

AI: Rượu Mới Bào Da Cũ - Thách Thức Và Cơ Hội Từ Quan Điểm Luân Lý Công Giáo
AI: New Wine in Old Wineskins – Challenges and Opportunities from a Catholic Moral Perspective

Lm. JB. Lê Đình Phương^{1*}, C.Ss.R.

¹Học Viện Công Giáo Việt Nam

¹Dòng Chúa Cứu Thế

*Email của tác giả liên hệ: hocvienconggiao@gmail.com

 <https://doi.org/10.54855/csl.25537>

® Copyright (c) 2025 Lê Đình Phương

Received: 29/06/2025

Revision: 19/08/2025

Accepted: 21/08/2025

Online: 22/08/2025

ABSTRACT

Keywords: Artificial Intelligence; Catholic Morality; Human Dignity; Common Good; Gospel

This article examines Artificial Intelligence (AI) from the perspective of Catholic moral theology, highlighting both the challenges and opportunities it presents. AI is presented as a highly intelligent and efficient tool, while also raising concerns about risks and the responsibility of its creators and users. The article emphasizes human moral responsibility toward AI—whether it is used for good or ill depends on human choices, not on AI itself—and stresses that it is the human person, not AI, who remains the ultimate goal of development. AI is created for humanity, not the other way around. Finally, the article issues a call for personal renewal, a transformation of mindset, and the pursuit of new approaches consistent with Catholic teaching in order to confront the challenges and embrace the opportunities that AI brings to both society and the Church today.

TÓM LƯỢC

Từ khóa: Trí tuệ nhân tạo; Luân lý Công giáo; Phẩm giá con người; Công ích; Tin Mừng

Bài viết khảo sát Trí Tuệ Nhân Tạo (AI) từ quan điểm luân lý Công giáo, cho thấy những thách thức và cơ hội mà AI có thể tạo ra. Trình bày AI như một công cụ rất thông minh và hiệu quả, đồng thời cảnh báo rủi ro cũng như trách nhiệm của người tạo ra và sử dụng AI. Bài viết nhấn mạnh trách nhiệm luân lý của con người đối với AI – tốt hay xấu là do con người, chứ không phải tại AI, và nhấn mạnh rằng chính con người, chứ không phải AI, mới là mục tiêu của sự phát triển. AI được tạo ra vì con người, chứ không phải ngược lại. Cuối cùng, là lời kêu gọi mỗi người canh tân chính mình, đổi mới tư duy và tìm kiếm những cách tiếp cận mới, phù hợp với giáo lý Công giáo, để đối mặt với những thách thức và cơ hội mà AI đem lại cho cả xã hội và Giáo hội hôm nay.

DẪN NHẬP

AI (Artificial Intelligence), trí tuệ nhân tạo, là một hiện tượng mới nổi gây nhiều chú ý. AI đang mở ra “một kỷ nguyên mới”¹ và được xem là đại diện tiêu biểu của “cuộc cách mạng công nghiệp lần IV”,² làm thay đổi nhiều khía cạnh khác nhau của đời sống con người, ví dụ như trong các lĩnh vực kinh tế, giáo dục, y tế, an ninh...³

AI tạo cơ hội nâng cao năng lực của con người, có thể giải quyết nhiều vấn đề phức tạp và mang lại nhiều lợi ích lớn lao cho xã hội, nhưng cùng lúc có thể gây ra những rủi ro cho nhân phẩm, tự do, quyền riêng tư, công lý và tương lai của cả loài người.⁴

Theo Stephen Hawking, “thành công trong việc tạo ra AI có thể là sự kiện lớn nhất trong lịch sử loài người. Thật không may, cũng có thể đó là sự kiện cuối cùng, trừ phi chúng ta học được cách tránh được những rủi ro. Như một công cụ, AI có thể tăng cường trí thông minh hiện có của chúng ta để mở ra những tiến bộ trong mọi lĩnh vực khoa học và xã hội. Tuy vậy, nó cũng sẽ mang đến nhiều nguy hiểm”.⁵

Dẫu rằng còn quá sớm để nói về tương lai của AI, nhưng từ bây giờ người ta có thể nói rằng AI là một bước phát triển không thể đảo ngược.⁶ Có người còn ví von rằng ngày nay người ta không thể nhốt “thần đèn AI” vào bình nữa rồi!

Vấn đề là cần phát triển và sử dụng AI thế nào từ quan điểm đức tin và luân lý? Làm thế nào bảo đảm rằng AI phục vụ lợi ích chung và vô hại đối với mọi người? Liệu AI có tương thích với luân lý Công giáo và các giá trị Tin Mừng không? Liệu các cấu trúc và khuôn khổ đạo đức hiện nay có phù hợp để xử lý những vấn đề phức tạp do AI đặt ra không? Phải chăng cần phát triển một cách tiếp cận mới dựa trên các chuẩn mực luân lý đã có? Hay AI chẳng qua chỉ là vấn đề “rượu mới bình cũ”?

¹ Theo Henry A. Kissinger, Eric Schmidt & Daniel Huttenlocher, “ngày nay, một kỷ nguyên mới đang vẫy gọi chúng ta. Trong đó, một lần nữa, công nghệ sẽ biến đổi tri thức, sự khám phá, giao tiếp và suy nghĩ cá nhân...” trong *The Age of AI and Our Human Future – Thời đại AI và tương lai loài người chúng ta*, Mai Chí dịch, Nxb Trẻ, 181.

² Xem Byron Reese, *Thời Đại Thứ Tư — The Fourth Age*, Lương Trọng Vũ dịch, Nxb Dân Trí, TPHCM 2023.

³ Teofilo Giovan S. Pugeda, “Artificial Intelligence and Ethical Reflections from the Catholic Church,” in *MST Review* 23/2 (2021): 52-93, accessed Jan 31, 2024, https://mstreview.com/wp-content/uploads/2021/12/Pugeda_Artificial-Intelligence-and-Ethical-Reflections-.pdf?ssp=1&setlang=en&cc=CA&safesearch=moderate

⁴ Xem “*Tương lai của trí tuệ nhân tạo: hứa hẹn và rủi ro*”, trong *Sứ điệp Trí Tuệ Nhân Tạo và Hòa Bình* của ĐTC Phanxicô nhân Ngày Hòa bình Thế giới lần thứ 57, ngày 1 tháng 1 năm 2024, số 2.

⁵ Stephen Hawking, “Will Artificial Intelligence Outsmart Us”, in *Brief Answers to the Big Questions*, John Murray, London 2018, 186.

⁶ Asha Gaud, “The Irreversible Rise of Artificial Intelligence: A Brief History with Long-lasting Implications”, in MechoMotive, July 14, 2021, accessed Jan 30, 2024, <https://mechomotive.com/artificial-intelligence-rise-of-ai-brief-history-machines-evolution-superintelligence-benefits-risks-ethical-dilemmas/?ssp=1&setlang=en&cc=CA&safesearch=moderate>

AI, MỘT CÔNG CỤ TUYỆT VỜI

Như một công cụ, AI thật tuyệt vời! AI mô phỏng trí khôn con người,⁷ có thể thực hiện những công việc phức tạp cách nhanh chóng và hiệu quả đến mức vượt xa khả năng của con người, ví dụ chơi cờ, lái xe, nhận dạng khuôn mặt, dịch thuật...

Theo Henry Kissinger và một số tác giả, “trí tuệ nhân tạo là một phát minh của con người, phản chiếu các quy trình do con người thiết kế lên những cỗ máy do con người tạo ra. Tuy nhiên, trong một số trường hợp, với quy mô và tốc độ đáng kinh ngạc, nó tạo ra những kết quả tiệm cận với những kết quả mà, cho đến nay, chỉ có thể đạt được thông qua lý trí con người. Đôi khi những kết quả mà nó tạo ra khiến chúng ta phải sửng sốt”.⁸

AI thật thông minh và mạnh mẽ! Với năng lực của các siêu máy tính, các hệ thống AI ngày nay có thể xử lý lượng dữ liệu khổng lồ (big data) trong chớp mắt, có khả năng “tự học” (học máy)⁹ và đạt tới những hiểu biết mới chưa từng có (AI tạo sinh).

Học máy (machine learning): Ngày 7/12/2017, siêu máy tính “AlphaZero” của Google đã đánh bại chương trình “Stockfish 8”, vốn là nhà vô địch cờ vua máy tính thế giới năm 2016. AlphaZero đã học chơi cờ từ số không, nó đã đi từ chỗ hoàn toàn không biết gì đến bậc thầy sáng tạo trong bốn giờ tự chơi với chính mình, không cần sự chỉ dẫn nào từ con người.¹⁰

AI tạo sinh (generative AI): Hiện nay, GPT-3, có khả năng tạo ra văn bản giống người, là một trong những AI tạo sinh đáng chú ý nhất. Nó mở rộng cách tiếp cận bằng cách biến việc dịch thuật ngôn ngữ thành tạo ra ngôn ngữ. Khi được cho một vài từ, nó có thể ‘ngoại suy’ để tạo ra một câu, hoặc khi được cho một câu chủ đề, nó có thể ‘ngoại suy’ để tạo ra một bài văn.¹¹

Chính vì thế, AI vừa hấp dẫn vừa đáng sợ! AI có thể chấp cánh cho con người, thúc đẩy những nỗ lực tìm kiếm, sáng tạo và khai phá những điều mới mẻ, nhưng đồng thời cũng có khả năng lừa dối (tạo tin giả, Deepfake), gây xáo trộn (mất việc làm), thậm chí đe dọa hủy diệt (robots, vũ khí tự động... nếu rơi vào tay kẻ xấu).

⁷ Chẳng hạn trong *phép thử Turning* (Turning test), “nếu một cỗ máy hoạt động hiệu quả đến mức người quan sát không thể phân biệt hành vi của nó với hành vi của con người, thì cỗ máy đó được coi là thông minh... Đó là những cỗ máy có thể thực hiện các nhiệm vụ đặc trưng cho trí thông minh của con người”, xem Henry A. Kissinger et al., *The Age of AI and Our Human Future...*, 50.

⁸ Henry A. Kissinger et al., *The Age of AI...*, 181.

⁹ *Học máy* (machine learning): “Các AI ‘học’ bằng cách sử dụng dữ liệu, sau đó rút ra các quan sát và kết luận dựa trên dữ liệu. Trong khi các hệ thống trước đó đòi hỏi đầu vào và đầu ra chính xác, thì các AI với chức năng ‘mơ hồ’ này lại không đòi hỏi như vậy... Ở máy móc, các thuộc tính này mang tính cách mạng. Ví dụ như đột phá của AlphaZero trong thế giới cờ vua...”, xem Henry A. Kissinger et al., *The Age of AI...*, 51-68.

¹⁰ Yuval N. Harari, *21 Bài Học Cho Thế Kỷ 21*, Dương Ngọc Trà dịch, Nxb Thế Giới, Hà Nội 2022, 38-66, esp. 53.

¹¹ Henry A. Kissinger et al., *The Age of AI...*, 65-66.

Nhưng, liệu AI có thông minh tới mức có thể thay thế con người không?¹² Câu trả lời có thể là không. Như một công cụ, tự nó AI không thể thay thế con người, cả với người tạo ra và/hay người sử dụng nó, trừ phi con người từ bỏ quyền tự chủ¹³ và để AI quyết định thay mình.¹⁴

Thực tế, Yuval N. Harari và một số tác giả khác cho thấy lo ngại trên là có cơ sở!

Điều gì sẽ xảy ra khi con người ngày càng dựa vào AI để quyết định? Hiện tại, ta để Netflix gợi ý phim cho mình và Google Maps chọn rẽ phải hay rẽ trái. Nhưng một khi ta bắt đầu dựa vào AI để quyết định học cái gì, làm ở đâu và cưới ai, cuộc đời con người sẽ thôi không còn là một vở kịch của việc ra quyết định nữa.¹⁵

Stephen Hawking cũng bày tỏ quan ngại rằng:

Trong khi tác động ngắn hạn của AI tùy thuộc vào người kiểm soát nó, tác động dài hạn lại phụ thuộc vào việc liệu chúng ta còn khả năng kiểm soát được nó nữa hay không. (...) Một AI siêu thông minh sẽ cực kỳ giỏi trong việc hoàn thành các mục tiêu của nó và nếu những mục tiêu của nó không phù hợp với mục tiêu của chúng ta thì chúng ta sẽ gặp rắc rối.¹⁶

Còn Steve Wozniak, người đồng sáng lập của Apple thời kỳ đầu, thì nhận định:

Ngày nay, bỗng nhiên chúng ta mất quá nhiều quyền kiểm soát. Chúng ta không thể tắt Internet, không thể tắt điện thoại thông minh, không thể tắt máy tính. Ngày xưa, chúng ta tìm đến một người thông minh để đặt câu hỏi. Còn bây giờ, chúng ta hỏi ai? Đối tượng chúng ta hỏi cũng có tên bắt đầu bằng chữ “G” và “o”, nhưng đó là Google, chứ không phải God (Thiên Chúa).¹⁷

AI, MỘT SẢN PHẨM CỦA CON NGƯỜI

Dù vậy, AI là một sản phẩm do con người tạo ra. AI phản ánh các giá trị, ý định, lợi ích và văn hóa của người tạo ra nó.¹⁸ Do đó, AI không trung lập và có chủ ý vì nó kế thừa những quan điểm, thiếu sót và giới hạn của con người.

¹² Xem S. Hawking, “Will Artificial Intelligence Outsmart Us...”, 181-196.

¹³ Theo Paula Boddington, “con người không bao giờ nên từ bỏ quyền tự chủ đạo đức cho máy móc và máy móc phải ‘phù hợp’ với các giá trị của con người; nhưng chúng ta cũng cần xem xét những giả định rộng rãi về năng lực đạo đức của chúng ta và khả năng của AI sẽ tác động như thế nào đến cách chúng ta nghĩ về AI và về đạo đức,” xem Paula Boddington, “AI and Moral Thinking: How Can We Live Well With Machines To Enhance Our Moral Agency?”, in *AI and Ethics* (2021) 1:109–111, accessed Feb 3, 2024, <https://doi.org/10.1007/S43681-020-00017-0>

¹⁴ “Một khi AI đưa ra các quyết định tốt hơn chúng ta về nghề nghiệp và có thể là cả các mối quan hệ, khi đó quan niệm của chúng ta về nhân tính và về cuộc đời sẽ thay đổi. Con người đã quen nghĩ về cuộc đời như một vở kịch của việc ra quyết định,” Yuval N. Harari, *21 Bài Học...*, 80-81.

¹⁵ Yuval N. Harari, *21 Bài Học...*, 80.

¹⁶ S. Hawking, “Will Artificial Intelligence Outsmart Us...”, 188.

¹⁷ Byron Reese, *Thời Đại Thứ Tư...*, 59.

¹⁸ “Chúng ta cần nhớ rằng nghiên cứu khoa học và đổi mới công nghệ không hề tách rời và “trung lập”, mà chịu ảnh hưởng của văn hóa. Với tư cách là những hoạt động hoàn toàn của con người, những hướng đi họ thực hiện phản ánh những lựa chọn được điều chỉnh bởi các giá trị cá nhân, xã hội và văn hóa ở bất kỳ thời đại nào”, xem *Sứ điệp Trí Tuệ Nhân Tạo và Hòa Bình số 2*.

Vì thế, trách nhiệm của những người tạo ra AI trước tiên phải bảo đảm rằng AI được thiết kế, phát triển và áp dụng theo những chuẩn mực tôn trọng phẩm giá, quyền lợi và giá trị của con người, với tiêu chí lấy con người làm trung tâm: AI phải được tạo ra vì con người, chứ không phải ngược lại.

Vào năm 2020, Học viện Giáo hoàng về Sự sống đã ký một tuyên bố kêu gọi sử dụng AI cách có đạo đức và có trách nhiệm. Tuyên bố này bao gồm 6 nguyên tắc đạo đức hướng dẫn sự phát triển của trí tuệ nhân tạo đó là:

1. Minh bạch: Về nguyên tắc, hệ thống AI phải có thể giải thích được;
2. Bao hàm: Không ai bị loại trừ. Nhu cầu của mọi người phải được tính đến để tất cả được hưởng lợi và mỗi cá nhân được tạo điều kiện tốt nhất để có thể hoàn thiện bản thân và phát triển;
3. Trách nhiệm giải trình: Những người thiết kế và sử dụng AI phải tiến hành cách có trách nhiệm và minh bạch;
4. Không thiên vị: AI không tạo ra hoặc hành động theo thành kiến, nhờ đó bảo vệ được sự công bằng và phẩm giá con người;
5. Đáng tin cậy: AI phải trung thực và đáng tin cậy.
6. Bảo mật và quyền riêng tư: AI phải được bảo mật và tôn trọng quyền riêng tư của người sử dụng.¹⁹

Thực tế, với nỗ lực thực dụng và duy hiệu quả, bận tâm của con người ngày nay, cả người tạo ra AI lẫn người sử dụng, thường chỉ chú ý đến cái-có-thể-tạo-ra-hiệu-quả hay tạo-ra-lợi-nhuận-tối-đa. Nhưng từ quan điểm luân lý, hiệu quả tối đa hay mục đích người ta nhắm đến không thể biện minh cho phương tiện, nếu nó không phù hợp với các chuẩn mực đạo đức của con người.

Ở đây, cần phân biệt “điều có thể” với “điều được phép”, vì không phải mọi thứ khả thi về kỹ thuật đều có thể được chấp nhận về phương diện đạo đức. Do đó, người sáng tạo, thiết kế lẫn sử dụng AI cần phải tuân theo các nguyên tắc và chuẩn mực đạo đức hướng dẫn các lựa chọn và hành động của họ.

TRÁCH NHIỆM CỦA NGƯỜI SỬ DỤNG

AI xuất hiện trong thời đại chúng ta vừa như một cơ hội vừa là một thách đố. Vậy, liệu có nên sử dụng AI? Chắc chắn rồi, trừ phi ta muốn chôn giấu nén vàng Chúa trao.

Theo Henry Kissinger, ngày nay “ở cấp độ nền văn minh, việc chối bỏ AI sẽ là một lựa chọn bất khả thi.” Tuy nhiên, “nhân loại lúc nào cũng sẽ có ba lựa chọn chính: (1) hạn chế AI, (2) cộng tác với nó hoặc (3) tuân theo nó. Ba lựa chọn này sẽ xác định việc ứng dụng AI cho các nhiệm vụ hoặc lĩnh vực cụ thể”.²⁰

¹⁹ Jonah McKeown, “Sentient AI?: Here's what the Catholic Church says about artificial intelligence,” in CNA, Catholic News Agency, accessed on Jan 30, 2024, <https://www.catholicnewsagency.com/news/251552/sentient-ai-heres-what-the-catholic-church-says-about-artificial-intelligence?ssp=1&setlang=en&cc=CA&safesearch=moderate>

²⁰ Henry A. Kissinger et al., *The Age of AI...*, 181-185.

Như thế, tất cả vấn đề hệ tại ở chỗ chúng ta chọn lựa sử dụng AI thế nào cho đúng đắn và hữu ích? Paula Boddington ví von rất cụ thể như sau:

AI là một công cụ. Chúng ta cần làm chủ các công cụ của mình. Chúng ta sử dụng các công cụ để bổ sung cho điểm mạnh và điểm yếu của bản thân. Chúng ta sử dụng búa để tăng cường sức mạnh của mình, nhưng chúng ta phải có một sức mạnh nhất định, một mức độ chính xác nhất định, một kỷ luật và kỹ năng nhất định để sử dụng nó.²¹

Do đó, trước tiên người sử dụng cần biết rõ khả năng và những giới hạn của AI. AI có thể là một công cụ rất thông minh và hiệu quả, nhưng nó chỉ là một cỗ máy chứ không phải là một con người hay chủ thể đạo đức có khả năng lựa chọn tốt xấu, trách nhiệm đó thuộc về con người.²² Người sử dụng chịu trách nhiệm những gì AI tạo ra, nên về luân lý người dùng không được để AI quyết định thay mình.²³

Hơn nữa, mặc dầu mô phỏng trí thông minh của loài người, AI không có trái tim, cũng không có khả năng nhận thức, suy tư hoặc cảm nhận như con người,²⁴ nó không sở hữu những điều chúng ta gọi là “lẽ thường tình”, và cũng không biết điều nó không biết. Sự “ngu dốt” của AI phản ánh mức độ nông cạn của những gì nó học được. Do đó, việc giám sát và phát triển các quy trình để đánh giá liệu AI có hoạt động đúng hay không là rất quan trọng.²⁵

Thực tế, các hệ thống AI dù rất thông minh và hiệu quả, vẫn có những hạn chế nhất định và có thể sai lầm.

Ví dụ, trên giao diện Web của ChatGPT hiện nay, Cty OpenAI lưu ý người sử dụng rằng “ChatGPT có thể mắc lỗi. Hãy xem xét kiểm tra các thông tin quan trọng”.²⁶ Google cũng khuyến cáo tương tự: “Gemini có thể hiển thị thông tin không chính xác,

²¹ Paula Boddington, “AI and Moral Thinking...”, in *AI and Ethics* (2021) 1:109–111, accessed Feb 3, 2024, <https://doi.org/10.1007/S43681-020-00017-0>

²² Theo Goertzel, rốt cuộc con người phải chịu trách nhiệm về mọi hậu quả của máy móc. Xem Ben Goertzel, “Artificial General Intelligence: Concept, State of the Art, and Future Prospects”, trong *Journal of Artificial General Intelligence* 5, số 1 (2014): 1-46.

²³ Ví dụ, Sứ điệp *Trí Tuệ Nhân Tạo và Hòa Bình* số 6 bảo rằng “hệ thống vũ khí tự động không bao giờ có thể là đối tượng chịu trách nhiệm về mặt đạo đức. Năng lực độc nhất của con người trong việc phán xét đạo đức và ra quyết định có đạo đức không chỉ là một tập hợp phức tạp các thuật toán, và khả năng đó không thể bị giảm xuống thành việc lập trình một cỗ máy, dù “thông minh” đến mức nào đi nữa thì vẫn là một cỗ máy. Vì lý do này, điều bắt buộc là phải đảm bảo sự giám sát đầy đủ, có ý nghĩa và sự nhất quán của con người đối với các hệ thống vũ khí.” Cũng xem “Những hạn chế của AI và cách quản lý”, trong Henry A. Kissinger et al., *The Age of AI...*, 68-75.

²⁴ “Trí tuệ nhân tạo không phải là con người. Nó không biết hy vọng, cầu nguyện hay cảm nhận. Nó không có khả năng nhận thức hoặc suy nghĩ. Trí tuệ nhân tạo là một phát minh của con người, phản chiếu các quy trình do con người thiết kế lên những cỗ máy do con người tạo ra,” xem Henry A. Kissinger et al., *The Age of AI...*, 181.

²⁵ Xem Henry A. Kissinger et al., *The Age of AI...*, 68-72.

²⁶ “ChatGPT can make mistakes. Consider checking important information.” Accessed Feb 10, 2024.

bao gồm cả thông tin về con người, vì vậy hãy kiểm tra kỹ các câu trả lời”.²⁷ Do đó, người sử dụng AI phải có khả năng kiểm tra các kết quả do AI tạo ra.²⁸

Cùng lúc, đừng quên rằng AI rất giỏi “quyên rũ” và có khả năng “gây nghiện”! Chỉ cần nhắc tới vài cái tên như Tiktok, Facebook, Twister hay smart phone là đủ. Chính vì thế, người sử dụng phải có khả năng làm chủ bản thân và công cụ mình sử dụng. Nếu quá lệ thuộc vào AI, cho phép AI suy nghĩ hay quyết định thay mình, người sử dụng dần dà sẽ không còn khả năng suy tư, phản biện và sẽ trở thành nô lệ của AI. Hãy nhớ rằng AI có thể là *một đấng tốt, nhưng là một ông chủ xấu!*

Theo Teofilo Giovan S. Pugeda, chúng ta càng dựa vào máy móc thì khả năng tự lực càng kém. Ví dụ: điện thoại thông minh cho phép chúng ta thu thập kiến thức nhanh hơn bất kỳ thời điểm nào khác trong lịch sử loài người. Tuy nhiên ngày nay, nếu một cá nhân bị mất điện thoại thông minh, người đó sẽ cảm thấy bất lực, bất an và không còn biết làm gì cả.²⁹

Hơn nữa, AI giới hạn như người sử dụng nó. Thực tế, khả năng của AI bị giới hạn bởi khả năng vận dụng của người dùng và cách nó được sử dụng. Chính vì thế, người sử dụng cần trang bị cho mình những kỹ năng cần thiết, nhất là biết vận dụng AI cách khôn ngoan và đạo đức.

Ví dụ, việc biết các ngoại ngữ, như tiếng Anh hay tiếng Pháp, là một lợi thế rất lớn, việc biết cách đặt vấn đề hay tạo prompt cách thông minh cũng là những kỹ năng quan trọng trong kỷ nguyên AI... Nhưng kỹ năng quan trọng hơn cả, không chỉ là biết sử dụng AI cách thành thạo, nhưng là sử dụng có kỷ luật và có đạo đức.

Thực tế, như bất kỳ phương tiện nào (tiền bạc, tài năng, sức khỏe...), AI có thể bị lạm dụng, thao túng hay bị chiếm đoạt (hacked) vì mục đích xấu. Do đó, như một công cụ, AI chỉ thông minh, hiệu quả và tốt lành nếu nó được sử dụng cách khôn ngoan, sáng suốt và có đạo đức.

Vì lý do này, trong Sứ điệp *Trí Tuệ Nhân Tạo và Hòa Bình*, ĐTC Phanxicô bảo:

Cần giáo dục cách sử dụng các hình thức trí tuệ nhân tạo, bằng cách huấn luyện tư duy phê phán. Người dùng ở mọi lứa tuổi, đặc biệt là giới trẻ, cần phát triển cách tiếp cận sáng suốt trong việc sử dụng dữ liệu và nội dung được thu thập trên các trang mạng hoặc do hệ thống trí tuệ nhân tạo tạo ra. Các trường trung học, đại học và hiệp hội khoa học cần giúp các sinh viên và các chuyên gia nắm bắt các khía cạnh xã hội và đạo đức của việc phát triển và sử dụng công nghệ.³⁰

Do đó, trách nhiệm của người dùng AI là sử dụng thế nào để AI thực sự là công cụ hữu ích và hiệu quả, tất cả tùy ở chọn lựa khôn ngoan của mỗi cá nhân!

²⁷ “Gemini may display inaccurate info, including about people, so double-check its responses.” Acc. Feb 10, 2024.

²⁸ Theo H. Kissinger, “khi đã có kết quả do AI tạo ra, con người – có thể là nhà nghiên cứu hay người sử dụng – phải kiểm chứng được liệu AI tạo ra những gì mình mong muốn,” trong Henry A. Kissinger *et al.*, *The Age of AI...*, 69-70.

²⁹ Còn theo Kissinger *et al.*, “AI cũng đang thực hiện một phép trừ. Nó đẩy nhanh những động lực làm xói mòn lý trí của con người (ví dụ: mạng xã hội làm giảm khả năng suy nghĩ và xuất sắc trong việc gây nghiện). Dần dà, vai trò của lý trí con người trong việc xem xét, kiểm tra và hiểu thông tin sẽ giảm đi, trong khi vai trò của AI sẽ được mở rộng... Trong kỷ nguyên AI, lý trí con người vừa được tăng cường vừa bị suy giảm,” xem *The Age of AI...*, 183-184.

³⁰ Sứ điệp *Trí Tuệ Nhân Tạo và Hòa Bình* số 7.

TỐT HAY XẤU ĐỀU DO CON NGƯỜI

Như thế, nguyên nhân của vấn đề không phải do AI mà do con người.

AI xuất hiện như một hiện tượng mới mẻ làm lộ ra phẩm giá của con người là loài thông minh (*homo sapiens*), được tạo dựng theo hình ảnh của Thiên Chúa, có trí khôn và sự tự do để chọn điều lành hay điều dữ. Do đó, chính con người phải chịu trách nhiệm về những gì mình làm chứ không thể đổ lỗi cho AI, cũng như Adong và Evà không thể đổ lỗi cho Thiên Chúa về việc tại sao lại có cái cây ở giữa vườn (St 2,9-17). AI chỉ ra khả năng và giới hạn của con người.

Quả thực, AI là một phát minh rất tuyệt vời của con người thời nay. Vấn đề là con người tạo ra và sử dụng AI để làm gì? Để tôn vinh Thiên Chúa, để phục vụ lợi ích chung của loài người và giúp nhau đạt tới ơn cứu độ, hay chỉ nhằm có thêm một phương tiện, một thứ vũ khí nữa, để tăng khả năng cạnh tranh giữa các cá nhân, phe nhóm, quốc gia... trên thương trường và cả trên chiến trường!

Chuyện kể rằng một ngày kia có hai người bạn thân vào rừng săn bắn. Bỗng có một vị thần hiện ra và cho cả hai một điều ước. Họ ước gì cũng được, vấn đề là nếu người ước được một thì người còn lại được hai. Cả hai đùn đẩy, không ai muốn ước trước cả, vì không ai muốn người kia hơn mình. Cuối cùng, vị thần thúc giục: “Nói mau, nếu không ta biến!” Một người vội giơ tay và ước: “Tôi ước chỉ mình... mù một mắt!!!” Tức thì người kia mù cả đôi!!! Gặp thần đèn tưởng là may, ai ngờ là họa, chung quy chỉ vì lòng ganh đua, ghen tị cách bệnh hoạn!

Do đó, AI hay bất kỳ điều gì xảy ra trong thế giới này là phúc hay họa, tất cả tùy thuộc ở sự chọn lựa và mục tiêu con người nhắm đến với ý thức và tự do. Do đó, “làm lành, lánh dữ” vẫn luôn nguyên tắc đạo đức căn bản và không bao giờ lỗi thời!

LIỆU AI CÓ THỂ CỨU ĐỘ CON NGƯỜI KHÔNG?

Với tư cách là công cụ hay phương tiện thì dù mạnh mẽ tới đâu, AI không phải là cứu cánh hay có thể cứu độ con người, vì tự mình AI không thể mang lại ý nghĩa tròn đầy và mục đích tối hậu cho sự hiện hữu của nó cũng như của con người.

Tuy vậy, như một công cụ mới mẻ và hiệu quả, có thể nói AI là một “dấu chỉ”, là “*Kairos*” của thời đại này, có thể tạo ra những cơ hội và khả năng giúp con người thăng tiến nhiều phương diện, có thể kiến tạo một cuộc sống tốt đẹp hơn, thậm chí có thể là một khí cụ tuyệt vời cho sứ vụ loan báo Tin Mừng của Hội thánh.³¹

Nhưng làm thế nào để AI có thể trở thành cơ hội, thành *Kairos* cho thời đại này? Chúng ta cần phải làm gì để AI có thể là cơ hội cho sự phát triển con người toàn diện, và là “dấu chỉ của thời đại” về sự hiện diện và hành động của Thiên Chúa trong thế giới hôm nay?

Một lần nữa, AI tiếp tục là thách đố và là cơ hội cho tất cả chúng ta, tất cả tùy thuộc vào chọn lựa của mỗi người, nhất là những chọn lựa trong đức tin.

RƯỢU MỚI, BẦU DA MỚI

Trước tiên, như một thứ rượu mới, AI có khả năng làm vỡ bầu da cũ là chính chúng ta, những con người được sinh ra và lớn lên trong một cấu trúc xã hội, văn hoá và luật lệ nhất định, với

³¹ Có thể nghĩ đến việc tạo ra các Chatbot tương tự như ChatGPT để cung cấp thông tin, dạy giáo lý, tư vấn mục vụ... như trường hợp của trang “Magisterium AI”, xem <https://www.magisterium.com>

những hiểu biết, định kiến, chuẩn mực và phương thế của một thời giờ đầy đường như đã khô cứng, lỗi thời.

Về phương diện kinh tế, bầu vờ sẽ dẫn tới nhiều hệ lụy nghiêm trọng, chẳng hạn mất việc, nhưng cùng lúc AI cũng hứa hẹn tạo ra nhiều cơ hội mới!

Theo báo cáo mới nhất của Diễn đàn Kinh tế Thế giới 2023 (WEF) tại Davos, nguy cơ mất việc do AI là một trong những thách thức lớn nhất đối với nền kinh tế toàn cầu: Khoảng 40% công việc của con người hiện nay sẽ rơi vào tay AI.³² Đàng khác, theo cái nhìn lạc quan của Byron Reese, “sự thật là công nghệ hầu như luôn làm *tăng* lượng công nhân chứ không thay thế họ”.³³

Nhưng dù thế nào, ngày nay người lao động phải nhanh chóng thích ứng bằng cách cập nhật kiến thức, nâng cao kỹ năng và đổi mới chính mình. Cách riêng, trong môi trường giáo dục, AI tạo ra rất nhiều thách đố và cơ hội cho cả thầy lẫn trò, cho nên đã đến lúc cần thích nghi và cải tiến cách dạy và học.³⁴

Dĩ nhiên, khi đã có máy vi tính rồi, chẳng ai còn muốn quay lại dùng máy đánh chữ cơ học nữa. Nhưng vấn đề không phải là việc có thêm các thiết bị hiện đại, nhưng là sử dụng chúng thế nào cho hữu ích. Người dạy và người học cần phải thích ứng và vận

³² World Economic Forum, *The Future of Jobs Report*, Inside Report May 2023, accessed Feb 3, 2024, https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf

³³ Byron Reese, *Thời Đại Thứ Tư...*, 151-152, 170-179.

³⁴ Việc cải tiến phương pháp dạy và học trong thời đại AI là một thách thức phức tạp và liên tục đòi hỏi sự sáng tạo, linh hoạt và tư duy phản biện. Không có giải pháp nào phù hợp cho tất cả. Sau đây là một số đề xuất:

- Coi AI như một công cụ chứ không phải một mối đe dọa: Thay vì lo sợ hoặc né tránh AI, các nhà giáo dục và người học nên khám phá cách sử dụng AI để nâng cao kinh nghiệm giáo dục của mình.
- Phát triển kiến thức về AI: Kiến thức về AI là khả năng hiểu, sử dụng và đánh giá các hệ thống AI cũng như ý nghĩa của chúng. Thầy và trò nên tìm hiểu về những kiến thức cơ bản về AI, chẳng hạn như cách AI hoạt động, những gì nó có thể và không thể làm cũng như các vấn đề đạo đức và xã hội liên quan.
- Tập trung vào các kỹ năng bền vững: Kỹ năng bền vững là những kỹ năng có thể chuyển giao, thích ứng và phù hợp trong một thế giới đang thay đổi. Chúng bao gồm tư duy phê phán, sáng tạo, giải quyết vấn đề, hợp tác và trí tuệ cảm xúc. Cần ưu tiên những kỹ năng này hơn là ghi nhớ và học thuộc lòng, là những điều AI có thể dễ dàng sao chép.
- Tham gia vào hoạt động học tập tích cực và hợp tác: Việc này đòi sinh viên tham gia, tương tác với nhau và áp dụng kiến thức của mình vào các tình huống thực tế. AI có thể tạo điều kiện thuận lợi cho các phương thức học tập này bằng cách cung cấp các công cụ tương tác, mô phỏng, phản hồi và kết nối toàn cầu.
- Liêm chính trong học tập: Tính liêm chính trong học thuật là cam kết về sự trung thực, công bằng và trách nhiệm trong công việc học tập. AI đặt ra những thách thức và cơ hội mới cho tính liêm chính trong học thuật, chẳng hạn như việc sử dụng các công cụ AI tổng hợp, phát hiện đạo văn và xác minh nguồn.

Đây là một số cách để cải thiện phương pháp dạy và học trong thời đại AI, nhưng vẫn còn nhiều khả năng khác để khám phá và thử nghiệm. AI không phải là một hiện tượng tĩnh mà là một hiện tượng năng động và đang phát triển, và các phương pháp dạy và học với nó cũng phải như vậy. Nguồn: Cuộc trò chuyện với Bing, 20/2/2024. Cf. “Teaching in the Age of AI” - Notre Dame Learning. <https://learning.nd.edu/resources/teaching-in-the-age-of-ai/>; “Teaching and Learning in the Age of AI: Considerations, Resources, and...” <https://ctl.columbia.edu/resources-and-technology/resources/teaching-learning-ai/>; “Practical Strategies for Teaching with AI” - MIT Sloan Teaching... <https://mitsloanedtech.mit.edu/ai/teach/practical-strategies-for-teaching-with-ai/>; “AI Tools in Teaching and Learning” | Teaching Commons. <https://teachingcommons.stanford.edu/news/ai-tools-teaching-and-learning>.

dụng những điều kiện mới mẻ mà AI tạo ra trong môi trường nghiên cứu học thuật, để có thể nghiên cứu và học tập cách trung thực, thông minh và hiệu quả.

Vì thế, AI phát triển thì con người cũng cần phải phát triển, nếu chúng ta không muốn bản thân rơi vào nguy cơ bị “tái mù chữ”, mất việc, bị đào thải, hoặc trở thành những kẻ vô dụng trong kỷ nguyên AI.³⁵

Theo Harari, “nguy cơ nằm ở chỗ nếu chúng ta đầu tư quá nhiều vào phát triển AI và quá ít vào phát triển con người thì trí tuệ nhân tạo rất phức tạp của máy tính có thể chỉ làm gia tăng sự ngu dốt bẩm sinh của con người mà thôi”.³⁶

Còn theo sự khôn ngoan của Tin Mừng: *Rượu mới cần những bầu da mới!* Do đó, người tạo ra hoặc sử dụng AI, không chỉ cần học hỏi, cải tiến, mà còn phải không ngừng hoán cải, đổi mới suy nghĩ và trái tim theo các giá trị Tin Mừng. Dù sống trong thời đại nào, mỗi người phải hành động theo lương tâm mình, biết làm lành lánh dữ, tôn trọng con người và phẩm giá của họ, và yêu thương tha nhân như chính mình, nhất là những người bé mọn. Cho nên, điều cần cải thiện và phát triển không phải là các hệ thống AI, mà chính là con người. Người tốt thì AI sẽ tốt!

Cách riêng trong kỷ nguyên AI, rất cần có những nỗ lực suy tư, tiếp cận và những chính sách phù hợp với các chuẩn mực đạo đức, pháp lý và tinh thần Tin Mừng, để có thể hướng dẫn và giải quyết các vấn đề mới do AI đặt ra, không chỉ ở mức độ cá nhân hay quốc gia, mà cả xã hội và toàn thể Hội thánh.

Theo Henry Kissinger, “chúng ta nhất thiết phải có một bộ quy tắc đạo đức AI. (...) Nếu việc quản trị con người cần được hướng dẫn bởi một bộ quy tắc đạo đức, thì AI cũng đòi hỏi một bộ quy tắc đạo đức của riêng nó – một bộ quy tắc phản ánh không chỉ bản chất của công nghệ này mà cả những thách thức do nó đặt ra”.³⁷

Do đó, thách thức lớn nhất hiện nay của chúng ta đó là cần phải làm gì để có thể tạo ra một bộ quy tắc như thế? Câu trả lời không đơn giản. Nhiều chuyên gia và tổ chức đã đề xuất một số hướng dẫn và nguyên tắc đạo đức về AI.³⁸

³⁵ Cf. “Công việc”, trong Yuval N. Harari, *21 Bài Học...*, 38-60.

³⁶ *Ibid.*, 97.

³⁷ Henry A. Kissinger *et al.*, *The Age of AI...*, 188-190.

³⁸ Chẳng hạn:

- AI phải được phát triển và sử dụng cho những mục đích tốt, ví dụ nâng cao phúc lợi của con người, công bằng xã hội và tính bền vững của môi trường.
- AI phải tôn trọng phẩm giá, quyền và giá trị của con người, đồng thời tránh phân biệt đối xử, thiên vị và gây tổn hại.
- AI phải minh bạch, có thể giải thích, để người dùng có thể hiểu cách AI hoạt động, chức năng của nó và chịu trách nhiệm về kết quả nó tạo ra.
- AI phải chịu sự giám sát và kiểm soát của con người, đồng thời cho phép con người can thiệp và sửa chữa.
- AI phải đáng tin cậy, an toàn và bảo mật, đồng thời tuân thủ các tiêu chuẩn cao về phẩm chất, an toàn và hiệu quả.³⁸

Và để thiết lập quy tắc đạo đức về AI, một số bước có thể thực hiện là:

- Tham gia vào cuộc đối thoại các bên liên quan bao gồm các nhà phát triển AI, người dùng, cơ quan quản lý, xã hội dân sự và các bên liên quan khác.

Cuối cùng, theo ĐTC Phanxicô:

Một quan điểm nhân bản đích thực và ước mơ về một tương lai tốt đẹp hơn cho thế giới của chúng ta chắc chắn cho thấy sự cần thiết phải có sự đối thoại liên ngành nhằm phát triển đạo đức của các thuật toán – *một đạo đức thuật toán (an algor-ethics)* – theo đó các giá trị sẽ xác định hướng đi của các công nghệ mới.³⁹

KẾT LUẬN

Để kết luận, có thể nói, AI là “dấu chỉ”, là “Kairos” hay “thần đèn” của Thế kỷ 21, vừa hấp dẫn vừa đáng sợ!

Tuy vậy, AI tốt hay xấu, là cơ hội hay nguy cơ, mang lại hạnh phúc hay bất hạnh... tất cả tùy vào chọn lựa và quyết định tự do của mỗi chúng ta. Để AI thực sự là công cụ hữu ích mang lại hạnh phúc cho mọi người, mỗi người cần cải thiện bản thân, biết hành động cách khôn ngoan và có trách nhiệm.

Hơn nữa, “rượu mới cần bầu da mới”, cần canh tân đổi mới cấu trúc xã hội và cần có những quy chế đạo đức và pháp lý phù hợp với những vấn đề mới do AI tạo ra. Nền tảng của những chuẩn mực hay quy định này có thể tìm thấy trong lương tri của mỗi người và trong mạc khải của Ngôi Lời nhập thể, “Đấng đã đến để cho con người được sống và sống dồi dào” (Ga 10,10).

-
- Phát triển và áp dụng các quy tắc ứng xử tự nguyện, các biện pháp thực hành tốt nhất và cơ chế tự điều chỉnh phản ánh các nguyên tắc và giá trị đạo đức của cộng đồng AI.
 - Thiết lập và thực thi các khuôn khổ pháp lý và quy định nhằm bảo vệ nhân quyền, quyền riêng tư và dữ liệu, đồng thời đảm bảo trách nhiệm giải trình và trách nhiệm pháp lý đối với các tác hại của AI.
 - Thúc đẩy việc giáo dục, nhận thức và hiểu biết về đạo đức AI trong công chúng và các chuyên viên.
 - Vun trồng văn hóa tin cậy, trách nhiệm và hợp tác giữa tất cả các bên tham gia vào hệ sinh thái AI.

Nguồn: Cuộc trò chuyện với Bing, 23/2/2024: (1) *Minister Champagne launches voluntary code of conduct relating to* <https://www.canada.ca/en/innovation-science-economic-development/news/2023/09/minister-champagne-launches-voluntary-code-of-conduct-relating-to-advanced-generative-ai-systems.html>. (2) *Canada's voluntary AI code of conduct is coming* - CBC.ca. <https://www.cbc.ca/news/business/ai-code-of-conduct-stopgap-1.6983064>. (3) *AI Ethics: What It Is and Why It Matters* | Coursera. <https://www.coursera.org/articles/ai-ethics>. (4) *Ethics of artificial intelligence* - Wikipedia. https://en.wikipedia.org/wiki/Ethics_of_artificial_intelligence.

³⁹ Sứ điệp *Trí Tuệ Nhân Tạo và Hòa Bình* số 6. Cũng xem José Roger Flahaux, Brian Patrick Green, and Ann Skeet, *Ethics in the Age of Disruptive Technologies: An Operational Roadmap*, The ITEC Handbook, Santa Clara University, Santa Clara 2023, accessed Feb 3, 2024: <https://www.scu.edu/institute-for-technology-ethics-and-culture/itec-handbook/>; Unesco, *Recommendation on the Ethics of Artificial Intelligence*. UNESCO's first-ever global standard on AI ethics – the ‘Recommendation on the Ethics of Artificial Intelligence’ – was adopted by all 193 Member States in November 2021, Unesco 2022, accessed Feb 3, 2024: <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>.

TÀI LIỆU THAM KHẢO

- Asha Gaud, “The Irreversible Rise of Artificial Intelligence: A Brief History with Long-lasting Implications”, in MechoMotive, July 14, 2021, accessed Jan 30, 2024, <https://mechomotive.com/artificial-intelligence-rise-of-ai-brief-history-machines-evolution-superintelligence-benefits-risks-ethical-dilemmas/?ssp=1&setlang=en&cc=CA&safesearch=moderate>
- Byron Reese, *Thời Đại Thứ Tư — The Fourth Age*, Lương Trọng Vũ dịch, Nxb Dân Trí, TPHCM 2023.
- “Gemini may display inaccurate info, including about people, so double-check its responses.” Acc. Feb 10, 2024.
- Henry A. Kissinger *et al.*, *The Age of AI...*, 181.
- Học máy (machine learning): “Các AI ‘học’ bằng cách sử dụng dữ liệu, sau đó rút ra các quan sát và kết luận dựa trên dữ liệu. Trong khi các hệ thống trước đó đòi hỏi đầu vào và đầu ra chính xác, thì các AI với chức năng ‘mơ hồ’ này lại không đòi hỏi như vậy... Ở máy móc, các thuộc tính này mang tính cách mạng. Ví dụ như đột phá của AlphaZero trong thế giới cờ vua...,” xem Henry A. Kissinger *et al.*, *The Age of AI...*, 51-68.
- Stephen Hawking, “Will Artificial Intelligence Outsmart Us”, in *Brief Answers to the Big Questions*, John Murray, London 2018, 186.
- Sứ điệp *Trí Tuệ Nhân Tạo và Hòa Bình* số 7.
- Theo H. Kissinger, “khi đã có kết quả do AI tạo ra, con người – có thể là nhà nghiên cứu hay người sử dụng – phải kiểm chứng được liệu AI tạo ra những gì mình mong muốn,” trong Henry A. Kissinger *et al.*, *The Age of AI...*, 69-70.
- Teofilo Giovan S. Pugeda, “Artificial Intelligence and Ethical Reflections from the Catholic Church,” in *MST Review* 23/2 (2021): 52-93, accessed Jan 31, 2024, https://mstreview.com/wp-content/uploads/2021/12/Pugeda_Artificial-Intelligence-and-Ethical-Reflections-.pdf?ssp=1&setlang=en&cc=CA&safesearch=moderate
- “*Tương lai của trí tuệ nhân tạo: hứa hẹn và rủi ro*”, trong Sứ điệp *Trí Tuệ Nhân Tạo và Hòa Bình* của ĐTC Phanxicô nhân Ngày Hòa bình Thế giới lần thứ 57, ngày 1 tháng 1 năm 2024, số 2.
- World Economic Forum, *The Future of Jobs Report*, Inside Report May 2023, accessed Feb 3, 2024, https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf
- Yuval N. Harari, *21 Bài Học Cho Thế Kỷ 21*, Dương Ngọc Trà dịch, Nxb Thế Giới, Hà Nội 2022, 38-66, esp. 53.

Biodata

Lm. JB. Lê Đình Phương là tu sĩ linh mục Dòng Chúa Cứu Thế và là Giáo sư tại Học Viện Công Giáo Việt Nam.

Fr. JB. Le Dinh Phuong is a priest of the Congregation of the Most Holy Redeemer (Redemptorists), and also a professor at the Catholic Institute of Vietnam.