

Cần Thơ, ngày 04 tháng 3 năm 2026

Số: 02/TB-KTPCPTCL

THÔNG BÁO
Về việc triển khai, phổ biến Luật Trí tuệ nhân tạo năm 2025

Thực hiện Kế hoạch số 185/KH-ĐHKTCN ngày 26 tháng 02 năm 2026 của Hiệu trưởng Trường Đại học Kỹ thuật - Công nghệ Cần Thơ về phổ biến, giáo dục pháp luật năm 2026.

Phòng Kiểm tra - Pháp chế - Phát triển chiến lược phổ biến nội dung các Luật đến các đơn vị, viên chức, người lao động như sau:

I. Luật Trí tuệ nhân tạo năm 2025

Được Quốc hội thông qua vào tháng 12/2025 và có hiệu lực từ 01/3/2026.

Những điểm nổi bật của Luật Luật Trí tuệ nhân tạo năm 2025, cụ thể như sau:

1. Lần đầu tiên xây dựng bộ khung định nghĩa pháp lý thống nhất về AI

Theo Điều 3 của Luật Trí tuệ nhân tạo 2025, “trí tuệ nhân tạo” được hiểu là việc thực hiện bằng điện tử các năng lực trí tuệ của con người như học tập, suy luận, nhận thức, phán đoán và hiểu ngôn ngữ tự nhiên

2. Nguyên tắc hoạt động AI: Nhấn mạnh “lấy con người làm trung tâm”

Điều 4 Luật Trí tuệ nhân tạo 2025 đưa ra bốn nguyên tắc cốt lõi định hướng toàn bộ quá trình phát triển và ứng dụng AI tại Việt Nam.

Thứ nhất, AI phải lấy con người làm trung tâm, bảo đảm quyền con người, quyền riêng tư, an ninh quốc gia và tuân thủ pháp luật. Điều này khẳng định AI không được vận hành ngoài khuôn khổ pháp lý hay gây tổn hại đến giá trị căn bản.

Thứ hai, AI không thay thế thẩm quyền và trách nhiệm của con người. Người vận hành luôn phải giữ quyền kiểm soát và khả năng can thiệp vào mọi quyết định do hệ thống AI tạo ra, đồng thời đảm bảo an toàn hệ thống và an ninh dữ liệu.

Thứ ba, AI phải bảo đảm công bằng, minh bạch và không thiên lệch. Tổ chức, cá nhân sử dụng AI có trách nhiệm phát hiện và ngăn ngừa các sai lệch, phân biệt đối xử, đồng thời phải có khả năng giải trình khi hệ thống đưa ra quyết định.

Cuối cùng, luật khuyến khích AI phát triển theo hướng xanh, tiết kiệm năng lượng và giảm tác động tiêu cực đến môi trường.

3. Các hành vi bị nghiêm cấm trong hoạt động trí tuệ nhân tạo

Điều 7 của Luật AI quy định một danh mục nghiêm ngặt các hành vi bị cấm nhằm bảo vệ an ninh, quyền con người và lợi ích chung.

Luật cấm mọi hành vi lợi dụng hoặc chiếm đoạt hệ thống AI để xâm phạm quyền và lợi ích hợp pháp của tổ chức, cá nhân.

Đặc biệt, nghiêm cấm phát triển hoặc sử dụng AI nhằm lừa dối, thao túng nhận thức, hoặc gây tổn hại nghiêm trọng đến con người và xã hội.

Điều này bao gồm hành vi tạo nội dung giả mạo như deepfake để lừa đảo hoặc gây hoang mang trong dư luận. Luật cũng cấm lợi dụng điểm yếu của nhóm dễ tổn thương như trẻ em, người già, người khuyết tật.

Ngoài ra, thu thập và sử dụng dữ liệu để huấn luyện hoặc vận hành AI trái pháp luật về dữ liệu cá nhân, sở hữu trí tuệ và an ninh mạng đều bị cấm.

Việc cản trở cơ chế kiểm soát của con người, che giấu thông tin bắt buộc phải minh bạch hoặc lợi dụng hoạt động nghiên cứu - thử nghiệm để thực hiện hành vi vi phạm cũng thuộc danh mục cấm.

4. Thiết lập Cổng thông tin một cửa và Cơ sở dữ liệu quốc gia về AI

Điều 8 Luật Trí tuệ nhân tạo 2025 quy định việc thiết lập hai hạ tầng quản lý quan trọng: Cổng thông tin điện tử một cửa về AI và Cơ sở dữ liệu quốc gia về hệ thống AI.

Cổng thông tin đóng vai trò tiếp nhận đăng ký thử nghiệm có kiểm soát, tiếp nhận thông báo phân loại rủi ro, báo cáo sự cố nghiêm trọng và báo cáo định kỳ; đồng thời công khai thông tin về hệ thống AI, kết quả đánh giá sự phù hợp và xử lý vi phạm.

Cổng còn kết nối các chương trình hỗ trợ, quỹ và hạ tầng dữ liệu dùng chung để tạo hệ sinh thái minh bạch.

Cơ sở dữ liệu quốc gia được xây dựng nhằm lưu trữ thống nhất thông tin về hệ thống AI phục vụ giám sát, hậu kiểm và minh bạch hóa việc vận hành hệ thống.

Việc công khai, chia sẻ dữ liệu phải bảo đảm an toàn thông tin, bảo vệ bí mật nhà nước, bí mật kinh doanh và dữ liệu cá nhân. Chính phủ sẽ quy định cơ chế vận hành, khai thác và kết nối hai hệ thống này.

5. 3 mức độ rủi ro của hệ thống trí tuệ nhân tạo

Theo Điều 9 Luật Trí tuệ nhân tạo 2025, hệ thống AI được phân loại theo ba mức rủi ro nhằm áp dụng nghĩa vụ quản lý tương ứng.

- AI rủi ro cao là các hệ thống có khả năng gây thiệt hại đáng kể đến tính mạng, sức khỏe, quyền con người, an ninh quốc gia hoặc lợi ích công cộng. Đây là nhóm phải chịu cơ chế giám sát và đánh giá nghiêm ngặt nhất.

- AI rủi ro trung bình là các hệ thống có khả năng gây nhầm lẫn hoặc thao túng người sử dụng, đặc biệt khi người dùng không nhận biết được rằng họ đang

tương tác với AI. Các ứng dụng có nguy cơ tác động đến hành vi hoặc nhận thức nằm trong nhóm này.

- AI rủi ro thấp là những hệ thống không thuộc hai loại trên, thường là công cụ hỗ trợ đơn giản hoặc có phạm vi tác động nhỏ.

6. Âm thanh, hình ảnh, video do AI tạo ra bắt buộc gắn dấu hiệu nhận biết

Điều 11 của Luật Trí tuệ nhân tạo 2025 quy định rõ trách nhiệm minh bạch khi phát triển và sử dụng AI. Nhà cung cấp phải bảo đảm người sử dụng nhận biết khi họ đang tương tác với hệ thống AI, trừ trường hợp pháp luật có quy định khác.

Tất cả âm thanh, hình ảnh, video do AI tạo ra bắt buộc phải được gắn dấu hiệu nhận biết ở định dạng máy đọc nhằm phân biệt với nội dung thật, đây được xem là một giải pháp quan trọng để ngăn chặn deepfake.

Khi bên triển khai cung cấp ra công chúng nội dung AI có khả năng gây nhầm lẫn về tính xác thực, họ phải thông báo rõ ràng và gắn nhãn phân biệt, đặc biệt với nội dung mô phỏng con người hoặc sự kiện thật.

Riêng các tác phẩm nghệ thuật hoặc điện ảnh, việc gắn nhãn có thể linh hoạt hơn nhưng vẫn phải đảm bảo phân biệt với nội dung thật. Trách nhiệm minh bạch phải được duy trì trong toàn bộ thời gian cung cấp hệ thống hoặc nội dung.

Chính phủ sẽ quy định hình thức thông báo và gắn nhãn cụ thể.

7. Báo cáo, xử lý sự cố AI thực hiện qua Cổng thông tin điện tử một cửa về AI

Theo quy định tại Điều 12 Luật Trí tuệ nhân tạo, mọi chủ thể liên quan đến hệ thống trí tuệ nhân tạo - gồm nhà phát triển, nhà cung cấp, bên triển khai và người sử dụng đều có trách nhiệm bảo đảm an toàn, an ninh, độ tin cậy của hệ thống và kịp thời phát hiện, khắc phục các sự cố có thể gây tổn hại đến con người, tài sản, dữ liệu hoặc trật tự xã hội.

Khi xảy ra sự cố nghiêm trọng, nhà phát triển và nhà cung cấp phải nhanh chóng áp dụng biện pháp kỹ thuật để khắc phục, tạm dừng hoặc thu hồi hệ thống và thông báo cho cơ quan có thẩm quyền; bên triển khai và người sử dụng phải ghi nhận, báo cáo kịp thời và phối hợp xử lý. Cơ quan quản lý nhà nước có thẩm quyền sẽ tiếp nhận, xác minh, hướng dẫn và khi cần thiết có thể yêu cầu tạm dừng, thu hồi hoặc đánh giá lại hệ thống.

Toàn bộ hoạt động báo cáo và xử lý sự cố được thực hiện thông qua Cổng thông tin điện tử một cửa về trí tuệ nhân tạo, và Chính phủ sẽ quy định chi tiết về quy trình báo cáo cũng như trách nhiệm của các cơ quan, tổ chức, cá nhân phù hợp với mức độ nghiêm trọng và phạm vi ảnh hưởng của sự cố.

8. Hệ thống AI rủi ro cao phải được đánh giá sự phù hợp chứng nhận trước khi vận hành

Điều 13 và 14 Luật Trí tuệ nhân tạo 2025 quy định một hệ thống quản lý nghiêm ngặt dành cho AI rủi ro cao. Trước khi đưa vào sử dụng, hệ thống này

phải được đánh giá sự phù hợp theo tiêu chuẩn, quy chuẩn kỹ thuật và các yêu cầu của Luật AI.

Một số hệ thống thuộc danh mục bắt buộc phải được tổ chức đánh giá sự phù hợp chứng nhận trước khi vận hành. Các hệ thống khác có thể tự đánh giá hoặc thuê tổ chức đánh giá được thừa nhận.

Kết quả đánh giá là điều kiện để hệ thống được phép hoạt động và phải được duy trì liên tục. Nhà cung cấp phải thiết lập cơ chế quản lý rủi ro, quản trị dữ liệu huấn luyện, lưu trữ hồ sơ kỹ thuật và nhật ký hoạt động; đồng thời bảo đảm khả năng can thiệp của con người và thực hiện nghĩa vụ minh bạch - xử lý sự cố.

Bên triển khai có trách nhiệm vận hành đúng phạm vi rủi ro, bảo đảm an toàn dữ liệu và giải trình khi được yêu cầu. Người sử dụng AI rủi ro cao phải tuân thủ hướng dẫn kỹ thuật và không được can thiệp trái phép vào hệ thống.

9. AI được hưởng mức ưu đãi cao nhất về khoa học công nghệ, công nghệ cao

Luật AI dành nhiều cơ chế ưu đãi cho doanh nghiệp để thúc đẩy phát triển hệ sinh thái công nghệ số.

Điều 20 Luật quy định doanh nghiệp AI được hưởng mức ưu đãi cao nhất theo pháp luật về khoa học công nghệ, công nghệ cao, chuyển đổi số và đầu tư; đồng thời được tạo điều kiện tiếp cận hạ tầng tính toán, dữ liệu và môi trường thử nghiệm.

Nhà nước phát triển thị trường AI thông qua việc ưu tiên sử dụng sản phẩm AI trong đấu thầu, xây dựng sản giao dịch công nghệ và bảo đảm môi trường thử nghiệm công bằng.

Điều 21 quy định cơ chế thử nghiệm có kiểm soát (sandbox), cho phép hệ thống AI mới được thử nghiệm trong môi trường rủi ro có kiểm soát, thậm chí được miễn hoặc giảm nghĩa vụ tuân thủ dựa trên kết quả thử nghiệm.

Điều 22 thành lập Quỹ Phát triển Trí tuệ Nhân tạo Quốc gia - một quỹ tài chính ngoài ngân sách, có cơ chế linh hoạt và chấp nhận rủi ro trong đổi mới sáng tạo.

Quỹ ưu tiên đầu tư hạ tầng AI, đào tạo nhân lực, phát triển công nghệ cốt lõi và hỗ trợ doanh nghiệp khởi nghiệp sáng tạo.

Điều 25 bổ sung chính sách hỗ trợ đặc thù cho startup và doanh nghiệp nhỏ, bao gồm hỗ trợ chi phí đánh giá sự phù hợp, công cụ tự đánh giá và quyền tiếp cận dữ liệu dùng chung.

10. Hạ tầng AI quốc gia là hạ tầng chiến lược, Nhà nước giữ vai trò điều phối

Điều 16 của Luật xác định hạ tầng AI quốc gia là hạ tầng chiến lược, được phát triển theo hướng mở, an toàn và có khả năng mở rộng, bao gồm năng lực tính toán, dữ liệu dùng chung, nền tảng huấn luyện - kiểm thử, mô hình nền tảng và môi trường thử nghiệm.

Nhà nước giữ vai trò điều phối và đầu tư các hạ tầng quan trọng phục vụ nghiên cứu, phát triển và quản lý nhà nước, đồng thời khuyến khích hợp tác công - tư và thúc đẩy doanh nghiệp, viện nghiên cứu, trường đại học tham gia.

Các ứng dụng AI trong lĩnh vực thiết yếu thuộc danh mục Thủ tướng ban hành phải triển khai trên hạ tầng AI quốc gia nhằm bảo đảm an toàn và khả năng kiểm soát.

Điều 17 quy định cơ sở dữ liệu phục vụ AI gồm dữ liệu quốc gia và dữ liệu của bộ ngành, địa phương, tổ chức và cá nhân.

Dữ liệu được phân thành dữ liệu mở, dữ liệu mở có điều kiện và dữ liệu thương mại, được kết nối thống nhất với cơ sở dữ liệu quốc gia và phải tuân thủ quy định về bảo vệ dữ liệu cá nhân. Thủ tướng sẽ ban hành danh mục dữ liệu ưu tiên như văn hóa, ngôn ngữ, y tế, giáo dục, nông nghiệp, kinh tế - xã hội.

11. Hệ thống cũ hoàn thành đầy đủ nghĩa vụ theo Luật AI trước ngày 01/3/2027

Theo Điều 34, Luật Trí tuệ nhân tạo 2025 có hiệu lực từ ngày 01/3/2026. Điều 35 quy định lộ trình chuyển tiếp dành cho hệ thống AI đã hoạt động trước thời điểm Luật có hiệu lực.

Theo đó, các hệ thống AI trong lĩnh vực y tế, giáo dục và tài chính có thời hạn 18 tháng để hoàn thành đầy đủ nghĩa vụ theo Luật AI, tức phải tuân thủ trước ngày 01/9/2027. Các hệ thống trong lĩnh vực khác có thời hạn 12 tháng, tức hoàn thành trước ngày 01/3/2027.

Trong thời gian chuyển tiếp, hệ thống vẫn được phép hoạt động, trừ trường hợp bị cơ quan quản lý xác định có nguy cơ gây thiệt hại nghiêm trọng và ra quyết định tạm dừng hoặc chấm dứt hoạt động.

Trên đây là Thông báo triển khai, phổ biến Luật Trí tuệ nhân tạo năm 2025, lãnh đạo các đơn vị triển khai thực hiện lồng ghép vào các buổi sinh hoạt tại đơn vị mình./.

(Đính kèm Luật Trí tuệ nhân tạo năm 2025).

Nơi nhận:

- Ban Giám hiệu (báo cáo);
- Các đơn vị trong Trường;
- VC, NLĐ;
- Lưu: VT, PC.

**TL. HIỆU TRƯỞNG
TRƯỞNG PHÒNG**



Lê Hồng Tuyên