

**APPROACHES TO APPLYING
CRITICAL THINKING IN THE USE
OF GENERATIVE AI
IN GENERAL EDUCATION**

Kieu Phuong Thuy*, Nguyen Khac An
and Vu Thai Giang

*Faculty of Information Technology, Hanoi National
University of Education, Hanoi city, Vietnam.*

*Corresponding author: Kieu Phuong Thuy,
e-mail: thuykp@hnue.edu.vn

Received March 17, 2025.

Revised April 28, 2025.

Accepted May 30, 2025.

**MỘT SỐ PHƯƠNG PHÁP ỨNG DỤNG
TƯ DUY PHẢN BIỆN KHI SỬ DỤNG AI
TẠO SINH TRONG GIÁO DỤC
PHỔ THÔNG**

Kiều Phương Thùy*, Nguyễn Khắc Ân
và Vũ Thái Giang

*Khoa Công nghệ thông tin, Trường Đại học Sư
phạm Hà Nội, thành phố Hà Nội, Việt Nam*

*Tác giả liên hệ: Kiều Phương Thùy,
e-mail: thuykp@hnue.edu.vn

Ngày nhận bài: 17/3/2025.

Ngày sửa bài: 28/4/2025.

Ngày nhận đăng: 30/5/2025.

Abstract. This paper investigates the role of critical thinking in the evaluation and application of Generative Artificial Intelligence (GenAI) within general education. While AI models such as ChatGPT, Microsoft Copilot, and Google Translate have significantly influenced teaching and learning processes, they also introduce challenges related to information bias, contextual inaccuracies, and algorithmic distortions. By analyzing the underlying mechanisms of these GenAI systems, the study underscores the necessity of applying critical thinking when interpreting and utilizing AI-generated content. The paper proposes practical strategies for integrating critical thinking into educational practices and suggests tools that enhance users' understanding of GenAI while mitigating risks associated with AI bias. The findings contribute to the development of a more systematic and ethically grounded framework for GenAI-based education, promoting both pedagogical effectiveness and responsible AI use.

Keywords: Generative AI, critical thinking, AI in education.

Tóm tắt. Bài báo này nghiên cứu vai trò của tư duy phản biện trong việc đánh giá và ứng dụng trí tuệ nhân tạo tạo sinh (Generative AI - GenAI) trong giáo dục phổ thông. Mặc dù các mô hình AI như ChatGPT, Microsoft Copilot và Google Translate đã tạo ra tác động đáng kể đến giảng dạy và học tập, nhưng chúng cũng đặt ra những thách thức liên quan đến sai lệch thông tin, thiên kiến và ngữ cảnh không chính xác. Bằng cách phân tích cơ chế hoạt động của các mô hình AI này, nghiên cứu nhấn mạnh sự cần thiết của tư duy phản biện khi sử dụng nội dung do AI tạo ra. Đồng thời, nghiên cứu đã đề xuất một số phương pháp ứng dụng tư duy phản biện vào thực tiễn giáo dục phổ thông cùng các ví dụ cụ thể nhằm nâng cao hiểu biết về GenAI và giảm thiểu rủi ro liên quan đến sai lệch của GenAI. Kết quả nghiên cứu đóng góp vào việc xây dựng một cách tiếp cận có hệ thống hơn đối với giáo dục ứng dụng GenAI, đảm bảo tính hiệu quả sư phạm và trách nhiệm đạo đức trong sử dụng GenAI.

Từ khóa: AI tạo sinh, tư duy phản biện, AI trong giáo dục.

1. Mở đầu

Trí tuệ nhân tạo (AI) ngày càng được ứng dụng rộng rãi trong nhiều lĩnh vực, đặc biệt là giáo dục, nhờ khả năng hỗ trợ trong việc quản lý học sinh (HS), cá nhân hóa học tập, kiểm tra đánh giá, và quản lý lớp học. Theo nghiên cứu của Brown và cộng sự (2020) [1], các mô hình ngôn ngữ lớn như GPT-3 có thể đạt được "hiệu suất mạnh mẽ trong nhiều tập dữ liệu NLP, bao gồm dịch thuật, hỏi-đáp, và các nhiệm vụ yêu cầu suy luận hoặc thích nghi theo ngữ cảnh". Điều này cho thấy tiềm năng của AI trong việc cung cấp hỗ trợ sáng tạo nội dung và xử lý thông tin phức tạp. Trong bối cảnh đó, AI tạo sinh (Generative AI - GenAI) đã nổi lên như một công cụ không thể thiếu. Các công cụ như ChatGPT, Microsoft Copilot và Google Translate không chỉ có khả năng tạo nội dung mà còn cho phép "các nhà phát triển lập trình được hỗ trợ bởi AI thông qua việc cung cấp gợi ý mã nhanh chóng và chính xác hơn" (Christian Bird và cộng sự (2022) [2]). Tuy nhiên, khả năng suy luận của AI tạo sinh vẫn còn những hạn chế, khi kết quả tạo ra đôi khi không chính xác hoặc không phù hợp với ngữ cảnh, như nhận định của Vaswani và cộng sự (2017) [3] cho rằng "mô hình Transformer có thể cải thiện chất lượng dịch thuật, nhưng sự phụ thuộc vào dữ liệu huấn luyện có thể dẫn đến thiên vị trong đầu ra". Những vấn đề này cho thấy các thách thức trong nội dung do AI tạo ra, khi kết quả đôi khi không chính xác hoặc không phù hợp với ngữ cảnh.

Ennis (1987) [4] định nghĩa tư duy phản biện là "tư duy hợp lý và phản ánh, tập trung vào việc quyết định điều cần tin tưởng hoặc thực hiện" nhấn mạnh các khả năng và phẩm chất cần thiết cho việc suy luận hiệu quả. Tương tự, Paul và Elder (2006) [5] lập luận rằng "tư duy phản biện là nghệ thuật phân tích và đánh giá suy nghĩ nhằm mục đích cải thiện nó", đòi hỏi phải đặt câu hỏi có hệ thống, đánh giá bằng chứng và duy trì kỉ luật tư duy. Halpern [6] bổ sung rằng tư duy phản biện bao gồm "việc sử dụng các kĩ năng hoặc chiến lược nhận thức để tăng khả năng đạt được kết quả mong muốn. Nó được sử dụng để mô tả tư duy có mục đích, có lí lẽ và hướng đến mục tiêu". Những quan điểm này nhấn mạnh tầm quan trọng của tư duy phản biện như một cách tiếp cận có cấu trúc để đánh giá, điều chỉnh và cải thiện các kết quả do hệ thống AI tạo ra.

Xem xét một vài ví dụ về câu trả lời của AI, như khi giao cho ChatGPT tạo một kế hoạch bài dạy nhưng không nói rõ cấu trúc thì kết quả trả lời thường là theo kiểu "Kiến thức - Kỹ năng - Thái độ" trong khi hiện nay kế hoạch bài dạy được soạn theo "Năng lực - Phẩm chất". Hay đôi khi câu trả lời của ChatGPT là tạo ra các liên kết tới video hoặc công trình nghiên cứu nhưng trong thực tế là không tồn tại. Một số sai lệch khác như viết code bị lỗi, giải toán sai hay đưa ra những câu trả lời không cập nhật kiến thức mới.

Những hạn chế này bắt nguồn từ cơ chế suy luận xác suất và thiếu khả năng hiểu ngữ nghĩa một cách sâu sắc. Facione (1990) [7] nhấn mạnh rằng "tư duy phản biện là một kĩ năng cần thiết để phân tích, đánh giá và cải thiện thông tin một cách có hệ thống". Điều này cho thấy tầm quan trọng của việc áp dụng tư duy phản biện khi sử dụng AI, nhằm đảm bảo tính chính xác và hiệu quả.

Các nghiên cứu đã có về tư duy phản biện trong giáo dục tập trung vào các phương pháp giảng dạy cổ điển của: Tariq et al (2024) [8]; Quan Wang et al (2024) [9] hoặc phương pháp giảng dạy trong thời đại số ở cấp đại học của: Barbara Z. Larson (2024) [10]; Brighid Golden (2023) [11]. Hiện tại vẫn còn thiếu những nghiên cứu phân tích cụ thể về việc ứng dụng tư duy phản biện vào việc sử dụng AI tạo sinh trong giáo dục phổ thông, một lĩnh vực đang thay đổi nhanh chóng dưới tác động của công nghệ. Trong bối cảnh AI tạo sinh ngày càng phổ biến và có ảnh hưởng sâu rộng tới cách thức học tập và đánh giá, việc khai thác tư duy phản biện không chỉ nhằm khắc phục các sai lệch và lỗi của AI mà còn giúp người học hình thành năng lực đánh giá độc lập, sáng tạo trước thông tin được cung cấp tự động. Vì vậy, nghiên cứu này là cần thiết để lấp đầy khoảng trống trong hệ thống lí luận hiện nay và cung cấp cơ sở thực tiễn cho việc ứng dụng AI tạo sinh một cách hiệu quả và có trách nhiệm trong giáo dục hiện đại.

Bài báo này tập trung vào phân tích các yếu tố dẫn đến sai lệch trong AI tạo sinh, nhấn mạnh sự cần thiết của tư duy phản biện khi sử dụng AI tạo sinh trong giáo dục phổ thông và đề xuất một số biện pháp giảm thiểu rủi ro. Qua đó, bài báo không chỉ góp phần vào việc nâng cao hiểu biết về AI mà còn khuyến khích ứng dụng tư duy phản biện trong việc sử dụng công nghệ này. Các ví dụ cụ thể sẽ được trình bày để minh họa cách AI hoạt động và vai trò của tư duy phản biện trong việc đánh giá và điều chỉnh kết quả.

2. Nội dung nghiên cứu

2.1 Mô hình hoạt động của Generative AI

Generative AI là một dạng AI được huấn luyện để tạo ra nội dung mới dựa trên dữ liệu mà nó đã được huấn luyện trước đó. Các mô hình như GPT (Generative Pre-trained Transformer), BERT (Bidirectional Encoder Representations from Transformers) và các phiên bản phát triển như ChatGPT, Microsoft Copilot và Google Translate là những ví dụ tiêu biểu của Generative AI.

* *GPT và ChatGPT*: Các mô hình GPT được phát triển bởi OpenAI sử dụng cấu trúc Transformer để xử lý ngôn ngữ tự nhiên (NLP). Chúng có khả năng học từ các mẫu ngữ pháp, từ vựng và ngữ cảnh từ khối lượng dữ liệu lớn để tạo ra phản hồi tương ứng với các câu hỏi hoặc yêu cầu từ người dùng.

GPT được phát triển dựa trên kiến trúc Transformer, một mạng nơ-ron sử dụng self-attention để mô hình hóa ngữ cảnh trong văn bản. Transformer cho phép mô hình xử lý thông tin theo cách song song thay vì tuần tự như RNN, từ đó cải thiện đáng kể hiệu suất và chất lượng đầu ra. Vaswani [12] đã “thử nghiệm trên các nhiệm vụ dịch máy cho thấy mô hình này đạt chất lượng cao hơn trong khi yêu cầu thời gian huấn luyện ít hơn đáng kể”.

Việc áp dụng Transformer giúp GPT và ChatGPT có khả năng mô hình hóa mối quan hệ dài hạn trong ngữ cảnh, từ đó tạo ra nội dung có tính mạch lạc và tự nhiên hơn. Tuy nhiên, một trong những hạn chế của Transformer là độ phức tạp tính toán cao, đặc biệt với những chuỗi văn bản dài. Cũng theo Vaswani (2023) [12], “một trong những thách thức lớn của Transformer là tính toán attention trên toàn bộ chuỗi đầu vào có độ phức tạp $O(n^2)$. Điều này có thể gây ra sự không nhất quán hoặc độ chệch trong các đầu ra của mô hình nếu không có chiến lược kiểm soát phù hợp”.

* *Microsoft Copilot*: Đây là một công cụ hỗ trợ lập trình viên, sử dụng các mô hình ngôn ngữ lớn để phân tích ngữ cảnh và đưa ra các gợi ý mã hoặc hoàn thiện mã theo yêu cầu của lập trình viên. Nghiên cứu của Rossettini và các cộng sự (2024) [13] cũng đã đánh giá hiệu suất của Microsoft Copilot so với ChatGPT-4 và Google Gemini trong bài kiểm tra đầu vào ngành khoa học sức khỏe tại Ý. Kết quả cho thấy “ChatGPT-4 và Microsoft Copilot đạt hiệu suất cao hơn đáng kể so với Google Gemini trong việc trả lời chính xác các câu hỏi trắc nghiệm, với sự khác biệt có ý nghĩa thống kê ($p < 0.001$)”.

* *Google Translate*: Mô hình dịch tự động của Google dựa trên các phiên bản Transformer, có khả năng dịch ngữ nghĩa từ ngôn ngữ này sang ngôn ngữ khác dựa trên dữ liệu dịch trước đó. Các mô hình này đều được huấn luyện trên các tập dữ liệu khổng lồ, sử dụng các kỹ thuật như học có giám sát, học tăng cường, và các thuật toán tối ưu hóa để cải thiện chất lượng đầu ra. Tuy nhiên, do tính chất của dữ liệu đầu vào và cơ chế suy luận, AI tạo sinh có thể đưa ra kết quả sai lệch hoặc thiên vị.

2.2. Nguyên nhân dẫn tới AI tạo sinh cho kết quả sai lệch

Có ba nguyên nhân chính dẫn đến việc Generative AI đưa ra kết quả sai lệch:

* *Chất lượng dữ liệu đầu vào*: Dữ liệu huấn luyện của AI thường đến từ nhiều nguồn khác nhau và không được kiểm soát hoàn toàn. Nếu dữ liệu chứa định kiến xã hội hoặc sai lệch ngữ cảnh,

AI có thể vô tình tái tạo các thiên vị đó trong quá trình phản hồi. Theo Bailin và các cộng sự (2010) [14], “một trong những vấn đề lớn nhất trong việc rèn luyện tư duy phản biện là cách mà thông tin được trình bày và tiếp nhận. Nếu dữ liệu đầu vào chứa định kiến, người học có xu hướng tiếp thu nó mà không nhận thức được sự sai lệch”. Tương tự, AI cũng học theo cách đó và nếu không có cơ chế lọc dữ liệu, nó có thể tái tạo định kiến hoặc thông tin sai lệch từ dữ liệu huấn luyện. Ví dụ, ChatGPT có thể trả lời theo các quan điểm định kiến nếu dữ liệu huấn luyện chứa nội dung này mà không có cơ chế lọc. Theo nghiên cứu của Chen và các cộng sự [15] (2020), “một vấn đề đáng lo ngại khi sử dụng AI trong giáo dục là việc thiếu sự kiểm duyệt nội dung trong dữ liệu đầu vào. Điều này dẫn đến các hệ thống AI có thể tạo ra các phản hồi mang tính thiên vị hoặc sai lệch, ảnh hưởng đến tính chính xác của kết quả”.

* *Cơ chế suy luận của AI*: AI tạo sinh dựa vào xác suất để tạo ra phản hồi, điều này có thể dẫn đến việc đưa ra các thông tin không chính xác hoặc không phù hợp với ngữ cảnh. AI không hiểu ngữ nghĩa theo cách con người hiểu, mà chỉ dự đoán từ tiếp theo dựa trên mẫu huấn luyện. Nghiên cứu của Krstić và các cộng sự (2022) [16] đã chỉ ra rằng: “Các mô hình AI như GPT sử dụng cơ chế tạo phản hồi dựa trên xác suất, điều này dẫn đến nguy cơ tạo ra thông tin không chính xác hoặc không có cơ sở thực tế”. Ví dụ, Microsoft Copilot đôi khi tạo ra mã không đúng hoặc không an toàn do sự không chắc chắn trong dữ liệu huấn luyện về cách sử dụng đúng các thư viện mã cụ thể. Rossetini và các cộng sự (2024) [13] cũng nhận định: “Mặc dù Copilot có thể tạo ra mã có vẻ hợp lý, nhưng không có cơ chế xác thực khiến nó dễ mắc lỗi, đặc biệt trong các trường hợp sử dụng liên quan đến bảo mật hoặc tối ưu hóa mã nguồn”. Điều này có nghĩa là nếu AI gặp phải một truy vấn yêu cầu thông tin cụ thể mà nó chưa từng thấy hoặc có thông tin không đầy đủ, phản hồi có thể thiếu chính xác hoặc sai hoàn toàn.

* *Lỗi mô hình hóa*: Các mô hình AI được huấn luyện để tối ưu hóa khả năng tạo ra nội dung phù hợp với yêu cầu đầu vào, nhưng không phải lúc nào cũng có cơ chế kiểm tra tính chính xác của thông tin được tạo ra. “Một trong những thách thức lớn nhất của AI trong giáo dục là đảm bảo rằng các mô hình không chỉ tái tạo dữ liệu mà còn có khả năng tự đánh giá và sửa lỗi. Nếu không, AI có thể trở thành một công cụ khuếch đại thông tin sai lệch thay vì hỗ trợ việc học tập có chất lượng”, UNESCO (2019) [17]. Ví dụ, Google Translate có thể dịch sai một số cụm từ chuyên ngành hoặc thành ngữ vì không hiểu ngữ cảnh đầy đủ hoặc thiếu dữ liệu ngôn ngữ đặc biệt. Nghiên cứu của Chen và các cộng sự (2020) [18] cũng nhận định: “Mặc dù AI có thể dịch ngôn ngữ với độ chính xác cao trong các trường hợp phổ biến, nhưng nó thường gặp khó khăn với ngôn ngữ chuyên ngành hoặc các cụm từ có ý nghĩa đa chiều, do thiếu cơ chế đánh giá ngữ cảnh đầy đủ”. Điều này nhấn mạnh rằng, dù AI có thể tạo ra nội dung có vẻ hợp lý, nhưng vẫn cần có sự giám sát của con người để đảm bảo độ chính xác và phù hợp của thông tin.

AI tạo sinh có tiềm năng lớn nhưng cũng đối mặt với nhiều thách thức, đặc biệt là về tính chính xác và kiểm soát sai lệch trong kết quả đầu ra. Chất lượng dữ liệu, cơ chế suy luận dựa trên xác suất, và lỗi mô hình hóa đều là những yếu tố ảnh hưởng đến hiệu suất của GenAI. Việc áp dụng các phương pháp kiểm duyệt dữ liệu, phát triển cơ chế kiểm tra lỗi và kết hợp với tư duy phản biện trong sử dụng AI là cần thiết để giảm thiểu sai lệch trong ứng dụng thực tế.

2.3. Tư duy phản biện khi sử dụng Generative AI trong giảng dạy và học tập

2.3.1. Tư duy phản biện và các thành phần của tư duy phản biện

Tư duy phản biện là một kỹ năng không thể thiếu trong quá trình giáo dục, nhất là trong việc đối mặt với khối lượng thông tin khổng lồ và đa chiều mà công nghệ hiện đại mang lại. Theo Ennis (1987) [4], “tư duy phản biện bao gồm cả kỹ năng và xu hướng suy nghĩ có lý lẽ, tập trung vào việc đánh giá và cải thiện niềm tin hoặc hành động của bản thân. Nó không chỉ đơn thuần là bác bỏ hoặc phủ nhận mà còn là một quá trình mang tính xây dựng”. Tư duy phản biện không chỉ là việc phản đối hay nghi ngờ, mà còn là quá trình suy nghĩ có cấu trúc, giúp cá nhân hiểu rõ hơn

về bản chất của thông tin và các lập luận. Các nghiên cứu từ Paul và Elder (2006) [5] đã làm nổi bật thêm những yếu tố cơ bản của tư duy phản biện, bao gồm khả năng lí luận logic, tư duy phân tích và tư duy sáng tạo. Theo nghiên cứu này, “một người có tư duy phản biện không chỉ đánh giá thông tin một cách khách quan mà còn phải biết đặt câu hỏi đúng, xem xét các quan điểm đối lập và nhận diện những thành kiến có thể có trong lập luận của mình”. Bên cạnh đó, Byrne và Johnstone (1987) [19] cũng chỉ ra rằng “tư duy phản biện không chỉ dừng lại ở khả năng phân tích thông tin mà còn đòi hỏi người học phải có khả năng đặt câu hỏi, suy ngẫm về các giá trị và kiểm tra các giả định của mình”.

Tư duy phản biện đóng vai trò quan trọng trong giáo dục, giúp người học không chỉ tiếp thu kiến thức một cách thụ động mà còn phát triển khả năng đánh giá và sáng tạo. Theo Howard (1982) [20], “tư duy phản biện giúp người học phát triển khả năng phân tích và đánh giá thông tin, thay vì chỉ tiếp thu một cách thụ động những gì họ được dạy. Đây là nền tảng của học tập độc lập và tư duy sáng tạo”.

Tư duy phản biện là một kĩ năng không thể thiếu trong quá trình học tập và nghiên cứu, đặc biệt khi đối mặt với lượng thông tin đa chiều mà công nghệ hiện đại mang lại. Facione (1990) [7] định nghĩa “tư duy phản biện là sự phán đoán có mục đích và tự điều chỉnh, dẫn đến việc diễn giải, phân tích, đánh giá và suy luận, cũng như giải thích các yếu tố bằng chứng, khái niệm, phương pháp luận, tiêu chí hoặc bối cảnh mà trên đó sự phán đoán đó được dựa vào”. Tư duy phản biện không chỉ dừng lại ở việc phân tích thông tin, mà còn đòi hỏi sự kiểm tra chặt chẽ về tính hợp lí và độ tin cậy của lập luận. Halpern (2014) [6] nhấn mạnh rằng tư duy phản biện không chỉ đơn thuần là một tập hợp các kĩ năng mà còn là một thái độ, một cách tiếp cận có hệ thống đối với các vấn đề phức tạp. Theo Halpern, “Tư duy phản biện là việc sử dụng các kĩ năng hoặc chiến lược nhận thức nhằm tăng khả năng đạt được kết quả mong muốn. Đây là một quá trình tư duy có mục đích, có lí lẽ và hướng đến mục tiêu. Nó liên quan đến việc giải quyết vấn đề, hình thành suy luận, tính toán xác suất và đưa ra quyết định”. Trong quá trình học tập, tư duy phản biện bao gồm nhiều thành phần khác nhau, đó là:

- *Nhận biết và hiểu biết về thông tin*: Người học cần nhận biết nguồn gốc, tính chính xác và độ tin cậy của thông tin trước khi sử dụng. Điều này bao gồm việc phân biệt giữa thông tin có căn cứ khoa học và những thông tin mang tính suy đoán hoặc không có căn cứ. Facione (1990) [7] đã chỉ ra rằng “một thành phần thiết yếu của tư duy phản biện là khả năng phân biệt giữa sự thật và ý kiến, cũng như khả năng đánh giá độ tin cậy của các nguồn thông tin và lập luận được trình bày”.

- *Phân tích thông tin*: Khả năng chia nhỏ các thông tin thành các phần nhỏ hơn để hiểu rõ hơn cấu trúc và ý nghĩa. Ví dụ, học sinh có thể phân tích các yếu tố khác nhau trong một bài viết để hiểu rõ hơn lập luận của tác giả. Nghiên cứu của Hatcher (2007) [21], cũng nhấn mạnh tầm quan trọng của tư duy phân tích, “học sinh được hướng dẫn cách đọc một đoạn văn tranh luận, xác định luận điểm chính (kết luận) và sau đó nêu các lí do (tiền đề) được đưa ra để hỗ trợ cho kết luận đó”. Điều này cho thấy rằng, để phát triển tư duy phản biện, người học cần có khả năng tách biệt thông tin, nhận diện cấu trúc lập luận, và đánh giá tính hợp lí của từng phần trong một bài viết hoặc nghiên cứu.

- *Đánh giá thông tin và lập luận*: Đây là khả năng xác định tính hợp lí của lập luận và phát hiện các lỗi logic. Việc phát hiện các ngụy biện trong lập luận giúp người học tránh bị ảnh hưởng bởi các thông tin sai lệch hoặc có mục đích thao túng. Hatcher (2007) [21] chỉ ra rằng “sau khi lập luận được tóm tắt, bước tiếp theo là đánh giá. Không phải tất cả các lập luận đều có giá trị như nhau. Phương pháp được sử dụng là “tái cấu trúc suy diễn”, cho phép HS chuyển đổi một lập luận thành dạng suy diễn hợp lệ để đánh giá tính hợp lí và độ vững chắc của nó”. Byrne và Johnstone (1987) [19] cũng nhấn mạnh “Chúng ta không thể dạy HS một quy trình để nhận diện các ngụy biện, vì không có cơ sở để tin rằng tồn tại một quy trình như vậy. Điều tốt nhất chúng ta có thể làm là định hướng cho họ bằng cách giảng dạy khái niệm về lập luận hợp lệ, giúp họ

nhận ra những nguy hiểm mà nếu không được hướng dẫn, họ có thể bỏ qua”. Điều này cho thấy rằng kỹ năng đánh giá lập luận không chỉ bao gồm việc phân biệt giữa lập luận hợp lệ và không hợp lệ, mà còn yêu cầu khả năng nhận diện các lỗi logic tiềm ẩn trong tranh luận.

- *Kết luận và phản biện*: Khả năng đưa ra quyết định dựa trên các phân tích và đánh giá đã thực hiện, đồng thời luôn sẵn sàng sửa đổi quan điểm khi có thông tin mới hoặc bằng chứng thuyết phục, là một yếu tố cốt lõi của tư duy phản biện. Điều này giúp người học tránh tư duy cố chấp, linh hoạt trong việc tiếp thu tri thức mới và nâng cao khả năng ra quyết định hợp lý. C. J. Bonk [16] cho rằng “Tư duy phản biện được định nghĩa là một nỗ lực bền bỉ để xem xét bất kỳ niềm tin hoặc hình thức tri thức giả định nào dưới ánh sáng của các bằng chứng hỗ trợ nó và các kết luận tiếp theo mà nó dẫn đến”. Byrne và Johnstone [19] cũng cho rằng “Tất cả các bằng chứng đều được đánh giá cẩn thận và khách quan, và việc đưa ra phản quyết sẽ tạm hoãn cho đến khi có đủ bằng chứng. Chính bằng chứng mới là tiêu chí quyết định, chứ không phải thẩm quyền của cá nhân hay tổ chức cung cấp bằng chứng đó”. Điều này cho thấy rằng, để phát triển tư duy phản biện hiệu quả, người học cần có thái độ cởi mở trước các quan điểm mới, không bảo thủ với những nhận định ban đầu mà thay vào đó phải liên tục đánh giá và điều chỉnh nhận thức của mình dựa trên bằng chứng đáng tin cậy.

Từ các nghiên cứu trên, tác giả cho rằng: Tư duy phản biện trong bối cảnh giáo dục, đặc biệt là khi tương tác với AI tạo sinh, nên được hiểu là quá trình nhận thức mang tính phân tích - đánh giá - phản biện - điều chỉnh, giúp người học không chỉ tiếp nhận mà còn chất vấn, kiểm tra và điều chỉnh các kết luận được cung cấp từ các hệ thống thông tin, bao gồm cả con người và máy móc.

Các thành tố cốt lõi của tư duy phản biện, theo quan điểm tổng hợp và phát triển của tác giả, gồm có: (1) Phân tích (Phân tách vấn đề hay phát biểu thành các phần nhỏ để hiểu rõ bản chất và mối liên hệ), (2) Đánh giá (Xác định tính đúng sai, hợp lý hay phi lý của thông tin và lập luận), (3) Suy luận (Đưa ra kết luận dựa trên bằng chứng và lập luận), (4) Tự điều chỉnh (Liên tục phản tỉnh, sẵn sàng thay đổi khi có bằng chứng mới) và (5) Sáng tạo (Đề xuất giải pháp mới hoặc cách nhìn khác biệt trong tình huống phức tạp).

Các thao tác trí tuệ đặc trưng trong quá trình tư duy phản biện bao gồm:

- Đặt câu hỏi và truy vấn ngược;
- Xác định và phân tích luận điểm;
- Đánh giá bằng chứng và phát hiện lỗi logic;
- So sánh các quan điểm trái chiều;
- Kết luận có cân nhắc;
- Kiểm nghiệm và điều chỉnh nhận thức.

2.3.2. Vai trò của tư duy phản biện khi sử dụng AI trong dạy học

Tư duy phản biện đóng vai trò quan trọng trong việc sử dụng AI tạo sinh để giảm thiểu rủi ro và đảm bảo tính chính xác của kết quả. Cần có sự kiểm chứng và đánh giá lại kết quả từ AI, thay vì tin tưởng hoàn toàn vào nó. Nghiên cứu của Andy Nguyen và cộng sự (2022) [22] về các nguyên tắc đạo đức của AI trong giáo dục cho thấy “Sự phức tạp và tính *thông minh* của công nghệ này đã dẫn đến các rủi ro đạo đức đáng kể, đòi hỏi phải có các quy trình đánh giá kỹ lưỡng để đảm bảo chất lượng”.

- *Trong giảng dạy*: Giáo viên (GV) cần xác minh thông tin trước khi áp dụng trong lớp học, vì kết quả từ AI có thể chứa các yếu tố không phù hợp hoặc sai lệch. Ví dụ, khi sử dụng ChatGPT để soạn bài giảng, GV cần kiểm tra thông tin kỹ càng để tránh trường hợp cung cấp thông tin sai lệch cho HS. Bên cạnh đó, Anil R Doshi và Oliver Hauser cho rằng [23] GenAI nếu sử dụng một cách phụ thuộc có thể làm cho người dùng giảm bớt sự sáng tạo trong công việc - một điều không mong muốn trong giáo dục.

Các công cụ AI tạo sinh như Copilot và ChatGPT có thể tăng năng suất viết bằng cách hỗ trợ các tác vụ như tạo nội dung, tạo ý tưởng và biên tập phong cách, giúp ích cho cả người sử dụng theo Oğuz "Oz" Buruk (2023) [24]; Tiago Lubiana (2023) [25] và Thiemo Wambsganss (2021) [26]. Tuy nhiên, có lo ngại rằng người sử dụng có thể trở nên quá phụ thuộc vào các công cụ này, có khả năng làm suy yếu sự phát triển kỹ năng lâu dài của họ bằng cách bỏ qua các quy trình viết quan trọng như xây dựng lập luận logic và hiểu vấn đề theo Ronald T. Kellogg (2008) [27]; Ronald T. Kellogg và Bascom A. Raulerson (2007) [28]. Việc phụ thuộc hoặc sử dụng GenAI một cách thụ động, không có phản biện có thể gây ra vấn đề được gọi là “mất trí nhớ kỹ thuật số” theo Chris Greenwood và Matthew Quinn (2017) [29], S. James Robert, S. Kathiravan và Dean McKay (2024) [30].

Hank và cộng sự, 2025 [31] đã thiết kế bộ thang đo, thực hiện khảo sát trên 319 đối tượng khảo sát để xác định xem khi nào và bằng cách nào những người làm công việc trí óc nhận thấy sự kích hoạt của tư duy phản biện khi sử dụng GenAI; Khi nào, tại sao những người làm công việc trí óc nhận thấy nỗ lực tăng/giảm đối với tư duy phản biện GenAI. Kết quả nghiên cứu chỉ ra rằng sự tin tưởng vào AI ảnh hưởng tới tư duy phản biện, vì các công cụ GenAI dường như làm giảm nỗ lực nhận thức cần thiết cho tư duy phản biện. Nghiên cứu cũng chỉ ra cách để tăng tư duy phản biện của người dùng, đó là tăng “động lực phản biện” khi người dùng thường có xu hướng bỏ qua tư duy phản biện nếu họ coi nó nằm ngoài phạm vi công việc của mình, nhưng lại tăng cường nó khi muốn cải thiện các kỹ năng chuyên môn của mình. Điều này gợi mở ra rằng việc áp dụng các phương pháp ứng dụng tư duy phản biện sẽ giúp giáo viên tăng cường “động lực” cần thiết.

- *Trong học tập*: GenAI có thể tăng cường việc học và các tư duy phản biện. Ví dụ, nếu HS được nhắc nhở yêu cầu AI chỉ ra lỗi trong văn bản hoặc nêu ra quan điểm đối với một lập luận theo Hyde và cộng sự (2024) [32].

Tuy nhiên, có một thay đổi đáng kể làm gia tăng nguy cơ GenAI sẽ ức chế tư duy phản biện của người dùng. Trước hết, GenAI cung cấp các câu trả lời nhanh chóng và thuyết phục, theo cách khiến người sử dụng dễ cho rằng các kết quả này là khách quan và mạch lạc, trong khi thực chất chúng vẫn có thể mang những thành kiến vốn có từ dữ liệu huấn luyện (Benber và cộng sự, 2021 [33]; Bianchi và cộng sự, 2023 [34]). Tiếp theo, nếu GenAI bắt chước được cách đối thoại của con người, nó có thể thực hiện nhiệm vụ này rất thành công. Tuy nhiên, GenAI lại thiếu khả năng tái tạo quá trình học tập và tư duy sâu sắc vốn cấu thành nên suy nghĩ và nhận thức của con người.

Theo Larson, Moser và cộng sự (2024) [10], cách GenAI trình bày thông tin có thể làm giảm khả năng người dùng sử dụng tư duy phản biện cá nhân. Điều này xảy ra bởi vì GenAI thường trình bày thông tin một cách tự tin và mạch lạc, khiến người dùng có xu hướng chấp nhận các câu trả lời mà không chủ động kiểm tra hay đặt câu hỏi ngược lại. Hậu quả của việc thiếu tư duy phản biện khi sử dụng GenAI có thể rất nghiêm trọng: HS có nhiều khả năng vô tình tiếp nhận thông tin sai lệch và ít chủ động tìm kiếm các nguồn thông tin bổ sung để kiểm chứng hoặc phản biện lại nội dung nhận được.

2.3.3. Một số phương pháp ứng dụng tư duy phản biện trong giáo dục phổ thông với AI

Việc khuyến khích HS phát triển tư duy phản biện là một trong những mục tiêu quan trọng của giáo dục hiện đại, đặc biệt trong bối cảnh AI ngày càng ảnh hưởng sâu rộng đến nhiều lĩnh vực. Tư duy phản biện giúp HS không chỉ tiếp thu kiến thức mà còn đánh giá, phân tích thông tin một cách logic và sáng tạo. Tư duy phản biện đóng vai trò quan trọng trong việc sử dụng AI tạo sinh để giảm thiểu rủi ro và đảm bảo tính chính xác của kết quả. Cần có sự kiểm chứng và đánh giá lại kết quả từ AI, thay vì tin tưởng hoàn toàn vào nó. Theo Nguyễn Thị Hòa (2017) [35], “năng lực tư duy phản biện có ảnh hưởng mang tính quyết định lên hiệu quả của các hoạt động giáo dục, kết quả học tập của sinh viên, của giảng viên và các nhà nghiên cứu trong môi trường giáo dục mang tính toàn cầu”.

Tuy nhiên, nếu không biết cách, những người có tư duy phản biện khi sử dụng GenAI vẫn có nguy cơ tạo ra các biện pháp can thiệp không giải quyết được nhu cầu thực sự của người dùng. Ví dụ, người sử dụng dùng kết quả của GenAI mà không chỉnh sửa nó vẫn có thể đưa ra một phán đoán phản biện. Tư duy phản biện như vậy là vô hình đối với các biện pháp chỉ tập trung vào kết quả cuối cùng được tạo ra, (Hank và cộng sự (2025) [31]). Vì thế, nghiên cứu này đề ra một số phương pháp để GV có thể lồng ghép các hoạt động và bài học vào chương trình giảng dạy, bao gồm:

- *Học qua trải nghiệm*: HS có thể tham gia vào các dự án thực tế và thử thách liên quan đến AI, nơi họ không chỉ tiếp thu lý thuyết mà còn có cơ hội trực tiếp trải nghiệm, phân tích và rút ra kết luận. Nghiên cứu của Mat Said và cộng sự (2020) [36] chỉ ra rằng “Ba loại chuyển giao tư duy phản biện đã được tìm thấy: chuyển giao gần, chuyển giao xa và chuyển giao tích hợp. Quá trình này giúp người học mở rộng việc áp dụng tư duy phản biện trong các bối cảnh khác nhau”.

- *Thảo luận nhóm*: Tổ chức các buổi thảo luận về các công nghệ AI khác nhau, trong đó học sinh được khuyến khích tranh luận và đánh giá các ý kiến trái chiều. Nguyễn Thị Nga (2018) [37] nhấn mạnh vai trò của tranh luận và biện luận trong mô hình trường học thông minh khi cho rằng “HS cần phải tự mình kiến tạo ra những tri thức mới một cách độc lập; xây dựng được chính kiến của bản thân; có khả năng đánh giá và phản biện các sự việc, các quan điểm, sự kiện một cách khoa học, sáng tạo”.

- *Phân tích tình huống thực tế*: Sử dụng các ví dụ từ thực tế để HS có thể thảo luận và phân tích cách mà AI ảnh hưởng đến cuộc sống hàng ngày, từ đó rèn luyện kỹ năng tư duy phản biện. Ví dụ, thông qua tình huống HS hỏi đáp với AI sẽ giúp HS rèn luyện tư duy phản biện về các mặt tích cực và tiêu cực của công nghệ. “Việc ứng dụng AI trong giáo dục đã làm gia tăng các rủi ro đạo đức, đặc biệt là liên quan đến dữ liệu cá nhân và quyền tự chủ của người học. Mặc dù đã có các hướng dẫn về AI đáng tin cậy, nhưng vẫn còn nhiều tranh cãi về các nguyên tắc đạo đức cốt lõi cần tuân thủ” (Andy Nguyen và cộng sự (2022) [22]).

Một số ví dụ minh họa cho các phương pháp trên:

Ví dụ 1 (Môn Tin học - Lớp 12): Học qua trải nghiệm: Dự án “Dự đoán điểm trung bình học kì bằng mô hình học máy”.

Thời điểm tổ chức: Sau khi HS học xong nội dung “Giới thiệu về học máy và khoa học dữ liệu”.

Dự án yêu cầu HS sử dụng mô hình học máy để dự đoán điểm trung bình học kì dựa trên dữ liệu như: điểm kiểm tra giữa kì, số giờ tự học, tần suất hoàn thành bài tập, số buổi đi học đầy đủ... HS sử dụng các công cụ AI đơn giản như Google Sheets với add-on học máy (AutoML) hoặc Teachable Machine.

Quy trình thực hiện:

- Bước 1: Thu thập và nhập liệu dữ liệu học tập ẩn danh từ lớp học (có sự đồng thuận).
- Bước 2: Phân tích dữ liệu bằng công cụ số hóa, lựa chọn biến độc lập phù hợp.
- Bước 3: Xây dựng mô hình dự đoán và chạy thử nghiệm.
- Bước 4: So sánh kết quả dự đoán với điểm thực tế, phản biện: "Vì sao kết quả lệch?", "Mô hình thiếu yếu tố gì?", "Dữ liệu có thiên lệch không?", từ đó đặt câu hỏi chất lượng cho mô hình AI.

Hoạt động này sẽ giúp HS tăng cường tư duy phản biện qua các thành tố như:

- *Tư duy phân tích*: HS cần phân tách các yếu tố ảnh hưởng đến kết quả học tập và phân tích dữ liệu để tìm ra mối liên hệ giữa các yếu tố này. Chẳng hạn, HS sẽ phải phân tích xem liệu có mối quan hệ nào giữa số giờ tự học và kết quả bài kiểm tra.

- *Tư duy đánh giá*: Sau khi thu thập và phân tích dữ liệu, HS sẽ đánh giá tính hợp lý và độ tin cậy của các dự đoán mà AI đưa ra. Liệu kết quả dự đoán có chính xác không? Liệu AI có thể đưa ra một dự đoán đúng khi dữ liệu đầu vào thiếu chính xác?

- *Tư duy suy luận*: HS sẽ sử dụng suy luận để giải thích kết quả và đề xuất cách cải thiện dự đoán của AI. Họ sẽ phải suy nghĩ về lý do tại sao mô hình AI có thể đưa ra các kết quả không chính xác và cách có thể cải thiện dữ liệu đầu vào.

- *Tư duy tự điều chỉnh*: HS sẽ tự đánh giá lại các giả thuyết của mình, kiểm tra xem mô hình AI có thể có sự thiên lệch nào không và tự điều chỉnh chiến lược thu thập dữ liệu nếu cần thiết.

Ví dụ 2 (Môn Giáo dục Kinh tế pháp luật - Lớp 11): Thảo luận nhóm: Chủ đề "AI và quyền riêng tư cá nhân trong giáo dục"

Thời điểm tổ chức: Sau khi HS học xong chủ đề “Một số quyền tự do cơ bản của công dân”.

GV tổ chức buổi thảo luận với tình huống: trường muốn lắp đặt hệ thống điểm danh bằng nhận diện khuôn mặt và AI sẽ phân tích hành vi HS trong lớp để hỗ trợ giảng dạy cá nhân hóa.

Quy trình thực hiện:

- Bước 1: Chia nhóm, mỗi nhóm tìm hiểu khái niệm về quyền riêng tư, quyền được bảo vệ dữ liệu cá nhân theo Luật An ninh mạng Việt Nam.

- Bước 2: Các nhóm đóng vai: nhóm ủng hộ (ban giám hiệu), nhóm phản đối (phụ huynh), nhóm trung lập (HS), rồi đưa ra lập luận, dẫn chứng cụ thể.

- Bước 3: Thảo luận, phản biện giữa các nhóm: “Liệu HS có quyền từ chối AI theo dõi không?”, “AI có giúp nâng cao hiệu quả học tập thật sự không?”, “Làm sao để đảm bảo minh bạch trong lưu trữ dữ liệu?”.

Hoạt động này sẽ giúp HS tăng cường tư duy phản biện qua các thành tố như:

Tư duy phân tích: HS cần phân tích các tình huống sử dụng AI và nhận diện các vấn đề đạo đức và pháp lý liên quan đến quyền riêng tư như: “AI thu thập dữ liệu cá nhân của HS có đáng tin cậy không? Dữ liệu này sẽ được bảo vệ như thế nào?”.

Tư duy đánh giá: HS sẽ đánh giá các lập luận và các yếu tố bảo vệ quyền riêng tư trong môi trường sử dụng AI. Chẳng hạn, khi nghiên cứu về quyền riêng tư và nhận diện khuôn mặt, HS sẽ đánh giá xem liệu việc sử dụng AI trong việc này có thực sự cần thiết và an toàn hay không.

Tư duy suy luận: HS sẽ phải sử dụng suy luận để đưa ra kết luận về việc áp dụng AI trong giáo dục. Ví dụ, nếu việc thu thập dữ liệu có thể gây ra các nguy cơ bảo mật, HS cần suy luận các biện pháp bảo vệ dữ liệu hoặc các giới hạn mà AI không nên vượt qua.

Tư duy tự điều chỉnh: Sau khi thảo luận, HS sẽ đánh giá lại các quan điểm của mình và sẵn sàng điều chỉnh khi có thêm bằng chứng mới. Nếu các bạn trong nhóm đưa ra một quan điểm hợp lý hơn, HS cần nhận thức và điều chỉnh quan điểm của mình.

Ví dụ 3: Phân tích tình huống thực tế: Sử dụng ChatGPT để làm dự án môn Lịch sử.

Thời điểm tổ chức: Khi học bài 6: “Cách mạng tháng Tám năm 1945” - Sách giáo khoa Lịch sử lớp 12 Bộ Kết nối tri thức.

Tình huống học tập:

GV giao cho HS bài tập yêu cầu phân tích các nguyên nhân dẫn đến cuộc cách mạng tháng Tám năm 1945 ở Việt Nam. HS quyết định sử dụng ChatGPT để hỗ trợ tìm hiểu và tạo ra dàn ý cho bài viết của mình.

Cách HS sử dụng ChatGPT:

- HS nhập câu hỏi vào ChatGPT: “Hãy liệt kê và giải thích các nguyên nhân dẫn đến cuộc cách mạng tháng Tám năm 1945 ở Việt Nam”.

- ChatGPT trả lời bằng một danh sách chi tiết các nguyên nhân như chính sách của thực dân Pháp, sự ảnh hưởng của các cuộc cách mạng thế giới, tình hình kinh tế xã hội Việt Nam, v.v. ChatGPT cũng cung cấp một số lý giải và ví dụ từ lịch sử.

GV tổ chức cho HS hoạt động để tăng cường tư duy phản biện qua các thành tố như sau:

Tư duy phân tích: Sau khi nhận được câu trả lời từ ChatGPT, HS không chấp nhận ngay mà phân tích lại thông tin. Họ kiểm tra xem những nguyên nhân mà ChatGPT liệt kê có phù hợp với kiến thức đã học trong sách giáo khoa không. HS có thể tự hỏi: "Có phải tất cả các nguyên nhân được liệt kê đều được xác thực trong các nguồn sử dụng trong học thuật không?" hoặc "Liệu có nguyên nhân nào quan trọng hơn mà ChatGPT chưa đề cập đến không?"

Tư duy đánh giá: HS đánh giá các câu trả lời của ChatGPT dựa trên các nguồn tài liệu học thuật mà họ đã được học. HS sẽ kiểm tra và đối chiếu với các nguồn tài liệu đáng tin cậy để xác định mức độ chính xác của thông tin ChatGPT cung cấp, đồng thời phê phán sự thiếu sót hoặc thiên lệch trong những thông tin này.

Tư duy suy luận: Sau khi đánh giá thông tin từ ChatGPT, HS sẽ suy luận về các yếu tố ảnh hưởng lẫn nhau giữa các nguyên nhân đã được liệt kê. Ví dụ, HS có thể tự hỏi: "Nếu tình hình kinh tế xã hội không thuận lợi, liệu phong trào đấu tranh có thể mạnh mẽ như vậy không?" hoặc "Có phải sự áp bức của thực dân Pháp là yếu tố quan trọng nhất trong việc làm dấy lên phong trào cách mạng, hay có yếu tố nào khác cần xem xét sâu hơn?" Qua đó, HS phát triển khả năng suy luận và kết nối các yếu tố trong một bức tranh lớn.

Tư duy tự điều chỉnh: HS nhận ra rằng ChatGPT chỉ là một công cụ và cần phải đánh giá lại các kết luận của mình dựa trên những phản hồi và câu hỏi mà công cụ này đưa ra. HS sẽ tự điều chỉnh quan điểm nếu thấy rằng những giải thích của ChatGPT không thuyết phục hoặc có sự thiếu sót trong việc cung cấp thông tin. Ví dụ, nếu ChatGPT đưa ra một nguyên nhân mà HS nhận thấy không chính xác hoặc thiếu thuyết phục thì cần tiếp tục tìm kiếm các nguồn tài liệu khác hoặc hỏi GV để làm rõ vấn đề. HS sẽ luôn sẵn sàng sửa đổi bài viết của mình khi có thêm thông tin hoặc lập luận mới.

Tư duy sáng tạo: Cuối cùng, HS sẽ sáng tạo trong cách tổ chức và trình bày dàn ý hoặc bài luận của mình, không chỉ dựa vào câu trả lời của AI mà còn phát triển thêm những luận điểm riêng dựa trên sự đánh giá và phân tích độc lập. HS có thể tạo ra một lập luận mới hoặc kết hợp các nguyên nhân theo một góc nhìn khác để tăng thêm chiều sâu cho bài viết.

3. Kết luận

Nghiên cứu này đã chỉ ra vai trò quan trọng của tư duy phản biện trong việc đánh giá và tối ưu hóa ứng dụng của trí tuệ nhân tạo tạo sinh (Generative AI) trong giáo dục. Qua các phân tích, bài báo nhấn mạnh rằng, mặc dù AI tạo sinh có tiềm năng đáng kể trong việc hỗ trợ giảng dạy và học tập, nhưng nó cũng đối mặt với các thách thức liên quan đến sai lệch thông tin, thiên kiến và thiếu chính xác trong ngữ cảnh. Các kết quả nghiên cứu chỉ ra rằng, để giảm thiểu các rủi ro và nâng cao hiệu quả ứng dụng AI trong giáo dục, việc kết hợp tư duy phản biện là rất cần thiết. Tư duy phản biện không chỉ giúp người dùng nhận diện và điều chỉnh các sai lệch trong kết quả do AI tạo ra mà còn nâng cao năng lực hiểu biết về AI trong cộng đồng học thuật và giáo dục.

Dựa trên kết quả nghiên cứu này, nghiên cứu tiếp theo có thể tập trung vào việc xây dựng thang đo đánh giá kết quả phát triển tư duy phản biện trong giáo dục phổ thông với AI theo từng phương pháp được đề xuất. Thực nghiệm, khảo sát, phân tích, đánh giá kết quả thu được để rút ra kết luận về các phương pháp lý thuyết đã được đề cập. Đề xuất ra các giải pháp nhằm ứng dụng, tăng cường tư duy phản biện của GV và HS khi học tập có sử dụng AI.

TÀI LIỆU THAM KHẢO

- [1] Brown T, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P & Amodei D, (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*. DOI: 10.48550/arXiv.2005.14165.

- [2] Bird C, Ford D, Zimmermann T, Forsgren N, Kalliamvakou E, Lowdermilk T, Gazit I, (2022). Taking flight with Copilot: Early insights and opportunities of AI-powered pair-programming tools. *Queue*, 20(6), 35-57. DOI: 10.1145/3582083.
- [3] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN & Polosukhin I, (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 5998-6008. DOI: 10.48550/arXiv.1706.03762.
- [4] Ennis RH, (1987). *A taxonomy of critical thinking dispositions and abilities*. In: Baron JB. & Sternberg RJ (eds.), "Teaching thinking skills: Theory and practice", W H Freeman/Times Books/Henry Holt & Co, p. 9-26.
- [5] Paul R & Elder L, (2006). *Critical thinking: Learn the tools the best thinkers use*. Pearson Education.
- [6] Halpern DF, (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). Psychology Press.
- [7] Facione PA, (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. *The Delphi Report*. ERIC Doc: ED 315 423. [researchgate.net/publication/242279575](https://www.researchgate.net/publication/242279575).
- [8] Tariq Mehmood Bhuttah, (2024). Enhancing student critical thinking and learning outcomes through innovative pedagogical approaches in higher education: The mediating role of inclusive leadership. *Scientific Reports* 14, 24362. DOI: 10.21203/rs.3.rs-4864296/v1
- [9] Qian Wang and Abdul Halim Abdullah, (2024). Enhancing Students' Critical Thinking Through Mathematics in Higher Education: A Systematic Review. *Sage Open*, 14(3). DOI: 10.1177/21582440241275651.
- [10] Barbara ZL, (2024). Critical Thinking in the Age of Generative AI. *Academy of Management Learning & Education*, 23(3). DOI: 10.5465/amle.2024.0338.
- [11] Brighid G, (2023). Enabling critical thinking development in higher education through the use of a structured planning tool. *Irish Educational Studies*, 949-969. DOI: 10.1080/03323315.2023.2258497.
- [12] Vaswani A, (2023). Attention is all you need. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA. DOI: 10.48550/arXiv.1706.03762.
- [13] Rossetini G & et al, (2024). Comparative accuracy of ChatGPT-4, Microsoft Copilot and Google Gemini in the Italian entrance test for healthcare sciences degrees: a cross-sectional study. *BMC Medical Education*, 24, DOI: 10.1186/s12909-024-05630-9.
- [14] Bailin S, Case R & Coombs JR, (2010). Common misconceptions of critical thinking. *Journal of Curriculum Studies*, 31(3), 269-283.
- [15] Chen X & Xie H, (2020). Application and theory gaps during the rise of artificial intelligence in education, *Computers and Education: Artificial Intelligence*. DOI: 10.1016/j.caeai.2020.100002.
- [16] Kristic L & Kristic M, (2022). Artificial intelligence in education: A review. *Conference: 9th International Scientific Conference Technics and Informatics in Education At: Čačak, Serbia*. DOI: 10.46793/TIE22.223K.
- [17] UNESCO, (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development.
- [18] Chen L & Chen P, (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264-75278.

- [19] Byrne M S & Johnstone AH, (1987). Critical thinking and science education. *Studies in Higher Education*, 12(3), 325-339.
- [20] Howard VA, (1982). *Artistry: The work of artists*. Hackett Publishing.
- [21] Hatcher DL, (2007). Stand-alone versus integrated critical thinking courses. *The Journal of General Education*, 55(3-4), 247-272.
- [22] Nguyen A & Ngo HN, (2022). Ethical principles for artificial intelligence in education. *Education and information technologies*, 28(4), 4221-4241.
- [23] Anil R Doshi, (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10, 28 (2024), eadn5290.
- [24] Oğuz "Oz" Buruk, (2023). Academic Writing with GPT-3.5: Reflections on Practices, Efficacy, and Transparency. *arXiv preprint arXiv:2304.11079*. DOI: 10.31224/2861.
- [25] Tiago Lubiana, (2023). Ten Quick Tips for Harnessing the Power of ChatGPT/GPT-4 in Computational Biology. *PLoS Computational Biology*, 19(8), e1011319.
- [26] Thiemo Wambsganss, (2021). ArgueTutor: An Adaptive Dialog-Based Learning System for Argumentation Skills. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, 1-13.
- [27] Ronald TK. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1(1), 1-26.
- [28] Ronald TK & Bascom AR, (2007). Improving the writing skills of college students. *Psychonomic Bulletin & Review*, 14, 237-242.
- [29] Chris G & Matthew Q, (2017). Digital amnesia and the future tourist. *Journal of Tourism Futures*, 3(1), 73-76.
- [30] James RS, Kathiravan S & Dean MK, (2024). The development and validation of the digital amnesia scale. *Current Psychology*, 43(22), 19594-19603.
- [31] Hao-Ping (Hank) Lee & et al., (2025). The impact of generative AI on critical thinking: self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *Association for Computing Machinery, New York, NY, USA, Article 1121*, 1-22. DOI: 10.1145/3706598.3713778.
- [32] Hyde S J, Busby A & Bonner RL, (2024). Tools or fools: Are we educating managers or creating tool-dependent robots?. *Journal of Management Education*, 48(4), 708-734.
- [33] Bender EM, Gebru T, McMillan-Major A & Shmitchell S, 2021. On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 610-623.
- [34] Bianchi F, Kalluri P, Durmus E, Ladhak F, Cheng M, Nozza D, Hashimoto T, Jurafsky D, Zou J & Caliskan A, 2023. Easily accessible text-to-image generation amplifies demographic stereotypes on a large scale. In *FAccT '23: Proceedings of the 2023 ACM conference on fairness, accountability, and transparency*.
- [35] NT Hòa, (2017). Bàn về tư duy phản biện trong giáo dục đại học. *Tạp chí Khoa học Đại học Đồng Nai*, (05), 23-30.
- [36] Mat Said N A & Bujang SM, (2020). Conceptualizing a critical thinking learning transfer model: A qualitative approach. *Malaysian Journal of Learning and Instruction*, 18, 111-130. DOI: 10.32890/mjli2021.18.1.5.
- [37] NT Nga, (2018). Phát triển tư duy phản biện cho học sinh trong mô hình trường học thông minh. *Kỷ yếu Hội thảo Quốc tế Giáo dục cho mọi người*. Nhà xuất bản Đại học Quốc gia Hà Nội.