THỰC TRẠNG PHÁP LUẬT VỀ HỆ THỐNG TRÍ TUỆ NHÂN TẠO Ở VIỆT NAM VÀ MỘT SỐ KIẾN NGHỊ HOÀN THIỆN

PHẠM THỊ HỒNG NGHĨA*

Tóm tắt: Cùng với sự bùng nổ của cuộc Cách mạng công nghiệp 4.0, tình trạng các đối tượng sử dụng hệ thống trí tuệ nhân tạo (TTNT) để thực hiện hành vi vi phạm pháp luật như công nghệ Deepfake để lừa đảo[1,2], ChatGPT để tấn công mạng[3]… ngày càng phổ biến. Trước thực trạng trên, thiết lập khung pháp lý vững chắc về TTNT đóng vai trò quan trọng trong kiểm soát những rủi ro phát sinh trong quá trình thiết kế, cung ứng và sử dụng TTNT. Do đó, bài viết tập trung làm rõ những vấn đề lý luận và thực trạng, đưa ra một số đề xuất hoàn thiện pháp luật điều chỉnh hệ thống TTNT ở Việt Nam.

Từ khoá: Trí tuệ nhân tạo; hệ thống trí tuệ nhân tạo; thực trạng pháp luật; hoàn thiện pháp luật; kiểm soát rủi ro; quy định về trí tuệ nhân tạo

Ngày nhận bài: 10/5/2024; Biên tập xong: 25/9/2025; Duyệt đăng: 22/11/2025

THE EXISTING LEGAL FRAMEWORK FOR ARTIFICIAL INTELLIGENCE SYSTEMS IN VIETNAM AND PROPOSED IMPROVEMENTS

Abstract: With the Fourth Industrial Revolution, criminals are increasingly using artificial intelligence systems to apply violations against regulations, such as Deepfake technology for duping and ChatGPT for cyberattack... In consideration of the foregoing, a strong legal framework for artificial intelligence is critical to control risks that are created throughout the creation, supply, and use of artificial intelligence. As a result, the paper focuses on analyzing theoretical issues, current legal situations, and makes some recommendations for enhancing the law governing artificial intelligence systems in Vietnam.

Keywords: Artificial intelligence; artificial intelligence system; legal framework; enhancing the law; risk control; artificial intelligence legislation

Received: May 10, 2024; Editing completed: Sep 25, 2025; Accepted for publication: Nov 22, 2025

1. Khái quát chung về hệ thống và pháp luật về hệ thống trí tuệ nhân tạo

1.1. Khái niệm hệ thống trí tuệ nhân tạo

TTNT (Artificial Intelligence - AI) được hiểu đơn giản là nghiên cứu và phát triển chương trình máy tính có thể sao chép hành vi trí tuệ của con người[4]. Dưới góc độ pháp lý, Liên minh châu Âu (EU) đưa ra định nghĩa hệ thống TTNT (AI system) là một hệ thống dựa trên máy tính được thiết kế để hoạt động với các mức độ tự động hóa khác nhau và có thể thích nghi sau khi triển khai, với các mục tiêu rõ ràng hoặc phải suy luận từ đầu vào mà nó nhận được để tạo ra các kết quả đầu ra như dự báo, nội dung, đề xuất hoặc quyết định có thể ảnh hưởng đến môi trường vật lý hoặc ảo[5]. Còn theo pháp luật Việt Nam, TTNT được hiểu là việc thực hiện bằng điện tử các năng lực trí tuệ của con người, bao gồm học tập, suy luận, nhận thức, phán đoán và hiểu ngôn ngữ tự nhiên; còn hệ thống TTNT là hệ thống dựa trên máy, được thiết kế để thực hiện các năng lực TTNT với các mức độ tự chủ khác nhau, có khả năng tự thích nghi sau khi được triển khai; căn cứ vào các mục tiêu được xác định rõ ràng hoặc được hình thành một cách ngầm định, hệ thống suy luận từ dữ liệu đầu vào để tạo ra các đầu ra như dự đoán, nội dung, khuyến nghị hoặc quyết định có thể

[1] Hùng Quân (2024), *Lừa đảo bằng công nghệ Deepfake lại bùng phát*, https://cand.com.vn/khoa-hoc-quan-su/lua-dao-bang-cong-nghe-deepfake-lai-bung-phat-i722392/, truy cập ngày 20/4/2024.

[2] Thái Sơn (2024), *Cảnh giác cao với tội phạm lừa đảo sử dụng TTNT*, https://nhandan.vn/canh-giac-cao-voi-toi-pham-lua-dao-su-dung-tri-tue-nhan-tao-post746314.html, truy cập ngày 20/4/2024.

[3] Tiểu Minh (2023), *ChatGPT được sử dụng để viết phần mềm độc hại*, https://plo.vn/ky-nguyen-so/chatgpt-duoc-su-dung-de-viet-phan-mem-doc-hai-post717954.html, truy cập ngày 20/4/2024.

[4] Từ điển Oxford: Artificial intelligence is the study and development of computer systems that can copy intelligent human behavior, https://www.oxfordlearnersdictionaries.com/definition/english/artificial-intelligence?q=artificial+intelligence, truy cập ngày 01/4/2024.

* *Email: Hongnghiapt87@gmail.com*

Tiến sĩ, Khoa luật, Học viện An ninh nhân dân

[5] Khoản 1 Điều 3 Đạo luật TTNT của EU năm 2024.

gây ảnh hưởng tới môi trường vật lý hoặc môi trường ảo. Các quy định trên cho thấy, mặc dù có sự khác biệt về diễn đạt nhưng cách định nghĩa của pháp luật Việt Nam về TTNT và hệ thống TTNT với cách giải thích tương đối tương đồng với quy định của EU[6]. Hiện nay, TTNT đã phát triển sang thế hệ mới với một số đặc trưng cơ bản như sau: (1) Tính tự động. Hệ thống TTNT ngày càng phát triển khả năng tự hành động, tự đưa ra quyết định mà không cần sự can thiệp, thậm chí cả sự hiểu biết của con người; (2) Khả năng học. TTNT có khả năng tự thu nạp, mở mang kiến thức thông qua dữ liệu và kinh nghiệm, đặc biệt, còn có khả năng học sâu thông qua dữ liệu lớn; (3) Tính không thể đoán trước. Do tính tự động và khả năng học, TTNT có thể tạo ra những sản phẩm con người khó có thể đoán trước được[7].

1.2. Khái quát chung về pháp luật về hệ thống trí tuệ nhân tạo

Với sự phát triển nhanh chóng của TTNT, việc thiết lập khung pháp lý cho lĩnh vực này là cần thiết để bảo đảm sự ổn định và phát triển xã hội, đồng thời hỗ trợ cho hoạt động quản lý của Nhà nước được tiến hành hiệu quả hơn. Tuy nhiên, TTNT cũng có thể gây ra những tác động tiêu cực, như việc sử dụng TTNT để vi phạm pháp luật và xâm phạm quyền lợi hợp pháp của tổ chức, cá nhân như quyền riêng tư, quyền đối với dữ liệu cá nhân, đặc biệt là các dữ liệu pháp lý nhạy cảm[8]. Do đó, pháp luật điều chỉnh hệ thống TTNT phải được thiết lập theo hướng hạn chế và loại bỏ các rủi ro mà hệ thống TTNT có thể gây ra cho toàn xã hội cũng như cho mỗi tổ chức, cá nhân trong xã hội[9].

Pháp luật về TTNT là hệ thống các nguyên tắc pháp lý và văn bản pháp luật được nhà nước ban hành hoặc thừa nhận nhằm điều chỉnh các quan hệ xã hội giữa các chủ thể trong quản lý nhà nước cũng như nghiên cứu, phát triển, cung cấp, triển khai và sử dụng hệ thống TTNT[10]. Pháp luật về hệ thống TTNT có một số đặc điểm gồm: (1) Trách nhiệm hạn chế rủi ro do hệ thống TTNT đầu tiên thuộc về bản thân mỗi tổ chức, cá nhân trong xã hội[11]. Theo đó, các tổ chức và cá nhân cần tự nguyện áp dụng tất cả các biện pháp có thể để giảm thiểu, loại bỏ các rủi ro mà hệ thống TTNT có thể gây ra. Pháp luật điều chỉnh TTNT cần quy định nguyên tắc trách nhiệm hạn chế rủi ro trong suốt quá trình thiết kế, triển khai và sử dụng hệ thống TTNT, đồng thời xác định trách nhiệm của Nhà nước trong việc hỗ trợ pháp lý và kỹ thuật, giúp các tổ chức, cá nhân thực hiện tốt trách nhiệm này; (2) Pháp luật điều chỉnh hệ thống TTNT hướng đến mục tiêu hạn chế và loại bỏ rủi ro trong quá trình thiết kế, cung ứng và sử dụng hệ thống TTNT song song với việc phát triển hệ thống TTNT[12]. Để thực hiện được mục tiêu này, pháp luật điều chỉnh hệ thống TTNT cần phân loại rõ các loại rủi ro phát sinh trong quá trình thiết kế, cung ứng và sử dụng hệ thống TTNT, từ đó đưa ra các biện pháp nhằm hạn chế, loại bỏ (như cấm) các loại rủi ro này.

Do đặc thù của hệ thống TTNT cũng như pháp luật về nó, hệ thống các văn bản pháp luật điều chỉnh hệ thống TTNT gồm nhóm các quy định sau: (1) Nhóm quy định điều chỉnh quan hệ giữa các cơ quan nhà nước trong hoạt động phân công, phối hợp để quản lý hệ thống TTNT; (2) Nhóm quy định điều chỉnh quan hệ giữa các cơ quan nhà nước có thẩm quyền với các chủ thể thiết kế, cung ứng và sử dụng hệ thống TTNT; (3) Nhóm quy định điều chỉnh quan hệ phát sinh giữa các chủ thể trực tiếp thiết kế, cung ứng và sử dụng hệ thống TTNT.

Đối chiếu các nhóm quy định trên, Luật TTNT năm 2025 quy định ba nhóm quan hệ chính: (1) Quan hệ giữa các cơ quan quản lý nhà nước, với quyền quản lý được giao cho Chính phủ, Bộ Khoa học và Công nghệ và các bộ khác như Bộ Công an, Bộ Quốc phòng…; (2) Quan hệ giữa cơ quan quản lý và các chủ thể thiết kế, cung cấp, sử dụng TTNT, trong đó việc quản lý được thực hiện theo mức độ rủi ro, đồng thời, pháp luật xác định nghĩa vụ cụ thể cho doanh nghiệp và trách nhiệm của các chủ thể (Chương II, V); cùng với đó

[6] Khoản 1, khoản 2 Điều 3 Luật TTNT năm 2025.

[7] Nicholas Berente, Bin Gu, Jan Recker, Radhika Santhanam, "Managing artificial intelligence", *MIS Quarterly*, Vol. 45, No. 3/2021, p. 1437, https://www.researchgate.net/publication/352400557_Managing_Artificial_Intelligence, truy cập ngày 05/4/2024.

[8] Joel Paul (2024), *Privacy and data security concerns in AI*, https://www.researchgate.net/publication/385781993_Privacy_and_data_security_concerns_in_AI, truy cập ngày 25/4/2025.

[9] European Parliament, *Shaping Europe's digital future*, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai, truy cập ngày 25/4/2024.

[10] Khoản 1 Điều 1 Luật TTNT năm 2025.

[11] Khoản 2 Điều 4 Luật TTNT năm 2025.

[12] Khoản 2, khoản 4 Điều 4 Luật TTNT năm 2025.

còn có các hoạt động hỗ trợ của Nhà nước (Chương III, IV); (3) Quan hệ giữa các chủ thể thiết kế, cung ứng và sử dụng hệ thống TTNT, với quy định về bồi thường thiệt hại phát sinh từ hoạt động TTNT (Điều 29).

2. Đánh giá thực trạng pháp luật về hệ thống trí tuệ nhân tạo ở Việt Nam

2.1. Những ưu điểm của pháp luật về hệ thống trí tuệ nhân tạo ở Việt Nam

Thứ nhất, Việt Nam đã ban hành văn bản pháp luật điều chỉnh trực tiếp hệ thống TTNT

Để tạo khung pháp lý điều chỉnh TTNT, Việt Nam đã ban hành những văn bản chính sách tạo cơ sở nền tảng để xây dựng hệ thống các văn bản pháp luật điều chỉnh TTNT. Có thể kể đến như: Quyết định số 127/QĐ-TTg ngày 26/01/2021 của Thủ tướng Chính phủ ban hành Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng TTNT đến năm 2030; Quyết định số 552/QĐ-BTP ngày 12/4/2021 của Bộ Tư pháp ban hành Kế hoạch thực hiện Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng TTNT đến năm 2030 và Chiến lược quốc gia về Cách mạng công nghiệp 4.0 đến năm 2030 của Bộ Tư pháp; Quyết định số 44/QĐ-BTC ngày 13/01/2022 của Bộ Tài chính ban hành kế hoạch hành động của Bộ Tài chính triển khai thực hiện Quyết định số 127/QĐ-TTG ngày 26/01/2021 của Thủ tướng Chính phủ về việc ban hành "Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng TTNT đến năm 2030"; Quyết định số 575/QĐ-UBND ngày 23/02/2021 của Ủy ban nhân dân (UBND) Thành phố Hồ Chí Minh phê duyệt Chương trình "Nghiên cứu và phát triển ứng dụng TTNT (AI) tại Thành phố Hồ Chí Minh giai đoạn 2020-2030";… Cùng với đó, sự ra đời của Quyết định số 3298/QĐ-BKHCN ngày 29/12/2023 của Bộ Khoa học Công nghệ về việc công bố tiêu chuẩn quốc gia (TCVN), trong đó công bố 02 TCVN về TTNT gồm: TCVN 13902:2023 ISO/IEC 22989:2022 Công nghệ thông tin - TTNT - Các khái niệm và thuật ngữ TTNT (TCVN số 02); và TCVN 13903:2023 ISO/IEC TR 24028:2020 Công nghệ thông tin - TTNT - Tổng quan về tính đáng tin cậy trong TTNT (TCVN số 03). Ngoài các văn bản trực tiếp, Việt Nam còn có những văn bản liên quan khác điều chỉnh hệ thống TTNT như Chỉ thị số 16/CT-TTg ngày 04/5/2017 của Thủ tướng Chính phủ về tăng cường năng lực tiếp cận cuộc cách mạng công nghiệp 4.0. Đặc biệt, Luật TTNT năm 2025 ra đời tập trung vào quản lý các rủi ro do việc ứng dụng hệ thống TTNT (theo 04 cấp độ rủi ro) như cách quy định của EU. Các văn bản pháp luật trên đã tạo nền tảng đầu tiên cho việc thiết lập khung pháp lý điều chỉnh, quản lý hệ thống TTNT ở Việt Nam.

Những văn bản pháp luật nhằm tạo lập các cơ quan hỗ trợ cho sự phát triển của hệ thống TTNT như Quyết định số 168/QĐ-UBND ngày 20/01/2024 của UBND Thành phố Đà Nẵng về việc quy định chức năng, nhiệm vụ và cơ cấu tổ chức đối với Trung tâm Nghiên cứu đào tạo thiết kế vi mạch và TTNT Đà Nẵng; Thủ tướng Chính phủ ban hành Quyết định số 1269/QĐ-TTg ngày 02/10/2019 về việc thành lập Trung tâm đổi mới sáng tạo Quốc gia với vị trí là đơn vị sự nghiệp công lập trong lĩnh vực sự nghiệp kinh tế tự bảo đảm chi thường xuyên và chi đầu tư, trực thuộc Bộ Tài chính nhằm thực hiện hỗ trợ, phát triển hệ sinh thái khởi nghiệp, đổi mới sáng tạo, góp phần đổi mới mô hình tăng trưởng trên nền tảng phát triển khoa học và công nghệ, trong đó có hệ thống TTNT. Sự ra đời của các đơn vị hỗ trợ trên đã tạo nền tảng về kỹ thuật và nguồn vốn cho việc phát triển hệ thống TTNT ở Việt Nam.

Thứ hai, nội dung các TCVN về TTNT đã phần nào tiếp cận tới pháp luật quốc tế

Nội dung tại TCVN số 02 và TCVN số 03, Luật TTNT năm 2025 đã đưa ra khá nhiều cách hiểu các thuật ngữ về hệ thống TTNT tương thích với Đạo luật TTNT của EU. Các thuật ngữ cơ bản, quan trọng nhất liên quan đến hệ thống TTNT đều được liệt kê trong Tiêu chuẩn với cách hiểu như những quy định tại Đạo luật TTNT của EU. Ví dụ: Thuật ngữ "hệ thống TTNT"[13] tương ứng với thuật ngữ "AI system"[14]; thuật ngữ "rủi ro"[15] và thuật ngữ "risk"[16]; thuật ngữ "dữ liệu huấn luyện"[17] với thuật ngữ "training data"[18]… Sự tương đối thống nhất trong cách hiểu các thuật ngữ liên quan đến hệ thống TTNT đã tạo tiền đề thuận lợi trong xây dựng khung pháp lý trực tiếp quản lý, cũng như quá trình quản lý của Nhà nước đối với hệ thống TTNT ở Việt Nam.

[13] Điểm 3.14 TCVN số 02.
[14] Khoản 1 Điều 3 Đạo luật TTNT của EU năm 2024.
[15] Điểm 3.31 TCVN số 03.
[16] Khoản 1(a) Điều 3 Đạo luật TTNT của EU năm 2024.
[17] Điểm 5.11.6 TCVN số 02.
[18] Khoản 29 Điều 3 Đạo luật TTNT của EU năm 2024.

TCVN số 03 đã tiếp cận nguyên tắc bảo đảm tính đáng tin cậy của hệ thống TTNT theo Đạo luật TTNT của EU năm 2024 (khoản 14 (a) lý do đưa ra Đạo luật). Cụ thể, các lớp tin cậy được xác định gồm: (1) Lớp tin cậy vật lý được đo đạc hoặc kiểm tra thông qua độ tin cậy và chế độ an toàn kỹ thuật, tương ứng với nguyên tắc bảo đảm độ bền và an toàn kỹ thuật của TTNT trong Đạo luật của EU; (2) Lớp tin cậy về không gian mạng dựa trên các yêu cầu bảo mật cơ sở hạ tầng thông tin như kiểm soát truy cập và các biện pháp bảo vệ tính toàn vẹn của hệ thống và dữ liệu, tương ứng với nguyên tắc quyền riêng tư và quản trị dữ liệu trong Đạo luật TTNT của EU; (3) Lớp tin cậy về xã hội dựa trên các yếu tố như tính cách, niềm tin của mỗi cá nhân trong xã hội, bảo đảm tính minh bạch và công bằng, tương ứng với các nguyên tắc về minh bạch, đa dạng, không phân biệt đối xử và phúc lợi xã hội trong pháp luật của EU[19]. Việc bảo đảm tính đáng tin cậy của hệ thống TTNT là rất quan trọng để bảo vệ an toàn cộng đồng và quyền lợi cá nhân, trong đó, các quy định trong Đạo luật TTNT của EU là tiêu chuẩn chặt chẽ và rõ ràng nhất. Vì vậy, việc áp dụng các nguyên tắc của pháp luật EU trong tiêu chuẩn đánh giá tính tin cậy của hệ thống TTNT ở Việt Nam thể hiện ưu điểm nổi bật trong việc điều chỉnh lĩnh vực này.

Thứ ba, Luật TTNT năm 2025 được xây dựng theo hướng tiếp thu những ưu điểm trong pháp luật về TTNT của các quốc gia phát triển trên thế giới là EU, Trung Quốc, Hoa Kỳ

Trước nguy cơ lạm dụng TTNT gây phương hại đến trật tự pháp lý và quyền, lợi ích hợp pháp của các chủ thể trong xã hội, đồng thời nhằm định hướng phát triển TTNT phục vụ đời sống, nhiều quốc gia đã ban hành các khung pháp lý chung để điều chỉnh lĩnh vực này. Có thể kể đến như: (1) Đạo luật TTNT năm 2024 của EU, hệ thống TTNT sẽ được quản lý ở các mức độ khác nhau trên cơ sở phân loại theo mức độ ảnh hưởng tiêu cực đến người dùng, gồm 04 cấp: Hệ thống bị cấm do rủi ro nghiêm trọng, hệ thống có rủi ro cao cần kiểm soát chặt chẽ, hệ thống có rủi ro trung bình yêu cầu minh bạch, và hệ thống rủi ro thấp. Đồng thời, EU cũng đưa ra cơ sở pháp lý cho việc thành lập cơ quan giám sát, thực thi các chế tài đối với các chủ thể vi phạm quy định về quản lý trong Đạo luật[20]; (2) Trung Quốc đã ban hành nhiều văn bản pháp luật để điều chỉnh TTNT, bao gồm: "Kế hoạch phát triển TTNT thế hệ mới" (2017), "Bộ nguyên tắc quản trị TTNT thế hệ mới" (2019), "Chuẩn mực đạo đức cho TTNT thế hệ mới" (2021), và "Dự thảo kiến nghị cho luật TTNT" (2024). Các quy định này cho thấy Trung Quốc áp dụng cơ chế điều chỉnh linh hoạt, cân bằng giữa đổi mới và an toàn, phòng ngừa rủi ro và khuyến khích phát triển. Mặc dù không áp đặt yêu cầu bắt buộc hay chế tài trong các văn bản như Kế hoạch và Bộ nguyên tắc, Dự thảo Luật TTNT năm 2024 của Trung Quốc dựa trên 04 nguyên tắc: Lấy con người làm trung tâm, tuân thủ pháp luật, công khai minh bạch, và bảo vệ quyền sở hữu trí tuệ và bảo mật thông tin người dùng[21]; (3) Ở Hoa Kỳ, Sắc lệnh Hành pháp ngày 23/01/2025 của Tổng thống Hoa Kỳ Donald J. Trump được ban hành để hủy bỏ Sắc lệnh Hành pháp của Joe Biden (được cho là cản trở sự đổi mới của TTNT và áp đặt các quy định kiểm soát chặt chẽ không cần thiết từ Chính phủ đối với sự phát triển của TTNT). Mục tiêu Sắc lệnh Hành pháp ngày 23/01/2025 là duy trì và củng cố vị thế dẫn đầu toàn cầu về TTNT, nhằm thúc đẩy sự phát triển của con người, nâng cao năng lực cạnh tranh kinh tế và bảo đảm an ninh quốc gia[22]. Như vậy, cách thức pháp luật của các quốc gia điều chỉnh hệ thống TTNT được thực hiện theo hai hướng: *Một là*, ghi nhận rủi ro phát sinh từ việc ứng dụng hệ thống TTNT và kiểm soát chặt chẽ các rủi ro theo cách quy định của châu Âu; *Hai là*, cân bằng giữa phát triển hệ thống TTNT và kiểm soát rủi ro như pháp luật của Trung Quốc và Hoa Kỳ.

So với Luật TTNT năm 2025 của Việt Nam, có thể thấy, mặc dù hướng tới việc cân bằng

[19] Khoản 14, khoản 14bb Phần mở đầu Đạo luật TTNT năm 2024 của EU.

[20] Đạo luật TTNT năm 2024 của EU,

[21] Nguyễn Ngọc Phương Hồng, Lưu Minh Sang (2024), *Luật về TTNT và tầm nhìn chiến lược của Trung Quốc*, https://thesaigontimes.vn/luat-ve-tri-tue-nhan-tao-va-tam-nhin-chien-luoc-cua-trung-quoc/, truy cập ngày 12/02/2025.

[22] Whitehouse, *Removing barriers to American leadership in artificial intelligence*, https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/, truy cập ngày 12/02/2025.

giữa phát triển và kiểm soát rủi ro như Trung Quốc và Hoa Kỳ, nhưng không áp dụng các biện pháp kiểm soát chặt chẽ như EU. Tại khoản 3 Điều 4 Luật TTNT năm 2025, nguyên tắc quản lý tương xứng cấp độ rủi ro và chỉ quản lý bắt buộc đối với các hệ thống có nguy cơ gây hại rõ ràng, thể hiện sự linh hoạt hơn trong việc quản lý rủi ro. Đồng thời, các Chương III và IV của Luật TTNT năm 2025 xây dựng khung pháp lý ưu đãi và khuyến khích thử nghiệm có kiểm soát, thể hiện quan điểm thúc đẩy phát triển AI; Chương II của Luật đã tiếp thu nhiều quy định của EU về phân loại và kiểm soát rủi ro. Như vậy, Việt Nam đã kết hợp các yếu tố từ pháp luật EU, Hoa Kỳ và Trung Quốc để vừa quản lý rủi ro, vừa thúc đẩy sự phát triển của hệ thống TTNT.

2.2. Những tồn tại của pháp luật về hệ thống trí tuệ nhân tạo ở Việt Nam

Thứ nhất, Việt Nam chưa có cơ quan nhà nước trực tiếp kiểm soát các rủi ro phát sinh trong quá trình nghiên cứu, phát triển, cung cấp, triển khai và sử dụng hệ thống TTNT

Như đã trình bày ở tiểu mục 1, các quy định điều chỉnh hệ thống TTNT sẽ liên quan đến công nghệ thông tin và bảo đảm an toàn trong hoạt động thiết kế, cung ứng và sử dụng hệ thống TTNT. Theo pháp luật hiện hành, chủ thể có thẩm quyền quản lý trực tiếp về công nghệ thông tin là Chính phủ và là Bộ Bưu chính, Viễn thông[23] (Bộ Thông tin và Truyền thông); còn chủ thể chịu trách nhiệm về an ninh mạng là Chính phủ, Bộ Công an[24]; Bộ Quốc phòng; Ban Cơ yếu Chính phủ; các Bộ, cơ quan ngang Bộ, cơ quan thuộc Chính phủ; UBND cấp tỉnh[25], đây đều là những cơ quan kiêm nhiệm chứ không phải cơ quan chuyên trách. Còn Điều 6 Dự thảo Luật TTNT lần hai ngày 21/11/2025 đề xuất thành lập Ủy ban Quốc gia về TTNT do Thủ tướng Chính phủ làm Chủ tịch để chỉ đạo, điều phối các hoạt động mang tính liên ngành và chiến lược về TTNT. Tuy nhiên, đến dự thảo lần 3 (ngày 10/12/2025) - Dự thảo được thông qua thì bỏ đi quy định về cơ quan này, trên thực tế Ủy ban Quốc gia về TTNT cũng chưa được thành lập gây khó khăn nhất định cho công tác quản lý, phát triển ứng dụng hệ thống TTNT ở Việt Nam.

Thứ hai, quy định về kiểm soát rủi ro phát sinh trong quá trình nghiên cứu, phát triển, cung cấp, triển khai và sử dụng hệ thống TTNT còn tồn tại hạn chế

Thực tế cho thấy, rủi ro phát sinh trong quá trình nghiên cứu, phát triển, cung cấp, triển khai và sử dụng hệ thống TTNT là rất đa dạng, do đó kiểm soát rủi ro phát sinh đóng vai trò quan trọng trong phát triển bền vững hệ thống TTNT cũng như bảo vệ quyền và lợi ích của các tổ chức, cá nhân trong xã hội. Hiện có nhiều đối tượng sử dụng hệ thống TTNT để thực hiện các hành vi vi phạm pháp luật, gây mất ổn định về trật tự xã hội, ảnh hưởng đến an ninh quốc gia cũng như gây hoang mang, lo lắng cho các tổ chức, cá nhân trong xã hội. Có thể kể đến như những vụ việc các đối tượng sử dụng công nghệ Deepfake để tạo ra các đoạn video với hình ảnh, khuôn mặt nhân vật giống hệt như hình ảnh của người dùng muốn giả mạo để thực hiện hành vi lừa đảo tài chính và gây ra những thiệt hại về vật chất và tinh thần cho các nạn nhân trong phạm vi cả nước[26,27]. Bên cạnh đó, các chuyên gia an ninh mạng tại Check Point Research đã phát hiện tin tặc đang sử dụng chatbot ChatGPT để viết phần mềm độc hại và cải thiện hoạt động tấn công, từ đó càng dẫn đến nguy cơ các tổ chức, cá nhân trong xã hội bị lừa đảo, bị tấn công mạng ngày càng cao hơn[28]. Theo Kaspersky thì hệ thống TTNT có thể bị tội phạm mạng sử dụng trong từng giai đoạn của một cuộc tấn công tinh vi hay còn được gọi là APT (Advanced Persistent Threat). APT sử dụng các kỹ thuật hack liên tục, bí mật và tinh vi để giành được quyền truy cập vào hệ thống và tồn tại trong đó một thời gian dài với những hậu quả rất nghiêm trọng đe doạ đến an ninh quốc gia, trật tự xã hội cũng như quyền và lợi ích hợp pháp của các tổ chức, cá nhân trong xã hội[29].

Ở Việt Nam hiện nay, quy định trực tiếp nhằm kiểm soát rủi ro phát sinh trong quá

[23] Khoản 1, khoản 2 Điều 7 Luật Công nghệ thông tin năm 2006.

[24] Khoản 2 Điều 27 Luật An ninh mạng năm 2025.

[25] Khoản 3, khoản 4, khoản 5, khoản 6 Điều 39 Luật An ninh mạng năm 2025.

[26] Hùng Quân, Tlđd.

[27] Thái Sơn, Tlđd.

[28] Tiểu Minh, Tlđd.

[29] Đào Công (2023), *Tội phạm mạng đã sử dụng AI cho các cuộc tấn công APT*, https://dientuungdung.vn/toi-pham-mang-da-su-dung-ai-cho-cac-cuoc-tan-cong-apt, truy cập ngày 02/4/2023.

trình nghiên cứu, phát triển, cung cấp, triển khai và sử dụng hệ thống TTNT mới chỉ được ghi nhận trong Luật TTNT năm 2025. Tuy nhiên, các quy định trong Luật vẫn còn tồn tại một số hạn chế so với pháp luật của EU làm giảm hiệu quả hoạt động kiểm soát rủi ro ở Việt Nam. Với tính chất là một văn bản điều chỉnh chung về TTNT nên Luật TTNT mới quy định chung về phân loại rủi ro mà không liệt kê cụ thể các lĩnh vực phải đăng ký và kiểm soát nghiêm ngặt (hạ tầng trọng yếu, giáo dục, việc làm, dịch vụ công, thực thi pháp luật, biên giới…) giống như pháp luật EU[30]. Bên cạnh đó, Luật TTNT năm 2025 cũng không quy định cụ thể về các thủ tục trong việc phân loại rủi ro, trách nhiệm trong kiểm soát rủi ro giống như pháp luật EU đặt ra đối với các nhà cung cấp. Bởi theo pháp luật EU nhà cung cấp có rủi ro cao phải có Hệ thống quản lý chất lượng (Quality Management System - QMS), hồ sơ kỹ thuật, đánh giá hợp chuẩn, giám sát trong toàn vòng đời hệ thống, báo cáo sự cố và cập nhật, và các tổ chức đánh giá độc lập có thể thực hiện kiểm tra[31]. Ngoài ra, Luật TTNT năm 2025 không xác định cụ thể mức phạt đối với các hành vi vi phạm như quy định của EU mà chỉ quy định mang tính nguyên tắc về xử lý hành vi vi phạm (Điều 29) và chờ Chính phủ quy định. Cụ thể, Đạo luật TTNT năm 2024 của EU đã xác định rõ các hành vi vi phạm (Điều 5) và đưa ra mức phạt đối với các hành vi vi phạm mang tính răn đe khá cao (từ 7.500.000 EUR hoặc 1% tổng doanh thu toàn cầu năm trước đến 35.000.000 EUR hoặc 7% tổng doanh thu trên phạm vi toàn thế giới trong năm tài chính trước).[32] Thực trạng không quy định mức phạt cụ thể của Việt Nam làm giảm hiệu lực răn đe của pháp luật, do doanh nghiệp và chủ thể công nghệ sẽ khó khăn trong đánh giá chi phí tuân thủ, nên phát sinh nguy cơ tăng hành vi vi phạm vì chi phí vi phạm thấp hoặc không rõ ràng, hoặc không lựa chọn thị trường Việt Nam.

Thực trạng hiện nay đặt ra thách thức lớn cho cơ quan nhà nước có thẩm quyền ở Việt Nam trong việc giám sát và kiểm tra hệ thống TTNT xuyên suốt cả quá trình từ nghiên cứu, phát triển, triển khai cho đến khi sản phẩm được đưa vào sử dụng rộng rãi. Đặc biệt, nếu Chính phủ không kịp thời ban hành Nghị định hướng dẫn thi hành Luật TTNT thì các tổ chức và cá nhân thực hiện pháp luật sẽ gặp khó khăn trong việc tuân thủ các quy định. Thêm vào đó, việc văn bản luật chưa ghi nhận đầy đủ các yếu tố quan trọng trong quản lý hệ thống AI, bao gồm các quy định cụ thể về trách nhiệm của các chủ thể thiết kế, cung cấp hệ thống và các hành vi vi phạm, mức phạt, có thể dẫn đến nguy cơ "lobby chính sách" từ các nhóm lợi ích trong Ngành. Điều này tạo ra khoảng trống pháp lý, làm giảm tính công khai, minh bạch trong quá trình quản lý và kiểm soát TTNT. Từ đó, hiệu quả kiểm soát rủi ro trong nghiên cứu, phát triển, cung cấp, triển khai và sử dụng hệ thống TTNT bị suy giảm, gây ra nguy cơ đe dọa quyền lợi của tổ chức, cá nhân và ảnh hưởng tiêu cực đến sự phát triển bền vững của đất nước.

3. Một số khuyến nghị cho Việt Nam trong xây dựng pháp luật về hệ thống trí tuệ nhân tạo

Thứ nhất, thành lập cơ quan chuyên trách có nhiệm vụ kiểm soát các rủi ro phát sinh trong quá trình thiết kế, cung ứng và sử dụng hệ thống TTNT

Như đã trình bày, hiện nhiệm vụ quản lý hệ thống TTNT thuộc về rất nhiều cơ quan có thẩm quyền chung nhưng chưa có cơ quan chuyên trách, trong khi Ủy ban Quốc gia về TTNT vẫn đang ở bước đề xuất thành lập. Điều này không chỉ hạn chế hoạt động phát triển hệ thống TTNT nhằm phục vụ đời sống mà còn gia tăng những rủi ro phát sinh trong quá trình thiết kế, cung ứng và sử dụng hệ thống TTNT do hiệu quả của hoạt động kiểm soát rủi ro không cao. Ở Hoa Kỳ - một trong các quốc gia phát triển hệ thống TTNT bậc nhất trên thế giới đã thành lập cơ quan riêng nhằm phát triển và bảo đảm an toàn của hệ thống TTNT. Từ ngày 01/11/2023, Bộ Thương mại Mỹ đã thông báo về việc thành lập Viện An toàn TTNT Mỹ (USAISI) với nhiệm vụ chính là đánh giá mức độ an toàn của các mô hình TTNT; phát triển các tiêu chuẩn liên quan để nhận biết và xác thực nội dung do hệ thống TTNT tạo ra; cung cấp nền tảng thử nghiệm cho các nhà nghiên cứu khám phá và thăm dò rủi ro của các mô hình TTNT[33]. Do đó, trên cơ

[30] Điều 6 Đạo luật TTNT năm 2024 của EU.

[31] Điều 11, Điều 16, Điều 17, Điều 43 Đạo luật TTNT năm 2024 của EU.

[32] Điều 71 Đạo luật TTNT năm 2024 của EU.

[33] H.Hà (2023), *Mỹ thành lập Viện An toàn TTNT*, https://dangcongsan.vn/the-gioi/tin-tuc/my-thanh-lap-vien-an-

sở nhận thức rõ những thành tựu cũng như những nguy cơ rủi ro cao mà hệ thống TTNT có thể tác động đến xã hội, Việt Nam cũng cần xem xét thành lập Ủy ban Quốc gia về TTNT, để quản lý và tạo động lực cho việc ứng dụng TTNT vào các lĩnh vực của đời sống.

Thứ hai, Việt Nam cần khắc phục hạn chế trong quy định về kiểm soát rủi ro của quá trình nghiên cứu, phát triển, cung cấp, triển khai và sử dụng hệ thống TTNT hiện hành

Để khắc phục hạn chế trong quy định về kiểm soát rủi ro của quá trình nghiên cứu, phát triển, cung cấp, triển khai và sử dụng hệ thống TTNT hiện hành, Việt Nam có thể học tập các quy định trong Đạo luật TTNT của EU năm 2024. Theo đó, Việt Nam nên xem xét để sớm ban hành Nghị định hướng dẫn Luật TTNT. Trong đó, Nghị định cần bổ sung quy định về các nội dung sau: (1) Danh mục cụ thể các lĩnh vực và loại hệ thống TTNT thuộc nhóm "rủi ro cao" (như hạ tầng trọng yếu, y tế, giáo dục, tài chính, an ninh mạng) cùng với tiêu chí định lượng rõ ràng và thủ tục phân loại bắt buộc ngay từ giai đoạn thiết kế thử nghiệm; (2) Trách nhiệm của nhà phát triển, nhà cung cấp hệ thống TTNT rủi ro cao trong: Thiết lập hệ thống quản lý chất lượng (QMS); soạn thảo hồ sơ kỹ thuật đầy đủ; thực hiện đánh giá phù hợp (conformity assessment) hoặc kiểm định độc lập; gắn dấu chứng nhận hợp chuẩn nếu cần; và chịu trách nhiệm về giám sát vòng đời hệ thống, cập nhật và báo cáo sự cố; (3) Cơ chế giám sát sau khi hệ thống TTNT đi vào sử dụng gồm: Lưu trữ nhật ký vận hành, báo cáo định kỳ, quyền yêu cầu đánh giá lại, thu hồi hoặc tạm dừng nếu có biến động lớn; thiết lập quy định xử phạt và trách nhiệm rõ ràng khi xảy ra sự cố từ thông báo tới nạn nhân, tới phối hợp điều tra, khắc phục hậu quả.

Ngoài ra, cơ quan nhà nước có thẩm quyền nên xem xét để ban hành văn bản dưới luật quy định về các hành vi vi phạm và các chế tài tương ứng với mức độ của hành vi vi phạm liên quan đến TTNT. Các chế tài có thể kể đến như phạt hành chính, thu hồi giấy phép hoặc phạt theo doanh thu hoặc giống như cách quy định của pháp luật EU thay vì việc đưa ra các mức phạt tiền cụ thể thường thấy trong các văn bản xử lý vi phạm hành chính hiện nay. Từ đó, giúp tăng tính tuân thủ, giảm việc lợi dụng khoảng trống pháp lý và nâng cao uy tín của Việt Nam trong quản lý TTNT công bằng, an toàn.

Có thể thấy, mặc dù Việt Nam đã ban hành Luật TTNT năm 2025, Việt Nam chưa có cơ quan quản lý chuyên trách với TTNT, đồng thời, Luật TTNT hiện còn nhiều quy định chưa cụ thể. Do đó, trong thời gian tới, Việt Nam cần sớm khắc phục những hạn chế để bảo vệ quyền và lợi ích hợp pháp của các tổ chức, cá nhân trong xã hội, thúc đẩy việc phát triển, ứng dụng hệ thống TTNT phục vụ đời sống./.

toan-tri-tue-nhan-tao-650994.html, truy cập ngày 05/5/2024.

TÀI LIỆU THAM KHẢO

1. Quyết định số 3298/QĐ-BKHCN ngày 29/12/2023 của Bộ Khoa học Công nghệ về việc công bố TCVN;

2. Đào Công (2023), *Tội phạm mạng đã sử dụng AI cho các cuộc tấn công APT*, https://dientuungdung.vn/toi-pham-mang-da-su-dung-ai-cho-cac-cuoc-tan-cong-apt, truy cập ngày 02/4/2023;

3. European Parliament, *Shaping Europe's digital future*, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai, truy cập ngày 25/4/2024;

4. EU, *Đạo luật TTNT năm 2024*, https://artificialintelligenceact.eu/the-act/, truy cập ngày 12/02/2025;

5. H.Hà (2023), *Mỹ thành lập Viện An toàn TTNT*, https://dangcongsan.vn/the-gioi/tin-tuc/my-thanh-lap-vien-an-toan-tri-tue-nhan-tao-650994.html, truy cập ngày 05/5/2024;

6. Nguyễn Ngọc Phương Hồng, Lưu Minh Sang (2024), *Luật về TTNT và tầm nhìn chiến lược của Trung Quốc*, https://thesaigontimes.vn/luat-ve-tri-tue-nhan-tao-va-tam-nhin-chien-luoc-cua-trung-quoc/, truy cập ngày 12/02/2025;

7. Joel Paul (2024), *Privacy and data security concerns in AI*, https://www.researchgate.net/publication/385781993_Privacy_and_data_security_concerns_in_AI, truy cập ngày 25/4/2025;

8. Tiểu Minh (2023), *ChatGPT được sử dụng để viết phần mềm độc hại*, https://plo.vn/ky-nguyen-so/chatgpt-duoc-su-dung-de-viet-phan-mem-doc-hai-post717954.html, truy cập ngày 20/4/2024;

9. Nicholas Berente, Bin Gu, Jan Recker, Radhika Santhanam, "Managing artificial intelligence", *MIS Quarterly*, Vol. 45, No. 3/2021, https://www.researchgate.net/publication/352400557_Managing_Artificial_Intelligence, truy cập ngày 05/4/2024;

10. Hùng Quân (2024), *Lừa đảo bằng công nghệ Deepfake lại bùng phát*, https://cand.com.vn/khoa-hoc-quan-su/lua-dao-bang-cong-nghe-deepfake-lai-bung-phat-i722392/, truy cập ngày 20/4/2024;

11. Thái Sơn (2024), *Cảnh giác cao với tội phạm lừa đảo sử dụng TTNT*, https://nhandan.vn/canh-giac-cao-voi-toi-pham-lua-dao-su-dung-tri-tue-nhan-tao-post746314.html, truy cập ngày 20/4/2024;

12. Whitehouse, *Removing barriers to American leadership in artificial intelligence*, https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/, truy cập ngày 12/02/2025.