



Cuộc đua công nghệ không tưởng hay con đường trung đạo với AI

ISSN: 2734-9195

18:02 23/02/2025

AI không thể tách rời khỏi giá trị nhân văn, cấu trúc xã hội và trách nhiệm đạo đức. Thay vì chờ đợi khủng hoảng xảy ra, chúng ta có cơ hội hành động ngay từ bây giờ với trí tuệ, nhãn nại và từ bi.

Hội nghị thượng đỉnh về AI tại Paris (10-11/02/2025) đã để lộ một thực tế quen thuộc: Các quốc gia và tập đoàn lớn ráo riết khai thác tiềm năng kinh tế của AI trong khi chỉ dừng lại ở mức tuyên bố mơ hồ về giám sát an toàn.



Bản tuyên bố cuối cùng của hội nghị này không đưa ra tiêu chuẩn an toàn rõ ràng, mà tập trung nhiều hơn vào các cơ hội kinh tế.

Trái ngược với các hội nghị trước đó, nơi các cam kết về thử nghiệm và điều chỉnh AI đã được xác lập, hội nghị lần này chỉ dừng lại ở những lời kêu gọi chung chung về việc xây dựng AI "đáng tin cậy."

Sự miễn cưỡng trong việc ưu tiên an toàn thay vì phát triển nhanh chóng là điều đáng lo ngại. Khi các quốc gia chạy đua thống lĩnh AI, họ đang bỏ qua những bài học lịch sử về rủi ro của tham vọng công nghệ không kiểm soát. Đây không

chỉ là vấn đề tăng trưởng ngành hay cạnh tranh thị trường, mà còn là câu chuyện về đạo đức, trách nhiệm và tương lai chung của nhân loại.

Khát vọng bất tử và ảo tưởng kiểm soát

Nhà giáo thọ Phật giáo đương đại - Thanissara nhận định rằng cơn sốt AI hiện nay không phải là điều mới mẻ, mà là một phần của khát vọng cổ xưa - mong muốn vượt lên giới hạn của tự nhiên và đạt đến sự bất tử.

Trong lịch sử, mong muốn này đã biểu hiện dưới nhiều hình thức khác nhau. Ở Ấn Độ cổ đại, tầng lớp tu sĩ Bà-la-môn thực hành hiến tế để cầu mong sự giải thoát vĩnh hằng. Tại Trung Hoa thời Trung cổ, các đạo sĩ luyện kim tìm kiếm phương thuốc trường sinh, nhưng nhiều người đã tử vong vì chính những hợp chất chứa thủy ngân mà họ chế tạo.

Ngày nay, những người đứng đầu ngành công nghệ - "tech bros" - đang theo đuổi một dạng bất tử khác: tải ý thức lên không gian số, kéo dài tuổi thọ bằng biến đổi gen, hay thậm chí chạy trốn lên sao Hỏa để thiết lập một nền văn minh liên hành tinh dành riêng cho tầng lớp tinh hoa.

Nhưng từ góc nhìn Phật giáo, đây là một sự hiểu lầm căn bản - ảo tưởng về sự thường hằng (anitya). Đức Phật dạy rằng tất cả pháp hữu vi, bao gồm cả sự sống và ý thức con người, đều vô thường. Không một công nghệ nào có thể thay đổi quy luật này. Thay vì bám chấp vào kiểm soát, Phật giáo khuyến khích con người chấp nhận và làm việc với tính vô thường, hướng đến trí tuệ và lòng từ bi, thay vì chạy theo những giấc mơ siêu việt hóa bằng công nghệ.

Hệ quả của công nghệ không kiểm soát

Lịch sử đã chứng minh rằng những công nghệ đột phá thường đi kèm với hậu quả khó lường. "Đại sử Mahavamsa" của Sri Lanka kể về vua Dutthagamani, người khởi binh vì lý tưởng cao đẹp, nhưng cuối cùng phải đối diện với cảm giác tội lỗi nặng nề sau chiến thắng. Câu chuyện này lặp lại nhiều lần trong lịch sử: con người theo đuổi quyền lực mà không lường trước hệ quả, để rồi sau đó phải tìm cách biện minh cho những sai lầm.

AI cũng đang bước vào một quỹ đạo tương tự. Những thử nghiệm nhỏ như dùng AI để đi chợ đã cho thấy nhiều sai sót nghiêm trọng. Khi AI được triển khai trong các lĩnh vực quan trọng như thực thi pháp luật, y tế hay tài chính, hậu quả sẽ còn lớn hơn. Một số thất bại thương mại gần đây đã cho thấy điều này:

+ *McDonald's phải từ bỏ hệ thống AI đặt món tự động tại các cửa hàng drive-through.*

+ *Air Canada bị buộc bồi thường vì chatbot AI đưa ra thông tin sai lệch cho khách hàng trong tình huống nhạy cảm.*

Nếu AI đã có thể hành xử khó đoán chỉ với các nhiệm vụ đơn giản, điều gì sẽ xảy ra khi nó được trao quyền quyết định trong những lĩnh vực sống còn?

Trí tuệ cần trọng - con đường Trung đạo cho AI

Đức Phật luôn cảnh báo về niềm tin mù quáng vào các giải pháp bên ngoài. Trong Kinh Kalama, Ngài khuyên không nên chấp nhận điều gì chỉ vì đó là truyền thống hay vì nó đến từ một nguồn có thẩm quyền. Nguyên tắc này cũng nên áp dụng cho AI: Chúng ta có thực sự kiểm soát được công nghệ này không? Hậu quả dài hạn đã được đánh giá kỹ chưa? Có cơ chế bảo vệ nào khi AI gây ra sai sót?

Thay vì chia cuộc tranh luận thành hai cực đoan - kiểm soát hay phát triển, an toàn hay tiến bộ - Phật giáo gợi ý một cách tiếp cận khác: Trung đạo.

Việc quản lý AI không phải là cản trở đổi mới, mà là một hình thức giới luật tinh thức (sila) giúp cân bằng lợi ích và đạo đức.

Tương tự như Luật Tỳ-kheo (Vinaya) - được đặt ra nhằm duy trì kỷ cương cho Tăng đoàn mà không kìm hãm sự thực hành - các quy tắc AI có thể tạo ra một môi trường ổn định, đáng tin cậy cho công nghệ phát triển.

5 nguyên tắc cho AI dựa trên tinh thần Phật giáo

Hội nghị AI tại Paris có thể chưa đạt được bước tiến quan trọng, nhưng sự kiện tiếp theo tại Kigali, Rwanda là cơ hội để thiết lập những nền tảng ý nghĩa hơn.

Một cách tiếp cận dựa trên Phật giáo có thể bao gồm:

1. Xây dựng AI có cam kết đạo đức rõ ràng. Như năm giới của Phật tử, AI cần tuân thủ các nguyên tắc cốt lõi: không gây hại, minh bạch, trách nhiệm giải trình, và công bằng. Điều này đòi hỏi những tiêu chuẩn thực tiễn thay vì những tuyên bố chung chung về "AI đáng tin cậy."

2. Thúc đẩy hợp tác quốc tế thay vì chạy đua cạnh tranh. Giống như các kỳ kết tập kinh điển sau khi Đức Phật nhập Niết-bàn, AI cần một nền tảng hợp tác toàn cầu để tránh rủi ro từ chủ nghĩa dân tộc công nghệ.

3. Chậm lại không phải là thất bại, mà là trí tuệ. Trong thiền định, chậm lại giúp trí tuệ sáng tỏ hơn. AI cũng vậy, sự cẩn trọng không phải là phản đối tiến bộ, mà là một cách tiếp cận khôn ngoan.

4. Thiết lập các ngưỡng kiểm soát AI rõ ràng. Khi AI đạt đến các cấp độ tự duy nhất định, cần có thêm các vòng kiểm tra và giám sát, tương tự như các giai đoạn đào luyện trong Phật giáo, khi trách nhiệm đạo đức tăng theo mức độ phát triển.

5. Chuyển hướng khỏi chủ nghĩa cá nhân cực đoan. Giấc mộng "thần thánh hóa" AI của giới tinh hoa công nghệ phản ánh sự chấp ngã sai lầm. Phật giáo nhấn mạnh vô ngã (anatman), sự giải thoát không nằm ở việc trốn chạy, mà ở nhận thức về tính tương tức của tất cả chúng sinh.

Hướng tới tương lai AI cân bằng

AI không thể tách rời khỏi giá trị nhân văn, cấu trúc xã hội và trách nhiệm đạo đức. Thay vì chờ đợi khủng hoảng xảy ra, chúng ta có cơ hội hành động ngay từ bây giờ với trí tuệ, nhẫn nại và từ bi.

Trung đạo chính là con đường khả thi nhất, dung hòa sáng tạo với trách nhiệm, tham vọng với cẩn trọng, và công nghệ với sự minh triết đạo đức.

Viết dịch: **Thích Vân Phong**

Nguồn: **www.buddhistdoor.net**