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

Thực hành Phật giáo trong thời đại trí tuệ nhân tạo (P.1/6)

ISSN: 2734-9195 15:15 28/10/2026

AI không tồn tại như một công nghệ biệt lập mà vận hành trong mạng lưới những mối quan hệ đan xen, tác động qua lại giữa dữ liệu, thuật toán, con người và các lĩnh vực của đời sống.

Thực hành Phật giáo trong thời đại trí tuệ nhân tạo (Phần 1/6): Nhìn nhận AI từ bức tranh toàn cảnh

1. Bức tranh toàn cảnh

Trí tuệ nhân tạo (AI) đang trở thành một trong những vấn đề được quan tâm hàng đầu hiện nay. Sự phát triển nhanh chóng của AI đặt ra câu hỏi: AI có thể làm được những gì, sẽ tác động ra sao đến đời sống con người và đâu là cái giá phải trả cho những lợi ích mà nó mang lại?

Mỗi người có thể đưa ra những nhận định khác nhau, nhưng tất cả đều đứng trước nhu cầu hiểu biết sâu sắc hơn về bản chất và hệ quả của công nghệ này. Nhìn nhận những biến chuyển ấy từ góc độ Phật giáo, trên cơ sở đó tìm kiếm phương thức ứng xử phù hợp, có trí tuệ và lòng từ bi, chính là nội dung xuyên suốt của loạt bài gồm sáu phần này.

Trước khi tiếp cận những vấn đề phức tạp của bức tranh toàn cảnh, cần bắt đầu từ thực tế ứng dụng AI trong lĩnh vực Phật giáo hiện nay.

Cho đến nay, việc **ứng dụng AI** trong Phật giáo chủ yếu tập trung vào hoạt động dịch kinh điển, văn bản Phật học sang tiếng Anh. Bên cạnh đó là những cuộc thảo luận về bản chất của ý thức và cách định nghĩa trí thông minh. Khi được tích hợp vào các hệ thống chatbot hoặc robot, AI có thể tái hiện tương đối chính xác nội dung giáo lý Phật giáo thông qua ngôn ngữ.

Tuy nhiên, những ứng dụng này nhìn chung chưa đạt đến trình độ phát triển thực sự đột phá và cũng chưa được triển khai rộng rãi.

Theo quan điểm của tác giả, nếu nhìn nhận một cách tích cực, đây mới chỉ là những nỗ lực còn đơn giản, thiếu chiều sâu. Ở góc độ đáng lo ngại hơn, chúng có thể trở thành biểu hiện của xu hướng né tránh thực tại, khi chỉ tập trung vào việc tái hiện và truyền đạt giáo lý mà chưa quan tâm đúng mức đến toàn bộ hệ sinh thái AI cùng những hệ lụy sâu rộng của nó.

Cách tiếp cận ấy phần nào phản ánh hình thức thực hành Phật giáo thiên về hoài niệm, thay vì chủ động xây dựng tầm nhìn dài hạn và vận dụng những phương tiện thiện xảo để ứng phó với các thách thức hiện hữu, cũng như những khổ đau có thể phát sinh trong tương lai gần.

Ở đây, cần làm rõ một nhận định của tác giả: "*Tôi không cho rằng nguy cơ AI vượt khỏi sự kiểm soát, dẫn đến sự tuyệt chủng của nhân loại, dù được một số ý kiến ước tính ở mức 10%, là mối quan ngại cần đặt lên hàng đầu, bất chấp việc kịch bản này đang thu hút sự chú ý đáng kể của truyền thông. Điều đáng lo ngại hơn, theo tôi, là một nguy cơ khác, được đánh giá ở mức trên 90% nhưng lại ít được truyền thông quan tâm tương xứng: chính con người, khi còn bị chi phối bởi tham dục, sân hận và vô minh trong vòng luân hồi, có thể trở nên nguy hiểm hơn rất nhiều nếu được tiếp sức bởi sức mạnh của AI. Trong lịch sử, con người không ít lần hành động thiếu lý trí, gây ra những hậu quả nghiêm trọng. Vì vậy, ít nhất chúng ta cần nỗ lực ngăn chặn việc những công cụ có sức mạnh lớn rơi vào tay những người có khả năng sử dụng chúng để gây tổn hại. Thế nhưng, thực tế đang diễn ra lại chưa cho thấy điều đó.*"

Hình ảnh chỉ mang tính chất minh hoạ. Ảnh sưu tầm.

Một số diễn biến được phản ánh trên **truyền thông** có thể minh họa cho những nguy cơ cần quan tâm.

Công nghệ nhận diện khuôn mặt đang được triển khai tại nhiều quốc gia, trong đó có việc sử dụng tại lối vào các cơ sở tôn giáo ở Trung Quốc. Tại Hoa Kỳ, hệ thống camera Flock, với hơn 120.000 thiết bị được lắp đặt, có khả năng theo dõi hoạt động của con người tại nhiều khu vực đô thị.

Phần mềm InquiryIQ của Clearview có thể rà soát hoạt động trực tuyến, từ đó xây dựng hồ sơ thông tin về một cá nhân và những người có liên hệ với họ.

Trong lĩnh vực an ninh mạng, công ty Calif được cho là đã phát triển WeWorm, một phương thức khai thác lỗ hổng bảo mật chưa được khắc phục, có khả năng xâm nhập và kiểm soát điện thoại thông minh sử dụng WeChat mà không cần người dùng thực hiện thao tác phản hồi.

Trong giáo dục đại học, nhiều trường ghi nhận tình trạng sinh viên bỏ học và gian lận học thuật gia tăng, đồng thời kết quả đánh giá năng lực học tập có xu hướng suy giảm khi việc sử dụng AI ngày càng phổ biến.

Sự xuất hiện của công nghệ tạo nội dung giả mạo bằng AI (deepfake) cùng lượng lớn sản phẩm AI kém chất lượng cũng đang khiến ranh giới giữa thông tin xác thực, nội dung hư cấu và thông tin sai lệch có chủ đích ngày càng khó phân biệt. Đáng chú ý, ngay cả hình ảnh và lời giảng của những người hướng dẫn tâm linh cũng có thể bị giả mạo hoàn toàn.

Không chỉ tác động đến đời sống xã hội, AI còn đặt ra những vấn đề môi trường đáng quan tâm. Các trung tâm dữ liệu đang phát triển nhanh chóng, trong khi nhiều cơ sở vẫn phụ thuộc vào nguồn năng lượng từ nhiên liệu hóa thạch và tiêu thụ lượng nước sạch rất lớn, thậm chí vượt mức sử dụng của toàn bộ cư dân tại những thị trấn nơi chúng được xây dựng.

Những hiện tượng trên cho thấy, để hiểu được bức tranh toàn cảnh về AI, cần nhận thức rằng quá trình thiết kế, phát triển và triển khai công nghệ này luôn gắn với hai câu hỏi căn bản: AI được tạo ra để phục vụ ai và được sử dụng để tác động lên những ai?

Cory Doctorow, một nhà bình luận có ảnh hưởng trong lĩnh vực công nghệ, đã sử dụng hình ảnh ẩn dụ "*nhân mã ngược*" (reverse centaur) để diễn giải mối quan hệ này.

Trong hình ảnh ấy, con người không còn là chủ thể điều khiển công nghệ mà trở thành bộ phận phải lao động và gánh chịu chi phí, trong khi quyền quyết định thực chất nằm trong tay những người thiết kế thuật toán. Nói cách khác, con người đảm nhận phần việc nặng nhọc và chịu những hệ quả phát sinh, còn hệ thống công nghệ quyết định phương thức vận hành.

Một cách gọi khác đang được sử dụng để mô tả tình trạng con người thực hiện mệnh lệnh của máy tính là "*người đại diện bằng xương thịt*" (meat puppet).

Một vấn đề nền tảng khác cần nhận thức là AI ngày nay không còn giới hạn ở các mô hình ngôn ngữ lớn (Large Language Models – LLMs).

Những thế hệ AI mới đang được phát triển theo hướng có khả năng chủ động thực hiện hành động, cho phép hệ thống không chỉ xử lý thông tin và đưa ra câu trả lời mà còn trực tiếp can thiệp vào thế giới thực.

Khả năng này có thể được sử dụng cho những công việc thông thường như đặt phòng khách sạn, nhưng cũng có thể bị khai thác để thực hiện các hành vi nguy hiểm, chẳng hạn can thiệp vào hoạt động của nhà máy xử lý nước tại một quốc gia khác.

Một vụ tấn công mạng gần đây được cho là có sự tham gia của AI đã khiến hơn 400 tủ lạnh thương mại tại các siêu thị trên khắp nước Pháp ngừng hoạt động.

Khả năng hành động trực tiếp của AI vì thế đặt ra những nguy cơ rõ ràng và hiện hữu. Đây cũng là một trong những nguyên nhân khiến nhiều tiếng nói cảnh báo về AI kêu gọi phải hết sức thận trọng trong quá trình phát triển và triển khai **công nghệ**.

Trong lĩnh vực này, vấn đề an toàn AI (AI safety) và sự căn chỉnh AI (AI alignment) đang được đặc biệt quan tâm. Tuy nhiên, theo nhận định của tác giả, những nỗ lực hiện nay vẫn chưa giải quyết hiệu quả khoảng cách giữa các giá trị nhân văn phổ quát với mục tiêu và định hướng của những doanh nghiệp phát triển AI.

Vấn đề nền tảng thứ ba là sự cạnh tranh giữa hai mô hình phát triển và phổ biến AI đang có ảnh hưởng lớn trên thế giới.

Hoa Kỳ đang tập trung vào việc triển khai những hệ thống AI có năng lực mạnh nhất, với công nghệ mang tính độc quyền và mức độ chuyên biệt hóa cao. Có thể hình dung mô hình này giống như một chiếc Rolls-Royce: công nghệ cao cấp, thiết kế riêng và đòi hỏi nguồn lực lớn.

Trong khi đó, Trung Quốc đang chú trọng hướng phát triển AI với trọng số mở (open-weight AI), có thể không đạt năng lực tương đương những hệ thống tiên tiến nhất nhưng tiết kiệm chi phí hơn và thuận lợi hơn cho việc phổ biến, ứng dụng rộng rãi. Mô hình này có thể được ví như một chiếc Volkswagen.

Từ phép so sánh đó, một câu hỏi được đặt ra: Mô hình nào sẽ có khả năng tiếp cận số đông người sử dụng hơn và tạo dựng được hệ thống hạ tầng rộng lớn hơn?

Cuộc cạnh tranh phát triển AI hiện chưa cho thấy một lối thoát rõ ràng, bất chấp những lời kêu gọi thận trọng và các nỗ lực xây dựng khuôn khổ pháp lý từ nhiều phía.

Có thể nói, công nghệ này đã bước ra khỏi những giới hạn ban đầu và rất khó quay trở lại trạng thái trước đây.

Trước thực tế ấy, mỗi người cần tự đặt câu hỏi: Việc thực hành Phật giáo của chính mình sẽ ứng xử như thế nào trước những biến chuyển này?

2. Siêu vật thể và Mạng lưới Indra

Timothy Morton, nhà triết học đương đại nổi tiếng với những nghiên cứu về sinh thái học, sử dụng khái niệm "*siêu vật thể*" (hyperobject) để chỉ những thực thể hoặc hiện tượng có quy mô phân bố rộng lớn trong không gian và thời gian, đến mức vượt khỏi khả năng nhận thức thông thường của con người về một đối tượng riêng biệt.

Hình ảnh chỉ mang tính chất minh hoạ về Mạng lưới Indra. Ảnh sưu tầm.

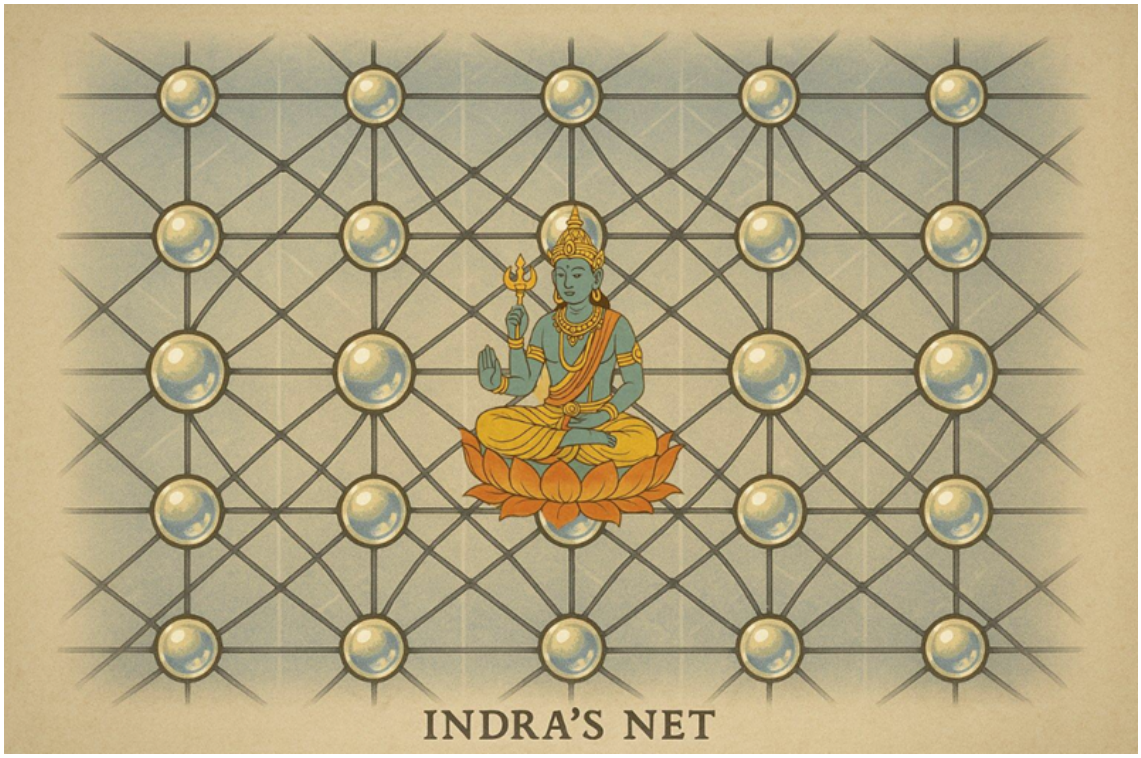
Theo Morton, các siêu vật thể có đặc tính mà ông gọi là "*tính nhớt*" (viscosity), nghĩa là chúng có khả năng tác động và gắn kết với những đối tượng mà chúng tiếp xúc.

Quy mô của các siêu vật thể lớn đến mức con người khó có thể tìm được một vị trí quan sát hoàn toàn khách quan để tách biệt và xác định chúng trong những giới hạn cụ thể của không gian và thời gian.

Tính toàn thể của chúng mang đặc điểm "*phi cục bộ*" (nonlocality), nghĩa là không thể hiểu đầy đủ bản chất của chúng chỉ thông qua một biểu hiện riêng lẻ.

Tùy thuộc vào vị trí và góc nhìn của người quan sát, siêu vật thể có thể được nhận thức theo những cách khác nhau, bởi bản thân chúng được hình thành từ mạng lưới quan hệ phức tạp giữa nhiều đối tượng.

Từ cách tiếp cận này, trí tuệ nhân tạo có thể được xem là một dạng siêu vật thể hoàn toàn mới.

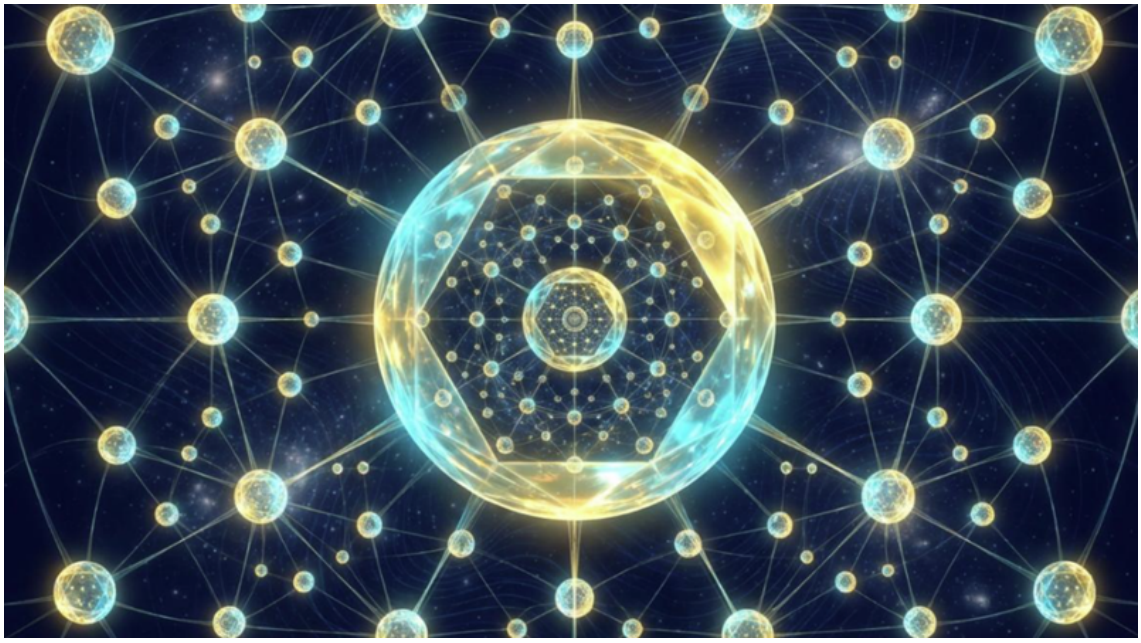


Hình ảnh chỉ mang tính chất minh hoạ về Mạng lưới Indra. Ảnh sưu tầm.

AI được hình thành trên nền tảng khối lượng dữ liệu khổng lồ, với nhiều loại hình và cấu trúc khác nhau, đồng thời có khả năng tác động đến hầu hết những lĩnh vực mà nó hiện diện.

Tính chất đa dạng và khó xác định ranh giới khiến AI không thể được giải thích đầy đủ bằng một cách tiếp cận đơn lẻ. Cấu trúc và cơ chế vận hành của nó phức tạp đến mức chỉ một số ít chuyên gia có thể hiểu sâu, trong khi những biểu hiện và tác động của AI lại ngày càng xuất hiện phổ biến trong đời sống.

Ở góc độ liên hệ với tư tưởng Phật giáo, hiện tượng này gợi liên tưởng đến hình ảnh Mạng lưới Indra trong truyền thống Phật giáo Đại thừa.



Hình ảnh chỉ mang tính chất minh hoạ về Mạng lưới Indra. Ảnh sưu tầm.

Trong Kinh Hoa Nghiêm thuộc truyền thống Phật giáo Đại thừa, hình tượng Mạng lưới Indra (Đế võng) được sử dụng để biểu đạt tính tương liên, tương tức và tương nhập giữa các hiện tượng. Mỗi mắt lưới được ví như một viên ngọc, trong đó phản chiếu vô số viên ngọc khác, tạo thành một chỉnh thể trùng trùng duyên khởi. Hình tượng này gợi mở cách nhìn về một thực tại mà các hiện tượng không tồn tại hoàn toàn biệt lập, mà luôn có quan hệ và tác động qua lại. Từ góc độ liên tưởng, hệ sinh thái trí tuệ nhân tạo cũng có thể được nhìn nhận như một mạng lưới phức hợp, trong đó dữ liệu, thuật toán, con người và các điều kiện xã hội liên tục tương tác, ảnh hưởng lẫn nhau.

Một trong những nỗ lực nhằm lý giải thực trạng này đến từ nhóm Phật giáo và Trí tuệ nhân tạo (Buddhism and AI – BAAI), một tổ chức mới được thành lập với bốn thành viên sáng lập có kinh nghiệm trong các lĩnh vực an toàn AI, lãnh đạo tổ chức phi lợi nhuận, nghiên cứu thiền định và hoạt động hướng dẫn, chăm sóc tinh thần theo Phật giáo.

Khung nghiên cứu do **BAAI** xây dựng tập hợp nhiều nguồn tư liệu, được phân chia thành sáu lĩnh vực có mối liên hệ và giao thoa với nhau: Trí tuệ nhân tạo hướng đến Phật giáo; Phật giáo hướng đến Trí tuệ nhân tạo; Nghiên cứu học thuật; Giảng dạy; Quản trị và Hệ sinh thái.

Mặc dù chưa phải là một hệ thống hoàn chỉnh, những tài liệu được công bố trên trang thông tin của BAAI vẫn cung cấp các gợi mở đáng chú ý, góp phần hình thành nền tảng cho những cuộc thảo luận sâu rộng hơn về mối quan hệ giữa Phật giáo và trí tuệ nhân tạo trong thời đại hiện nay.

Tác giả: **John Harvey Negru**
Chuyển ngữ: **Nguyễn Minh Anh**
Nguồn: **https://www.buddhistdoor.net/**

***

Mạng lưới Indra trong truyền thống Phật giáo Đại thừa

Mạng lưới Indra (tiếng Phạn: Indrajāla, Hán ngữ: □□□□ , thường gọi là Đế võng hoặc Lưới Đế Thích) là một hình ảnh biểu tượng nổi bật trong tư tưởng Phật giáo Đại thừa, đặc biệt gắn với triết học Hoa Nghiêm tông.

Hình ảnh này diễn tả tính tương tức, tương nhập và mối quan hệ tương liên giữa mọi hiện tượng trong vũ trụ.

Qua đó, làm sáng tỏ một trong những tư tưởng trọng yếu của triết học Hoa Nghiêm: Một hiện hữu không tồn tại hoàn toàn biệt lập, mà luôn có quan hệ với vô số hiện hữu khác.

1. Hình tượng Mạng lưới Indra

Theo hình tượng truyền thống, tại cung điện của trời Đế Thích (Indra) có một tấm lưới rộng lớn, trải dài vô tận. Tại mỗi mắt lưới đều có một viên ngọc sáng trong suốt. Điều đặc biệt là mỗi viên ngọc không chỉ phản chiếu những viên ngọc xung quanh mà còn phản chiếu tất cả những viên ngọc khác trong toàn bộ tấm lưới.

Trong mỗi hình ảnh phản chiếu ấy lại có vô số hình ảnh phản chiếu khác, tạo thành một hệ thống trùng trùng vô tận. Hình ảnh này biểu đạt ba tầng ý nghĩa:

- Tính tương liên: Mỗi sự vật đều tồn tại trong quan hệ với những sự vật khác.

- Tính tương tức, tương nhập: Một hiện tượng có thể phản ánh và liên hệ với toàn thể mà không đánh mất tính đặc thù của chính nó.

- Tính trùng trùng duyên khởi: Các hiện tượng hình thành và tác động qua lại trong những mạng lưới nhân duyên nhiều tầng, nhiều lớp, không thể quy giản thành một nguyên nhân đơn lẻ.

Trong triết học Hoa Nghiêm, tư tưởng này được diễn đạt qua những khái niệm như sự sự vô ngại pháp giới, nghĩa là các hiện tượng tuy có sự sai biệt nhưng không hoàn toàn ngăn ngại nhau trong tính tương dung, tương nhập. Điều cần lưu ý là Mạng lưới Indra không đơn thuần mô tả một mạng lưới kết nối vật lý.

Đây là một ẩn dụ triết học nhằm biểu đạt cách nhìn về thực tại theo tư tưởng Hoa Nghiêm.

2. Nguồn gốc kinh điển và cơ sở học thuật

Hình tượng Mạng lưới Indra gắn với Kinh Hoa Nghiêm (Avataṃsaka Sūtra) và được phát triển sâu sắc trong hệ thống triết học Hoa Nghiêm, đặc biệt qua tư tưởng của ngài Pháp Tạng (643–712).

Theo Stanford Encyclopedia of Philosophy, Mạng lưới Indra là một ẩn dụ tiêu biểu để giải thích sự phụ thuộc lẫn nhau và tính tương nhập giữa các hiện tượng.

Tuy nhiên, về phương diện nghiên cứu văn bản, cần phân biệt hình tượng được đề cập trong kinh điển với những cách diễn giải, hệ thống hóa về sau của các luận sư Hoa Nghiêm. Không nên mặc nhiên coi toàn bộ câu chuyện về những viên ngọc phản chiếu vô tận là một đoạn kinh văn nguyên văn trong Kinh Hoa Nghiêm.

3. Liên hệ với trí tuệ nhân tạo

Trong bài chuyển ngữ Thực hành Phật giáo trong thời đại trí tuệ nhân tạo, tác giả sử dụng hình tượng Mạng lưới Indra để liên hệ với tính phức hợp và sự phụ thuộc lẫn nhau trong hệ sinh thái AI.

Chẳng hạn, một hệ thống AI không chỉ bao gồm thuật toán mà còn phụ thuộc vào dữ liệu, hạ tầng điện toán, năng lượng, con người, doanh nghiệp và những điều kiện xã hội.

Những yếu tố này liên tục tương tác, tạo nên các hệ quả vượt ra ngoài phạm vi của một ứng dụng riêng lẻ. Tuy vậy, đây là sự liên hệ mang tính ẩn dụ, không phải sự đồng nhất giữa triết học Hoa Nghiêm và khoa học công nghệ hiện đại.