

Danh mục hệ thống trí tuệ nhân tạo có rủi ro cao

06:21 03/07/2026

Phó Thủ tướng Hồ Quốc Dũng ký Quyết định số 33/2026/QĐ-TTg ban hành Danh mục hệ thống trí tuệ nhân tạo có rủi ro cao.



Ảnh minh họa

Theo Quyết định, hệ thống trí tuệ nhân tạo phải đáp ứng các nguyên tắc hoạt động cơ bản theo quy định của pháp luật về trí tuệ nhân tạo; việc sử dụng hệ thống trí tuệ nhân tạo không làm thay đổi, chuyển giao hoặc loại trừ thẩm quyền, trách nhiệm của cơ quan, tổ chức, cá nhân có thẩm quyền theo quy định của pháp luật; bảo đảm sự giám sát, kiểm soát và khả năng can thiệp của con người trong quá trình vận hành hệ thống; không thực hiện các hành vi bị nghiêm cấm theo quy định của pháp luật.

Tiêu chí xác định hệ thống trí tuệ nhân tạo rủi ro cao

Các hệ thống trí tuệ nhân tạo được xác định là có rủi ro cao căn cứ trên các tiêu chí sau đây:

Hệ thống có thể gây thiệt hại đáng kể đến tính mạng, sức khỏe, quyền và lợi ích hợp pháp của tổ chức, cá nhân, lợi ích quốc gia, lợi ích công cộng, an ninh quốc gia theo quy định tại Điều 9 Luật Trí tuệ nhân tạo; khoản 1 Điều 8 của Nghị định số 142/2026/NĐ-CP quy định chi tiết một số điều và biện pháp thi hành Luật Trí tuệ nhân tạo.

Hệ thống không được loại trừ theo quy định tại khoản 2 Điều 8 của Nghị định số 142/2026/NĐ-CP quy định chi tiết một số điều và biện pháp thi hành Luật Trí tuệ nhân tạo.

Danh mục hệ thống trí tuệ nhân tạo có rủi ro cao

Quyết định nêu rõ, danh mục các hệ thống trí tuệ nhân tạo được xác định là có rủi ro cao ở các lĩnh vực: Giáo dục, dân tộc và tôn giáo, y tế, ngân hàng, tổ tụng, giao thông vận tải. Trong đó:

1. Lĩnh vực giáo dục: Hệ thống trí tuệ nhân tạo cung cấp nội dung tự học theo chương trình giáo dục, có sử dụng các nguồn dữ liệu không kiểm soát; Hệ thống có sử dụng AI tự động kiểm tra, đánh giá kết quả, xếp hạng người học; Hệ thống có sử dụng AI để giám sát, phân tích hành vi người học.

2. Lĩnh vực dân tộc và tôn giáo: Hệ thống trí tuệ nhân tạo tự động chấm điểm, phân loại, xếp hạng hồ sơ để xác định đối tượng, địa bàn thụ hưởng chính sách dân tộc; Hệ thống trí tuệ nhân tạo tự động xác thực và quyết định tính hợp lệ của thông tin liên quan đến dân tộc và tôn giáo; Hệ thống trí tuệ nhân tạo tự động quyết định chấp thuận hoặc từ chối hồ sơ đăng ký, gia hạn hoặc hủy bỏ các quyết định quản lý nhà nước trong lĩnh vực dân tộc và tôn giáo...

3. Lĩnh vực y tế: Hệ thống trí tuệ nhân tạo hỗ trợ phẫu thuật/rô-bốt phẫu thuật; Hệ thống trí tuệ nhân tạo hỗ trợ phẫu thuật/rô-bốt phẫu thuật (Hệ thống máy móc, thiết bị rô-bốt sử dụng trí tuệ nhân tạo điều khiển tự động).

4. Lĩnh vực ngân hàng: Hệ thống trí tuệ nhân tạo tự động thực hiện giao dịch điện tử trong hoạt động ngân hàng; Hệ thống trí tuệ nhân tạo tự động quyết định cấp tín dụng.

5. Lĩnh vực tố tụng: Hệ thống trí tuệ nhân tạo nhận dạng sinh trắc học diện rộng phục vụ giải quyết vụ án dân sự công ích.

6. Lĩnh vực giao thông vận tải gồm 31 hệ thống như: Hệ thống trí tuệ nhân tạo điều khiển tự hành cấp độ cao trên phương tiện giao thông; Hệ thống trí tuệ nhân tạo điều khiển hoặc tự động vận hành tín hiệu, điều độ giao thông đường bộ và đường sắt; Hệ thống trí tuệ nhân tạo điều hành, kiểm soát vận hành công

trình giao thông và hạ tầng kỹ thuật trọng yếu...

Đối với các hệ thống trí tuệ nhân tạo thuộc Danh mục ban hành kèm theo Quyết định này đã được đưa vào hoạt động **trước ngày 15/8/2026**, nhà cung cấp và bên triển khai có trách nhiệm hoàn thành các nghĩa vụ tuân thủ theo quy định của pháp luật về trí tuệ nhân tạo theo lộ trình sau:

Trước ngày 1/9/2027 đối với các hệ thống trí tuệ nhân tạo trong lĩnh vực y tế, giáo dục và tài chính.

Trước ngày 1/3/2027 đối với các hệ thống trí tuệ nhân tạo thuộc các lĩnh vực còn lại trong Danh mục.

Trong thời hạn chuyển tiếp, các hệ thống trí tuệ nhân tạo nêu trên được phép tiếp tục hoạt động, trừ trường hợp cơ quan nhà nước có thẩm quyền xác định hệ thống có nguy cơ gây thiệt hại nghiêm trọng và yêu cầu tạm dừng hoặc chấm dứt hoạt động theo quy định tại khoản 2 Điều 35 Luật Trí tuệ nhân tạo.

Theo chinhphu.vn