

BẢO ĐẢM QUYỀN CON NGƯỜI TRONG QUÁ TRÌNH ỨNG DỤNG TRÍ TUỆ NHÂN TẠO VÀO HOẠT ĐỘNG XÉT XỬ TẠI VIỆT NAM

LÊ THỊ THÚY KIỀU
NGUYỄN THANH HUY
Trường Đại học Mở Thành phố Hồ Chí Minh

Nhận bài ngày 08/6/2026. Sửa chữa xong 08/7/2026. Duyệt đăng 04/7/2026.

Abstract

Amid Vietnam's accelerating digital transformation, artificial intelligence (AI) is increasingly being studied and gradually applied across various areas of social life, including judicial and adjudicative activities. The integration of AI into adjudication is expected to enhance court efficiency by enabling the processing of large volumes of data, facilitating legal research, supporting evidence analysis, and reducing case-resolution times. Despite these technical and administrative benefits, the application of AI in the judicial sector raises significant legal and human rights concerns that require careful consideration. This article examines the theoretical and legal foundations for protecting human rights in adjudicative activities under the 2013 Constitution of Vietnam and the country's current procedural laws. Drawing on international experience and comparative legal approaches, the article also identifies legal gaps and challenges associated with the application of AI in Vietnam's judicial system. Based on this analysis, it proposes directions for strengthening the legal framework to ensure that the use of AI in adjudicative activities remains consistent with constitutional principles and international human rights standards.

Keywords: Adjudication, artificial intelligence, human rights, protection of rights and legitimate interests

1. Đặt vấn đề

Chuyển đổi số đã trở thành một hướng đi mới cho sự phát triển trong nhiều lĩnh vực của đời sống kinh tế và xã hội, nhưng hiện nay nó đang dần trở thành một yêu cầu bắt buộc. Khi các công nghệ số trưởng thành, AI đã được xem như một động lực tiềm năng cho quá trình thử nghiệm. Nhiều quốc gia, AI đã được ứng dụng trong nghiên cứu hợp tác và pháp lý, thu thập dữ liệu để diễn giải pháp luật, tính toán xác suất tái phạm hoặc cung cấp hỗ trợ cho các quy trình khác nhau. Thực tiễn này, cho thấy AI có thể góp phần nâng cao hiệu quả hoạt động của hệ thống tư pháp và giảm áp lực cho đội ngũ thẩm phán (Nguyễn Trần Quế Trân, 2025). Tuy nhiên, tại Việt Nam, vấn đề ứng dụng AI trong hoạt động xét xử hiện vẫn chưa được điều chỉnh bằng một cơ chế pháp lý đầy đủ và thống nhất.

Về mặt pháp lý, quyền con người và quyền công dân là nền tảng cơ bản của Nhà nước pháp quyền xã hội chủ nghĩa Việt Nam. Hiến pháp năm 2013 thể hiện sự ghi nhận nguyên tắc tôn trọng, bảo vệ và bảo đảm quyền con người trong mọi lĩnh vực của đời sống xã hội, bao gồm cả hoạt động tố tụng và xét xử. Đặc biệt, các quyền như quyền được xét xử công bằng, quyền bào chữa, quyền tranh tụng tại phiên tòa, quyền được bảo vệ đời tư và dữ liệu cá nhân đều là những quyền hiến định có ý nghĩa quan trọng trong việc bảo vệ công lý và bảo đảm tính khách quan của quá trình tố tụng. Đồng thời, các bộ luật tố tụng hiện hành như Bộ luật Tố tụng hình sự năm 2015 và Bộ luật Tố tụng dân sự năm 2015 cũng ghi nhận nhiều nguyên tắc nhằm bảo đảm quyền con người trong quá trình giải quyết vụ việc tại Tòa án (Nguyễn Thị Lan Anh, Lê Xuân Lục, 2023).

Tuy nhiên, việc sử dụng AI trong các quy trình tư pháp đã đặt ra hàng loạt câu hỏi pháp lý mới mà

Email: lekieu0408@gmail.com.

khôn khổ pháp luật hiện hành vẫn chưa giải quyết. Lời kết: Nó có nguy cơ làm suy giảm tính minh bạch, có thể khó theo dõi hơn hoặc có thể dẫn đến định kiến trong cách thức tiến hành xử lý. Điều này đưa chúng ta đến khái niệm rằng nếu một phần việc phân tích hoặc đánh giá được thực hiện bởi một hệ thống công nghệ, liệu điều đó có vi phạm quyền của các bên được xét xử công bằng hay không? Hơn nữa, khi AI đóng vai trò trong việc thu thập, xử lý và lưu trữ khối lượng lớn dữ liệu cá nhân liên quan đến các thủ tục tố tụng, các quyền liên quan đến quyền riêng tư, bảo vệ dữ liệu cá nhân và quyền tiếp cận thông tin cũng có thể bị đe dọa.

Vì vậy, mối quan hệ giữa việc ứng dụng AI trong xét xử và mục tiêu bảo vệ quyền con người là có ý nghĩa cả về phương diện lý luận lẫn thực tiễn (Phạm Liên, 2025). Đây không chỉ là vấn đề năng lực trong cải cách tư pháp, mà còn liên quan đến việc xây dựng các cơ chế pháp lý phù hợp để bảo đảm việc sử dụng và kiểm soát công nghệ trong các hoạt động xét xử (Nguyễn Trần Quốc Trân, 2025). Bài viết được thực hiện nhằm làm rõ các khoảng trống pháp lý này trong quá trình áp dụng AI vào xét xử trên cơ sở các phương pháp phân tích kỹ thuật xác định các quy định của pháp luật Việt Nam hiện hành kết hợp với kinh nghiệm quốc tế. Sau đó, đề xuất sẽ đưa ra định hướng cho các cải thiện pháp lý tiếp theo nhằm hài hòa hai sức ép song song là đổi mới công nghệ và đáp ứng quyền con người trong bối cảnh chuyển đổi số hiện nay.

2. Nội dung nghiên cứu

2.1. Quy định pháp luật về bảo đảm quyền con người trong hoạt động tố tụng khi ứng dụng trí tuệ nhân tạo

Việc ứng dụng trí tuệ nhân tạo trong tranh tụng và xét xử cần được triển khai trên cơ sở bảo đảm các quyền con người cơ bản, đây là nền tảng cốt lõi của nền tư pháp hiện đại (Ngô Anh Cường, 2023). Mặc dù AI có thể hỗ trợ phân tích dữ liệu, đánh giá chứng cứ hoặc dự đoán rủi ro pháp lý, nhưng loại công nghệ này không được xâm phạm các nguyên tắc hiến định và các quyền tố tụng mà pháp luật Việt Nam thừa nhận.

Thứ nhất, quyền được xét xử công bằng và đúng trình tự pháp luật. Hiến pháp năm 2013 và các điều ước quốc tế mà Việt Nam là thành viên đều thừa nhận mọi cá nhân có quyền được xét xử khách quan, công khai và tuân theo pháp luật. Nội dung này được cụ thể hóa trong Bộ luật Tố tụng hình sự năm 2015 thông qua nguyên tắc về tính độc lập trong xét xử và việc tòa án chỉ có nghĩa vụ tuân theo pháp luật (Nguyễn Thị Lan Anh, Lê Xuân Lục, 2023). Trong trường hợp AI được sử dụng như một công cụ hỗ trợ trong việc giải quyết vụ án, công nghệ chỉ có thể đảm nhiệm vai trò hỗ trợ kỹ thuật và không thể thay thế cho thẩm phán hoặc hội đồng xét xử. Tuy nhiên, nếu các cơ quan truy tố hoặc xét xử quá phụ thuộc vào các kết quả phân tích do AI tạo ra (Ngô Anh Cường, 2023), thì có nguy cơ thực sự rằng tính khách quan và sự công bằng của quá trình tố tụng tư pháp có thể bị ảnh hưởng.

Thứ hai, quyền bào chữa và quyền tiếp cận công lý. Pháp luật tố tụng hiện hành ghi nhận người bị buộc tội có quyền tự bảo vệ hoặc nhờ người khác bảo vệ quyền và lợi ích hợp pháp của mình. Đồng thời, Luật Trợ giúp pháp lý năm 2017 cũng tạo cơ sở để các nhóm yếu thế có điều kiện tiếp cận công lý. Tuy nhiên, trong trường hợp AI được sử dụng để phân tích chứng cứ hoặc đánh giá hành vi, người tham gia tố tụng có thể gặp khó khăn nếu không được tiếp cận thông tin về cách thức hệ thống AI đưa ra kết quả. Khi thuật toán vận hành thiếu minh bạch, quyền tranh luận, phản biện và kiểm tra chứng cứ của luật sư hoặc bị can, bị cáo có nguy cơ bị hạn chế, từ đó ảnh hưởng đến nguyên tắc tranh tụng trong tố tụng hiện đại.

Thứ ba, quyền bình đẳng trước pháp luật. Một trong những yêu cầu quan trọng khi ứng dụng AI trong hoạt động tư pháp là phải bảo đảm tính khách quan và không phân biệt đối xử. Trên thực tế, các hệ thống AI hoạt động dựa trên dữ liệu đầu vào và cơ chế học máy; vì vậy, nếu dữ liệu chứa định kiến hoặc thiếu tính đại diện thì kết quả xử lý có thể dẫn đến sự thiên lệch trong quá trình đánh giá vụ việc. Điều này làm phát sinh nguy cơ ảnh hưởng đến nguyên tắc bình đẳng giữa các chủ thể tham gia tố tụng. Do đó, việc kiểm soát chất lượng dữ liệu, cơ chế vận hành thuật toán và tính minh bạch của AI là yêu cầu cần thiết nhằm bảo đảm quyền bình đẳng trước pháp luật của mọi cá nhân.

Thứ tư, quyền bảo vệ đời tư và dữ liệu cá nhân. Tại điều 21 của Hiến pháp năm 2013 cùng các quy định của Bộ luật Dân sự năm 2015, Luật mạng năm 2018 và Nghị định của Chính phủ số 13/2023/NĐ-CP ngày 17 tháng 4 năm 2023 về bảo vệ dữ liệu cá nhân đã ghi nhận quyền bất khả xâm phạm đối với đời sống riêng tư và thông tin cá nhân. Trong khi đó, AI thường cần khai thác lượng lớn dữ liệu để phục vụ hoạt động phân tích và xử lý thông tin. Việc thu thập, lưu trữ hoặc chia sẻ dữ liệu liên quan đến người tham gia tố tụng nếu không có cơ chế kiểm soát phù hợp có thể dẫn đến nguy cơ lộ lọt thông tin hoặc sử dụng dữ liệu vượt quá phạm vi cho phép. Vì vậy, việc ứng dụng AI trong xét xử phải đi kèm với cơ chế bảo mật nghiêm ngặt, giới hạn mục đích sử dụng dữ liệu và xác định rõ trách nhiệm của cơ quan quản lý dữ liệu.

Thứ năm, nguyên tắc suy đoán vô tội trong tố tụng hình sự. Đây là nguyên tắc quan trọng đã được ghi nhận trong Bộ luật Tố tụng Hình sự năm 2015: Người bị buộc tội được coi là không có tội cho đến khi có bản án kết tội của Tòa án đã có hiệu lực pháp luật. Tuy nhiên, khả năng hỗ trợ phát hiện dấu hiệu vi phạm hoặc phân tích xác suất một hành vi phạm tội được thực hiện vẫn còn bị giới hạn, vì vậy không thể diễn giải rằng AI là cơ sở duy nhất để xác định trách nhiệm hình sự của một người. (Nguyễn Thị Lan Anh, Lê Xuân Lục, 2023) Ngoài ra, việc sử dụng sai kết quả đánh giá của AI có thể tạo ra sự thiên lệch hoặc làm suy giảm tầm quan trọng của việc đánh giá chứng cứ bởi các cơ quan tiến hành tố tụng có thẩm quyền (Ngô Anh Cường, 2023). Do đó, trong mọi trường hợp, vẫn cần thiết rằng yếu tố con người cuối cùng phải đóng vai trò quyết định để đạt được công lý và bảo vệ quyền con người trong các hoạt động tố tụng tại tòa án.

2.2. Kinh nghiệm quốc tế về bảo đảm quyền con người khi ứng dụng trí tuệ nhân tạo trong hoạt động tư pháp

Thực tiễn trên thế giới cho thấy việc ứng dụng trí tuệ nhân tạo (AI) trong lĩnh vực tư pháp đang trở thành xu hướng phổ biến trong bối cảnh chuyển đổi số. Tuy nhiên, đa số quốc gia đều tiếp cận vấn đề này theo hướng thận trọng, đặt yêu cầu kiểm soát pháp lý và bảo vệ quyền con người lên hàng đầu thay vì ưu tiên tuyệt đối cho yếu tố công nghệ (Ngô Anh Cường, 2023). Việc xây dựng cơ chế quản lý AI trong hoạt động xét xử thường gắn với nguyên tắc minh bạch, trách nhiệm giải trình và duy trì vai trò quyết định cuối cùng của con người trong quá trình tố tụng.

Liên minh Châu Âu hiện nay là một trong những chủ thể quốc tế lớn đầu tiên nỗ lực xây dựng khung pháp lý về AI thông qua Đạo luật AI năm 2024¹. Một trong những khía cạnh quan trọng nhất của văn bản pháp luật này là các hệ thống AI được phân loại theo mức độ rủi ro của chúng để triển khai cơ chế quản trị tương ứng. Các ứng dụng trí tuệ nhân tạo liên quan đến các quy trình tư pháp-những ứng dụng hỗ trợ việc đánh giá trọng lượng của chuyên môn, đưa ra khuyến nghị xét xử, hoặc xác định mức độ đe dọa do một cá nhân gây ra thông thường được xếp vào nhóm đặc biệt hoặc nhóm rủi ro cao. Điều này hàm ý rằng các hệ thống như vậy phải tuân thủ các tiêu chí nghiêm ngặt liên quan đến tính minh bạch, khả năng giải thích kết quả, tính toàn vẹn của dữ liệu và cơ chế giám sát của con người. Liên minh Châu Âu mở rộng xu hướng này, cho thấy rằng ngoài việc thúc đẩy đổi mới công nghệ, Liên minh Châu Âu cũng hướng tới sự công bằng và tôn trọng các quyền cơ bản của cá nhân trong không gian số (European Parliament & Council of the European Union, 2024).

Ngoài ra, bên cạnh EU, một quốc gia khác cũng có khá nhiều chuyên môn trong việc thử nghiệm và quản lý AI trong khu vực công là Singapore (Anh Thư, 2019). Cơ chế “pháp lý sandbox” (Abbinaya Kuzhanthaivel, 2023) được áp dụng tại quốc gia này, theo đó các công nghệ mới có thể hoạt động trong phạm vi và thời hạn giới hạn, dưới sự kiểm tra, giám sát của Cơ quan thẩm quyền như một cơ chế thử nghiệm. Thông qua cơ chế này, các cơ quan quản lý có thể thu thập hiểu biết về tính an toàn, hiệu quả và tiềm năng tác động của AI trước khi triển khai rộng rãi trong thực tiễn. Trong lĩnh vực tư pháp, mô hình sandbox là một giải pháp có thể giảm thiểu các rủi ro liên quan đến việc ứng dụng công nghệ một cách vội vàng, đồng thời vẫn tạo điều kiện cho đổi mới (Abbinaya Kuzhanthaivel, 2023). Cách tiếp

1) European Parliament & Council of the European Union (2024). *Regulation (EU) 2024/1689 (AI Act)*. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>.

cận này được ca ngợi như một phương án dung hòa giữa nhu cầu hiện đại hóa và việc bảo vệ quyền con người trong tòa án.

Đối với Việt Nam, việc nghiên cứu kinh nghiệm quốc tế có ý nghĩa quan trọng trong bối cảnh hệ thống pháp luật hiện nay vẫn chưa có quy định chuyên biệt về AI trong lĩnh vực tư pháp. Thay vì xây dựng ngay một đạo luật hoàn chỉnh về AI khi điều kiện thực tiễn chưa đầy đủ, Việt Nam có thể cân nhắc áp dụng cơ chế thử nghiệm có kiểm soát đối với một số hoạt động hỗ trợ tư pháp như tra cứu án lệ, phân tích văn bản pháp luật hoặc quản lý dữ liệu tố tụng. Việc triển khai theo lộ trình thử nghiệm sẽ giúp cơ quan nhà nước có điều kiện đánh giá hiệu quả vận hành, mức độ tác động cũng như những rủi ro pháp lý có thể phát sinh trong thực tế.

Ngoài ra, quá trình thử nghiệm còn tạo cơ sở thực tiễn quan trọng để từng bước hoàn thiện khung pháp lý điều chỉnh AI trong hoạt động xét xử. Trên cơ sở đó, Việt Nam có thể nghiên cứu sửa đổi, bổ sung các đạo luật liên quan như Bộ luật Tố tụng hình sự, Bộ luật Tố tụng dân sự, Luật An ninh mạng hoặc pháp luật về bảo vệ dữ liệu cá nhân nhằm thiết lập cơ chế kiểm soát phù hợp đối với công nghệ AI (Nguyễn Thị Lan Anh, Lê Xuân Lục, 2023). Về lâu dài, việc xây dựng quy định pháp lý riêng về AI trong lĩnh vực tư pháp không chỉ góp phần nâng cao hiệu quả hoạt động xét xử mà còn thể hiện cam kết của Nhà nước trong việc bảo vệ quyền con người và bảo đảm công lý trong thời đại số.

2.3. Thực tiễn chuyển đổi số và ứng dụng trí tuệ nhân tạo trong ngành Tư pháp Việt Nam

Trong những năm gần đây, cùng với xu hướng chuyển đổi số quốc gia, ngành tư pháp Việt Nam đã từng bước đẩy mạnh việc ứng dụng công nghệ thông tin vào hoạt động quản lý và giải quyết công việc chuyên môn. Đặc biệt, sau khi ban hành các chủ trương về xây dựng Nhà nước pháp quyền và thúc đẩy chuyển đổi số trong hoạt động công vụ, nhiều cơ quan tư pháp đã triển khai các nền tảng số phục vụ hoạt động tố tụng và quản lý tư pháp. Hệ thống xét xử trực tuyến, cơ sở dữ liệu bản án điện tử, phần mềm quản lý hồ sơ vụ án và các cổng thông tin tra cứu pháp luật được đưa vào vận hành nhằm nâng cao hiệu quả giải quyết công việc, tăng tính minh bạch và tạo thuận lợi cho người dân tiếp cận công lý. Trong quá trình đó, các yếu tố liên quan đến trí tuệ nhân tạo cũng bắt đầu xuất hiện ở mức độ hỗ trợ kỹ thuật và hành chính. Một số hệ thống bước đầu được sử dụng để tự động hóa việc phân loại hồ sơ, nhận diện văn bản, hỗ trợ tìm kiếm dữ liệu hoặc thống kê hoạt động tư pháp. Bên cạnh đó, các mô hình phiên tòa trực tuyến và nền tảng quản lý án điện tử tại một số địa phương lớn đã cho thấy xu hướng chuyển dịch từ phương thức làm việc truyền thống sang môi trường số trong hoạt động tư pháp. Tuy nhiên, cho đến nay, Việt Nam vẫn chưa triển khai chính thức hệ thống AI có khả năng tham gia trực tiếp vào việc đưa ra quyết định xét xử hoặc thay thế vai trò đánh giá của người tiến hành tố tụng.

Mặc dù chuyển đổi số mang lại nhiều lợi ích trong việc nâng cao hiệu quả quản lý và rút ngắn thời gian xử lý công việc, song thực tiễn cũng cho thấy nhiều khó khăn và thách thức pháp lý mới đang phát sinh. Trước hết, hệ thống pháp luật hiện hành chưa có quy định cụ thể về trách nhiệm pháp lý trong trường hợp AI gây ra sai sót trong quá trình hỗ trợ hoạt động tố tụng (Ngô Anh Cường, 2023). Điều này dẫn đến khoảng trống trong việc xác định trách nhiệm của cơ quan quản lý, đơn vị vận hành hoặc cá nhân sử dụng hệ thống công nghệ khi xảy ra hậu quả ảnh hưởng đến quyền và lợi ích của người tham gia tố tụng.

Ngoài ra, vấn đề bảo vệ dữ liệu cá nhân và quyền riêng tư cũng trở thành thách thức lớn trong môi trường tư pháp số. Việc tổ chức phiên tòa trực tuyến, lưu trữ hồ sơ điện tử hoặc xử lý dữ liệu thông qua nền tảng số làm gia tăng nguy cơ lộ lọt thông tin cá nhân, đặc biệt đối với các dữ liệu liên quan đến nhân thân, hình ảnh, giọng nói hoặc thông tin vụ án. Trong khi đó, cơ chế kiểm soát dữ liệu và bảo mật thông tin trong hoạt động tư pháp hiện nay vẫn chưa thật sự đồng bộ, dẫn đến nguy cơ xâm phạm quyền riêng tư của cá nhân khi công nghệ được ứng dụng ở quy mô lớn.

Bên cạnh các vấn đề pháp lý, hạn chế về năng lực công nghệ và kỹ năng số của đội ngũ cán bộ tư pháp cũng là một trở ngại đáng chú ý. Thực tế cho thấy nhiều thẩm phán, kiểm sát viên, luật sư hoặc cán bộ chuyên môn vẫn chưa được trang bị đầy đủ kiến thức về công nghệ, an toàn dữ liệu và cơ chế vận hành của AI. Điều này có thể dẫn đến sự phụ thuộc vào hệ thống kỹ thuật mà thiếu khả năng kiểm tra, đánh giá hoặc

kiểm soát kết quả do công nghệ tạo ra. Đồng thời, việc thiếu hiểu biết về rủi ro thuật toán và bảo mật thông tin cũng làm gia tăng nguy cơ ảnh hưởng đến quyền con người trong quá trình giải quyết vụ việc.

Một vấn đề khác cần được quan tâm là hệ thống pháp luật Việt Nam hiện vẫn chưa hình thành đầy đủ cơ chế pháp lý chuyên biệt điều chỉnh việc ứng dụng AI trong hoạt động tư pháp. Sự thiếu hụt này có thể tạo ra khoảng trống pháp lý trong việc kiểm soát hoạt động công nghệ, nhất là khi AI tham gia vào quá trình xử lý dữ liệu hoặc hỗ trợ ra quyết định tố tụng. Nếu không có cơ chế quản lý phù hợp, nguy cơ ảnh hưởng đến quyền được xét xử công bằng, quyền bào chữa hoặc quyền bảo vệ dữ liệu cá nhân là điều hoàn toàn có thể xảy ra.

Từ thực tiễn trên có thể thấy rằng, ngành tư pháp Việt Nam hiện đang đứng trước giai đoạn chuyển đổi quan trọng giữa yêu cầu hiện đại hóa hoạt động xét xử và nhiệm vụ bảo vệ quyền con người trong môi trường số. Điều này đòi hỏi cơ quan lập pháp và các cơ quan tư pháp cần sớm xây dựng cơ chế pháp lý phù hợp để kiểm soát việc ứng dụng AI trong hoạt động tố tụng, bảo đảm mọi tiến trình đổi mới công nghệ đều được thực hiện trên cơ sở minh bạch, an toàn và tôn trọng các giá trị cốt lõi của công lý.

2.4. Giải pháp hoàn thiện pháp luật và nâng cao hiệu quả bảo vệ quyền con người trong bối cảnh ứng dụng trí tuệ nhân tạo vào hoạt động xét xử tại Việt Nam

Từ thực tiễn chuyển đổi số của ngành tư pháp Việt Nam cùng kinh nghiệm của nhiều quốc gia trên thế giới, có thể nhận thấy việc ứng dụng trí tuệ nhân tạo trong hoạt động xét xử là xu hướng khó tránh khỏi trong tương lai. Tuy nhiên, để AI thực sự trở thành công cụ hỗ trợ hiệu quả cho hoạt động tư pháp, quá trình triển khai cần được đặt dưới sự kiểm soát chặt chẽ của pháp luật và phải lấy con người làm trung tâm. Mục tiêu của việc ứng dụng AI không chỉ nhằm nâng cao hiệu quả quản lý, giảm tải áp lực công việc cho cơ quan tư pháp mà còn phải bảo đảm các nguyên tắc công bằng, minh bạch và tôn trọng quyền con người trong mọi giai đoạn tố tụng.

Trước hết, Việt Nam cần xây dựng cơ chế thử nghiệm có kiểm soát đối với AI trong lĩnh vực tư pháp theo mô hình “*sandbox pháp lý*” (Abbinaya Kuzhanthaivel, 2023). Đây được xem là giải pháp phù hợp trong bối cảnh pháp luật hiện nay chưa có quy định chuyên biệt điều chỉnh AI trong hoạt động xét xử. Thông qua cơ chế này, cơ quan tư pháp có thể triển khai thí điểm một số ứng dụng AI trong phạm vi giới hạn như hỗ trợ quản lý hồ sơ, tra cứu án lệ, thống kê dữ liệu hoặc nhận diện văn bản pháp lý dưới sự giám sát của cơ quan chuyên môn. Việc thử nghiệm trước khi áp dụng rộng rãi sẽ giúp đánh giá tính khả thi, mức độ an toàn và nguy cơ phát sinh đối với quyền con người, từ đó tạo cơ sở thực tiễn để hoàn thiện khung pháp lý trong tương lai. Đây cũng là hướng tiếp cận đã được nhiều quốc gia như Singapore (Anh Thư, 2019) áp dụng nhằm dung hòa giữa yêu cầu đổi mới công nghệ và nhiệm vụ bảo vệ công lý.

Ngoài ra, cần đặc biệt nhấn mạnh tầm quan trọng của việc phát triển năng lực công nghệ và nâng cao nhận thức pháp luật đối với đội ngũ cán bộ tư pháp. Khi đó, trên cơ sở vượt qua các kỳ thi, thẩm phán, kiểm sát viên, luật sư và điều tra viên không chỉ cần có kiến thức pháp luật mà còn phải nắm được nguyên lý vận hành của AI; liệu có thể tồn tại thiên lệch dữ liệu và cần biết các rủi ro của thuật toán. Việc đào tạo các kỹ năng đánh giá và kiểm chứng sự thật đối chiếu với các kết quả do AI tạo ra là vô cùng quan trọng để giảm thiểu sự phụ thuộc hoàn toàn vào công nghệ. Đồng thời, việc đó cũng tạo cơ sở để khẳng định nguyên tắc “*con người quyết định hay không xét xử, và bảo đảm các cơ quan tư pháp độc lập, khách quan*”.

Một giải pháp lớn khác là hoàn thiện đầy đủ cơ chế bảo vệ dữ liệu cá nhân trong quá trình tố tụng trên môi trường số. Tuy nhiên, Nghị định số 13/2023/NĐ-CP về bảo vệ dữ liệu cá nhân đã được ban hành nhưng các quy định hiện hành chưa bao quát đầy đủ tình huống này khi dữ liệu được xử lý thông qua các hệ thống AI trong lĩnh vực tư pháp. Vì vậy, cần bổ sung các quy định cụ thể hơn về phạm vi thu thập dữ liệu, mục đích sử dụng, thời hạn lưu trữ cũng như trách nhiệm bảo mật đối với thông tin của những người tham gia tố tụng (Trang thông tin điện tử phổ biến, giáo dục pháp luật Thành phố Hải Phòng, 2025). Đặc biệt, đối với các dữ liệu nhạy cảm như thông tin cá nhân, hình ảnh, bản ghi âm giọng nói hoặc hồ sơ vụ án... cần có một cơ chế kiểm soát nghiêm ngặt để hạn chế rò rỉ và việc khai thác trái pháp luật trong môi trường số.

Thứ hai, Việt Nam cần tiên phong và thí điểm các thủ tục trong hoạt động tư pháp với việc ứng dụng AI. Nhưng chỉ những nền tảng thực tiễn vững chắc và kết quả thử nghiệm mới có thể làm cơ sở để ban hành một đạo luật chuyên biệt. Đạo luật này nên được soạn thảo theo định hướng lấy con người làm trung tâm, quy định rõ rằng AI chỉ hỗ trợ nhưng không thay thế cho hoạt động xét xử của con người. Đồng thời, đạo luật cần điều chỉnh nhiệm vụ của các cơ quan tiến hành phiên tòa sử dụng AI, như các nghĩa vụ kiểm thử và xác minh thuật toán; và trao cho những người tham gia tố tụng quyền yêu cầu được giải thích trong trường hợp bị ảnh hưởng bởi kết quả do AI tạo ra.

Có thể thấy rằng, AI mang lại nhiều tiềm năng trong quá trình hiện đại hóa nền tư pháp, nhưng việc ứng dụng công nghệ này phải luôn gắn liền với yêu cầu bảo vệ quyền con người và duy trì bản chất công bằng của hoạt động xét xử. Vì vậy, Việt Nam cần xây dựng cơ chế pháp lý phù hợp, vừa tạo điều kiện thúc đẩy đổi mới công nghệ, vừa bảo đảm các giá trị cốt lõi của Nhà nước pháp quyền trong bối cảnh chuyển đổi số hiện nay.

3. Kết luận

Trong điều kiện chuyển đổi số mạnh mẽ, việc ứng dụng trí tuệ nhân tạo vào hoạt động tư pháp nói chung và xét xử nói riêng là xu hướng tất yếu nhằm nâng cao hiệu quả quản lý; hỗ trợ xử lý dữ liệu, số hóa, chuyển đổi số, hiện đại hóa hệ thống pháp luật. AI có thể hỗ trợ các cơ quan tiến hành tố tụng và tranh tụng trong nhiều hoạt động tư pháp như quản lý hồ sơ vụ án, phân tích thông tin từ các sự kiện pháp lý, tra cứu thông tin pháp lý và xây dựng hệ thống quản lý vụ án để xử lý hiệu quả các biện pháp khắc phục. Tuy nhiên, bên cạnh các ưu thế công nghệ, việc ứng dụng AI trong xét xử đặt ra nhiều thách thức liên quan đến quyền con người-đó là quyền được xét xử công bằng, quyền bào chữa và tư vấn pháp lý, quyền chống lại sự can thiệp trái pháp luật vào hoạt động xử lý dữ liệu cá nhân và quyền riêng tư đối với những người tham gia vào các thủ tục tố tụng tư pháp.

Phân tích trên, có thể thấy rằng hệ thống pháp luật hiện hành của Việt Nam, mặc dù đã có những quy định cơ bản về bảo vệ quyền con người bằng pháp luật trong hoạt động tư pháp, vẫn còn thiếu một cơ chế pháp lý chuyên biệt để điều chỉnh AI. Khoảng trống này có thể gây ra nhiều rủi ro như dữ liệu hoặc thuật toán không chính xác, thiếu minh bạch trong quá trình áp dụng thuật toán hoặc xác định trách nhiệm pháp lý khi xảy ra vi phạm (Phạm Liên, 2025). Do đó, Việt Nam cần sớm hoàn thiện khung pháp lý để kiểm soát và quản lý chặt chẽ việc sử dụng AI trong xét xử nhưng đồng thời bảo đảm nguyên tắc con người phải đóng vai trò quyết định cuối cùng trong hoạt động tư pháp. Nếu không có khung pháp lý phù hợp, AI không bao giờ có thể phát huy tiềm năng của mình như một công cụ hỗ trợ vô giá cho nền tư pháp hiện đại, đồng thời vẫn bảo đảm đầy đủ và bảo vệ hiệu quả các quyền con người trong Nhà nước pháp quyền.

Tài liệu tham khảo

- Anh Thư (2019). *Singapore thiết lập khung pháp lý đạo đức khi sử dụng AI*. <https://tuoitre.vn/singapore-thiet-lap-khung-phap-ly-dao-duc-khi-su-dung-ai-2019012508045414.htm>.
- Kuzhanthaivel, A. (2023). *Singapore launches generative AI evaluation sandbox*. ITNews Asia. <https://www.itnews.asia/news/singapore-launches-generative-ai-evaluation-sandbox-601943>.
- Ngô Anh Cường (2023). Sự trỗi dậy của trí tuệ nhân tạo: Ngành luật Việt Nam có cần phải đổi mới? *Tạp chí điện tử Pháp lý*. <https://phaply.net.vn/su-noi-day-cua-tri-tue-nhan-tao-ai-nganh-luat-viet-nam-co-can-phai-doi-moi-pho-a256672.html>.
- Nguyễn Thị Lan Anh & Lê Xuân Lục (2023). Quyền được xét xử công bằng nhìn từ góc độ pháp luật quốc tế và pháp luật tố tụng hình sự Việt Nam. *Tạp chí Tòa án nhân dân điện tử*. <https://tapchitoaan.vn/quyen-duoc-xet-xu-cong-bang-nhin-tu-goc-do-phap-luat-quoc-te-va-phap-luat-to-tung-hinh-su-viet-nam7945.html>.
- Nguyễn Trần Quế Trân (2025). Tiềm năng và thách thức sử dụng trí tuệ nhân tạo (AI) trong hoạt động đoán định tư pháp – Một số kiến nghị cho Việt Nam. *Tạp chí Tòa án nhân dân điện tử*. <https://tapchitoaan.vn/tiem-nang-va-thach-thuc-su-dung-tri-tue-nhan-tao-ai-trong-hoat-dong-don-dinh-tu-phap-mot-so-kien-nghi-cho-viet-nam14157.html>.
- Phạm Liên (2025). *Ứng dụng trí tuệ nhân tạo (AI) và một số vấn đề pháp lý đặt ra ở Việt Nam trong giai đoạn hiện nay*. <https://pbgdpl.haiphong.gov.vn/CCHC---Chuyen-doi-so/Chuyen-doi-so/Ung-dung-tri-tue-nhan-tao-AI-va-mot-so-van-de-phap-ly-dat-ra-o-Viet-Nam-trong-giai-doan-hien-nay-177083.html>, ngày 8/10/2025.