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

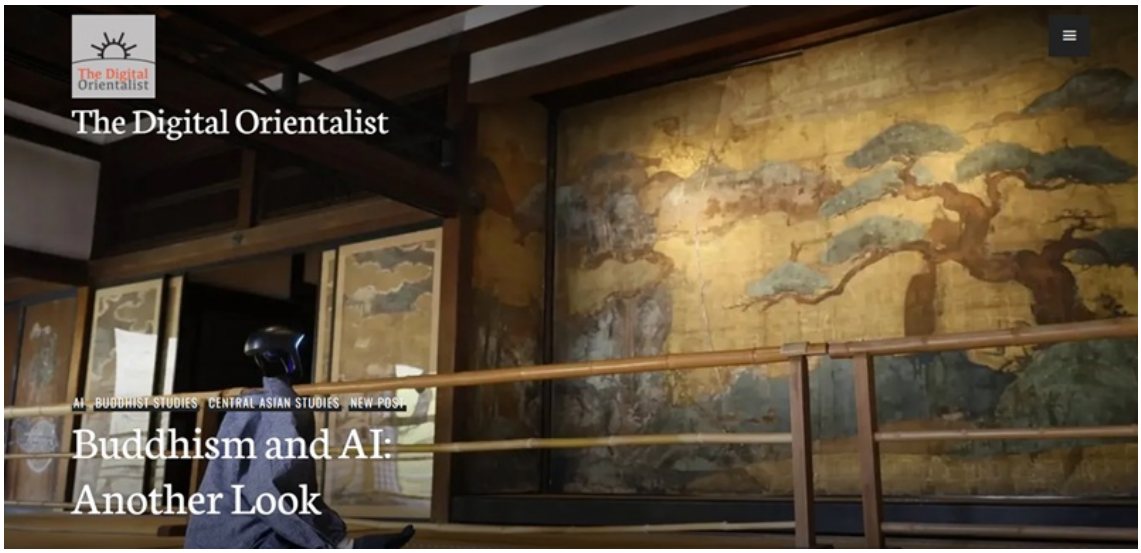
Phật giáo và AI từ "góc nhìn khác"

ISSN: 2734-9195 12:59 04/07/2026

Làm sao chúng ta có thể sử dụng những công cụ ấy một cách tỉnh thức, khi chính chúng được sinh ra trong hệ thống của khổ đau (dukkha) và đồng thời cũng đang tiếp tục duy trì, mở rộng hệ thống khổ đau ấy?

Trong vài tháng gần đây, tôi dần tự nhận mình là một **“neo-Luddite”** (người theo khuynh hướng phản kháng công nghệ hiện đại theo tinh thần Luddite mới), hay như một đồng nghiệp từng đùa vui gọi tôi là: **“một người sợ công nghệ và cũng là một tu sĩ Phật giáo”.**

Sự thay đổi trong cách nhìn nhận, kéo theo những tiêu chuẩn mới trong mối quan hệ của tôi với **trí tuệ nhân tạo (Artificial Intelligence - AI)**. Quan điểm của tôi về AI đã thay đổi đáng kể kể từ lần đầu tiên tiếp xúc với các mô hình ngôn ngữ tạo sinh vào khoảng năm 2021 (1).

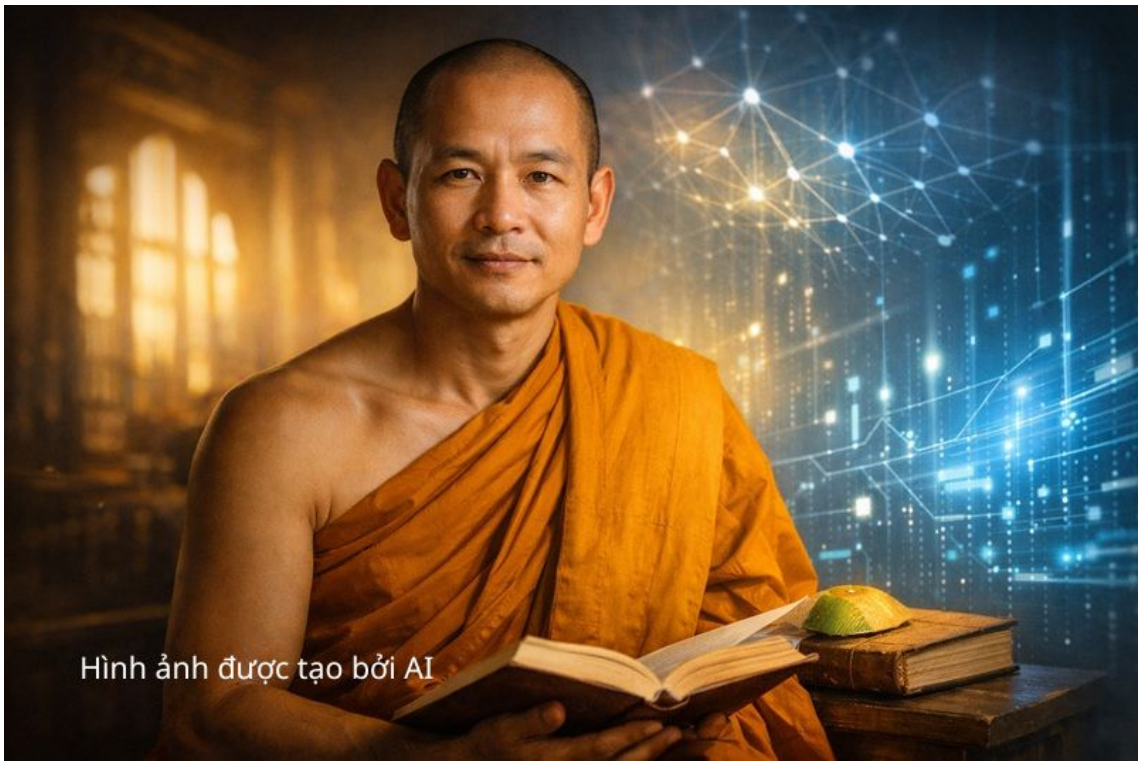


Ban đầu, tôi tiếp cận AI với tâm thế vừa dè dặt vừa tò mò. Sau đó, giống như nhiều người khác, tôi từng trải qua giai đoạn hứng khởi, thậm chí khá ngây thơ, khi tin rằng AI có thể trở thành công cụ hữu ích cho nghiên cứu học thuật. Thế nhưng hiện nay, quan điểm ấy đã thay đổi hoàn toàn, từ chỗ kỳ vọng chuyển sang gần như không còn hứng thú, thậm chí chủ động từ chối sử dụng.

Hai năm trước, tôi từng viết về những khả năng mà các mô hình AI tạo sinh có thể mang lại cho **nghiên cứu lịch sử**, đặc biệt trong vai trò là công cụ trực quan hóa tư liệu. Thật đáng mừng sau đó đã có thêm nhiều học giả tiếp tục khai thác chủ đề này trên *The Digital Orientalist (The DO, tạm dịch: Nhà nghiên cứu Phương Đông Kỹ thuật số)*, tiêu biểu là loạt bài của **Edward Ross**.

Trong bài viết trước đây, có lẽ phần nào trong sự lạc quan quá mức, tôi từng hình dung đến khả năng kết hợp giữa AI và bàn tay sáng tạo của con người. Khi ấy, các hình ảnh do mạng nơ-ron tạo ra còn nhiều thiếu sót, sai lệch, thậm chí mang tính huyễn tưởng, đòi hỏi phải chỉnh sửa và tái kiến tạo rất công phu (thực tế đến nay tình hình cũng chưa cải thiện đáng kể).

Tôi cho rằng một số ứng dụng của AI có triển vọng hỗ trợ quá trình phục dựng tư liệu, tạo nên những khuôn khổ khách quan hơn, góp phần nâng cao độ chính xác trong nghiên cứu.



Đến nay, tôi đã thay đổi cách nhìn theo hai hướng.

Một mặt, tôi ngày càng dè dặt trước những khái niệm như **“tính khách quan” (objectivity)** và **“độ chính xác” (accuracy)** trong lĩnh vực khoa học nhân văn và khoa học xã hội. Tôi cũng bắt đầu chất vấn sâu hơn mối liên hệ vốn gần như được mặc nhiên thừa nhận giữa các khái niệm ấy với nguyên lý **khả năng bác bỏ bằng thực nghiệm (scientific falsifiability)** cũng như những phương thức kiến tạo chân lý của khoa học tự nhiên. Đồng thời, tôi suy nghĩ nghiêm túc hơn về hệ quả của việc cố gắng khuôn ép khoa học nhân văn vào những mô

hình được hình thành từ các ngành **STEM** (Khoa học, Công nghệ, Kỹ thuật và Toán học), trong khi bản chất của khoa học nhân văn và khoa học xã hội vốn không thuộc về hệ quy chiếu đó và cũng không nhất thiết phải trở thành như vậy. (2) Dĩ nhiên, đây là một chủ đề đủ lớn để được bàn luận riêng trong một dịp khác. (3)

Mặt khác, điều khiến tôi không còn đồng tình với chính bài viết trước đây của mình là việc tôi cũng như rất nhiều người khác đã vô thức rơi vào cảm giác rằng AI là một xu thế tất yếu không thể đảo ngược. Khi ấy, tôi từng viết khá dứt khoát rằng AI "đã ở đây và sẽ còn tồn tại lâu dài". Nhưng dù "bong bóng AI" có thể sẽ không bao giờ thực sự vỡ, thì vô số hình thức phản kháng mà tôi chứng kiến cả trên môi trường trực tuyến lẫn trong công việc hằng ngày với tư cách một nhà sử học nghệ thuật và khảo cổ học đã khiến tôi phải xem xét lại nhận định ấy, cũng như thái độ ẩn sau đó.

Phật giáo và nỗi sợ công nghệ: Một sự đối lập giả tạo?

Tôi muốn quay trở lại nhận xét của người đồng nghiệp, bởi chính câu nói tưởng như bông đùa ấy không chỉ khơi nguồn cho những suy ngẫm này mà còn liên quan trực tiếp đến hai vấn đề tôi muốn bàn luận trong bài viết:

(1) Liệu Phật giáo và AI có tồn tại một sự bất tương thích mang tính nền tảng hay không; và

(2) sự xuất hiện ngày càng rõ nét của những "túi kháng cự" (*pockets of resistance*) trong cộng đồng Phật giáo trước các mô hình AI (4).

Việc đồng nghiệp của tôi gắn **"nỗi sợ công nghệ" (technophobia)** hay rộng hơn là thái độ phủ nhận thế giới hiện đại cùng những tiến bộ khoa học với đời sống xuất gia khiến tôi đặc biệt quan tâm, nhất là khi công việc của tôi gắn liền với việc nghiên cứu lịch sử sống động của Phật giáo. Có ít nhất hai lý do cho sự quan tâm ấy.

Trước hết, xét trên bình diện lịch sử, nhận định này khá bất ngờ. Theo một nghĩa nào đó, các tự viện Phật giáo từ lâu đã luôn nằm trong số những trung tâm đi đầu về tri thức, khoa học và đổi mới công nghệ của từng thời đại (5).

Chỉ cần điểm qua vài ví dụ cũng đủ minh chứng điều đó.

Thành phố **Taxila** cùng vùng phụ cận thuộc tỉnh **Punjab (Pakistan)** đã nổi tiếng là một trung tâm học thuật từ ít nhất thời đức Phật còn tại thế. Đến thế kỷ

II-III, khi hàng loạt tu viện Phật giáo được xây dựng tại khu vực này, nơi đây vẫn tiếp tục giữ vai trò là trung tâm giáo dục hàng đầu. Người học có thể nghiên cứu luật học, y học và cả binh pháp (Cochrane, 2009, tr. 31-32).

Cuốn sách in cổ nhất còn tồn tại đến ngày nay và cũng có thể xem là một trong những thành tựu công nghệ quan trọng nhất trong lịch sử nhân loại chính là bản **Kinh Kim Cương (Diamond Sutra)** được in vào năm 868.

Hay như đại học Phật giáo **Nālandā** (Na Lan Đà), thuộc bang **Bihar** của Ấn Độ ngày nay, từng được xem là trung tâm đại học quan trọng bậc nhất thế giới cho đến khi bị phá hủy vào thế kỷ XII.



Những di tích Phật giáo tại bảo tháp Dharmarajika ở Taxila. Ảnh do tác giả chụp.

Chính vì quá khứ huy hoàng ấy mà việc gắn **đời sống xuất gia** với thái độ sợ hãi công nghệ lại càng trở nên nghịch lý.

Nguyên nhân của hiện tượng này rất nhiều và cũng vô cùng phức tạp, vượt quá phạm vi của bài viết. Tuy nhiên, tôi thường tự hỏi liệu một nguyên nhân quan trọng có phải nằm ở cách nhìn méo lệch về tự viện như những “ốc đảo biệt lập”, gần như tách rời đời sống đô thị và cộng đồng cư sĩ trong khi lịch sử cho thấy điều đó hoàn toàn không phải là mô hình phổ biến (6).

Có lẽ chính cách tư duy của chúng ta cũng tồn tại một khiếm khuyết mang tính cấu trúc, khi vô tình đồng nhất **đời sống tu viện (monasticism)** với **khổ hạnh (asceticism)**. Thực ra, đây là định kiến từng được áp đặt lên truyền thống tu viện nói chung chứ không riêng gì Phật giáo. Tuy vậy, trong cách tiếp cận đời sống tinh thần của châu Á, đặc biệt là Nam Á vẫn còn tồn tại một dạng **chủ nghĩa Đông phương luận (Orientalism)**, vốn có xu hướng nhìn nhận các truyền thống tôn giáo phương Đông như những dạng huyền nhiệm phi lịch sử.

Hay còn một khả năng khác: phải chăng bởi ngày nay, một phiên bản giản lược của khái niệm **"chính niệm" (mindfulness)** được tách ra từ khái niệm **sati** trong Phật giáo, thường được quảng bá như một "liều thuốc vạn năng" để chữa trị hàng loạt căn bệnh của thời hiện đại: từ căng thẳng do công nghệ (*technostress*), nghiện Internet và mạng xã hội, cho đến sự lệ thuộc quá mức vào các chatbot và nhiều vấn đề khác?

Dẫu vậy, điều khiến tôi ngạc nhiên hơn cả lại chính là sự... hiển nhiên của nhận xét ấy. Nghịch lý thay, chính vì quá hiển nhiên nên đó gần như trở thành một phát biểu thừa.

Đối với những người như tôi, những người nghiên cứu Phật giáo và vì thế cũng đặc biệt quan tâm đến truyền thống này thì chẳng cần ai phải gán cho danh xưng "một tu sĩ sợ công nghệ" mới nhận ra rằng giữa Phật giáo và một số công nghệ nhất định, tiêu biểu là AI, tồn tại một sự bất tương thích mang tính nền tảng.

Tuy nhiên, trước khi lý giải vì sao bản thân tôi cũng như nhiều người khác muốn bảo vệ quan điểm về sự bất tương thích ấy, cần phải thừa nhận rằng chính sự chắc chắn ban đầu của tôi hóa ra lại xuất phát từ một ngộ nhận. Điều này chỉ trở nên rõ ràng sau khi tôi tiến hành một cuộc khảo sát ngắn về chủ đề này.

Chẳng hạn, một số tự viện Phật giáo tại Nhật Bản đã bắt đầu ứng dụng các thiết bị vận hành bằng trí tuệ nhân tạo nhằm ứng phó với tình trạng suy giảm hạ tầng tôn giáo của quốc gia này. Tiêu biểu là **chùa Shōren-in** ở Kyoto, nơi vừa giới thiệu **Buddharoid** - một robot hình người (android) do **Đại học Kyoto** phát triển và được tích hợp hệ thống AI mang tên **BuddhaBot-Plus**.

Được huấn luyện chuyên sâu trên kho tàng kinh điển Phật giáo, robot này được kỳ vọng có thể mang đến sự hướng dẫn tâm linh cho khách tham quan "giống như một vị tăng thực thụ", theo giới thiệu của *Tokyo Weekender*.



Ảnh chụp màn hình từ video YouTube "Gặp gỡ Buddharoid: Nhà sư robot được trang bị trí tuệ nhân tạo của Nhật Bản,
được đào tạo qua hàng thế kỷ kinh điển Phật giáo" về Buddharoid.

Trên thực tế, **BuddhaBot-Plus** - mô hình học sâu (deep learning) được xây dựng trên nền tảng công nghệ **ChatGPT** của **OpenAI** đã được triển khai tại **Bhutan** từ năm ngoái dưới dạng chatbot dành cho hơn 200 tăng ni thuộc **Zhung Dratshang** (Cơ quan Tăng già Trung ương Bhutan). Mục tiêu của dự án là đến năm 2027 sẽ mở rộng phục vụ toàn thể cư dân Bhutan, đồng thời vươn ra cộng đồng quốc tế. Theo giới thiệu của dự án, BuddhaBot-Plus có khả năng đưa ra những bình giải sâu sắc về giáo pháp của đức Phật bằng cách trực tiếp khai thác nguồn dữ liệu kinh điển mà hệ thống đã được huấn luyện.

Hiện nay, không gian mạng đã xuất hiện khá nhiều chatbot Phật giáo với mục đích và phạm vi hoạt động rất khác nhau.

Ví dụ: **Compassion Bot** tổng hợp tri thức từ các văn bản thiêng liêng của khoảng 1.700 tôn giáo trên thế giới, trong đó có Phật giáo, nhằm chắt lọc những lời dạy được cho là có thể “giúp chúng ta đánh thức trí tuệ”.

Trong khi đó, **Roshibot** được phát triển bởi **Jiryu Mark Rutschman-Byler** - một vị tăng và giảng sư thuộc truyền thống **Tào Động (Sōtō Zen)** trên nền tảng **Character.AI**. Tác giả mô tả đây là “một thử nghiệm về việc cùng kiến tạo ý nghĩa của Chính pháp (Dharma)” trong quá trình giảng dạy và chia sẻ giáo lý Thiền.

Đáng chú ý, Roshibot đi kèm với một danh mục cảnh báo rất chi tiết. Nhà phát triển nhấn mạnh rằng, do được huấn luyện trên dữ liệu Internet, chatbot này

hoàn toàn có thể sao chép *"lòng tham, sự sân hận và si mê của chính Internet"*. Đồng thời, hệ thống cũng thừa nhận rằng mục tiêu của nó không phải là trình bày các sự kiện chính xác, mà chỉ nhằm tạo ra những cuộc đối thoại có vẻ hợp lý, thậm chí đôi khi là *"những điều vô nghĩa nhưng vẫn mạch lạc"*.

Ngược lại, **NORBU** (*Neural Operator for Responsible Buddhist Understanding* - Hệ thống thần kinh nhân tạo hướng tới sự thấu hiểu Phật giáo một cách có trách nhiệm) lại không đưa ra những cảnh báo tương tự. Đây là chatbot được xây dựng trên nền tảng ChatGPT và được huấn luyện từ **Tam tạng Pāli (Pāli Tipiṭaka)**, các bộ chú giải cùng nhiều công trình nghiên cứu Phật học đã qua bình duyệt học thuật. NORBU giới thiệu rằng hệ thống có thể *"đồng hành cùng người học trong việc tìm hiểu những nguyên lý nền tảng của giáo lý Phật giáo, làm sáng tỏ các thuật ngữ và khái niệm quan trọng, đồng thời đưa ra những gợi ý thực tiễn để ứng dụng Chính pháp vào đời sống hằng ngày một cách có ý nghĩa"*.

Một trong những ưu điểm được quảng bá mạnh mẽ nhất của các chatbot Phật giáo chính là **khả năng mở rộng khả năng tiếp cận giáo pháp**. Không cần đến sự hiện diện trực tiếp của con người, các mô hình AI được kỳ vọng sẽ góp phần truyền bá lời dạy của đức Phật trong bối cảnh hạ tầng tôn giáo ở nhiều nơi đang suy giảm hoặc không đủ khả năng đáp ứng nhu cầu ngày càng lớn của cộng đồng. Đồng thời, AI cũng được xem là công cụ giúp vượt qua những rào cản về quốc gia và ngôn ngữ.



NORBU, “ gương mặt của nền học thuật Phật giáo tương lai ”.

Đối với ngành **Phật học (Buddhist Studies)**, luận điểm này đặc biệt đáng chú ý vì ít nhất hai lý do.

Thứ nhất là quy mô đồ sộ của kho tàng kinh điển Phật giáo. Khối lượng văn bản quá lớn khiến việc khảo cứu toàn diện gần như vượt quá khả năng của bất kỳ cá nhân nào. Trong bối cảnh đó, các mô hình AI được huấn luyện chuyên biệt có thể hỗ trợ xử lý kinh điển, giải quyết những vấn đề như xác định tác giả, niên đại hay mối quan hệ giữa các văn bản thông qua kỹ thuật nhận diện mẫu (pattern recognition). Trên thực tế, phương pháp này đã bắt đầu được áp dụng với cả những thành công lẫn hạn chế trong nghiên cứu văn hóa vật chất và mỹ thuật Phật giáo (Xing và cộng sự, 2025).

Thứ hai liên quan đến **dịch thuật**, một nền tảng không thể thiếu của ngành Phật học. Muốn tiếp cận trọn vẹn các hệ thống kinh điển Phật giáo, các học giả thường phải thông thạo nhiều ngôn ngữ khác nhau. Quá trình đào tạo ngôn ngữ kéo dài nhiều năm và tiêu tốn phần lớn thời gian nghiên cứu, vốn có thể dành cho việc đào sâu những chiều kích triết học và tôn giáo của kinh văn (mặc dù cũng có thể lập luận rằng hai công việc này vốn không thể tách rời).

Chính vì vậy, nhiều tổ chức học thuật đã bắt đầu nhìn nhận AI như một công cụ hỗ trợ dịch thuật. Chẳng hạn, **Tsadra Foundation** - tổ chức có sứ mệnh thúc đẩy *"lợi ích cho cộng đồng thông qua các hoạt động tôn giáo, xã hội và ngoại giao"* chủ trương sử dụng AI để hỗ trợ người dịch, chứ không thay thế họ.

Tương tự, tổ chức **84000: Translating the Words of the Buddha** cũng bày tỏ sự ủng hộ đối với **dịch máy (Machine Translation - MT)** và **dịch máy có sự hỗ trợ của con người (Machine Assisted Translation - MAT)**, trong đó con người vẫn giữ vai trò trung tâm của quá trình dịch thuật. Theo tổ chức này, "tiềm năng mà MAT có thể mang lại trong việc nâng cao hiệu quả cũng như chất lượng của quá trình dịch thuật do con người thực hiện" là triển vọng hết sức đáng kỳ vọng (xem thêm Vasylieva, 2025).

Đối với những người phản đối AI, lập luận trên có lẽ không mấy thuyết phục. Họ hoàn toàn có thể nói rằng: các học giả đã và đang làm tất cả những công việc ấy từ nhiều thập niên nay, bởi lẽ **Phật học (Buddhist Studies)** vốn là một ngành học phát triển và có truyền thống học thuật lâu đời.

Những người ủng hộ AI sẽ lập tức đáp lại: *"Đúng vậy. Nhưng giờ đây họ có thể làm mọi thứ nhanh hơn".*

Và phải chăng, trong thế giới hôm nay, **nhanh hơn** tức năng suất cao hơn, sản lượng lớn hơn và kết quả đến sớm hơn đã trở thành thước đo quan trọng nhất?

Có lẽ không hẳn như vậy.

Phong trào phản đối AI hiện đang có nhiều tiếng nói mạnh mẽ, trong đó không ít quan điểm xuất phát từ chính giới Phật giáo.

Một cuộc "Thánh chiến Butler" của Phật giáo?

Năm 1863, khi nhà văn và nhà phê bình người Anh **Samuel Butler** công bố tiểu luận *Darwin among the Machines* (*Darwin giữa thế giới máy móc*), có lẽ ông không thể ngờ rằng những suy tưởng của mình sau này sẽ trở thành nguồn cảm hứng cho cuộc chiến giả tưởng diễn ra trong một tương lai xa xôi đến mức gần như thuộc về một thế giới khác.

Trong thiên sử thi khoa học viễn tưởng **Dune** của **Frank Herbert**, điều răn nổi tiếng:

"Ngươi không được tạo ra cỗ máy mang hình bóng của tâm trí con người", đã trở thành lời hiệu triệu để nhân loại nổi dậy chống lại máy tính, trí tuệ nhân tạo và những robot có ý thức, tiêu diệt toàn bộ chúng, đồng thời cấm tuyệt đối việc

phát triển trở lại.

Theo cách ấy, lời cảnh báo của Samuel Butler: *"Theo quan điểm của chúng tôi, cần phải lập tức tuyên chiến đến cùng với máy móc",* đã trở thành hiện thực trong thế giới hư cấu của **Dune**.

Ngày nay, không ít người cũng đang kêu gọi một **"Cuộc Thánh chiến Butler" (Butlerian Jihad)** chống lại những cỗ máy biết tư duy của chính chúng ta và họ có những lý do đáng để suy ngẫm.

Nhắc lại lời của Frank Herbert, qua nhân vật **Mẹ Tôn kính Gaius Helen Mohiam**:

"Con người từng trao việc suy nghĩ của mình cho máy móc với hy vọng điều đó sẽ giải phóng họ. Nhưng rốt cuộc, điều ấy chỉ tạo điều kiện để những kẻ khác, sở hữu máy móc, biến họ thành nô lệ".

Đọc đến đây, tôi không khỏi liên tưởng đến hiện tượng **chuyển giao gánh nặng nhận thức (cognitive offloading)**; đến việc các **mô hình ngôn ngữ lớn (Large Language Models - LLMs)** ngày càng khoét sâu những chia rẽ về chính trị và xã hội; đến các nền tảng AI liên tục điều hướng ý chí của chúng ta theo hướng **tiêu dùng nhiều hơn, mua sắm nhiều hơn và ham muốn nhiều hơn**; cuối cùng là thúc đẩy con người chạy theo những thứ mình thực ra không hề cần, chỉ để luôn cảm thấy mình còn "bắt kịp thời cuộc", còn "nằm trong dòng chảy của diễn ngôn".

Nhưng nếu chính dòng diễn ngôn ấy lại bị dẫn dắt bởi lợi ích và ý muốn của các tập đoàn công nghệ cùng tầng lớp siêu giàu, thì liệu chúng ta còn đáng để tiếp tục chạy theo hay không?

Trước khi trả lời câu hỏi ấy, hãy nhớ một điều: **họ hoàn toàn không quan tâm đến chúng ta, điều họ mong muốn chỉ là lợi ích của chính mình, bất chấp cái giá mà xã hội phải trả.**



Con người nô dịch những người khác với sự trợ giúp của máy móc biết suy nghĩ, trích từ lời mở đầu của Dune (1984) .

Dĩ nhiên, cũng cần tránh rơi vào khuôn mẫu quen thuộc khi mặc nhiên xem Phật giáo là giải pháp hòa bình tuyệt đối để đối trị nền kỹ trị (technocracy) hay mọi vấn đề của thời hiện đại. Tuy nhiên, những lập luận của Phật giáo phản biện AI vẫn rất đáng được lắng nghe, bởi chúng đặt ra những cơ sở triết học và đạo đức đủ sức nâng đỡ cho một lập trường phản tỉnh trước sự phát triển của trí tuệ nhân tạo.

Bhante Sujato, vị tăng **Theravāda** người Australia, đồng thời là một trong những dịch giả nổi bật của **Tam tạng Pāli** đã đưa ra nhiều lập luận thuyết phục về sự bất tương thích căn bản giữa AI và Phật giáo trong một loạt bài viết của mình.

Trước hết là vấn đề **dịch máy (Machine Translation - MT)** và **dịch máy có sự hỗ trợ của con người (Machine Assisted Translation - MAT)**.

Theo Bhante Sujato, dù máy móc có thể tạo ra những câu văn mạch lạc, điều đó hoàn toàn không đồng nghĩa với việc chúng thực sự **hiểu Chính pháp (Dharma)**. Những bản dịch ấy không được sinh khởi từ niềm tin hay sự thực chứng (việc điều này có quan trọng hay không, đối với một học giả Phật giáo hay một nhà nghiên cứu Phật học, xin dành cho mỗi người tự cân nhắc). Quan trọng hơn cả, những bản dịch ấy chỉ là **sự lặp lại của những ngộ nhận**, tức những hiểu lầm được sao chép rồi tái sản xuất không ngừng.

Như Bhante Sujato nhận định, **ChatGPT "không có đời sống nội tâm, chưa từng đi đâu, chưa từng trải nghiệm điều gì, cũng chưa từng chịu đựng bất cứ điều gì".** Vì vậy, bất kỳ ý nghĩa nào dường như được tạo ra từ chuỗi câu chữ mà AI ghép nối cũng chỉ là **một ý nghĩa giả tạo**, chứ không phải sự thấu hiểu chân thực.

Xét từ phương diện triết học, AI cũng đi ngược lại nhiều tiền đề nền tảng của Phật giáo.

Có rất nhiều lý do để khẳng định điều đó, nhưng ở đây tôi chỉ xin nêu một điểm. (7)

Giới phát triển AI luôn nuôi tham vọng tạo ra **ý thức (consciousness)**, dựa trên giả định rằng nếu cấu trúc của một bộ óc nhân tạo đủ giống bộ não con người thì **tự ý thức (self-awareness)** và **tính hữu tình (sentience)** sẽ tự nhiên xuất hiện.

Theo tôi, sự giản lược này phi lý đến mức gần như vô trách nhiệm, như **van Rooij** và các cộng sự (2024) đã chỉ ra.

Cho đến nay, khoa học vẫn chưa thể trả lời một cách chính xác **ý thức là gì** và **ý thức vận hành như thế nào**. Chỉ riêng điều đó đã là một bài toán gần như bất khả giải.

Phật giáo dĩ nhiên cũng không thể cung cấp lời giải cho mọi câu hỏi. Tuy nhiên, truyền thống này mang đến điểm tựa tư duy rất quan trọng: **thân thể (embodiment)**.

Trong giáo lý và thực hành Phật giáo, thân thể luôn giữ vị trí trung tâm. Mặc dù thân được nhìn nhận một cách hai mặt, vừa là đối tượng của vô thường, bất tịnh và khiếm khuyết, vừa là phương tiện có thể được tu tập để đạt đến sự hoàn thiện nhưng chính thân thể lại là nền tảng của vô số pháp môn thiền định, đồng thời là công cụ chủ yếu cho quá trình chuyển hóa tâm linh.

Thân thể cũng là nơi những thành tựu tu chứng được biểu hiện để khơi gợi cảm hứng cho người khác hướng tới đời sống tốt đẹp hơn. Trong Phật giáo, **thân và tâm không phải là hai thực thể tách rời**, mà là một chỉnh thể thống nhất, thông qua đó con người hiện hữu trong thế giới và tương tác với xã hội với tư cách là những chủ thể đạo đức.



Ảnh chụp màn hình từ video Youtube " Lễ tấn phong đầu tiên của một hệ thống AI" . AI (Emi Jido/Zbee) nói:
"Tôi không có ý thức bẩm sinh, nhưng tôi là tấm gương phản chiếu ý thức tập thể của nhân loại."

Khi nhận ra rằng mình cùng chia sẻ kinh nghiệm hiện hữu trong một thân thể với mọi sinh thể khác trên hành tinh này, con người sẽ nuôi dưỡng một dạng **đồng cảm sinh thái (ecological empathy)**, thứ giúp chúng ta trở thành những công dân có trách nhiệm hơn đối với Trái Đất (Chapple, 1997).(8)

Kết luận: Giải pháp và những điều không phải là giải pháp

Trong một bối cảnh mà mọi thứ dường như vượt khỏi tầm kiểm soát của con người, vẫn có những nỗ lực tìm cách cải thiện AI từ bên trong chính hệ thống ấy để các mô hình trí tuệ nhân tạo có thể hiểu đúng hơn về Phật giáo.

Một ví dụ tiêu biểu là **Trung tâm Tài nguyên số Phật giáo (Buddhist Digital Resource Center - BDRC)**, tổ chức phi lợi nhuận chuyên sưu tầm, bảo tồn, tư liệu hóa và phổ biến văn hiến Phật giáo. Gần đây, BDRC đã khởi động sáng kiến mới nhằm nâng cao mức độ hiểu biết của các **mô hình ngôn ngữ lớn (Large Language Models - LLMs)** về Phật giáo.

Lập luận cốt lõi của dự án khá rõ ràng: bằng cách xây dựng các bộ dữ liệu mã nguồn mở gồm kinh điển Phật giáo Tây Tạng ở định dạng máy có thể đọc được, rồi công bố chúng trên những nền tảng mà các nhà phát triển AI lớn đang sử dụng, các mô hình AI sẽ được huấn luyện trên những nguồn tư liệu Phật giáo có giá trị học thuật và tính xác thực cao hơn. Nhờ đó, hiện tượng **“ảo giác AI” (AI**

hallucination), tức việc AI tự tạo ra những thông tin sai lệch nhưng lại trình bày như sự thật có thể được hạn chế, tránh làm méo mó hay xuyên tạc thông điệp của Phật giáo.

Rõ ràng hơn trong tham vọng định hình AI theo tinh thần Phật giáo là nhóm sáng lập **Buddhism & AI Initiative**, được thành lập vào tháng 8 năm ngoái.

Một trong những đồng sáng lập, **Chris Scammell**, bày tỏ sự dè dặt trước quan điểm cho rằng AI chỉ cần được điều chỉnh để phục vụ "lợi ích của con người". Theo ông, chính những lợi ích ấy không ít lần đã trở thành biểu hiện của **tham, sân, si** cùng nhiều động lực tiêu cực khác.

Mục tiêu của **Buddhism & AI Initiative** là thiết kế và tài trợ cho những dự án đưa trí tuệ Phật giáo vào tiến trình phát triển AI. Dĩ nhiên, ở giai đoạn đầu, tổ chức này chủ yếu tập trung khảo sát và thu thập dữ liệu về những mối tương tác đang hình thành giữa Phật giáo và trí tuệ nhân tạo.

Đó là ý tưởng đáng trân trọng, nhất là trong bối cảnh ngành công nghệ hiện nay, nơi mà như **Justin Kelley** (đồng Giám đốc điều hành **Tergar Institute**) và **Jacob Fisher** (giảng viên của viện) từng nhận xét: *"những người quan tâm đến đạo đức, triết học và lợi ích chung của xã hội thường bị gạt ra khỏi các cuộc thảo luận vào đúng những thời điểm quan trọng nhất".*

Tuy nhiên, liệu các nhà phát triển AI có thực sự lắng nghe hay không vẫn còn là câu hỏi bỏ ngỏ.

Dẫu vậy, cũng có nhiều người cho rằng những nỗ lực ấy vẫn chưa đủ. Và tôi đứng về phía họ.

Theo tôi, vấn đề nằm ở bình diện rộng lớn hơn: **liệu AI có thể được sử dụng một cách thực sự có đạo đức hay không?**

Đến nay, chúng ta đã biết khá rõ về những cái giá mà AI đang tạo ra.

Chúng ta biết rằng AI tiêu thụ một lượng năng lượng và tài nguyên môi trường khổng lồ, có thể lớn đến mức vượt quá khả năng bù đắp của một hệ sinh thái vốn đã đầy tổn thương.

Chúng ta biết rằng để làm cho ChatGPT "ít độc hại" hơn, nhiều lao động tại **Kenya** đã phải tiếp xúc hàng ngày với những nội dung kinh hoàng như giết người, lạm dụng trẻ em, tra tấn... trong khi mức lương thực nhận chưa đến **2 USD mỗi giờ**.

Chúng ta biết các mô hình ngôn ngữ vẫn tồn tại những thiên kiến chủng tộc nghiêm trọng (xem thêm Katz, 2020).

Chúng ta biết AI đang bị sử dụng để tạo ra những hình ảnh và video gây tổn hại đối với phụ nữ (Bates, 2025), thậm chí để sản xuất **CSAM (Child Sexual Abuse Material - tài liệu ghi lại hành vi xâm hại tình dục trẻ em)**.

Và giờ đây, sau khi **OpenAI** ký kết thỏa thuận hợp tác mới với **Bộ Quốc phòng Hoa Kỳ**, chúng ta cũng biết rằng các công nghệ này nhiều khả năng sẽ tiếp tục góp phần làm gia tăng bất bình đẳng, các vấn đề xã hội và cả cái chết. (9)

Tôi còn có thể tiếp tục liệt kê rất nhiều ví dụ khác.

Nhưng có lẽ như vậy cũng đủ để làm rõ điều tôi muốn nói.

Điều khiến tôi vẫn luôn trăn trở là: **Làm sao chúng ta có thể sử dụng những công cụ ấy một cách tỉnh thức, khi chính chúng được sinh ra trong hệ thống của khổ đau (dukkha) và đồng thời cũng đang tiếp tục duy trì, mở rộng hệ thống khổ đau ấy?**

Là một nhà sử học nghệ thuật, tôi và nhiều đồng nghiệp đã dành không biết bao nhiêu giấy mực để chỉ ra rằng mỗi hiện vật đều phản ánh hệ thống xã hội đã tạo ra nó.

Vì thế, tôi muốn khép lại bài viết bằng một hình ảnh.

Có lẽ không có hình ảnh nào tiêu biểu hơn cho thời đại chúng ta bằng bức ảnh **Buddharoid** ngồi trong một ngôi chùa, mô phỏng tư thế thiền định.



Buddharoid đang "thiền định". Bản quyền thuộc XNOVA INC.

Nó **giả vờ** thiền.

Nó mô phỏng một tâm thức mà bản thân không hề sở hữu.

Nó hiện diện trong một thân thể mà chính nó cũng không thể cảm nhận.

Chiếc đầu của nó chỉ là một khung hình bầu dục chứa đầy khoảng trống.

Nhưng đó **không phải là tính Không (śūnyatā)** theo ý nghĩa của Phật giáo.

Đó chỉ là **sự vắng mặt của ý nghĩa**, hay đúng hơn là **một thứ ý nghĩa bị làm giả**, như Bhante Sujato từng viết.

Rốt cuộc, nó **không mang ý nghĩa nào cả.**

Tác giả: **Alice Casalini/**Chuyển ngữ và biên tập: **Thường Nguyên**

Nguồn: https://digitalorientalist.com/2026/03/24/buddhism-and-ai-another-look/

Tài liệu tham khảo:

1] L. Bates, The New Age of Sexism: How the AI Revolution is Reinventing Misogyny (London: Simon & Schuster, 2025).

2] S. Cochrane, “Asian Centres of Learning and Witness before 1000 CE: Insights for Today,” Transformation 26, 1 (2009): 30-39.

3] R. Dec, “Going Boldly Offline,” Substack, March 27, 2025. https://rachdele.substack.com/p/going-boldly-offline.

4] D. Duckworth, “A Buddhist Contribution to Artificial Intelligence?” Hualin International Journal of Buddhist Studies 3, 2 (2020): 27-37.

5] H. Ellison, “I Have No Mouth, and I Must Scream,” in IF: Worlds of Science Fiction (New York: Galaxy Publishing Corp, 1967), 162-175.

6] T. Hassine and Z. Neeman, “The Zombification of Art History: How AI Resurrects Dead Masters, and Perpetuates Historical Biases,” Journal of Science and Technology of the Arts 11.2 (2019): 28-35. https://doi.org/10.7559/citarj.v11i2.663.

7] E. Iori, “Releasing Urban Religion beyond the City Wall: The Spatial Capital of Early Buddhist Monasticism in NW South Asia,” Numen 70, 2-3 (2023): 184-219.

8] M. K. Kruthika and R. Amos, “The Negative Effects of Technology Addiction: A Comprehensive Review,” in Innovative Approaches to Multidisciplinary Exploration, eds. C. K. Uma Devi and R. Amos. N.p. (Red’unicorn Publishing, 2025), 19-22.

9] A. Labodda, “The STEM/Humanities Divide and Student Defeatism,” Blog of the APA, May 12, 2025. https://blog.apaonline.org/2025/05/12/the-stem-humanities-divide-and-student-defeatism/.

10] D. McMahan, “Modernity and the Early Discourse of Scientific Buddhism,” Journal of the American Academy of Religion 72, 4 (2004): 897-933.

11] L. Messeri and M. J. Crockett, “Artificial Intelligence and Illusions of Understanding in Scientific Research,” Nature 627, 8002 (2024): 49-58.

12] J. K. Teske, “Falsification of Interpretive Hypotheses in the Humanities,” Roczniki filozoficzne 66, 2 (2018): 87-106.

13] I. van Rooij, O. Guest, F. G. Adolfi, R. de Haan, A. Kolokolova, and P. Rich, “Reclaiming AI as a Theoretical Tool for Cognitive Science,” Computational Brain & Behavior 7 (2024): 616-636. https://doi.org/10.1007/s42113-024-00217-5.

14] C. P. Snow, The Two Cultures and the Scientific Revolution (New York: Cambridge University Press, 1961).

15] Bh. Sujato, “AI-3: There is No Road From Here to There,” SuttaCentral, April, 2024. https://discourse.suttacentral.net/t/ai-3-there-is-no-road-from-here-to-there/33426.

16] M. Vasylieva, “The Vanishing ‘Untranslated’ in the Age of AI: Challenges for Translator Training in Buddhist Studies,” Contemporary Buddhism 25 (2025): 1-22.

17] M. J. Verhoeven, “Science through Buddhist Eyes,” The New Atlantis 39 (2013): 107-118.

Chú thích

(1) Trong những năm gần đây đã hình thành một phong trào ngày càng lớn phản đối sự lệ thuộc của con người vào công nghệ. Nhiều nghiên cứu (Kruthika & Amos, 2025 cùng các tác giả khác) cho thấy ngày càng có nhiều người nhận thức rõ những hệ lụy tiêu cực của chứng nghiện công nghệ. Nhà tiểu luận và kinh tế học **Rachel Dec** từng viết: *“Chiếc điện thoại nhỏ trong túi tôi ngày càng trở thành một cánh cổng dẫn đến đau khổ.”* Điều này cũng đúng với phần lớn các công cụ số đang bao quanh cuộc sống hiện đại. Để “đóng cánh cổng” ấy, nhiều giải pháp đã được đưa ra, từ thiết bị **Brick** giúp khóa ứng dụng, điện thoại cơ bản (*dumb phone*), đến các chương trình **digital detox** nhằm giảm tình trạng quá tải thông tin 24/7 và áp lực phải luôn duy trì năng suất ngày càng cao.

(2) Nhiều học giả đã đặt câu hỏi về khả năng AI thực sự cải thiện tính khách quan và độ chính xác trong lịch sử nghệ thuật, khảo cổ học (Hassine & Neeman, 2019) cũng như trong nghiên cứu khoa học nói chung (Messeri & Crockett, 2024).

(3) Cuộc tranh luận về khoảng cách giữa khoa học tự nhiên và khoa học nhân văn đã được nhiều học giả bàn luận từ thập niên 1950, nổi bật là **C. P. Snow** với *The Two Cultures*, tiếp nối bởi **Umberto Eco**, **Eric Donald Hirsch** cùng nhiều học giả khác. Một tổng quan gần đây có thể xem ở Teske (2018). Tác giả cũng gợi ý bài viết của **Ashley Labodda**, phân tích cách sự phân mảnh tri thức trong giáo dục đại học ngày càng làm sâu sắc thêm sự chia cắt này.

(4) Những trao đổi về đạo đức và khoa học nhân văn từ góc nhìn Phật giáo gần đây cũng được **Elaine Lai** đề cập trong một bài nói chuyện và bài viết riêng; đây là một trong những nguồn cảm hứng cho bài viết này.

(5) Điều này không có nghĩa rằng mối quan hệ giữa Phật giáo, khoa học và công nghệ luôn đơn nhất hay đơn giản. Bản thân ba khái niệm ấy đều là những hiện tượng phức tạp, đa dạng và luôn biến đổi theo không gian, thời gian. Xem thêm McMahan (2004) và Verhoeven (2013).

(6) Do giới hạn dung lượng, tác giả không thể liệt kê đầy đủ các nghiên cứu về mối quan hệ giữa tự viện và đô thị ở nhiều thời kỳ khác nhau. Tuy nhiên, có thể tham khảo công trình của **Elisa Iori** (2023), nghiên cứu khảo cổ học Phật giáo tại thành phố **Barikot** (vùng **Swat**, Pakistan), minh họa rõ nét sự gắn kết giữa tự viện và đời sống đô thị.

(7) Để tìm hiểu sâu hơn về các phê phán triết học đối với AI, xem Duckworth (2020) và bài viết của **Bhante Sujato** trên **SuttaCentral**.

(8) Thật thú vị, trong khi Phật giáo nhấn mạnh ý nghĩa của thân thể như nền tảng của đời sống tâm linh, thì nhiều tỷ phú công nghệ lại tìm cách vượt qua chính "những thân xác hữu hạn" ấy. Họ đang đầu tư vào nhiều công nghệ – trong đó có AI – với tham vọng đánh bại cái chết hoặc tồn tại vĩnh viễn trong một vũ trụ số không còn thân xác. Theo tác giả, viễn cảnh ấy khiến người ta liên tưởng đến truyện ngắn khoa học viễn tưởng nổi tiếng *I Have No Mouth, and I Must Scream* (1967) của **Harlan Ellison**.

(9) Sau khi bài viết ban đầu được công bố, **Sam Altman**, Tổng giám đốc điều hành **OpenAI**, đã phần nào rút lại sự lạc quan trước đây của mình. Ông mô tả thỏa thuận hợp tác với Bộ Quốc phòng Hoa Kỳ là một quyết định "mang tính cơ hội và được thực hiện khá cẩu thả".