



AI và Phật giáo, điểm gặp gỡ của "ước mơ"?

ISSN: 2734-9195

15:25 23/09/2024

Robot tại chùa Kodaiji có tên Quán Âm (Kannon), được làm dựa trên hình dáng Bồ tát Quán Thế Âm, Quán Âm (Kannon) đang hàng ngày thuyết giảng các bài kinh cho du khách thập phương.

Tác giả: **Nữ Tiến sĩ Nitasha Kaul**

Viết dịch: **Thích Vân Phong**

Nguồn: Kuensel Online

Nhà văn, nhà báo, nhà thơ và là cựu chiến binh trong cuộc nội chiến ở Hoa Kỳ, Ambrose Bierce (1842-1914) lần đầu tiên xuất bản trên tờ The San Francisco Examiner vào ngày 16 tháng 4 năm 1899 một truyện ngắn suy đoán về bản chất của sự sống và trí thông minh có tựa đề 'Moxon's Master'. *Truyện ngắn* mô tả một người máy chơi cờ vua, giết chết người sáng tạo ra nó, đây là một trong những mô tả đầu tiên về một người máy trong văn học tiếng Anh, được viết trước khi từ 'robot' được sử dụng.

Trong đó một người đàn ông tên là Moxon tạo ra người máy chơi cờ vua với chủ nhân của mình. Khi Moxon thắng ván cờ, người máy đó nổi cơn giận, và cố gắng giết chết người phát minh ra nó. Người kể trong câu chuyện này nhắc đến định nghĩa 'sự sống' của triết gia thế kỷ 19, nhà lý thuyết chính trị tự do cổ điển; nhà lý thuyết xã hội học Anh, Herbert Spencer (1820-1903).

Triết gia Herbert Spencer suy ngẫm rằng cuộc sống là sự kết hợp rõ ràng của những thay đổi không đồng nhất, cả đồng thời và liên tiếp, tương ứng với sự cùng tồn tại và trình tự bên ngoài. Nói cách khác, không có gì về nó là duy nhất của con người.

Tôi hiểu rằng có thể chỉ đơn giản khi nghĩ về con người như những dạng sống dựa trên carbon, trong khi máy móc là những dạng sống dựa trên silicon.

Khi nhà văn Ambrose Bierce viết câu chuyện này, từ 'robot' vẫn chưa xuất hiện (nhà văn người Séc đã đặt ra nó vào năm 1920 trong một vở kịch), hơn một trăm năm sau, vào tháng 7 năm 2022, người ta đưa tin rằng một cậu robot chơi cờ vua ở Moscow, Nga đã mất bình tĩnh trước những động thái của một đối thủ

7 tuổi và phản ứng với cách túm lấy ngón tay của cậu bé, và làm gãy ngón tay đó.



Ảnh: St

Ngày nay, chúng ta đang chứng kiến cuộc cách mạng công nghiệp lần thứ 4 (4IR) hay công nghiệp 4.0 là một phương thức mô tả sự kết nối giữa thế giới vật lý và kỹ thuật số hiện đại, (sau năng lượng hơi nước, năng lượng điện, phát triển và CNTT), và trí tuệ nhân tạo (AI) hiện diện khắp mọi nơi. Từ công trình nghiên cứu khoa học đến kinh nghiệm của con người cho đến báo chí truyền thông, bối cảnh, ứng dụng ở khắp mọi nơi đều có mối quan tâm về trí tuệ nhân tạo (AI).

Các ứng dụng trí tuệ nhân tạo (AI) là một phần của quá trình ra quyết định trong chăm sóc sức khỏe, giao thông, hậu cần, tiếp thị, phương tiện truyền thông xã hội, tuyển dụng, giải trí, thực thi pháp luật, tài chính, quân đội, cảnh sát, an ninh, giáo dục, v.v.

Trong bài thuyết trình của tôi vào năm 2022 tại Hội nghị Vajrayana quốc tế lần thứ tư về 'Phật giáo hiện đại' được tổ chức tại Trung tâm Nghiên cứu Bhutan và Nghiên cứu GNH (trước đây là Trung tâm Nghiên cứu Bhutan), tôi đã lập luận về mối liên hệ cấp bách và có liên quan giữa Phật giáo và trí tuệ nhân tạo (AI).

Về bản chất, trí tuệ nhân tạo (AI) là một lĩnh vực có mục tiêu, có hệ thống là xây dựng cỗ máy có thể hoạt động theo những cách thông minh. Ngược lại với 'AI yếu' (hay AI hẹp), loại AI đặt ra mục tiêu nhỏ hơn, hướng tới một ứng dụng cụ thể và để hoạt động hiệu quả cần có sự tham gia của con người, AI mạnh (Strong AI), một lý thuyết về việc có thể tạo ra máy móc có trí thông minh ngang bằng với trí tuệ của con người hay Trí tuệ nhân tạo tổng quát (AGI), một

lĩnh vực nghiên cứu AI lý thuyết nhằm cố gắng tạo ra phần mềm có trí thông minh giống con người và có khả năng tự học và thậm chí Siêu trí tuệ nhân tạo (ASI - Artificial Super Intelligence): ASI là hệ thống AI có thể vượt xa trí thông minh của con người. Giai đoạn đầu tiên của quá trình phát triển trí tuệ nhân tạo (AI) mang tính biểu tượng, và nó dựa trên các quy tắc, logic và ký hiệu, đồng thời tìm ra các giải pháp mà con người có thể hiểu được.

Tuy nhiên, trong vài thập kỷ qua, đã được ưu tiên các mô hình kết nối có nghĩa là Học máy không giám sát (unsupervised machine learning - ML). Các Mạng nơ-ron sâu (DNN), một kỹ thuật học máy cho phép máy tính, bằng cách đào tạo nó, thực hiện các tác vụ rất khó thực hiện bằng các kỹ thuật lập trình thông thường, học từ lượng dữ liệu khổng lồ để đưa ra các dự đoán.

Điều này làm nảy sinh hai vấn đề quan trọng liên quan đến thành kiến và giải thích. Bởi trí tuệ nhân tạo (AI) được đào tạo dựa trên dữ liệu của con người, nên nó có thể (và thực sự) nhập các thành kiến mà con người thể hiện, chẳng hạn như liên quan đến giới tính và chủng tộc.

Ví dụ như, tại Hoa Kỳ, COMPAS là một trong những công nghệ trí tuệ nhân tạo được sử dụng trong quá trình giải quyết các vụ án hình sự. Chức năng chính của ứng dụng này nhằm giúp Thẩm phán xác định liệu có cần thiết áp dụng biện pháp giam giữ đối với một đối tượng phạm tội hay cho phép họ được tại ngoại chờ ngày xét xử, và đã duy trì thành kiến phân biệt chủng tộc có hệ thống thông qua các đánh giá rủi ro của nó.

Hơn nữa, máy dự đoán chức năng của chuỗi Mạng nơ-ron sâu DNA, gọi quá trình ẩn này như một “hộp đen - black box” - được gọi là vấn đề về bản thể hộp đen trong AI - dựa trên phép tính phức tạp và nhiều lớp mà con người không thể tiếp cận được. Vì thế, chính những nhà thiết kế của chúng chứ đừng nói đến người dùng cuối, không thể giải thích được cho quá trình ra quyết định tự động như thế.

Máy móc tạo ra các kết quả quyết định có thể hữu ích, nhưng chúng ta không biết chúng được đưa ra như thế nào hoặc tại sao.

Điều này dẫn đến nhu cầu cần về trí tuệ nhân tạo (AI) có thể giải thích được hoặc Explainable AI (XAI), một tập hợp quy trình và phương pháp cho phép người dùng con người hiểu rõ hơn và tin tưởng vào kết quả và đầu ra được tạo ra bởi các thuật toán học máy, bởi việc giải thích các kết quả là rất quan trọng theo quan điểm chính sách và quy định cũng như đối với trách nhiệm giải trình, xây dựng lòng tin và quy trách nhiệm trong các lĩnh vực có rủi ro cao như chẩn đoán y tế, xe tự hành, tài chính hoặc quốc phòng. Khi có điều gì đó không ổn, ai

sẽ chịu trách nhiệm?

Bất chấp những vấn đề này, trí tuệ nhân tạo (AI) đang ngày càng định hình kiến thức của chúng ta về bản thân và thế giới khi chúng ta giao tiếp với các thực thể kỹ thuật số và thực, quyết định những gì chúng ta có thể truy cập và quản lý những gì chúng ta có thể và thực sự quan tâm.

Các ví dụ như tôi đưa ra trong bài diễn giảng của mình, bao gồm các thuật toán có phiếu bầu trong hội đồng quản trị tập đoàn đầu tư, con người kết tình cảm như keo sơn với AI trực tuyến vì tin rằng họ giống mình, AI có thể giải mã giọng nói từ hoạt động của não với độ chính xác khá cao, các thẩm phán được yêu cầu tham khảo các khuyến nghị AI, và khả năng nhận thức của các thuật toán có khả năng xuất hiện trong tương lai gần.

Có thể Phật giáo đi vào diễn ngôn AI ở đâu? Trong số những thứ khác, Phật giáo là nghiên cứu về tâm linh. Có nhiều điểm tương đồng trong mối bận tâm của các nhà khoa học công nghệ trí tuệ nhân tạo (AI) và các triết gia Phật giáo - một điểm khác biệt chính là trong hàng thiên niên kỷ các triết gia Phật giáo đã tư duy về các thắc mắc về tâm linh. Hai lĩnh vực này không thường được cùng nhau nghĩ đến, nhưng văn hóa đại chúng đã phản ánh sự chồng chéo này.

Hãy nghĩ đến Neo (tên khai sinh là Thomas A. Anderson, còn được gọi là The One, một từ đảo chữ của Neo) là một nhân vật hư cấu và là nhân vật chính của loạt phim The Matrix, do Wachowskis tạo ra. Anh được Keanu Reeves miêu tả là một tội phạm mạng và lập trình viên máy tính trong các bộ phim, cũng như có một vai khách mời trong phim ngắn Kid's Story của The Animatrix, người nhận ra rằng mình bị giam hãm trong một thế giới vật chất là mô phỏng máy tính do AI thiết kế, phải cố gắng được tự do và dạy người khác để cách được tự do, giống như đức Phật.

Trong khi những cậu robot giết người thống trị trí tưởng tượng của phương Tây về AI, ở Nhật Bản, những cậu robot lại là một phần của đời sống công cộng.

Tôi cho rằng có một số cách mà chúng ta có thể khái niệm hoá mối giao thoa giữa Phật giáo và AI.

Thứ nhất, có thể trí tuệ nhân tạo (AI) giúp gì cho Phật giáo, tức là cách mà có thể AI tác động đến việc thực hành và diễn giảng Phật học. Công nghệ cung cấp tôn giáo một phương tiện sáng tạo để truyền bá thông điệp, đặc biệt là vào thời điểm mà theo báo cáo của Trung tâm nghiên cứu Pew (Pew Research Center, Pew) năm 2017, số lượng phật tử sẽ giảm về số lượng tuyệt đối, giảm 7% từ khoảng 500 triệu vào năm 2015 xuống còn 462 triệu vào năm 2060.

Ngôi chùa Kodaiji 400 năm tuổi ở Kyoto, Nhật Bản đang gây nhiều chú ý với tuyên bố sẽ “thay đổi bộ mặt của Phật giáo”. Ngôi chùa này vừa sử dụng một nhà sư robot trong dự án truyền đạo mang tên Mindar trị giá 1 triệu USD. Robot tại chùa Kodaiji có tên Quán Âm (Kannon), được làm dựa trên hình dáng Bồ tát Quán Thế Âm, Quán Âm (Kannon) đang hàng ngày thuyết giảng các bài kinh cho du khách thập phương.

Robot Quán Âm (Kannon) bắt đầu công việc tại chùa vào đầu năm 2019. Quán Âm (Kannon) có thể cử động thân, tay và đầu. Quán Âm (Kannon) cũng có thể chấp tay làm động tác như cầu nguyện. Tuy nhiên, chỉ có tay, mặt và vai của Quán Âm (Kannon) là được silicon bao phủ để mô phỏng da người, cao hơn 6 feet, nặng 70 pound và có giá 1 triệu đô la. Phần còn lại để nguyên cho thấy đây là máy móc.

Quán Âm (Kannon) có thể diễn giảng giáo lý đạo Phật về nhiều chủ đề khác nhau. Những bài thuyết pháp được giảng bằng tiếng Nhật và được dịch ra tiếng Anh và tiếng Trung Quốc trên các màn hình lắp đặt gần đó. Nó thuyết giảng các bài giảng Phật giáo và hiện chưa được trang bị Thuật toán ML nhưng những người sáng tạo ra nó hy vọng rằng một ngày nào đó nó sẽ được trang bị.

Một robot hình người tên là ‘Pepper’ có thể làm dịch vụ tang lễ ở Nhật Bản, có thể nó sử dụng các loại nhạc khí, chuông, mõ và đọc tụng kinh cầu nguyện giống như một nhà sư Phật giáo. Ngoài ra còn có sự bùng nổ của con người ảo và tại Vương quốc Phật giáo Thái Lan, một nhà sư kỹ thuật số đạo hiệu là Phra Maha AI giảng dạy Phật pháp cho mọi người trên Facebook và Instagram.

Người ta đã đề xuất rằng AI có thể giúp hiển cơ chế não bộ của việc thực hành thiền định và tạo điều kiện thuận lợi cho nó thông qua việc theo dõi hệ thần kinh sinh lý. AI trong thế giới nghệ thuật đang mở rộng và điều này bao gồm tiềm năng cho nghệ thuật Phật giáo. Sau đó, có bản đồ hỗ trợ AI của các văn bản và con đường Phật giáo lịch sử.

Thứ hai, Phật giáo giúp gì được cho AI, tức là công trình nghiên cứu khoa học về việc phát triển các chiều kích của tâm trí, lấy từ triết học Phật giáo có thể được sử dụng cho các mô hình AI, cách triết học Phật giáo cung cấp những hiểu biết quan trọng cho sự phát triển của AI Mạnh (Strong AI). Trong Phật giáo có thuyết minh về 5 hợp chất - trong đó các hoạt động của chủ thể có tâm trí được chia thành 5 hợp chất: vật chất và cơ thể vật lý (Sắc - Rūpa), sự cảm nhận, sự hưởng cảnh của tâm (Thụ - Vedanā), sự nhận biết do ấn tượng (Tưởng - Saññā), hành động của tâm (Hành - Saṅkhāra), quá trình tâm lý phát sinh từ các hiện tượng tâm lý khác (Thức - Vijñāna) - có thể đóng vai trò là khuôn khổ cho các nghiên cứu AI.

Các nhà triết học Phật giáo có thể giúp các chuyên gia phần mềm xây dựng AI thực sự thông minh qua một ý tưởng Phật giáo tinh vi về nhân quả (điều đó không hoàn toàn mang tính xác định cũng không hoàn toàn mang tính xác suất) có thể hiểu rõ hơn về sự thay đổi tư duy và phương thức hoạt động của tổ chức bằng công nghệ kỹ thuật số và sự phụ thuộc lẫn nhau.

Triết học Phật giáo có thể bổ sung giá trị cho cuộc tranh luận về AI mạnh mẽ về ý thức. Đạo Phật là một tôn giáo độc đáo ở khả năng coi sự sống không chỉ là con người, mà còn là động vật, thậm chí là máy móc.

Trái ngược với các triết gia phương Tây như Triết gia người Mỹ, Giáo sư John Searle, Đức Đạt Lai Lạt Ma đã khẳng định rằng nhận thức của máy móc và sự sống của máy móc không phải là không thể loại trừ. Khi được hỏi liệu robot bao giờ có thể trở thành chúng sinh có tri giác hay không, Ngài trả lời: *"Nếu cơ sở vật lý của máy tính có được tiềm năng hoặc khả năng làm cơ sở cho một chuỗi ý thức. . . thì một luồng ý thức thực sự có thể nhập vào máy tính"*. Đối với những người phật tử, không có bản ngã ổn định và liên tục thiết yếu nào và trọng tâm là thay đổi trạng thái tinh thần bị tác động bởi các mối liên hệ ràng buộc tạo nên trải nghiệm về sự liên tục đó.

Đối với học giả Phật giáo về trí tuệ nhân tạo (AI) mà tôi tham khảo, *Tính không* (một kinh nghiệm cảm nhận về thực thể tối hậu của mọi hiện tượng) có thể áp dụng cho các bộ phận và toàn bộ, có thể chúng ta nghĩ đến lời Tư duy phản biện của Bồ tát Long Thụ (Nagarjun) về sự vật hóa sức mạnh nhân quả của con người mà họ thúc đẩy. Một lần nữa, đối với các nhà logic học Phật giáo như triết gia Phật giáo Ấn Độ của trường phái Yogācāra và trường phái nhận thức luận, nhà logic học Phật giáo Ratnakirti, không có luồng tâm trí cá nhân nào được phân định và nhận thức được khái niệm hóa theo thuyết nhị nguyên bị bác bỏ, vì thế, câu hỏi về việc liệu một hệ thống tính toán có 'thực sự có ý thức' hay không bị giảm bớt.

Cuối cùng, một cách tiếp cận của Phật giáo rất hữu ích để nghĩ về cách tạo ra AI từ bi và robot vị tha. Để mô hình hoá cảm xúc, các mô hình về tâm trí là điều cần thiết và các truyền thống Phật giáo phân biệt các phiên bản khác nhau của từ bi tâm (metta, karuna, mudita, uppekkha) và các đức tính bền bờ của giác ngộ, của không còn sợ hãi, an nhiên tự tại, an bình (paramitas - ba la mật) có thể hỗ trợ xây dựng tính cách người máy (robot) và khả năng học hỏi thông qua các tương tác. Vì Phật giáo tưởng tượng con người không chỉ là những cỗ máy vô tri, nó có thể là một biển báo cho những cỗ máy có trí tuệ.

Thứ ba, những gì trí tuệ nhân tạo (AI) Phật giáo sẽ làm cho quản trị toàn cầu, tức là vượt xa hơn nhiều so với tính toán tình cảm, một đạo đức nhân văn Phật

giáo tập trung vào từ bi tâm và chuyển hoá, chữa lành những nỗi khổ niềm đau phải là một phần thiết yếu của tầm nhìn toàn cầu về AI, sự quan tâm phải được coi là động lực quan trọng của trí thông minh.

Sự phát triển không đồng bộ của công nghệ trên toàn cầu, đã liên tục đảm bảo sự khác biệt có hệ thống giữa việc thiết lập kết nối so với quan hệ. Trí tuệ nhân tạo định hình lý luận chính trị và sắp xếp lại tổ chức kinh tế ngay cả khi năng lực điều chỉnh và quản lý trong thời đại AI thay đổi có hệ thống trên toàn cầu. Năng lực của AI mang đến cơ hội, nhưng cũng có những rủi ro riêng biệt liên quan đến lao động, quyền riêng tư xung đột, v.v.

AI được gọi là 'trò chơi lớn mới'; và AI vũ khí hoá liên quan đến tất cả. Báo cáo năm 2021 của Ủy ban An ninh Quốc gia Hoa Kỳ về AI nêu rõ, *"Khả năng của máy móc trong việc nhận thức, đánh giá và hành động nhanh hơn và chính xác hơn con người thể hiện lợi thế cạnh tranh trong mọi lĩnh vực - dân sự hay quân sự. Công nghệ AI sẽ là nguồn động lực cho các công ty và quốc gia khai thác chúng"*. Năm 2017, Tổng thống Nga Putin tuyên bố *"bất kỳ ai trở thành người dẫn đầu trong lĩnh vực này sẽ trở thành người thống trị thế giới"*.

Với sự bất cân xứng toàn cầu trong khả năng tham gia vào các công nghệ như thế, chưa nói đến việc điều chỉnh chúng, khiến kiến trúc chuẩn mực xung quanh AI bị lệch rất nhiều khỏi một số lượng lớn các nước đang phát triển. Điều quan trọng là đạo đức phải được đặt lên hàng đầu trong AI và quản trị toàn cầu, đạo đức nhân văn Phật giáo có thể giúp lưu tâm đến nhu cầu phải có các mục tiêu như chuyển hoá chữa lành những nỗi khổ niềm đau và tham gia vào sự đa dạng hoá của các giá trị văn hoá vượt ra ngoài phương Tây.

Mọi công nghệ đều mang tính chính trị và được đặt trong bối cảnh đương đại của nền dân chủ đang suy yếu và bất bình đẳng ngày càng trầm trọng, cần có những cách diễn đạt thay thế mong muốn, đau khổ và mối quan hệ. Ở đây, việc tập trung vào Phật giáo và sự giao thoa giữa trí tuệ nhân tạo (AI) có thể giúp chúng ta hiểu rõ hơn về cách nhận thức, ý thức, sự đồng thuận, lương tâm và lòng từ bi trong tư duy về công nghệ AI. Những gì chúng ta ước mơ, chúng ta có cơ hội hiện thực hoá và chúng ta có quyền và trách nhiệm ước mơ tốt hơn - trong 100 năm nữa, chúng ta muốn cùng tồn tại với AI trí tuệ và từ bi, thách thức chúng ta làm tốt hơn và thể hiện điều hoàn hảo hơn nữa chứ không phải điều tệ nhất của chúng ta với tư cách là một con người.

Tác giả: **Nữ Tiến sĩ Nitasha Kaul**

Viết dịch: **Thích Vân Phong**

Nguồn: *Kuensel Online*