SỬ DỤNG TRÍ TUỆ NHÂN TẠO TRONG NGHIÊN CỨU KHOA HỌC: CƠ HỘI, THÁCH THỨC VÀ VẤN ĐỀ LIÊM CHÍNH HỌC THUẬT

Đào Trung Hiếu[1]
Email: hieubaocand@gmail.com

Ngày tòa soạn nhận được bài báo: 02/06/2025
Ngày phản biện đánh giá: 15/08/2025
Ngày bài báo được duyệt đăng: 29/08/2025

DOI: 10.59266/houjs.2025.667

Tóm tắt: *Bài viết phân tích một cách hệ thống cơ hội, thách thức và các vấn đề liêm chính học thuật phát sinh từ việc ứng dụng trí tuệ nhân tạo (AI) trong nghiên cứu khoa học hiện đại. Trên cơ sở khảo sát các xu hướng công nghệ và chuẩn mực học thuật quốc tế, bài viết làm rõ những tác động đa chiều của AI tới chu trình sản xuất tri thức, đồng thời đề xuất giải pháp nhằm sử dụng AI có trách nhiệm trong nghiên cứu, bảo đảm tính trung thực và minh bạch khoa học trong bối cảnh chuyển đổi số.*

Từ khóa: *trí tuệ nhân tạo, nghiên cứu khoa học, liêm chính học thuật, đạo văn, chuẩn mực quốc tế*

I. Mở đầu

Trí tuệ nhân tạo (AI) đang trở thành một công nghệ đột phá, tác động sâu sắc đến nhiều lĩnh vực của xã hội hiện đại, đặc biệt là giáo dục và nghiên cứu khoa học. Sự phát triển của các mô hình ngôn ngữ lớn (LLM) như ChatGPT (OpenAI), Gemini (Google), Claude (Anthropic) và Copilot (Microsoft) đã mở ra khả năng mới cho việc tự động hóa, truy xuất thông tin, mô hình hóa và trình bày kết quả nghiên cứu *(Floridi & Chiriatti, 2023).*

AI hiện hỗ trợ nhà nghiên cứu trong nhiều khâu: từ tổng quan tài liệu, phân tích dữ liệu định tính - định lượng, hình thành giả thuyết, đến soạn thảo và hiệu đính văn bản khoa học. Các công cụ như Elicit, Semantic Scholar AI, Scite, Writefull hay Grammarly góp phần cải thiện hiệu suất viết học thuật và gia tăng cơ hội công bố quốc tế *(Springer Nature, 2023).*

Tuy nhiên, song hành với cơ hội là các rủi ro: nguy cơ đạo văn tinh vi, trích dẫn giả mạo, sự lệ thuộc vào tư duy máy móc và thiếu minh bạch trong khai báo công cụ hỗ trợ. Nhiều tạp chí và tổ chức quốc tế đã khẳng định: AI không được coi là tác giả; mọi sử dụng phải khai báo minh bạch; và trách nhiệm cuối cùng thuộc về

[1] Khoa Luật, Trường Đại học Mở Hà Nội

con người *(COPE, 2023; Nature, 2023; Science, 2023)*. Ở Việt Nam, cảnh báo đã xuất hiện nhưng vẫn thiếu một khung chính sách thống nhất từ cấp quản lý nhà nước, ngoài Chiến lược quốc gia về AI đến 2030 *(QĐ 127/QĐ-TTg, 2021)*.

Với trải nghiệm trực tiếp trong công tác giảng dạy và nghiên cứu pháp lý, tác giả nhận thấy sự xuất hiện của AI không chỉ là vấn đề kỹ thuật mà còn đặt ra thách thức căn bản về chuẩn mực học thuật. Nhiều sinh viên của tác giả khi tiếp cận công cụ AI đã lúng túng trong việc phân định đâu là hỗ trợ hợp pháp và đâu là hành vi vi phạm liêm chính. Điều này cho thấy tính cấp thiết của một khung hướng dẫn rõ ràng.

II. Câu hỏi, mục tiêu

Bài viết đặt ra các câu hỏi khoa học cốt lõi: *Thứ nhất,* việc ứng dụng AI mang lại lợi ích gì đối với hiệu quả, chất lượng và tính công bằng trong sản xuất tri thức? *Thứ hai,* những nguy cơ nào về đạo đức, pháp lý và liêm chính học thuật có thể phát sinh từ việc sử dụng AI trong nghiên cứu? *Thứ ba,* cộng đồng khoa học quốc tế đã đưa ra các chuẩn mực gì để định hướng việc ứng dụng AI một cách có trách nhiệm? *Và cuối cùng,* Việt Nam cần có những chính sách gì để bảo đảm liêm chính học thuật trong bối cảnh số *(UNESCO, 2023; WAME, 2023)*.

Trên cơ sở đó, mục tiêu nghiên cứu của bài viết là: (i) Hệ thống hóa các ứng dụng tiêu biểu của AI trong nghiên cứu khoa học hiện đại, qua đó làm rõ những giá trị và cơ hội mà công nghệ này mang lại; (ii) Nhận diện các thách thức liên quan đến đạo đức, pháp lý và học thuật; (iii) Đối chiếu kinh nghiệm quốc tế để đề xuất khuyến nghị cho Việt Nam.

III. Phương pháp nghiên cứu

Về phương pháp, bài viết sử dụng kết hợp: phân tích - tổng hợp tài liệu; so sánh luật học giữa chính sách quốc tế và Việt Nam; khảo sát bàn giấy (desk-scan) trên các nguồn chính thống trong nước giai đoạn 2020 - 2025; và bình luận khoa học nhằm đưa ra đánh giá, kiến nghị *(Stanford HAI, 2023; Turnitin, 2023)*.

IV. Tổng quan nghiên cứu và cơ sở lý luận

Trong những năm gần đây, AI trở thành chủ đề nổi bật trong học thuật quốc tế. Theo *AI Index Report 2024* của Stanford HAI, số lượng công bố có từ khóa liên quan đến "AI in science" tăng mạnh mẽ, phản ánh sự lan tỏa toàn cầu của công nghệ này *(Stanford HAI, 2024)*.

Về lý luận, có hai trường phái chính. Trường phái lạc quan cho rằng AI là công cụ dân chủ hóa tri thức, nâng cao hiệu suất và giảm bất bình đẳng tiếp cận học thuật. Hutson (2021) mô tả AI như "trợ lý tri thức" có thể xử lý công việc lặp lại, giải phóng thời gian cho nghiên cứu sáng tạo.

Ngược lại, trường phái cảnh báo lo ngại AI có thể làm xói mòn tư duy phản biện, phổ biến sản phẩm học thuật "trơn tru" nhưng rỗng nội dung, và tiềm ẩn nguy cơ đạo văn tinh vi, dữ liệu giả mạo *(Floridi & Chiriatti, 2023; Crothers & Zhang, 2023)*.

Song song với đó, các tổ chức quốc tế đã ban hành khung quy định: COPE (2023), WAME (2023), Nature (2023), Science (2023), Elsevier (2023), Springer Nature (2023), Wiley (2023) đều thống

nhất rằng: AI không được đứng tên tác giả; phải khai báo minh bạch; trách nhiệm cuối cùng thuộc về con người.

Tại Việt Nam, các nghiên cứu liên quan đến tác động của AI trong lĩnh vực học thuật còn khá sơ khai và phân tán. Một số bài viết chỉ dừng ở việc nhận diện hiện tượng, chưa phân tích sâu về khung đạo đức, chính sách hoặc khuyến nghị hành động cụ thể. Bộ Giáo dục và Đào tạo đã tổ chức nhiều hội thảo về chuyển đổi số và ứng dụng AI trong giáo dục *(MOET, 2024)*, trong khi một số trường đại học lớn như Đại học Quốc gia Hà Nội *(VNU, 2024)* hay *RMIT Việt Nam (2023)* đã ban hành khuyến cáo và triển khai đào tạo kỹ năng AI cho sinh viên. Các cơ quan báo chí như *VOV (2023)* hay *VnExpress (2023)* cũng đã lên tiếng cảnh báo về nguy cơ đạo văn và kêu gọi tăng cường kỹ năng số cho người học. Tuy nhiên, cho đến nay vẫn chưa tồn tại một hành lang pháp lý hoặc khung chuẩn quốc gia để điều chỉnh vấn đề này một cách hệ thống.

Từ góc nhìn đào tạo luật, tác giả đặc biệt quan tâm đến nguy cơ "suy giảm năng lực tư duy phản biện". Không ít luận văn mà tôi phản biện trong hai năm qua có phong cách diễn đạt quá "trơn tru", khiến hội đồng phải đặt câu hỏi: đâu là giọng văn của người viết, đâu là sản phẩm của máy móc?

V. Kết quả nghiên cứu và thảo luận

5.1. Cơ hội và giá trị AI mang lại cho nghiên cứu khoa học

Đầu tiên, ứng dụng AI đang mở ra những triển vọng tích cực cho toàn bộ vòng đời nghiên cứu. Một trong những lợi ích rõ rệt nhất là khả năng rút ngắn thời gian xử lý công việc. Các công cụ như Elicit, Scite hay Semantic Scholar AI có thể phân tích tài liệu, gợi ý giả thuyết và tổng hợp văn bản một cách hệ thống, giúp nhà nghiên cứu tiết kiệm đáng kể thời gian cho các công đoạn tổng quan và phân tích *(Elsevier, 2023).*

Không chỉ vậy, AI còn góp phần nâng cao chất lượng viết học thuật. Các phần mềm như Grammarly, Writefull hay DeepL Write đã hỗ trợ nhiều học giả, đặc biệt là những người không dùng tiếng Anh như ngôn ngữ mẹ đẻ, cải thiện văn phong và độ mạch lạc của bản thảo, từ đó tăng khả năng được chấp nhận tại các tạp chí uy tín (*Springer Nature*, 2023).

Ngoài ra, các nhà xuất bản quốc tế như Elsevier và Springer Nature đã đưa AI vào thử nghiệm trong quy trình biên tập để phát hiện đạo văn, kiểm tra trích dẫn và nhận diện sai sót kỹ thuật. Điều này hứa hẹn góp phần chuẩn hóa hoạt động xuất bản và giảm tải cho hệ thống phản biện học thuật vốn đang chịu áp lực lớn *(Wiley, 2023).*

5.2. Thách thức và vấn đề liêm chính học thuật

Song song với cơ hội, AI cũng đem đến nhiều nguy cơ. Trước hết, hiện tượng "ảo tưởng tri thức" (hallucination) của các mô hình ngôn ngữ lớn có thể tạo ra dữ liệu hoặc trích dẫn giả mạo. *Nature* (2023) và *Science* (2023) đã cảnh báo rằng tác giả phải kiểm chứng toàn bộ tài liệu tham khảo để tránh sai lệch nghiêm trọng.

Bên cạnh đó, sự thiếu minh bạch trong khai báo sử dụng AI dễ dẫn đến vi phạm đạo đức học thuật. *COPE (2023)*

và *WAME (2023)* đều quy định rõ rằng mọi sự can thiệp của AI phải được công khai, đồng thời nghiêm cấm tải bản thảo hoặc phản biện kín lên các công cụ AI công khai.

Thêm vào đó, AI đặt ra rủi ro pháp lý mới. Đạo luật AI của Liên minh châu Âu *(EU AI Act, 2024)* yêu cầu các tổ chức minh bạch khi triển khai AI trong bối cảnh rủi ro cao, kể cả trong nghiên cứu và phát triển. Điều này đặt ra yêu cầu xây dựng cơ chế quản trị nội bộ tại các cơ sở học thuật.

Ngoài ra, hiệu quả của các công cụ phát hiện AI còn nhiều hạn chế. OpenAI đã phải dừng "AI Classifier" vì độ chính xác thấp *(OpenAI, 2023)*, trong khi Turnitin khuyến cáo không nên dựa vào AI detector làm bằng chứng duy nhất để xử lý vi phạm *(Turnitin, 2023; Stanford HAI, 2023)*. Thực tế, đã có nhiều bài báo trong hệ thống Wiley/Hindawi bị rút lại vì nghi ngờ do "paper mills" sử dụng AI, cho thấy sự phụ thuộc quá mức vào công cụ mà thiếu kiểm chứng học thuật có thể gây hậu quả nghiêm trọng *(Wired, 2024)*.

Cá nhân tác giả cho rằng, vấn đề đáng lo không chỉ nằm ở "ảo tưởng tri thức" do AI sinh ra, mà còn ở chỗ một bộ phận người học có xu hướng "khoán trắng" tư duy cho công cụ. Điều này, nếu không được kiểm soát, sẽ làm xói mòn nền tảng sáng tạo - vốn là linh hồn của nghiên cứu khoa học."

5.3. So sánh chính sách quốc tế - gợi ý cho Việt Nam

Nhìn từ kinh nghiệm quốc tế, có thể thấy sự đồng thuận cao trong các nguyên tắc quản lý AI. *COPE (2023)* nhấn mạnh rằng AI không thể đứng tên tác giả và tác giả con người phải chịu trách nhiệm cuối cùng. *WAME (2023)* khẳng định việc khai báo là bắt buộc và cấm AI tham gia phản biện kín. *Nature* (2023) và *Science* (2023) ban hành hướng dẫn chi tiết về nghĩa vụ của người viết, trong khi *Elsevier (2023)*, *Springer Nature (2023)* và *Wiley (2023)* bổ sung thêm quy trình kỹ thuật và khuyến nghị "human-in-the-loop".

Để làm rõ hơn, Bảng 1 dưới đây tổng hợp một số điểm chính trong chính sách quốc tế về AI:

Bảng 1. So sánh ngắn các nguyên tắc sử dụng AI trong học thuật quốc tế

Tổ chức / Nhà XB	AI là tác giả?	Khai báo bắt buộc?	Quy định nổi bật
COPE (2023)	Không	Có	Tác giả chịu trách nhiệm cuối cùng
WAME (2023)	Không	Có	Cấm AI tham gia phản biện kín
Nature (2023)	Không	Có	Người viết chịu trách nhiệm toàn bộ
Science (2023)	Không	Có	Cấm AI làm tác giả, phản biện
Elsevier (2023)	Không	Có	Minh bạch, trách nhiệm rõ
Springer (2023)	Không	Có	Xác định rõ phạm vi AI
Wiley (2023)	Không	Có	Khuyến nghị "human-in-the-loop"

Từ thực tiễn này, Việt Nam có thể tham khảo để xây dựng bộ nguyên tắc thống nhất, vừa hội nhập với chuẩn mực quốc tế, vừa phù hợp bối cảnh trong nước.

5.4. Thực trạng tại Việt Nam

Ở Việt Nam, mặc dù đã có những động thái ban đầu, việc quản lý AI trong nghiên cứu vẫn chưa thành hệ thống. Bộ Giáo dục và Đào tạo *(MOET, 2024)* đã tổ chức nhiều hội thảo về AI trong giáo

dục, song chưa ban hành quy định mang tính toàn quốc. Đại học Quốc gia Hà Nội đã công bố chương trình đào tạo AI cho sinh viên *(VNU News, 2024),* trong khi *RMIT Việt Nam (2023)* đã công bố khuyến nghị về chính sách AI trong giáo dục đại học, nhấn mạnh sự cân bằng giữa đổi mới và liêm chính học thuật. Ngoài ra, báo chí chính thống như *VOV (2023)* và *VnExpress (2023)* cũng đã nhiều lần cảnh báo về nguy cơ đạo văn, đồng thời kêu gọi tăng cường kỹ năng số cho người học.

Tuy vậy, có thể thấy Việt Nam vẫn thiếu một khung chuẩn quốc gia điều chỉnh việc sử dụng AI trong nghiên cứu, dẫn đến sự phân tán trong cách tiếp cận của từng cơ sở đào tạo và nghiên cứu.

VI. Khuyến nghị chính sách

Từ những phân tích trên, có thể thấy rằng việc sử dụng trí tuệ nhân tạo trong nghiên cứu khoa học là một xu thế tất yếu, song đi kèm cả cơ hội và rủi ro. Để bảo đảm tính liêm chính học thuật, Việt Nam cần triển khai các khuyến nghị chính sách ở cả cấp quốc gia và cơ sở đào tạo.

Trước hết, cần sớm ban hành khung chuẩn quốc gia về sử dụng AI trong nghiên cứu và đào tạo, dựa trên kinh nghiệm quốc tế (*COPE, 2023; Nature*, 2023; *Science*, 2023) nhưng được điều chỉnh phù hợp với điều kiện Việt Nam. Khung chuẩn này sẽ giúp thống nhất ngôn ngữ chính sách, quy trình xử lý và cơ chế khai báo giữa các cơ sở đào tạo và nghiên cứu.

Bên cạnh đó, mỗi nhà trường/viện nghiên cứu cần chủ động xây dựng **quy chế nội bộ** về sử dụng AI. Trong đó, cần quy định rõ phạm vi cho phép, trách nhiệm khai báo và cơ chế kiểm soát. Đặc biệt, nên khuyến khích áp dụng "Nhật ký AI" - một biểu mẫu ghi lại quá trình sử dụng AI trong nghiên cứu, giúp tăng tính minh bạch và trách nhiệm giải trình. Đồng thời, cần tổ chức các khóa tập huấn kỹ năng cho giảng viên, nghiên cứu sinh và sinh viên về sử dụng AI có trách nhiệm, nhấn mạnh kỹ năng thẩm định kết quả đầu ra do AI tạo ra, cũng như cách phát hiện và kiểm chứng trích dẫn.

Để hiện thực hóa các khuyến nghị trên, tác giả đề xuất Ma trận "đèn giao thông" như một công cụ tham chiếu nhanh. Khung tham chiếu này vừa ngắn gọn, vừa dễ áp dụng, có thể được triển khai ngay trong quy chế nội bộ của các cơ sở đào tạo, đồng thời giúp định hướng người học và nhà nghiên cứu sử dụng AI một cách có trách nhiệm và minh bạch. *(Xem Phụ lục A).*

Từ kinh nghiệm giảng dạy ngành luật, tác giả đề nghị các cơ sở không chỉ dừng ở quy định cấm - cho phép, mà cần tạo "văn hóa khai báo AI". Mọi người viết, từ sinh viên đến giảng viên, nên coi việc ghi rõ mức độ sử dụng AI trong công trình khoa học là một biểu hiện của sự trung thực và lòng tự trọng nghề nghiệp.

VII. Kết luận

Sự phát triển nhanh chóng của trí tuệ nhân tạo đang tạo ra một bước ngoặt lớn trong hoạt động nghiên cứu khoa học toàn cầu. Với khả năng hỗ trợ xử lý dữ liệu, sinh văn bản, phát hiện tri thức mới và tối ưu hóa quy trình nghiên cứu, AI không chỉ nâng cao hiệu suất mà còn góp phần dân chủ hóa tiếp cận tri thức học thuật. Tuy nhiên, đi kèm với cơ hội là những thách thức nghiêm trọng về đạo đức, pháp lý

và liêm chính học thuật. Nếu thiếu khung quy định rõ ràng và văn hóa sử dụng AI có trách nhiệm, nguy cơ đạo văn tinh vi, trích dẫn giả mạo hay sự lệ thuộc tư duy máy móc có thể làm xói mòn giá trị khoa học.

Với bối cảnh Việt Nam, nhu cầu cấp thiết hiện nay là ban hành một khung chuẩn quốc gia về sử dụng AI trong nghiên cứu, làm cơ sở cho các trường/viện xây dựng quy chế nội bộ, đồng thời tổ chức tập huấn kỹ năng số cho giảng viên và sinh viên. Những gợi ý trong bài - từ bảng so sánh chính sách quốc tế, ma trận "đèn giao thông" cho đến kế hoạch triển khai 6 - 12 tháng - có thể trở thành nền tảng để xây dựng một môi trường học thuật minh bạch, trung thực và hội nhập.

Với tư cách là một nhà nghiên cứu, tác giả tin rằng AI chỉ thực sự là "trợ thủ đắc lực" khi được đặt dưới sự kiểm soát của liêm chính và trách nhiệm con người. Nếu không, nó sẽ trở thành "con dao hai lưỡi", gây tổn thương cho chính nền khoa học mà chúng ta đang nỗ lực xây dựng./.

Tài liệu tham khảo:

[1]. Committee on Publication Ethics (COPE). (2023). *Authorship and AI tools: How should editors handle "AI-generated" submissions*. Retrieved August 23, 2025, from https://publicationethics.org

[2]. Crothers, B., & Zhang, L. (2023). AI misuse in higher education. *Nature Human Behaviour, 7*(5), 615-618. https://doi.org/10.1038/s41562-023-01678-2

[3]. Elsevier. (2023). *Publishing ethics & policies on generative AI*. Elsevier. Retrieved August 23, 2025, from https://www.elsevier.com/about/policies-and-standards/publishing-ethics

[4]. Elsevier Research Intelligence. (2024). *Artificial intelligence in research: Global trends and insights*. Elsevier. Retrieved August 23, 2025, from https://www.elsevier.com/research-intelligence

[5]. European Parliament. (2024). *Artificial Intelligence Act adopted - key provisions and timeline*. European Union. Retrieved August 23, 2025, from https://www.europarl.europa.eu/news/en/press-room/20240313IPR20339/artificial-intelligence-act

[6]. EUR-Lex. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council on artificial intelligence (AI Act)*. Retrieved August 23, 2025, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689

[7]. Floridi, L., & Chiriatti, M. (2023). GPT and the ethical challenges of AI in science. *AI & Society, 38*(2), 231-240. https://doi.org/10.1007/s00146-023-01564-6

[8]. Hutson, M. (2021). Robo-writers: The rise and risks of language-generating AI. *Nature, 591*(7848), 22-25. https://doi.org/10.1038/d41586-021-00530-0

[9]. OpenAI. (2023). *AI Classifier is no longer available*. OpenAI. Retrieved August 23, 2025, from https://openai.com/blog/new-ai-classifier

[10]. RMIT Vietnam. (2023). *Complexities of generative AI policies in higher education*. RMIT University Vietnam. Retrieved August 23, 2025, from https://www.rmit.edu.vn/news

[11]. Springer Nature. (2023). *Advice on the use of AI tools in manuscript preparation & peer review*. Springer Nature. Retrieved August 23, 2025, from https://www.springernature.com/gp/policies/editorial-policies/ai

[12]. Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2023).

On the (un)reliability of AI-generated text detectors. Stanford University. Retrieved August 23, 2025, from https://hai.stanford.edu/news/reliability-ai-generated-text-detectors

[13]. Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2024). *AI Index Report 2024*. Stanford University. Retrieved August 23, 2025, from https://aiindex.stanford.edu

[14]. Turnitin. (2023). *AI writing detection: Accuracy, false positives, and guidance for educators*. Turnitin. Retrieved August 23, 2025, from https://www.turnitin.com/products/features/ai-writing-detection

[15]. United Nations Educational, Scientific and Cultural Organization (UNESCO). (2023). *Guidance for generative AI in education and research*. UNESCO Publishing. Retrieved August 23, 2025, from https://unesdoc.unesco.org/ark:/48223/pf0000387242

[16]. Vietnam National University (VNU). (2024). *VNU to introduce AI courses for students*. VNU News. Retrieved August 23, 2025, from https://news.vnu.edu.vn

[17]. Voice of Vietnam (VOV). (2023, March 15). *ChatGPT vào đại học: dùng sao cho đúng?* VOV Online. Retrieved August 23, 2025, from https://vov.vn/giao-duc/chatgpt-vao-dai-hoc-dung-sao-cho-dung

[18]. VnExpress. (2023, April 2). *ChatGPT bùng nổ - trường học cảnh báo đạo văn và kỹ năng số*. VnExpress. Retrieved August 23, 2025, from https://vnexpress.net/chatgpt-bung-no-truong-hoc-canh-bao-dao-van-va-ky-nang-so-4593561.html

[19]. Wiley Author Services. (2023). *Best practice guidelines on generative AI for authors and editors*. Wiley. Retrieved August 23, 2025, from https://authorservices.wiley.com/author-resources/Journal-Authors/Prepare/writing-and-ethics/best-practice-guidelines-on-generative-ai.html

[20]. Wired. (2024, February 21). *A top science journal retracted a paper after AI concerns highlighted deeper editorial flaws*. Wired. Retrieved August 23, 2025, from https://www.wired.com/story/ai-science-journal-retraction

[21]. World Association of Medical Editors (WAME). (2023). *Chatbots, generative AI, and scholarly publications (recommendations)*. WAME. Retrieved August 23, 2025, from https://wame.org/page3.php?id=106

Phụ lục:

Phụ lục A - Ma trận "đèn giao thông" sử dụng AI trong nghiên cứu:

Màu xanh (Được phép): Các hoạt động như kiểm tra chính tả, sửa lỗi ngữ pháp, dịch thô, gợi ý dàn ý, tạo mã minh họa đều được chấp nhận, miễn là có khai báo trong Lời cảm ơn hoặc Phương pháp.

Màu vàng (Hạn chế): Các hoạt động như sinh văn bản học thuật dài, tóm tắt kết quả nghiên cứu, gợi ý giả thuyết hoặc sinh mã phân tích dữ liệu chỉ nên dùng trong phạm vi hỗ trợ. Khi sử dụng, tác giả phải kiểm chứng nguồn, lưu lại nhật ký AI và có xác nhận của đồng tác giả.

Màu đỏ (Cấm): Nghiêm cấm để AI viết thay phần kết quả và thảo luận, tạo dữ liệu giả, sinh trích dẫn không kiểm chứng, tải bản thảo hoặc phản biện kín lên công cụ AI công khai. Đồng thời, không được dùng công cụ phát hiện AI như bằng chứng duy nhất để kỷ luật (COPE, 2023; WAME, 2023; *Science*, 2023; Stanford HAI, 2023; Turnitin, 2023).

Phụ lục B - Kế hoạch triển khai khung AI có trách nhiệm (6 - 12 tháng)

Tháng 1 - 2: Thành lập Ban Công nghệ & Liêm chính học thuật; rà soát quy chế hiện hành; ban hành chính sách tạm thời dựa trên ma trận "đèn giao thông".

Tháng 3 - 4: Tập huấn giảng viên - nghiên cứu sinh; triển khai biểu mẫu khai báo AI; quy định cấm tải bản thảo hoặc phản biện kín lên công cụ AI công khai.

Tháng 5 - 6: Tích hợp công cụ hỗ trợ (kiểm lỗi, quản lý trích dẫn); xây dựng Sổ tay kiểm chứng trích dẫn; thử nghiệm tổ kiểm định chéo "human-in-the-loop".

Tháng 7 - 12: Đánh giá - cập nhật chính sách; ban hành quy chế chính thức; công bố báo cáo thường niên về liêm chính học thuật & AI.

Phụ lục C - Một số công cụ AI tiêu biểu dùng trong học thuật

Hỗ trợ đọc - tổng quan: Semantic Scholar (AI), scite_, Elicit.

Viết - biên tập: Grammarly, Writefull, LanguageTool, DeepL Write.

Phân tích dữ liệu/mã: Jupyter Notebook với Copilot hoặc CodeWhisperer (chỉ cho mã không nhạy cảm).

Quản lý trích dẫn: Zotero, EndNote (cảnh giác chức năng "tự sinh trích dẫn").

Kiểm tra tương đồng & AI: Turnitin (không dùng làm bằng chứng duy nhất); không khuyến nghị dựa vào "AI detector" thuần túy.

APPLICATION OF ARTIFICIAL INTELLIGENCE IN SCIENTIFIC RESEARCH: OPPORTUNITIES, CHALLENGES, AND ISSUES OF ACADEMIC INTEGRITY

Dao Trung Hieu[2]

Abstract: *This article systematically analyzes the opportunities, challenges, and academic integrity issues arising from the application of artificial intelligence (AI) in modern scientific research. Based on a review of global technology trends and international academic standards, the paper clarifies the multifaceted impacts of AI on the knowledge production cycle. It also proposes policy recommendations to promote the responsible use of AI in academia, ensuring transparency, honesty, and scientific integrity in the digital age.*

Keywords: *artificial intelligence, academic integrity, ethics in research, plagiarism; responsible AI use*

[2] Faculty of Law, Hanoi Open University