đạt thông tin một cách dễ hiểu để công chúng nhận biết khi và cách thức AI được sử dụng, cung cấp giải thích có ý nghĩa về cơ chế hoạt động, đặc biệt với các quyết định có rủi ro cao, đồng thời liên tục cập nhật thông tin khi hệ thống có thay đổi. *Nguyên tắc thứ năm*, nhấn mạnh đảm bảo vai trò và quyền can thiệp của con người, quy định quyền từ chối nhằm cho phép người dân lựa chọn quy trình do con người thực hiện thay vì hệ thống tự động, thiết lập cơ chế hỗ trợ và khiếu nại đối với các sai sót hoặc thiệt hại do AI gây ra, và duy trì yêu cầu giám sát bắt buộc của con người trong các lĩnh vực trọng yếu như tư pháp, y tế và giáo dục.

Nhìn chung, khung chính sách này thể hiện cách tiếp cận “quản trị rủi ro đa tầng”, năm nguyên tắc hoạt động như các lớp bảo vệ liên kết: từ khâu thiết kế kỹ thuật an toàn, đảm bảo công bằng, bảo vệ dữ liệu cá nhân, tăng cường minh bạch đến duy trì quyền kiểm soát của con người.

2.2. Vương Quốc Anh

Theo Bộ Khoa học, Đổi mới và Công nghệ Vương quốc Anh (2023), chính sách quản trị rủi ro nhấn mạnh yêu cầu duy trì cơ chế giám sát xuyên suốt vòng đời hệ thống AI, trong đó các chủ thể phải thực hiện thẩm định và thử nghiệm định kỳ để bảo đảm ba yếu tố cốt lõi: tính năng, khả năng phục hồi và an ninh mạng. Vương quốc Anh đã tận dụng những tiêu chuẩn hiện có – ví dụ như các nguyên tắc bảo mật học máy của Trung tâm An ninh mạng Quốc gia (NCSC) – để làm cơ sở tham chiếu cho việc đo lường và đánh giá. Cách tiếp cận này thể hiện nhận thức rõ ràng về rủi ro tiềm ẩn trong AI khi xem xét yêu cầu giải trình và minh bạch phải tương ứng với mức độ rủi ro.

Vương quốc Anh đánh giá rủi ro AI không chỉ ở khía cạnh kỹ thuật, mà còn ở khía cạnh xã hội và kinh tế. Rủi ro được nhìn nhận là sự bất đối xứng thông tin giữa nhà phát triển và người sử dụng, nguy cơ bóp méo thị trường hoặc xâm phạm quyền lợi cá nhân. Từ đó, quốc gia này luôn gắn việc nhận diện rủi ro với sự can thiệp ở mức độ phù hợp – minh bạch hóa, giải thích thuật toán, bảo vệ quyền lợi và khả năng khiếu nại – đảm bảo cân bằng giữa phát triển công nghệ và an toàn hệ thống.

Để quản trị và giảm thiểu rủi ro, Anh xây dựng bốn nhóm nguyên tắc xử lý chính: (i) Giảm bất đối xứng thông tin: chia thành hai tầng “minh bạch” và “khả năng giải thích”, giúp công chúng hiểu được cách thức hoạt động và logic vận hành của AI; (ii) Bảo vệ quyền lợi và duy trì thị trường công bằng: tích hợp yêu cầu kiểm soát AI vào các luật hiện hành như Đạo luật Bình đẳng, Đạo luật Nhân quyền và luật bảo vệ dữ liệu, nhằm ngăn chặn các thuật toán gây phân biệt đối xử hoặc tổn hại lợi ích hợp pháp; (iii) Thiết lập cơ chế ủy thác và giám sát: yêu cầu xác định rõ trách nhiệm của từng tổ chức trong chuỗi nghiên cứu, phát triển, vận hành; đồng thời khuyến khích áp dụng chuẩn kỹ thuật và công cụ đánh giá tác động để phát hiện rủi ro sớm; (iv) Cơ chế phục hồi và xử lý tranh chấp: cho phép cá nhân hoặc bên thứ ba khiếu nại các quyết định do AI gây tổn hại, đảm bảo lộ trình giải quyết minh bạch và dễ tiếp cận.

Cách triển khai của Vương quốc Anh theo hướng linh hoạt: khuyến khích thị trường tự điều chỉnh, nhưng vẫn chuẩn bị phương án an toàn pháp lý để chuyển sang quy định bắt buộc nếu hệ thống quản trị

mềm không đủ hiệu quả. Đây là mô hình quản trị AI chủ động – vừa thúc đẩy sáng tạo, vừa duy trì khả năng kiểm soát rủi ro và bảo đảm sự phát triển bền vững, công bằng cho xã hội.

2.3. Trung Quốc

Theo Cục Quản lý Không gian mạng Trung Quốc (2023), trong văn bản “Các biện pháp tạm thời về quản lý dịch vụ trí tuệ nhân tạo sinh”, các giá trị cốt lõi và tiêu chuẩn nội dung được xác định và quy định một cách cụ thể, chặt chẽ. Trong đó, rủi ro không được coi đơn thuần là vấn đề kỹ thuật mà được định nghĩa như một yếu tố mang tính hệ thống, liên quan trực tiếp đến an ninh quốc gia, định hướng tư tưởng và ổn định xã hội. Chính phủ nhìn nhận AI là công nghệ có sức ảnh hưởng xuyên lĩnh vực, có thể tạo ra ngoại ứng mạnh mẽ trong chính trị, truyền thông và kinh tế, do đó mọi rủi ro đều phải được kiểm soát ngay từ khâu thiết kế và triển khai. Rủi ro lớn nhất được xác định nằm ở ba tầng: thể chế, xã hội và cá nhân. Ở tầng thể chế, AI có nguy cơ xâm hại nền tảng ý thức hệ, lan truyền nội dung chống đối, hoặc làm suy yếu khối thống nhất quốc gia; đây là lý do khiến yếu tố an ninh tư tưởng và trật tự xã hội trở thành yếu tố ảnh hưởng lớn trong mọi quy định. Ở tầng xã hội, Trung Quốc coi truyền thông AI là một công cụ hai mặt – vừa giúp lan tỏa giá trị cộng đồng, vừa tiềm ẩn khả năng tạo ra thông tin sai lệch, kích động thù hận hoặc đe dọa đạo đức tập thể; vì thế, việc kiểm soát nội dung được xem như hình thức giảm thiểu rủi ro đạo đức. Còn ở tầng cá nhân, sự phát triển của AI tạo sinh làm gia tăng bất đối xứng thông tin và nguy cơ xâm phạm quyền nhân thân, khiến người dân có thể bị tổn hại danh dự, hình ảnh và quyền riêng tư. Ngoài các khía cạnh trên, Trung Quốc cũng nhận diện rủi ro kinh tế – thị trường xuất phát từ độc quyền thuật toán, lạm dụng dữ liệu và cạnh tranh không lành mạnh, có thể làm biến dạng cấu trúc thị trường và ảnh hưởng đến công bằng xã hội. Từ những nhận thức này, quốc gia này xác định AI là lĩnh vực chịu rủi ro tổng hợp, không chỉ cần giám sát kỹ thuật mà còn phải được gắn với yêu cầu bảo đảm an ninh chính trị, văn hóa và đạo đức cộng đồng, xây dựng một trật tự quản trị mang tính phòng ngừa sâu rộng.

Để kiểm soát các dạng rủi ro nói trên, Trung Quốc xây dựng một khung chính sách mang tính bao trùm với phương thức thực thi tập trung và cưỡng chế cao gồm năm trụ cột điều chỉnh: (i) Ranh giới về thể chế – quy định mọi hệ thống AI phải tuân thủ định hướng chính trị quốc gia và bảo vệ nền tảng xã hội chủ nghĩa; mọi hành vi tạo ra sản phẩm gây tổn hại cho an ninh, chủ quyền hay hình ảnh quốc gia đều bị cấm và xử lý nghiêm. (ii) AI không chỉ bị buộc loại bỏ nội dung tiêu cực mà còn được định hướng tạo ra thông tin tích cực, lành mạnh, qua đó Nhà nước chủ động xây dựng môi trường mạng có trật tự đạo đức cao. (iii) Nguyên tắc công bằng và không phân biệt đối xử – cơ chế can thiệp được đặt từ giai đoạn thiết kế thuật toán, yêu cầu các nhà phát triển phải loại bỏ thiên lệch dữ liệu và đảm bảo quyền bình đẳng cho mọi nhóm dân cư, nhằm ngăn ngừa bất công cấu trúc. (iv) Quản trị thị trường và đạo đức kinh doanh – tập trung vào việc hợp pháp hóa nguồn dữ liệu, bảo vệ quyền sở hữu trí tuệ, chống độc quyền thuật toán và hành vi cạnh tranh không lành mạnh, qua đó tạo sân chơi cân bằng giữa các doanh nghiệp công nghệ. (v) Bảo vệ cá nhân và minh bạch thông tin – đặt ra nghĩa vụ bảo vệ quyền nhân thân và yêu cầu dán nhãn nhận diện cho tất cả sản phẩm do AI tạo ra, giúp người dùng phân biệt giữa nội dung của con người và máy, đồng thời giảm bất đối xứng thông tin trong giao dịch. Toàn bộ năm nhóm nguyên tắc được củng cố bởi hệ thống chế tài theo cấp độ: từ xử phạt hành chính, yêu cầu khắc phục kỹ thuật, cho đến đình chỉ hoạt động hay truy cứu trách nhiệm hình sự đối với vi phạm nghiêm trọng. Chính sách này thể hiện mô hình quản trị AI có tính điều tiết mạnh. So với các quốc gia theo mô hình thị trường linh hoạt, cách tiếp cận của Trung Quốc mang đặc trưng chủ động kiểm soát, lấy sự an toàn và thống nhất hệ thống làm ưu tiên, qua đó đảm bảo sự phát triển của AI không đi chệch khỏi lợi ích và tầm nhìn quản trị quốc gia.

3. Bài học kinh nghiệm cho Việt Nam

Trên cơ sở phân tích quản trị rủi ro trong phát triển trí tuệ nhân tạo của các quốc gia trên, có thể rút ra một số bài học sau:

Thứ nhất, xây dựng khung quản trị rủi ro đa tầng và có hệ thống. Từ mô hình của Hoa Kỳ, bài học là cần thiết lập các lớp bảo vệ liên kết trong quản lý AI, bao gồm bảo đảm an toàn kỹ thuật, công bằng

thuật toán, bảo vệ dữ liệu cá nhân, tăng minh bạch và duy trì quyền kiểm soát của con người. Cấu trúc quản trị nên phân biệt rõ giai đoạn trước và sau triển khai - tức kiểm thử, thẩm định độc lập trước khi ứng dụng và giám sát, đánh giá sau khi vận hành - giúp nhận diện và xử lý rủi ro một cách chủ động.

Thứ hai, kết hợp nguyên tắc pháp lý với linh hoạt điều tiết. Kinh nghiệm của Vương quốc Anh cho thấy quản trị rủi ro đối với AI không nên quá cứng nhắc mà phải gắn với mức độ rủi ro cụ thể của từng ứng dụng. Việt Nam có thể áp dụng cách tiếp cận tiệm tiến, bảo đảm sự phát triển công nghệ song hành với quy định pháp luật, đồng thời duy trì khả năng điều chỉnh pháp lý nếu cơ chế tự quản của thị trường chưa đủ hiệu quả. Việc sử dụng các tiêu chuẩn hiện có, như tiêu chuẩn an ninh mạng và bảo vệ dữ liệu, cũng giúp tránh thiết lập rào cản kỹ thuật mới gây cản trở đổi mới.

Thứ ba, làm rõ trách nhiệm và cơ chế giải trình trong toàn bộ vòng đời hệ thống.

Từ kinh nghiệm của Vương quốc Anh, một bài học quan trọng là xác định rõ vai trò, nghĩa vụ và trách nhiệm giải trình của các chủ thể tham gia chuỗi phát triển – triển khai – vận hành AI. Cơ chế này tạo cơ sở cho việc giám sát, xử lý tranh chấp và bảo vệ quyền lợi hợp pháp của cá nhân, góp phần nâng cao tính minh bạch và công bằng của thị trường.

Thứ tư, gắn quản trị rủi ro AI với bảo đảm an ninh, đạo đức và trật tự xã hội.

Kinh nghiệm của Trung Quốc cho thấy AI có thể tạo tác động sâu rộng đến ổn định thể chế và đời sống xã hội, vì vậy Việt Nam cần nhận thức rủi ro không chỉ ở góc kỹ thuật mà cả ở khía cạnh văn hóa, đạo đức và an ninh quốc gia. Việc đặt ra ranh giới pháp lý rõ ràng, cùng yêu cầu tuân thủ chuẩn mực đạo đức – từ thiết kế thuật toán đến sử dụng dữ liệu – là cần thiết để phòng ngừa rủi ro mang tính hệ thống.

4. Kết luận

Kinh nghiệm của Hoa Kỳ, Vương quốc Anh và Trung Quốc cho thấy quản trị rủi ro trong phát triển trí tuệ nhân tạo phải được xem là một cấu phần quan trọng trong phát triển công nghệ quốc gia, không chỉ dừng ở kiểm soát kỹ thuật mà còn bao hàm yếu tố pháp lý, đạo đức và xã hội. Đối với Việt Nam, việc học hỏi các mô hình quốc tế giúp hình thành cách tiếp cận chủ động, đa tầng và linh hoạt, trong đó mục tiêu trọng tâm là bảo đảm AI phát triển vì lợi ích của con người, duy trì niềm tin xã hội và hỗ trợ quá trình chuyển đổi số theo hướng bền vững, an toàn và nhân văn.

Tài liệu tham khảo

1. Thủ tướng Chính phủ. (2021). Quyết định số 127/QĐ-TTg ngày 26/01/2021 về việc ban hành chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng trí tuệ nhân tạo đến năm 2030
2. Quốc hội. (2025). Luật số 134/2025/QH15 ngày 10/12/2025 về luật trí tuệ nhân tạo.
3. Cyberspace Administration of China. (2023). *Interim measures for the management of generative artificial intelligence services*. https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm
4. Department for Science, Innovation and Technology. (2023). *A pro-innovation approach to AI regulation*. <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>
5. McKinsey & Company. (2023). *The economic potential of generative AI: The next productivity frontier*. <https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>
6. White House Office of Science and Technology Policy. (2022). *Blueprint for an AI bill of rights*. <https://www.govinfo.gov/app/details/GOVPUB-PREX23-PURL-gpo193638>
7. World Economic Forum. (2024). *Global risks report 2024*. <https://www.weforum.org/publications/global-risks-report-2024/in-full/>