TẠP CHÍ NGHIÊN CỨU
PHẬT HỌC
CƠ QUAN NGÔN LUẬN CỦA TRUNG ƯƠNG GIÁO HỘI PHẬT GIÁO VIỆT NAM
ISSN: 2734-9195

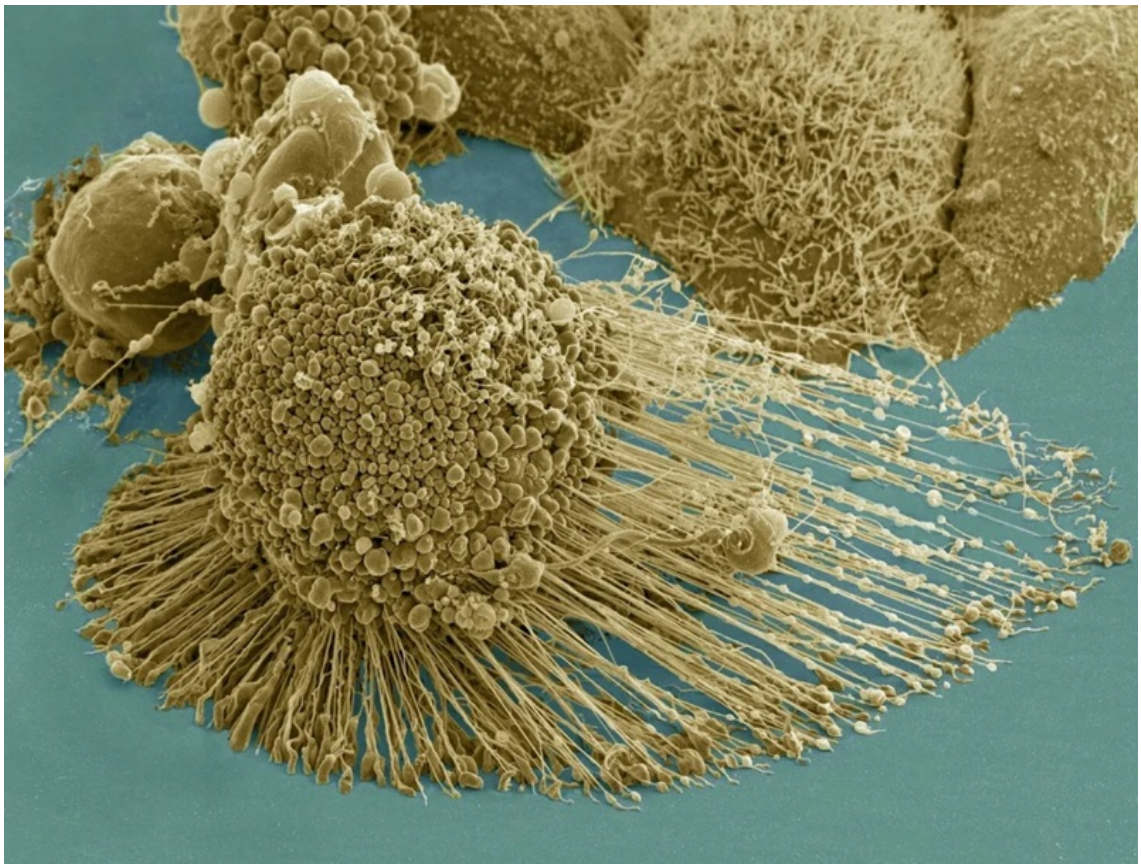
AI và bài toán bảo mật trong kỷ nguyên lượng tử

ISSN: 2734-9195 15:10 31/03/2026

Một hệ thống AI thiếu bảo mật không chỉ là lỗi kỹ thuật, mà còn là biểu hiện của: Thiếu tỉnh thức. Thiếu trách nhiệm với tha nhân. Và đôi khi là lòng tham (khai thác dữ liệu vô độ).

Lời tòa soạn

*Sự phát triển nhanh chóng của **trí tuệ nhân tạo (AI)** không chỉ mở ra những khả năng chưa từng có trong xử lý dữ liệu và tối ưu hóa hệ thống, mà còn đặt ra những thách thức sâu sắc về bảo mật, quyền riêng tư và đạo đức.*



Nguồn ảnh: “Ảnh hiển vi điện tử quét của một tế bào HeLa đang trải qua quá trình chết theo chương trình (apoptosis)” do Viện Y tế Quốc gia Hoa Kỳ (NIH) thực hiện, được cấp phép theo giấy phép CC BY-NC

Trong bối cảnh công nghệ lượng tử đang dần hình thành, những chuẩn mực bảo mật hiện nay có nguy cơ trở nên lỗi thời. Bài viết dưới đây không chỉ cung cấp một cái nhìn kỹ thuật về an ninh AI, mà còn gợi mở những suy tư từ góc độ Phật học, nơi trách nhiệm, chính niệm và trí tuệ đóng vai trò nền tảng trong việc kiến tạo tương lai công nghệ bền vững.

Bảo mật AI hiện tại và tương lai gần

Một nghiên cứu được công bố trong eBook AI Quantum Resilience của Utimaco cho thấy, các tổ chức hiện nay xem rủi ro bảo mật là rào cản lớn nhất đối với việc triển khai AI trên dữ liệu nội bộ.

(Ảnh: Internet)

Giá trị của AI phụ thuộc trực tiếp vào khối dữ liệu mà tổ chức tích lũy. Tuy nhiên, chính nguồn dữ liệu này lại trở thành điểm yếu dễ bị khai thác trong quá trình xây dựng và huấn luyện mô hình.

(Ảnh: Internet)

Các nguy cơ chính bao gồm:

+ Dữ liệu huấn luyện bị thao túng: khiến kết quả đầu ra sai lệch một cách khó phát hiện.
+ Mô hình bị sao chép hoặc đánh cắp: làm suy giảm giá trị sở hữu trí tuệ.

+ Dữ liệu nhạy cảm bị lộ: trong quá trình huấn luyện hoặc suy luận (inference).

Đáng chú ý, những rủi ro này không chỉ xuất hiện ở giai đoạn vận hành mà trải dài suốt vòng đời của hệ thống AI từ thu thập dữ liệu đến triển khai thực tế.

Nguy cơ từ điện toán lượng tử: “cơn sóng ngầm”

Theo các chuyên gia, trong vòng khoảng 10 năm tới, mã hóa khóa công khai hiện nay có thể bị phá vỡ khi các hệ thống lượng tử đủ mạnh xuất hiện.

(Ảnh: Internet)

Thậm chí, một số tổ chức đã bắt đầu:

+ Thu thập dữ liệu đã mã hóa hôm nay.

+ Lưu trữ lại.

+ Chờ đến khi có công nghệ lượng tử để giải mã.

Điều này đặt ra một vấn đề nghiêm trọng: Những dữ liệu “an toàn hôm nay” có thể trở thành “minh bạch ngoài ý muốn” trong tương lai.

(Ảnh: Internet)

Các loại dữ liệu cần bảo vệ dài hạn bao gồm:

+ Dữ liệu huấn luyện AI.

+ Hồ sơ tài chính.

+ Tài sản trí tuệ.

+ Thông tin cá nhân.

Giải pháp kỹ thuật: Từ “crypto-agility” đến phần cứng tin cậy

Để đối phó với những rủi ro này, Utimaco đề xuất một số hướng đi quan trọng:

Crypto-agility (tính linh hoạt mật mã): Cho phép thay đổi thuật toán mã hóa mà không cần thiết kế lại toàn bộ hệ thống.

(Ảnh: Internet)

Cách tiếp cận này thường sử dụng: Mã hóa truyền thống; Kết hợp với các thuật toán hậu lượng tử (post-quantum).

Thiết bị phần cứng đáng tin cậy

Các thiết bị này giúp:

+ Cô lập khóa mã hóa và dữ liệu nhạy cảm

+ Ngăn chặn truy cập ngay cả từ quản trị viên hệ thống

+ Xác minh tính toàn vẹn của mô hình trước khi triển khai

+ Tạo “chuỗi niềm tin” (chain of trust) từ phần cứng đến ứng dụng

Ngoài ra, chúng còn:

+ Ghi lại nhật ký chống giả mạo

+ Hỗ trợ tuân thủ các khung pháp lý như EU AI Act

Bảo mật toàn bộ vòng đời AI

Từ:

+ Thu thập dữ liệu

+ Huấn luyện mô hình

+ Triển khai

+ Suy luận trong môi trường thực

Tất cả đều cần được bảo vệ nhất quán.

Góc nhìn Phật học: Bảo mật không chỉ là kỹ thuật, mà là “giữ giới” trong kỷ nguyên số

Hình minh họa tạo bởi AI

Nếu nhìn từ Phật học, vấn đề bảo mật không đơn thuần là bài toán công nghệ, mà còn là biểu hiện của:

+ Chính niệm: nhận biết rõ hậu quả của hành vi công nghệ

+ Trách nhiệm (nghiệp): mọi hành động xử lý dữ liệu đều để lại hệ quả

+ Giới (sīla): bảo vệ dữ liệu cũng là bảo vệ niềm tin của con người.

Trong Kinh Pháp Cú, phẩm Song Yếu, kệ 01 có dạy:

“Ý dẫn đầu các pháp

Ý làm chủ, ý tạo...”

Trong bối cảnh AI, có thể hiểu rằng: Tâm ý của người thiết kế hệ thống sẽ định hình “đạo đức” của công nghệ đó.

Một hệ thống AI thiếu bảo mật không chỉ là lỗi kỹ thuật, mà còn là biểu hiện của:

+ Thiếu tỉnh thức.

+ Thiếu trách nhiệm với tha nhân.

+ Và đôi khi là lòng tham (khai thác dữ liệu vô độ).

Hàm ý cho hiện tại: Chuẩn bị từ hôm nay

Hình minh họa tạo bởi AI

Dù rủi ro từ điện toán lượng tử chưa xảy ra ngay lập tức, nhưng những quyết định về dữ liệu và hạ tầng hôm nay sẽ quyết định mức độ an toàn trong tương lai.

Các tổ chức cần:

+ Tăng cường kiểm soát toàn bộ vòng đời AI

+ Triển khai crypto-agility

+ Áp dụng các cơ chế bảo mật dựa trên phần cứng

Thay lời kết: Khi công nghệ đối diện với tâm thức

Trong kỷ nguyên AI, câu hỏi không còn là “có thể bảo mật hay không”, mà là:

+ Chúng ta có đủ tỉnh thức để bảo vệ dữ liệu như bảo vệ chính mình?

+ Khi AI ngày càng mạnh, ai sẽ chịu trách nhiệm cho những hệ quả vô hình mà AI tạo ra?

+ Liệu một hệ thống an toàn về kỹ thuật có thể thiếu an toàn về đạo đức?

Theo: **AI News**/Chuyển ngữ và biên tập: **Hoa Mạn**