CƠ CHẾ ĐÁNH GIÁ SỰ PHÙ HỢP ĐỐI VỚI HỆ THỐNG AI RỦI RO CAO THEO ĐẠO LUẬT AI CỦA EU VÀ HÀM Ý CHO VIỆT NAM

ĐÀO KIM ANH
NGUYỄN PHƯƠNG ANH
TRỊNH NGỌC ANH
HOÀNG THỊ CHÂM
NGUYỄN THỊ TRANG NHUNG
Trường Đại học Ngoại thương

Nhận bài ngày 11/02/2026. Sửa chữa xong 19/03/2026. Duyệt đăng 20/03/2026.

Abstract

In the context of Vietnam's gradual development of a legal framework for the digital economy, managing the potential risks arising from artificial intelligence has become an urgent requirement. This article examines the experience of the European Union through the EU AI Act, with a particular focus on conformity assessment mechanisms for high-risk AI systems. By analyzing the management principles and regulatory processes of the European Union, the article identifies gaps in Vietnam's current legal framework and, on that basis, proposes directions for developing control and quality assessment mechanisms for AI systems that are appropriate to the domestic context.

Keywords: *AI Law, conformity assessment, EU AI Act, high-risk AI.*

1. Đặt vấn đề

Trong bối cảnh AI đang trở thành động lực tăng trưởng then chốt nhưng cũng đặt ra những thách thức pháp lý chưa từng có về tính "hộp đen" và rủi ro thiên kiến, xu hướng lập pháp toàn cầu đang chuyển dịch mạnh mẽ sang các cơ chế kiểm soát có tính ràng buộc. Tiên phong trong làn sóng này, Đạo luật AI của EU đã thiết lập quy trình "đánh giá sự phù hợp" như một chốt chặn pháp lý bắt buộc đối với các hệ thống AI rủi ro cao, yêu cầu nhà cung cấp chứng minh sự tuân thủ các chuẩn mực an toàn trước khi gia nhập thị trường. Tại Việt Nam, sự ra đời của Luật Trí tuệ nhân tạo năm 2025 tuy đã nội luật hóa nguyên tắc quản lý dựa trên rủi ro nhưng vẫn còn những khoảng trống lớn trong cơ chế thực thi kỹ thuật. Do đó, bài viết tập trung phân tích bản chất và quy trình đánh giá sự phù hợp theo pháp luật EU để rút ra những hàm ý chính sách thiết thực, nhằm hỗ trợ Việt Nam xây dựng một hành lang pháp lý cân bằng giữa mục tiêu bảo đảm an ninh công cộng và thúc đẩy đổi mới sáng tạo trong kỷ nguyên số.

2. Nội dung nghiên cứu

2.1. Khái quát về hệ thống AI rủi ro cao và yêu cầu đánh giá sự phù hợp

2.1.1. Khái niệm và đặc điểm của hệ thống AI rủi ro cao

Trong bối cảnh xây dựng hành lang pháp lý cho nền kinh tế số, việc xác định phạm vi điều chỉnh đòi hỏi sự phân định rõ ràng giữa các phần mềm truyền thống và các hệ thống trí tuệ nhân tạo. Kế thừa định nghĩa được Tổ chức Hợp tác và Phát triển Kinh tế (OECD), hệ thống AI được hiểu là một hệ thống dựa trên máy móc, hoạt động với các mức độ tự chủ và khả năng thích ứng khác nhau, có thể suy luận từ dữ liệu đầu vào để tạo ra các đầu ra phục vụ mục tiêu xác định [9].

Về việc xác định hệ thống AI rủi ro cao, OECD đánh giá linh hoạt dựa trên tương tác giữa bối cảnh

Email: k62.2311610006@ftu.edu.vn

và tác động tiềm tàng nhằm bao quát kịp thời các công nghệ mới nổi. Trong khi đó, Đạo luật AI của EU áp dụng hướng tiếp cận kiểm soát nghiêm ngặt hơn đối với các hệ thống đe dọa sức khỏe, an toàn và quyền cơ bản, được cụ thể hóa bằng cơ chế phân loại kép tại Điều 6 [1].

Một là, nhóm AI tích hợp trong sản phẩm (như thiết bị y tế, máy móc, phương tiện giao thông), nơi rủi ro của thuật toán gắn liền trực tiếp với rủi ro vật lý của sản phẩm đó. *Hai là, n*hóm AI hoạt động độc lập có nguy cơ gây hậu quả xã hội nghiêm trọng, được quy định cụ thể trong 8 lĩnh vực thiết yếu: sinh trắc học, hạ tầng quan trọng, giáo dục, việc làm, dịch vụ công, thực thi pháp luật, quản lý di cư và hỗ trợ tư pháp [11].

Tổng hợp từ hai cách tiếp cận trên, nghiên cứu này xác định khái niệm hệ thống AI rủi ro cao như sau: *"Hệ thống AI rủi ro cao là các hệ thống trí tuệ nhân tạo được triển khai trong các lĩnh vực thiết yếu của đời sống xã hội (y tế, giáo dục, an ninh, hạ tầng kỹ thuật), có khả năng gây ra những tác động tiêu cực đáng kể, mang tính hệ thống hoặc không thể khắc phục đối với sức khỏe, sự an toàn, quyền sở hữu tài sản hoặc các quyền cơ bản của con người".*

Khác với rủi ro của hàng hóa vật lý thông thường - vốn thường hữu hình và cục bộ, đặc tính của các hệ thống AI rủi ro cao thể hiện qua ba yếu tố cốt lõi: *Thứ nhất,* tính "hộp đen" và thiếu minh bạch, gây cản trở việc giải thích quyết định, truy cứu trách nhiệm và tạo ra lỗ hổng bảo mật; *Thứ hai,* tính tự chủ, dẫn đến tình trạng con người quá tin tưởng vào kết quả của máy móc, đồng thời AI dễ tự động khuếch đại các định kiến dữ liệu cũ, gây phân biệt đối xử có hệ thống [1]; *Thứ ba*, khả năng gây thiệt hại diện rộng và khó đảo ngược (như sự cố y tế, lưới điện), dẫn đến những hậu quả nghiêm trọng về sinh mạng và tài sản mà các biện pháp xử lý hậu kiểm không thể khắc phục được [11].

2.1.2. Vai trò của quy trình đánh giá sự phù hợp hệ thống AI rủi ro cao

Theo định nghĩa chuẩn mực tại tiêu chuẩn quốc tế ISO/IEC 17000:2020, bản chất của quy trình đánh giá sự phù hợp là hoạt động chứng minh rằng các yêu cầu quy định liên quan đến một sản phẩm, quy trình, hệ thống hoặc tổ chức đã được đáp ứng hoàn toàn. Vai trò cốt lõi của quy trình này là chuyển hóa các nguyên tắc đạo đức và pháp lý trừu tượng thành các tiêu chuẩn kỹ thuật có thể kiểm chứng.

Thứ nhất, vai trò xây dựng "niềm tin số" và đảm bảo an ninh hệ thống. Quy trình đánh giá sự phù hợp không chỉ là thủ tục hành chính mà là "chốt chặn" kỹ thuật để đảm bảo khả năng phục hồi của hệ thống. Đối với các lĩnh vực nhạy cảm như y tế hay phương tiện tự hành, quy trình này hoạt động như một cơ chế sàng lọc các lỗ hổng bảo mật trong chuỗi cung ứng. Tem chứng nhận hợp quy đóng vai trò là "tín hiệu niềm tin", giúp xã hội và người dùng cuối an tâm chấp nhận sử dụng các công nghệ AI rủi ro cao [7].

Thứ hai, quy trình này buộc các nhà phát triển phải thiết lập các quy trình bài bản ngay từ giai đoạn thiết kế. Các khái niệm định tính như "công bằng", "minh bạch" hay "không phân biệt đối xử" được chuyển hóa thành các chỉ số kỹ thuật cụ thể trong tài liệu hồ sơ. Ví dụ: yêu cầu về dữ liệu huấn luyện phải "phù hợp và đại diện" được cụ thể hóa thông qua việc kiểm tra các tập dữ liệu về các sai số đối với các nhóm nhân khẩu học khác nhau.

2.2. Quy định về đánh giá sự phù hợp theo Đạo luật AI của EU

Đạo luật Trí tuệ nhân tạo của Liên minh châu Âu (Đạo Luật AI của EU) được ký kết vào tháng 6/2024 và có hiệu lực từ tháng 8/2024, được xem là khung pháp lý toàn diện đầu tiên điều chỉnh trực tiếp và có hệ thống về AI. Trong bối cảnh nhiều quốc gia vẫn đang dừng lại ở chiến lược hoặc khung thử nghiệm, đạo luật này thể hiện vai trò tiên phong của EU trong việc luật hóa việc quản trị AI ở quy mô khu vực và toàn cầu. Đồng thời phản ánh hiện tượng "Brussels Effect", theo đó các tiêu chuẩn pháp lý của EU có xu hướng lan tỏa và trở thành chuẩn mực tham chiếu toàn cầu [3].

Đạo luật AI lựa chọn mô hình quản lý dựa trên rủi ro (risk-based approach), trong đó đánh giá sự phù hợp (conformity assessment) là công cụ thực thi trung tâm đối với các hệ thống AI rủi ro cao. Cơ chế này nhằm kiểm soát rủi ro trước khi đưa hệ thống ra thị trường, đồng thời gắn với hệ thống tiêu chuẩn hài hòa của EU nhằm cung cấp các yêu cầu kỹ thuật và quy trình cụ thể để thực hiện việc đánh giá sự phù hợp và tạo ra cơ chế "giả định tuân thủ".

2.2.1. Hai cơ chế đánh giá sự phù hợp chủ đạo

Việc Liên minh châu Âu lựa chọn đánh giá sự phù hợp làm công cụ thực thi trung tâm trong Đạo luật AI xuất phát từ đặc thù rủi ro của trí tuệ nhân tạo. Nhiều nghiên cứu cho rằng các cơ chế hậu kiểm truyền thống không đủ để kiểm soát rủi ro AI khi thiệt hại có thể xảy ra nhanh và trên diện rộng, đặc biệt trong bối cảnh tính "hộp đen" của các mô hình AI khiến việc truy vết sai sót trở nên phức tạp. Vì vậy, đánh giá sự phù hợp được thiết kế như một cơ chế tiền kiểm nhằm chứng minh tính an toàn và tuân thủ trước khi hệ thống được đưa vào vận hành. Về bản chất, cơ chế này đặt ra yêu cầu các hệ thống AI rủi ro cao phải đáp ứng yêu cầu pháp lý trước khi được gắn chứng nhận CE. Tuy nhiên, để bảo đảm tính linh hoạt và không cản trở sự đổi mới, Điều 43 của đạo luật thiết lập hai lộ trình tuân thủ: cơ chế kiểm soát nội bộ và cơ chế đánh giá sự phù hợp với sự tham gia của bên thứ ba.

Thứ nhất là cơ chế kiểm soát nội bộ. Cơ chế này được quy định tại khoản 2 Điều 43 của Đạo luật AI của EU, đối với đa số hệ thống AI rủi ro cao, nhà cung cấp được phép thực hiện đánh giá sự phù hợp thông qua cơ chế kiểm soát nội bộ, tức tự đánh giá mức độ tuân thủ của hệ thống AI đối với các yêu cầu pháp lý của đạo luật trên cơ sở đã áp dụng các tiêu chuẩn hài hòa hoặc đặc tả chung do EU ban hành. Quy định này nhằm giảm gánh nặng tuân thủ và chi phí pháp lý cho doanh nghiệp trong bối cảnh AI là lĩnh vực đổi mới nhanh. Tuy nhiên, việc để nhà cung cấp tự đánh giá cũng làm phát sinh lo ngại về tính khách quan và hiệu quả kiểm soát, nhất là khi hệ thống AI có thể tác động trực tiếp đến quyền con người. Vì vậy, dù được pháp luật cho phép, kiểm soát nội bộ không đồng nghĩa với việc giảm trách nhiệm pháp lý, và nhà cung cấp phải chịu trách nhiệm trước cơ quan giám sát nếu hệ thống bị phát hiện không phù hợp.

Thứ hai là cơ chế đánh giá sự phù hợp với sự tham gia của bên thứ ba được chỉ định. Khoản 3 Điều 43 Đạo luật AI của EU quy định cơ chế đánh giá nghiêm ngặt hơn đối với một số hệ thống AI rủi ro cao có mức độ tác động nghiêm trọng, đặc biệt là hệ thống nhận dạng sinh trắc học từ xa thuộc Phụ lục III khi không có hoặc không áp dụng đầy đủ các tiêu chuẩn hài hòa và đặc tả chung. Trong các trường hợp này, việc đánh giá phải được thực hiện bởi bên thứ ba độc lập nhằm tăng cường mức độ bảo đảm an toàn và tính khách quan của quá trình đánh giá. Từ góc độ thị trường, cơ chế này được xem như một "hàng rào bảo hiểm pháp lý", giúp giảm rủi ro tranh chấp và chế tài trong các lĩnh vực nhạy cảm như y tế, tài chính hay thực thi pháp luật. Tuy nhiên, chi phí tuân thủ cao và quy trình phức tạp vẫn là thách thức đối với doanh nghiệp, đặc biệt khi số lượng các bên thứ ba đủ năng lực còn hạn chế. Bên cạnh đó, theo ước tính của Trung tâm Đổi mới dữ liệu (Center for Data Innovation - CDI), tổng chi phí tuân thủ đối với một hệ thống AI rủi ro cao, bao gồm đánh giá sự phù hợp bởi bên thứ ba, có thể lên tới khoảng 400.000 euro, tương đương gần 40% lợi nhuận của một doanh nghiệp có doanh thu khoảng 10 triệu euro [2].

2.2.2. Quy định về tiêu chuẩn hài hòa và giả định tuân thủ

Tiêu chuẩn hài hòa là các tiêu chuẩn châu Âu do các tổ chức tiêu chuẩn hóa được công nhận xây dựng, được sử dụng để chứng minh sản phẩm, dịch vụ hoặc quy trình tuân thủ pháp luật EU.

Trong khuôn khổ Đạo luật AI của EU, tiêu chuẩn hài hòa được quy định tại Điều 40 và giữ vai trò cụ thể hóa các yêu cầu pháp lý mang tính nguyên tắc thành các tiêu chí kỹ thuật có thể áp dụng và kiểm chứng trong thực tiễn. Cách tiếp cận này phản ánh mô hình quản lý dựa trên rủi ro của EU. Theo đó pháp luật không trực tiếp quy định cứng các yêu cầu kỹ thuật mà trao không gian cho hệ thống tiêu chuẩn hóa nhằm thích ứng với tốc độ phát triển nhanh của công nghệ AI, đồng thời bảo đảm tính thống nhất của thị trường nội khối.

Theo Điều 40 của đạo luật, khi hệ thống AI rủi ro cao tuân thủ các tiêu chuẩn hài hòa đã được tham chiếu trên Công báo EU, hệ thống đó được hưởng cơ chế "giả định tuân thủ" (presumption of conformity) đối với các yêu cầu pháp lý tương ứng. Khi đó, nhà cung cấp có thể dựa vào việc áp dụng tiêu chuẩn như một căn cứ để chứng minh sự phù hợp, thay vì phải chứng minh lại từng nghĩa vụ pháp lý ở cấp độ luật. Các tiêu chuẩn hóa do Ủy ban châu Âu yêu cầu bao gồm nhiều khía cạnh kỹ thuật của hệ thống AI như: 1) Quản lý rủi ro, 2) Quản trị và chất lượng dữ liệu, 3) Lưu giữ hồ sơ, 4)

Minh bạch, 5) Giám sát con người, 6) Sự chính xác, 7) Độ bền, 8) An ninh mạng, 9) Quản lý chất lượng, 10) Đánh giá sự phù hợp [10].

Đặc biệt, cơ chế giả định tuân thủ còn được cụ thể hóa thông qua Tuyên bố tuân thủ của EU (EU Declaration of Conformity) và chứng nhận CE (CE marking), được quy định lần lượt tại Điều 47 và Điều 48 của Đạo luật. Theo đó, nhà cung cấp phải lập tuyên bố tuân thủ bằng văn bản đối với mỗi hệ thống AI rủi ro cao trước khi đưa ra thị trường hoặc đưa vào sử dụng, Trên cơ sở đó, chứng nhận CE đóng vai trò là biểu thị pháp lý bên ngoài cho thấy hệ thống được coi là đáp ứng các yêu cầu của pháp luật EU và có thể lưu thông hợp pháp trên thị trường nội khối.

Từ góc độ thị trường, cơ chế tiêu chuẩn hài hòa và giả định tuân thủ mang lại ưu điểm về tính chắc chắn pháp lý và khả năng áp dụng thực tiễn. Chúng cho phép chuyển hóa các yêu cầu pháp lý thành các tiêu chí kỹ thuật cụ thể, giúp nhà cung cấp chứng minh sự phù hợp. Tuy nhiên, hiệu quả của cơ chế này phụ thuộc vào tiến độ và chất lượng của các tiêu chuẩn hài hòa, trong khi nhiều tiêu chuẩn AI vẫn đang trong quá trình xây dựng. Ngoài ra, chi phí tuân thủ các tiêu chuẩn kỹ thuật phức tạp có thể tạo gánh nặng cho doanh nghiệp nhỏ và vừa, và giả định tuân thủ cũng không mang tính tuyệt đối khi vẫn có thể bị bác bỏ thông qua cơ chế hậu kiểm của cơ quan có thẩm quyền.

2.3. Quy định về đánh giá sự phù hợp đối với hệ thống AI rủi ro cao tại Việt Nam và một số kiến nghị thực thi

2.3.1. Quy định pháp luật Việt Nam về đánh giá sự phù hợp đối với hệ thống AI rủi ro cao

Ngày 10/12/2025, Quốc hội thông qua Luật Trí tuệ nhân tạo 2025, thiết lập khung quản lý đối với trí tuệ nhân tạo tại Việt Nam, có hiệu lực từ ngày 01/3/2026. Luật đặt ra các quy định về đánh giá sự phù hợp nhằm bảo đảm các hệ thống AI rủi ro cao đáp ứng các yêu cầu về pháp lý, kỹ thuật và an toàn trước khi được đưa vào triển khai.

Liên quan đến việc xác định hệ thống AI rủi ro cao, Điều 9 quy định về phân loại mức độ rủi ro của hệ thống trí tuệ nhân tạo. Theo điểm a khoản 1 Điều 9, hệ thống AI rủi ro cao được hiểu là hệ thống có thể gây thiệt hại đáng kể đến quyền và lợi ích hợp pháp của tổ chức, cá nhân hoặc lợi ích công cộng và an ninh quốc gia. Trên cơ sở đó, Luật yêu cầu hệ thống AI rủi ro cao phải được đánh giá sự phù hợp trước khi đưa vào sử dụng hoặc khi có thay đổi đáng kể trong quá trình sử dụng (khoản 1 Điều 13), thể hiện cách tiếp cận tiền kiểm dựa trên rủi ro.

Theo khoản 2 Điều 13, đánh giá sự phù hợp là việc xác nhận hệ thống AI đáp ứng các yêu cầu tại Điều 14, bao gồm an toàn, quản lý rủi ro, quản trị dữ liệu, minh bạch và giám sát con người. Luật đồng thời phân biệt hai hình thức đánh giá: đối với các hệ thống AI rủi ro cao thuộc danh mục phải chứng nhận sự phù hợp trước khi đưa vào sử dụng, việc đánh giá phải do tổ chức đã được đăng ký hoặc thừa nhận thực hiện (điểm a khoản 2 Điều 13). Đối với các hệ thống còn lại, nhà cung cấp được phép tự đánh giá hoặc thuê tổ chức đánh giá đã được đăng ký hoặc thừa nhận (điểm b khoản 2 Điều 13). Khoản 3 Điều 13 cũng khẳng định kết quả đánh giá sự phù hợp là điều kiện pháp lý bắt buộc để hệ thống AI rủi ro cao được phép đưa vào sử dụng và là căn cứ cho hoạt động kiểm tra, giám sát của cơ quan nhà nước.

Trên cơ sở các quy định này, có thể thấy Luật Trí tuệ nhân tạo 2025 bước đầu thiết lập cơ chế đánh giá sự phù hợp đối với hệ thống AI rủi ro cao theo hướng tiền kiểm và phân tầng dựa trên mức độ rủi ro. Quy định này phản ánh cách tiếp cận quản lý phù hợp với xu hướng pháp luật quốc tế. Các nghiên cứu so sánh pháp lý cho thấy luật học hỏi mô hình tiếp cận rủi ro tương tự Đạo luật AI EU và các khuôn khổ tiên tiến nhằm cân bằng giữa kiểm soát và thúc đẩy đổi mới [4]. Tuy nhiên, Luật hiện hành vẫn bộc lộ một số hạn chế thực tiễn. 1) Các quy định chưa xác định rõ tiêu chí kỹ thuật định lượng cho việc phân loại và đánh giá AI rủi ro cao, khi Điều 9 và Điều 13 chủ yếu nêu khái niệm mang tính mô tả. 2) Luật chưa có hướng dẫn cụ thể về cơ chế công nhận kết quả đánh giá từ các tổ chức quốc tế hoặc tổ chức kỹ thuật uy tín, hạn chế khả năng công nhận chéo và hội nhập với các tiêu chuẩn kỹ thuật quốc tế. 3) Việt Nam hiện chưa có bộ tiêu chuẩn kỹ thuật quốc gia về AI và hệ thống tổ chức đánh giá độc lập với năng lực chuyên sâu còn hạn chế, gây khó khăn cho cả cơ quan quản lý và doanh nghiệp khi thực hiện đánh giá trong thực tiễn.

2.3.2. Một số kiến nghị cho Việt Nam

Thứ nhất, về việc xác định phạm vi quản lý, Việt Nam cần xây dựng danh mục "hệ thống AI rủi ro cao" phù hợp với đặc thù kinh tế - xã hội trong nước thay vì áp dụng nguyên bản quy định của EU. Dựa trên nguyên tắc phân loại rủi ro tại Điều 6 và Phụ lục III của Đạo luật AI EU [11], cần ưu tiên kiểm soát ba lĩnh vực trọng yếu: 1) Y tế - do rủi ro trực tiếp đến sinh mạng; 2) Giao thông và hạ tầng trọng yếu - đòi hỏi tiêu chuẩn kiểm định đặc thù phù hợp với mô hình giao thông hỗn hợp; 3) Tài chính - nhằm ngăn chặn lỗi phân biệt đối xử của AI và bảo đảm tính công bằng.

Thứ hai, về quy trình thực hiện, cần áp dụng cơ chế đánh giá linh hoạt theo mức độ rủi ro. Với nhóm rủi ro trung bình - cao, nên cho phép doanh nghiệp tự đánh giá kết hợp tham gia "Khung thử nghiệm pháp lý" nhằm giảm bớt thủ tục hành chính, tạo điều kiện đổi mới sáng tạo nhưng vẫn giữ được sự giám sát của cơ quan quản lý. Ngược lại, đối với các hệ thống AI nhạy cảm liên quan đến an ninh quốc gia hoặc dữ liệu sinh trắc học, bắt buộc phải do một bên thứ ba độc lập kiểm định. Tham khảo kinh nghiệm từ EU, yêu cầu bắt buộc này là vô cùng cần thiết để đảm bảo tính khách quan, bảo vệ quyền riêng tư và ngăn chặn nguy cơ lạm dụng theo dõi người dân [11].

Thứ ba, về nền tảng kỹ thuật, Nhà nước cần đẩy mạnh việc nội hóa các tiêu chuẩn quốc tế làm cơ sở tham chiếu cho hoạt động đánh giá sự phù hợp. Cụ thể, cần nhanh chóng nghiên cứu và chuyển dịch bộ tiêu chuẩn ISO/IEC 42001 thành Tiêu chuẩn Việt Nam. Việc áp dụng tiêu chuẩn này mang lại hai lợi ích thiết thực: vừa giúp doanh nghiệp minh bạch hóa quy trình phát triển AI và có cơ sở rõ ràng để giải trình trách nhiệm khi được kiểm tra; vừa giúp sản phẩm AI của Việt Nam đáp ứng các yêu cầu kỹ thuật chung của thế giới, từ đó dễ dàng vượt qua rào cản thương mại để xuất khẩu sang các thị trường khắt khe như EU hay Hoa Kỳ.

3. Kết luận

Trong bối cảnh AI phát triển nhanh, việc quản lý rủi ro đối với các hệ thống AI rủi ro cao là cần thiết. Kinh nghiệm từ Đạo luật AI của EU cho thấy đánh giá sự phù hợp đóng vai trò như một công cụ tiền kiểm quan trọng nhằm bảo đảm an toàn và trách nhiệm của các hệ thống AI trước khi đưa vào sử dụng. Đối với Việt Nam, Luật Trí tuệ nhân tạo 2025 là bước tiến quan trọng trong việc hình thành khung pháp lý quản trị AI. Tuy nhiên, để vận hành hiệu quả, Việt Nam cần tiếp tục hoàn thiện tiêu chuẩn kỹ thuật, phát triển hệ thống tổ chức đánh giá độc lập và tiếp thu có chọn lọc kinh nghiệm quốc tế phù hợp với điều kiện trong nước.

Tài liệu tham khảo

[1] Ali Sunyaev, Alexander Benlian, Jella Pfeiffer, Ekaterina Jussupow, Scott Thiebes, Alexander Maedche, Joshua Gawlitza (2025). *High-Risk Artificial Intelligence.* https://link.springer.com/content/pdf/10.1007/s12599-025-00942-6.pdf, truy cập ngày 11/01/2026.

[2] Benjamin Mueller (2021). *How much will the Artificial Intelligence Act Cost Europe?.* Center for Data Innovation. https://www2.datainnovation.org/2021-aia-costs.pdf, truy cập ngày 3/2/2026.

[3] Charlotte Siegmann, Markus Anderljung (2022). *The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market.* https://arxiv.org/pdf/2208.12645, truy cập ngày 3/2/2026.

[4] Dao Gia Phuc, Nguyen Van Duong (2025). *A Risk-Based Approach to Artificial Intelligence Regulation: Lessons from the European Union for Vietnam.* VNU Journal of Science, https://js.vnu.edu.vn/LS/article/view/4799, truy cập ngày 3/2/2026.

[5] EDRi (2025). *Undermining the GDPR through 'simplification': EDRi pushes back against dangerous deregulation*. https://edri.org/our-work/undermining-the-gdpr-through-simplification-edri-pushes-back-against-dangerous-deregulation, truy cập ngày 3/2/2026.

[6] ISO/IEC 17000:2020 (2020). *Conformity assessment — Vocabulary and general principles.* https://www.iso.org/obp/ui/en/#iso:std:73029:en, truy cập ngày 31/01/2026.

[7] Kitty Kioskli, Elisavet Grigoriou, Shareeful Islam, Andrianos M. Yiorkas, Loizos Christofi, Haralambos Mouratidis (2025). *A risk and conformity assessment framework to ensure security and resilience of healthcare systems and medical supply chain.* https://link.springer.com/article/10.1007/s10207-025-01009-z, truy cập ngày 01/02/2026.

[8] Michael Veale, Frederik Zuiderveen Borgesius (2021). *Demystifying the Draft EU Artificial Intelligence Act.* Computer Law Review International, https://pure.uva.nl/ws/files/154819168/VealeBorgesius2021_typeset.pdf, truy cập ngày 3/2/2026.

[9] OECD. *The Recommendation on Artificial Intelligence.* https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449, truy cập ngày 10/01/2026.

[10] SOLER GARRIDO Josep (2024). *Harmonised Standards for the European AI Act.* https://publications.jrc.ec.europa.eu/repository/handle/JRC139430, truy cập ngày 19/1/2026.

[11] Theodoros Karathanasis. *Guidance on Classification and Conformity Assessments for High-Risk AI Systems under EU AI Act.* https://ai-regulation.com/wp-content/uploads/2023/02/Article-AI-Classification.pdf, truy cập ngày 11/01/2026.