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

Quản trị AI cần được soi chiếu bởi "trí tuệ tỉnh thức"

ISSN: 2734-9195 09:05 10/04/2026

Điều này gợi mở liên tưởng sâu sắc trong Phật học: Giống như "chính niệm", sự tỉnh thức liên tục trong từng hành vi, hệ thống AI cũng cần một cơ chế "tỉnh thức kỹ thuật số" để không trượt khỏi mục đích ban đầu.

Lời toà soạn

*Sự phát triển nhanh chóng của **trí tuệ nhân tạo** (AI), đặc biệt là các hệ thống có khả năng tự chủ (agentic AI), đang đặt ra những câu hỏi không chỉ về kỹ thuật mà còn về đạo đức, trách nhiệm và nhận thức. Khi máy móc không còn chỉ "trả lời" mà bắt đầu "hành động", ranh giới giữa công cụ và chủ thể trở nên mong manh hơn bao giờ hết.*



(Ảnh: Internet)

Bài viết dưới đây, dựa trên phân tích từ thực tiễn doanh nghiệp, đồng thời được soi chiếu dưới góc nhìn Phật học, nhằm góp phần gợi mở một cách tiếp cận cân bằng: giữa công nghệ và con người, giữa năng lực và trách nhiệm, giữa trí tuệ nhân tạo và trí tuệ tỉnh thức.

Từ công cụ đến “tác nhân”: AI đang thay đổi bản chất vận hành

Các hệ thống AI đang dần vượt ra khỏi vai trò truyền thống là những công cụ phản hồi đơn thuần. Trong nhiều tổ chức, AI không chỉ phân tích dữ liệu hay tạo nội dung, mà còn có thể:



(Ảnh: Internet)

+ Lập kế hoạch công việc

+ Tự đưa ra quyết định

+ Thực thi hành động với mức độ can thiệp của con người rất hạn chế.

Sự chuyển dịch này đánh dấu bước tiến sang AI tác nhân (agentic AI), nơi hệ thống có thể tự phân tách mục tiêu thành các bước, lựa chọn phương án và tương tác với các hệ thống khác để hoàn thành nhiệm vụ.

Tuy nhiên, khi AI được trao quyền “hành động”, câu hỏi không còn là “AI có trả lời đúng không?” mà là: “Điều gì sẽ xảy ra khi AI tự quyết định và hành động trong thế giới thực?”

Quản trị AI: Không phải lựa chọn, mà là điều kiện tất yếu

Theo nghiên cứu của Deloitte, khoảng 23% doanh nghiệp đã triển khai AI tác nhân, và con số này có thể đạt 74% trong vòng hai năm tới. Tuy nhiên, chỉ 21% trong số đó có hệ thống kiểm soát đủ mạnh.

Điều này cho thấy một thực tế đáng lưu ý: Sự phát triển của AI đang nhanh hơn khả năng kiểm soát của con người.



(Ảnh: Internet)

Quản trị AI (AI governance) vì vậy không thể là bước "bổ sung sau cùng", mà cần được tích hợp xuyên suốt vòng đời hệ thống:

1. Giai đoạn thiết kế

+ Xác định rõ phạm vi hành động của AI

+ Thiết lập giới hạn truy cập dữ liệu

+ Định nghĩa cách xử lý trong tình huống bất định

2. Giai đoạn triển khai

+ Kiểm soát quyền truy cập

+ Xác định kết nối hệ thống

+ Phân quyền người dùng

3. Giai đoạn vận hành

+ Giám sát liên tục

+ Ghi lại hành vi và quyết định

+ Phát hiện sai lệch theo thời gian

Một hệ thống AI nếu không được "định giới" rõ ràng, dù được huấn luyện tốt, vẫn có thể tạo ra những hệ quả khó phát hiện hoặc không thể đảo ngược.

Minh bạch và trách nhiệm: Ai chịu trách nhiệm khi AI hành động?



(Ảnh: Internet)

Khi AI bắt đầu đảm nhận vai trò ra quyết định, việc truy vết nguồn gốc hành động trở nên khó khăn hơn. Điều này dẫn đến nhu cầu cấp thiết về: Tính minh bạch (transparency). Trách nhiệm giải trình (accountability).

Các hệ thống cần có khả năng: Ghi lại toàn bộ hành vi, Lưu trữ quyết định, Cho phép truy vết khi có sự cố.

Câu hỏi cốt lõi đặt ra: Nếu AI sai, trách nhiệm thuộc về ai - người lập trình, tổ chức hay chính hệ thống?

Giám sát thời gian thực: “Chính niệm công nghệ” trong vận hành AI

Một trong những hướng tiếp cận quan trọng là giám sát theo thời gian thực (real-time monitoring). Thay vì chỉ dựa vào các quy tắc tĩnh, hệ thống được theo dõi liên tục khi hoạt động.



(Ảnh: Internet)

Trong thực tế, AI đã được ứng dụng để:

+ Theo dõi thiết bị

+ Phát hiện sớm sự cố

+ Kích hoạt quy trình bảo trì tự động

Các hành động này có thể diễn ra xuyên suốt nhiều hệ thống, nhưng với người dùng, chỉ như một thao tác duy nhất.

Điều này gợi mở liên tưởng sâu sắc trong Phật học: Giống như “chính niệm”, sự tỉnh thức liên tục trong từng hành vi, hệ thống AI cũng cần một cơ chế “tỉnh thức kỹ thuật số” để không trượt khỏi mục đích ban đầu.

Khi trí tuệ không đi cùng tỉnh thức

Từ quan điểm Phật học, vấn đề cốt lõi không nằm ở việc AI mạnh đến đâu, mà ở chỗ: Năng lực hành động có đi kèm với trí tuệ và từ bi hay không?



(Ảnh: Internet)

Một số điểm phản biện có thể đặt ra:

1. AI có “trí” nhưng không có “tuệ”

AI có thể xử lý dữ liệu và tối ưu quyết định, nhưng không có khả năng:

+ Quán chiếu hậu quả nghiệp

+ Nhận thức khổ đau của con người

+ Phát sinh lòng từ bi



(Ảnh: Internet)

2. Thiếu “chủ thể chịu trách nhiệm nghiệp”

Trong Phật giáo, mọi hành động đều gắn với nghiệp và chủ thể tạo nghiệp.

AI, dù hành động, lại không phải là chủ thể đạo đức thực sự. → Điều này khiến việc quy trách nhiệm trở nên mơ hồ.

3. Nguy cơ “phó thác tâm trí”

Khi con người dần giao quyền quyết định cho AI, có thể phát sinh:

+ Sự lệ thuộc

+ Suy giảm năng lực tự quán chiếu

+ Đánh mất vai trò của “chánh kiến”.

Thay lời kết: Trí tuệ nào dẫn đường cho tương lai?

Khi AI ngày càng “thông minh” và tự chủ, câu hỏi không còn nằm ở công nghệ, mà nằm ở con người:

+ Chúng ta đang trao quyền cho AI, hay đang vô tình trao đi trách nhiệm của chính mình?

+ Liệu một hệ thống không có từ bi có thể phục vụ một thế giới đầy khổ đau?

+ Sau cùng, trí tuệ nào mới thực sự dẫn đường: trí tuệ nhân tạo hay trí tuệ tỉnh thức?

Tác giả: **Muhammad Zulhusni/**Chuyển ngữ và biên tập: **Hoa Mạn**

Nguồn: artificialintelligence-news.com