BỘ CÔNG AN
CÔNG AN TP HÀ NỘI

Số: 2129 /CAHN-ANM&PCTPSDCNC
V/v cảnh báo nguy cơ mất an ninh mạng, an ninh dữ liệu liên quan đến hoạt động sử dụng trí tuệ nhân tạo (AI)

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

Hà Nội, ngày 31 tháng 3 năm 2026

Kính gửi:

- Các Sở, ban, ngành, cơ quan ngang Sở;
- UBND các xã, phường thành phố Hà Nội.

Căn cứ Luật Trí tuệ nhân tạo số 134/2025/QH15 *(có hiệu lực từ ngày 01/3/2026)*; thực hiện chức năng quản lý nhà nước về an toàn thông tin mạng trên địa bàn thành phố Hà Nội; trước tình hình tội phạm mạng lợi dụng trí tuệ nhân tạo *(AI)* để xâm phạm an ninh quốc gia và trật tự an toàn xã hội, Công an Thành phố *(CATP)* thông báo và kiến nghị các đơn vị triển khai một số nội dung trọng tâm sau:

1. Qua công tác giám sát, bảo đảm an ninh mạng, an ninh dữ liệu, CATP xác định 03 nhóm nguy cơ chủ yếu khi sử dụng AI tại các cơ quan nhà nước, cụ thể như sau:

(1) Nguy cơ lộ mất tài liệu nội bộ, bí mật nhà nước trên không gian mạng đến từ việc cán bộ, công chức, viên chức có thói quen đưa dự thảo văn bản, dữ liệu, đề án, báo cáo vào các AI công cộng trên mạng Internet để tóm tắt, nghiên cứu, biên tập dẫn đến dữ liệu bị lưu trữ tại máy chủ nước ngoài, nằm ngoài tầm kiểm soát, tiềm ẩn nguy cơ vi phạm quy định về bảo vệ dữ liệu cá nhân và bí mật nhà nước.

(2) Các ứng dụng AI tự vận hành *(AI Agent)* có chức năng tự động rà soát, đọc tệp tin, khởi chạy ứng dụng trên thiết bị máy tính và thực hiện theo các yêu cầu của người dùng. Việc khai thác lỗ hổng bảo mật trên AI Agent *(như lỗ hổng nghiêm trọng CVE-2026-25253 tồn tại trên OpenClaw)* cho phép tin tặc chiếm quyền điều khiển hệ thống và ra lệnh cho AI Agent đánh cắp thông tin, dữ liệu, ảnh hưởng trực tiếp đến hoạt động của hệ thống thông tin. Đáng lo ngại hơn, AI Agent thường cho phép mở rộng tính năng của hệ thống thông qua các phần mềm không rõ nguồn gốc và kho ứng dụng *(App)* như ClawHub. Theo thống kê, khoảng 12-20% app trên kho ứng dụng này là phần mềm độc hại có khả năng cài đặt mã độc đánh cắp thông tin như Atomic Stealer lên thiết bị của người dùng, gây mất an ninh mạng, an ninh dữ liệu.

(3) Nguy cơ từ các cuộc tấn công lừa đảo Deepfake, trong đó phổ biến nhất là thủ đoạn sử dụng AI để tạo ra hình ảnh, video và giọng nói giả mạo với

độ chân thực cao để mạo danh cơ quan Công an, công chức nhà nước hoặc người thân để lừa đảo chiếm đoạt tài sản, ảnh hưởng trực tiếp đến an ninh trật tự.

2. Từ tình hình trên, để tránh nguy cơ lộ mất tài liệu nội bộ, bí mật nhà nước trong quá trình sử dụng AI và tăng cường công tác bảo đảm an ninh mạng, an ninh dữ liệu trên địa bàn thành phố Hà Nội, CATP đề nghị các đơn vị quán triệt cán bộ, công chức, viên chức thực hiện:

(1) Không cung cấp dữ liệu cá nhân, thông tin tài liệu nội bộ, hồ sơ công vụ, tài liệu bí mật nhà nước cho các hệ thống AI công cộng trên mạng Internet.

(2) Chấp hành nghiêm các quy định của pháp luật về bảo đảm an ninh mạng, bảo vệ dữ liệu cá nhân và ứng dụng trí tuệ nhân tạo, đặc biệt là các quy định tại Điều 7, Luật Trí tuệ nhân tạo, trong đó nghiêm cấm một số hành vi như sau: [(1)]Tạo ra hoặc phổ biến nội dung giả mạo có khả năng gây nguy hại nghiêm trọng đến an ninh quốc gia, trật tự, an toàn xã hội; [(2)]Thu thập, xử lý hoặc sử dụng dữ liệu để phát triển, huấn luyện, kiểm thử hoặc vận hành hệ thống AI trái với quy định của pháp luật về dữ liệu, bảo vệ dữ liệu cá nhân, an ninh mạng…

(3) Tuyệt đối không cài đặt các AI Agent mã nguồn mở trên máy tính có kết nối mạng nội bộ của đơn vị khi chưa được cơ quan chức năng kiểm tra, đánh giá an ninh, an toàn.

(4) Phổ biến, quán triệt tới cán bộ, công chức, viên chức áp dụng cơ chế xác thực đa lớp cho thiết bị điện tử và cần xác minh thông tin khi nhận được các yêu cầu, chỉ đạo nhạy cảm qua ứng dụng mạng xã hội nhằm phòng, chống lừa đảo bằng Deepfake.

(5) Tổ chức rà soát toàn bộ các ứng dụng AI đang triển khai tại cơ quan, đơn vị *(nếu có)* và gửi danh mục về CATP trước ngày 20 **/4/2026** *(qua Phòng An ninh mạng & phòng, chống tội phạm sử dụng công nghệ cao; địa chỉ: Số 55 Lý Thường Kiệt, phường Cửa Nam, TP Hà Nội; đầu mối liên hệ: Đồng chí Hoàng Triều Dương, số điện thoại: 0986.837.519)* để tổng hợp, đánh giá cấp độ an toàn hệ thống thông tin theo quy định.

Nhận được Công văn này, đề nghị Thủ trưởng các cơ quan, đơn vị khẩn trương triển khai thực hiện./.

Nơi nhận:
- Như trên;
- Đ/c Giám đốc CATP (để báo cáo);
- Lưu: VT, ANM&PCTPSDCNC(Đ5).D.04b.

KT. GIÁM ĐỐC
PHÓ GIÁM ĐỐC

CỘNG HOÀ X.H.C.N VIỆT NAM
CÔNG AN THÀNH PHỐ HÀ NỘI

Đại tá Nguyễn Tiến Đạt