

Hà Nội, ngày tháng 02 năm 2026

BẢN THUYẾT MINH

Dự thảo Thông tư quy định về an toàn, quản lý rủi ro và điều kiện triển khai ứng dụng trí tuệ nhân tạo trong ngành Ngân hàng

I. Sự cần thiết ban hành Thông tư

1. Cơ sở pháp lý

Tại khoản 4 Điều 6 Luật Trí tuệ nhân tạo số 134/2025/QH15 quy định: “Chính phủ, Bộ, cơ quan ngang Bộ trong phạm vi chức năng, nhiệm vụ và quyền hạn của mình quy định chi tiết yêu cầu về an toàn, quản lý rủi ro và điều kiện triển khai đối với việc ứng dụng trí tuệ nhân tạo trong ngành, lĩnh vực quản lý, bảo đảm phù hợp với quy định của Luật này”. Theo đó, Ngân hàng Nhà nước (NHNN) cần phải ban hành văn bản thuộc thẩm quyền (Thông tư) để quy định chi tiết yêu cầu về an toàn, quản lý rủi ro và điều kiện triển khai đối với việc ứng dụng trí tuệ nhân tạo trong ngành Ngân hàng.

Vì vậy, việc xây dựng và ban hành Thông tư là yêu cầu cấp thiết nhằm cụ thể hóa kịp thời quy định của Luật Trí tuệ nhân tạo trong lĩnh vực ngân hàng, đáp ứng tiến độ theo chỉ đạo và bảo đảm triển khai thống nhất, an toàn, hiệu quả trong toàn ngành.

2. Thực tiễn ứng dụng AI trong ngành Ngân hàng và yêu cầu quản lý rủi ro

Trong thời gian qua, các đơn vị trong ngành Ngân hàng đã đẩy mạnh ứng dụng AI trong nhiều nghiệp vụ như: AI dùng để chấm điểm tín dụng (Credit Scoring), thẩm định cho vay, phát hiện gian lận thanh toán, tư vấn tự động (chatbot), eKYC...

Tuy nhiên, việc triển khai AI cũng phát sinh những rủi ro, thách thức về an ninh, an toàn hệ thống thông tin; rủi ro về mô hình thuật toán có thể dẫn đến sai lệch/thiên lệch dữ liệu đầu ra; có thể xâm phạm quyền riêng tư dữ liệu khách hàng từ đó tác động đến quyền, lợi ích hợp pháp của tổ chức, cá nhân.

Như vậy, việc xây dựng Thông tư quy định về yêu cầu an toàn, quản lý rủi ro và điều kiện triển khai ứng dụng trí tuệ nhân tạo trong ngành Ngân hàng là cần thiết nhằm:

- Thiết lập khung yêu cầu tối thiểu về an toàn, quản lý rủi ro và điều kiện triển khai đối với việc ứng dụng AI trong hoạt động ngân hàng, làm cơ sở thống nhất để

các tổ chức tín dụng triển khai và cơ quan quản lý giám sát, đồng thời thúc đẩy việc ứng dụng AI trong hoạt động ngân hàng.

- Kiểm soát, giảm thiểu rủi ro phát sinh từ các ứng dụng AI trong các hoạt động nghiệp vụ ngân hàng (như chấm điểm tín dụng, thẩm định cho vay, eKYC, phát hiện gian lận, tư vấn/hỗ trợ ra quyết định), qua đó góp phần bảo đảm an toàn hoạt động ngân hàng và bảo vệ quyền, lợi ích hợp pháp của tổ chức, cá nhân có liên quan.

- Đáp ứng yêu cầu pháp lý và tiến độ ban hành văn bản hướng dẫn thi hành Luật Trí tuệ nhân tạo theo chỉ đạo của cơ quan có thẩm quyền, bảo đảm triển khai đồng bộ, kịp thời trong toàn ngành.

II. Quan điểm và phương pháp tiếp cận khi xây dựng Thông tư

1. Quan điểm xây dựng dự thảo Thông tư

Bảo đảm phù hợp với đường lối, chủ trương của Đảng, chính sách của Nhà nước; thống nhất, đồng bộ với hệ thống pháp luật hiện hành.

2. Phương pháp tiếp cận khi xây dựng dự thảo Thông tư

- Tuân thủ quy định pháp luật: Cục CNTT đã rà soát, đối chiếu các quy định của Luật Trí tuệ nhân tạo, các dự thảo văn bản triển khai Luật trí tuệ nhân tạo và các văn bản pháp luật liên quan để cụ thể hóa các yêu cầu đặc thù trong lĩnh vực ngân hàng, bảo đảm không chồng chéo, mâu thuẫn và thuận lợi khi tổ chức thực hiện.

- Tham chiếu chuẩn mực, thông lệ quốc tế: Trong quá trình xây dựng Thông tư, nghiên cứu, tham chiếu và tiếp thu có chọn lọc các khung, tiêu chuẩn và thông lệ quốc tế phù hợp với điều kiện Việt Nam và đặc thù hoạt động ngân hàng, làm cơ sở hoàn thiện các yêu cầu về an toàn, quản lý rủi ro và điều kiện triển khai ứng dụng AI; trong đó xem xét các tài liệu như Khung quản lý rủi ro AI của NIST (NIST AI RMF 1.0 - NIST.AI.100-1), cách tiếp cận quản lý theo rủi ro của EU AI Act (Regulation (EU) 2024/1689), các tiêu chuẩn ISO/IEC liên quan đến hệ thống quản lý AI và quản lý rủi ro AI (ISO/IEC 42001:2023; ISO/IEC 23894:2023; ISO/IEC 22989:2022),...

- Quản lý theo mức độ rủi ro: Quy định yêu cầu an toàn, quản lý rủi ro và điều kiện triển khai tương ứng mức độ rủi ro và phạm vi tác động của ứng dụng AI; ưu tiên kiểm soát chặt chẽ các trường hợp tác động lớn đến khách hàng, an toàn hệ thống và ổn định tài chính.

- Bảo đảm khả thi và hiệu quả tuân thủ: Thiết kế quy định tối thiểu, rõ ràng, có thể kiểm chứng; phù hợp đặc thù hoạt động ngân hàng và điều kiện triển khai thực tế; hạn chế phát sinh nghĩa vụ không cần thiết.

III. Một số nội dung chính của dự thảo Thông tư

Dự thảo Thông tư được xây dựng với kết cấu gồm 06 Chương, 27 Điều, tiếp

cần theo hướng quản lý dựa trên mức độ rủi ro, quy định một cách có hệ thống các yêu cầu về an toàn, quản lý rủi ro và điều kiện triển khai ứng dụng trí tuệ nhân tạo trong ngành Ngân hàng, nhằm thiết lập khung quản lý thống nhất trong toàn ngành, bảo đảm an toàn hoạt động ngân hàng và bảo vệ quyền, lợi ích hợp pháp của khách hàng:

1. Chương I. Quy định chung

(gồm 03 Điều, từ Điều 1 đến Điều 3)

Chương I quy định các nội dung mang tính nền tảng, bao gồm phạm vi điều chỉnh, đối tượng áp dụng, giải thích thuật ngữ và các nguyên tắc ứng dụng trí tuệ nhân tạo trong ngành Ngân hàng, làm cơ sở thống nhất cho việc tổ chức thực hiện Thông tư. Việc quy định Chương này nhằm xác định rõ phạm vi quản lý của Ngân hàng Nhà nước đối với hoạt động ứng dụng trí tuệ nhân tạo trong các nghiệp vụ ngân hàng; thống nhất cách hiểu, cách áp dụng các thuật ngữ, khái niệm liên quan; đồng thời thiết lập các nguyên tắc chung xuyên suốt, bảo đảm trí tuệ nhân tạo chỉ là công cụ hỗ trợ, không thay thế trách nhiệm và thẩm quyền của con người trong hoạt động ngân hàng.

2. Chương II. Yêu cầu bảo đảm an toàn hoạt động trong ứng dụng trí tuệ nhân tạo

(gồm 06 Điều, từ Điều 4 đến Điều 9)

Trên cơ sở các nguyên tắc chung được quy định tại Chương I, Chương II tập trung quy định các yêu cầu bảo đảm an toàn trong toàn bộ vòng đời của hệ thống trí tuệ nhân tạo, từ giai đoạn thiết kế, thu thập và xử lý dữ liệu, phát triển và huấn luyện mô hình, kiểm thử, triển khai, vận hành, giám sát cho đến ngừng sử dụng hệ thống. Các yêu cầu này bao gồm an toàn hệ thống thông tin, an toàn dữ liệu, quản lý thay đổi, xử lý sự cố, ghi nhật ký, truy vết và quản lý rủi ro chuỗi cung ứng công nghệ,... Đồng thời, Chương này tăng cường kiểm soát rủi ro phát sinh từ dữ liệu - yếu tố cốt lõi quyết định chất lượng và độ tin cậy của hệ thống trí tuệ nhân tạo trong bối cảnh dữ liệu phục vụ hệ thống trí tuệ nhân tạo ngày càng gia tăng, có nguồn gốc đa dạng, được thu thập và xử lý từ nhiều hệ thống, nhiều giai đoạn khác nhau, với mức độ nhạy cảm và tác động ngày càng lớn đối với hoạt động ngân hàng và quyền lợi khách hàng.

Bên cạnh đó, Chương II thiết lập cơ chế quản lý thay đổi, xử lý sự cố, ghi nhận và truy vết đầy đủ nhằm bảo đảm tính liên tục hoạt động, khả năng khôi phục hệ thống và phục vụ hiệu quả công tác kiểm tra, thanh tra, giám sát khi cần thiết. Các quy định về quản lý rủi ro chuỗi cung ứng công nghệ trí tuệ nhân tạo cũng được đặt trong mối liên hệ xuyên suốt vòng đời hệ thống, đặc biệt đối với trường hợp sử dụng phần mềm, mô hình hoặc dịch vụ của bên thứ ba, qua đó hạn chế nguy cơ lan truyền rủi ro và phụ thuộc công nghệ trong hoạt động ngân hàng.

3. Chương III. Quản lý rủi ro trong ứng dụng trí tuệ nhân tạo

(gồm 09 Điều, từ Điều 10 đến Điều 18)

Bên cạnh các yêu cầu về an toàn kỹ thuật và vận hành, Chương III quy định các nội dung cốt lõi về quản lý rủi ro trong ứng dụng trí tuệ nhân tạo theo cách tiếp cận dựa trên mức độ rủi ro. Chương này quy định việc phân loại hệ thống trí tuệ nhân tạo theo mức độ rủi ro; yêu cầu về kiểm soát của con người; đánh giá tác động đối với hệ thống trí tuệ nhân tạo rủi ro cao; trách nhiệm quản lý rủi ro của các cấp có thẩm quyền; quản lý rủi ro mô hình, dữ liệu và hoạt động thuê ngoài... Việc quy định các nội dung này nhằm bảo đảm biện pháp quản lý tương xứng với phạm vi tác động của từng hệ thống, tăng cường kiểm soát đối với các hệ thống có ảnh hưởng trực tiếp đến quyền lợi khách hàng, an toàn hoạt động ngân hàng và ổn định hệ thống tài chính; đồng thời khẳng định trách nhiệm cuối cùng vẫn thuộc về đơn vị triển khai, không chuyển giao trách nhiệm cho công nghệ hoặc bên thứ ba.

4. Chương IV. Điều kiện triển khai ứng dụng trí tuệ nhân tạo

(gồm 02 Điều, từ Điều 19 đến Điều 20)

Trên cơ sở các yêu cầu về an toàn và quản lý rủi ro, Chương IV quy định các điều kiện cụ thể trước khi đưa hệ thống trí tuệ nhân tạo vào vận hành, bao gồm hồ sơ phê duyệt nội bộ, yêu cầu về nhân sự triển khai. Việc quy định Chương này nhằm bảo đảm các hệ thống trí tuệ nhân tạo chỉ được đưa vào vận hành khi đã được đánh giá đầy đủ về an toàn, rủi ro và mức độ sẵn sàng; tăng cường trách nhiệm của đơn vị triển khai trong việc chuẩn bị phương án vận hành, phương án dự phòng và quản lý rủi ro; đồng thời hạn chế tình trạng triển khai trí tuệ nhân tạo theo phong trào, thiếu chuẩn bị và thiếu kiểm soát trong thực tiễn.

5. Chương V. Bảo vệ quyền lợi của khách hàng

(gồm 02 Điều, từ Điều 21 đến Điều 22)

Chương V quy định các nội dung nhằm bảo đảm quyền và lợi ích hợp pháp của khách hàng trong quá trình các tổ chức tín dụng ứng dụng trí tuệ nhân tạo vào hoạt động ngân hàng, gắn với yêu cầu nâng cao mức độ bảo vệ khách hàng và tăng cường niềm tin đối với dịch vụ ngân hàng số.

Theo đó, Chương này quy định các yêu cầu về cung cấp thông tin đối với việc sử dụng trí tuệ nhân tạo, quyền của khách hàng được yêu cầu xem xét lại bởi con người đối với các quyết định có tác động đến quyền và lợi ích hợp pháp của mình, cơ chế tiếp nhận và giải quyết khiếu nại, cũng như các biện pháp phòng ngừa phân biệt đối xử và bảo vệ nhóm khách hàng yếu thế. Thông qua các quy định này, Thông tư hướng tới việc bảo đảm khách hàng được đối xử công bằng, được bảo vệ phù hợp và được sử dụng dịch vụ ngân hàng an toàn trong bối cảnh ứng dụng trí tuệ nhân tạo ngày càng mở rộng.

6. Chương VI. Điều khoản thi hành

(gồm 05 Điều, từ Điều 23 đến Điều 27)

Chương VI quy định các nội dung về chế độ báo cáo, trách nhiệm của các đơn vị thuộc Ngân hàng Nhà nước, hiệu lực thi hành và quy định chuyển tiếp. Việc quy định Chương này nhằm làm rõ trách nhiệm tổ chức thực hiện, thiết lập cơ chế theo dõi, giám sát và đánh giá việc triển khai Thông tư; đồng thời bảo đảm việc áp dụng Thông tư có lộ trình phù hợp đối với các hệ thống trí tuệ nhân tạo đã được triển khai trước thời điểm Thông tư có hiệu lực, tránh gây gián đoạn hoạt động ngân hàng.

IV. Các nội dung chính cần xin ý kiến

1. Về phạm vi, đối tượng áp dụng

Đối tượng chính chịu tác động của Thông tư là: các tổ chức tín dụng, chi nhánh ngân hàng nước ngoài, tổ chức cung ứng dịch vụ trung gian thanh toán, công ty thông tin tín dụng, Công ty Quản lý tài sản của các tổ chức tín dụng Việt Nam, Bảo hiểm tiền gửi Việt Nam.

Thông tư quy định áp dụng bắt buộc đối với việc ứng dụng trí tuệ nhân tạo trong các hoạt động nghiệp vụ cung cấp dịch vụ cho khách hàng, bao gồm: (i) Hoạt động ngân hàng và các hoạt động kinh doanh khác của tổ chức tín dụng, chi nhánh ngân hàng nước ngoài; (ii) Hoạt động cung ứng dịch vụ trung gian thanh toán; (iii) Hoạt động thông tin tín dụng.

Đối với việc ứng dụng trí tuệ nhân tạo trong hoạt động quản trị, điều hành nội bộ phải tuân thủ quy định của pháp luật về trí tuệ nhân tạo; đồng thời khuyến khích áp dụng các quy định tại Thông tư này nhằm tăng cường quản lý rủi ro và bảo đảm an toàn hệ thống.

2. Về yêu cầu quản lý rủi ro trong ứng dụng trí tuệ nhân tạo

Dự thảo Thông tư quy định nguyên tắc việc ứng dụng trí tuệ nhân tạo trong ngành ngân hàng phải tuân thủ quy định của pháp luật về trí tuệ nhân tạo và các quy định về quản lý rủi ro của Ngân hàng Nhà nước. Trong đó có quy định một số nội dung cụ thể:

- Việc phân loại rủi ro của hệ thống trí tuệ nhân tạo thực hiện theo quy định của pháp luật về trí tuệ nhân tạo.

- Thiết lập cơ chế giám sát và can thiệp của con người phù hợp với mức độ rủi ro của hệ thống trí tuệ nhân tạo trên nguyên tắc trí tuệ nhân tạo không thay thế thẩm quyền và trách nhiệm của con người; đơn vị chịu trách nhiệm cuối cùng trước pháp luật và khách hàng đối với các quyết định do hệ thống trí tuệ nhân tạo đưa ra hoặc hỗ trợ đưa ra.

- Trách nhiệm quản lý rủi ro trong ứng dụng trí tuệ nhân tạo. Trong đó hội đồng quản trị chịu trách nhiệm cao nhất về việc ứng dụng trí tuệ nhân tạo tại đơn vị mình.

- Xây dựng Khung quản lý rủi ro đối với ứng dụng trí tuệ nhân tạo và các yêu cầu cụ thể đối với quản lý rủi ro đặc thù đối với ứng dụng trí tuệ nhân tạo như Rủi ro mô hình trí tuệ nhân tạo; Rủi ro chuỗi cung ứng; Rủi ro từ bên thứ ba...

3. Về yêu cầu bảo đảm an toàn thông tin, an ninh mạng và quản lý dữ liệu trong ứng dụng trí tuệ nhân tạo

Dự thảo Thông tư quy định hệ thống trí tuệ nhân tạo phải tuân thủ quy định về an toàn hệ thống thông tin trong hoạt động ngân hàng; được kiểm tra, đánh giá an toàn thông tin trước khi đưa vào vận hành và định kỳ theo mức độ rủi ro, hình thức triển khai; áp dụng biện pháp phòng, chống các rủi ro đặc thù đối với hệ thống trí tuệ nhân tạo, đặc biệt đối với hệ thống rủi ro cao.

Đơn vị phải giám sát liên tục hiệu năng và hành vi của hệ thống; thực hiện ghi nhật ký và lưu trữ dữ liệu vận hành theo quy định.

Đối với dữ liệu, Thông tư yêu cầu bảo đảm dữ liệu huấn luyện, kiểm thử và vận hành đầy đủ, chính xác, có nguồn gốc hợp pháp; thiết lập quy trình kiểm soát chất lượng dữ liệu, phát hiện và xử lý dữ liệu bị sai lệch, bị thao túng hoặc bị đầu độc; giám sát hiện tượng trôi dữ liệu và tuân thủ quy định về bảo vệ dữ liệu và dữ liệu cá nhân.

4. Về điều kiện triển khai

Dự thảo Thông tư quy định trước khi đưa hệ thống trí tuệ nhân tạo vào vận hành chính thức, đơn vị phải tuân thủ quy định của pháp luật về trí tuệ nhân tạo và hoàn thành hồ sơ được phê duyệt theo thẩm quyền, bao gồm: hồ sơ phân loại rủi ro; kết quả kiểm tra, đánh giá an toàn thông tin; đánh giá tác động đối với hệ thống rủi ro cao; kế hoạch vận hành an toàn và điều kiện dừng, khôi phục hệ thống. Đồng thời, đơn vị phải thiết lập các tiêu chí triển khai (độ chính xác, độ tin cậy, hiệu năng, kết quả đánh giá an toàn thông tin) và xây dựng kế hoạch triển khai chi tiết, bao gồm phương án chuyển đổi dữ liệu và phục hồi khi xảy ra sự cố.

5. Về bảo vệ quyền lợi của khách hàng

Dự thảo Thông tư quy định các yêu cầu về minh bạch thông tin khi triển khai hệ thống trí tuệ nhân tạo tương tác với khách hàng; bảo đảm quyền được xem xét lại bởi con người đối với các quyết định được đưa ra hoàn toàn tự động; thiết lập cơ chế tiếp nhận, xử lý khiếu nại và cung cấp thông tin giải thích phù hợp đối với các quyết định của hệ thống trí tuệ nhân tạo.

6. Lộ trình triển khai

Dự thảo Thông tư quy định Thông tư có hiệu lực thi hành từ ngày 01 tháng 03 năm 2026. Đối với các hệ thống trí tuệ nhân tạo đã được triển khai trước thời điểm Thông tư có hiệu lực, Thông tư quy định thời hạn chuyển tiếp đến ngày 01 tháng 09 năm 2027 để các đơn vị thực hiện rà soát, đánh giá và hoàn thiện việc tuân thủ theo quy định mới. Quy định này bảo đảm thống nhất với lộ trình thực hiện của pháp luật về trí tuệ nhân tạo.

Riêng quy định tại điểm b khoản 2 Điều 4 có hiệu lực từ ngày 01 tháng 01 năm 2027 nhằm tạo điều kiện cho các đơn vị chuẩn bị, đáp ứng các yêu cầu kỹ thuật chuyên sâu theo lộ trình phù hợp.

Trên đây là một số nội dung thuyết minh của Đơn vị soạn thảo về việc xây dựng dự thảo Thông tư Quy định về an toàn, quản lý rủi ro và điều kiện triển khai ứng dụng trí tuệ nhân tạo trong ngành Ngân hàng để các cá nhân, tổ chức nghiên cứu và góp ý cho dự thảo Thông tư./.

NGÂN HÀNG NHÀ NƯỚC VIỆT NAM