VJE Tạp chí Giáo dục (2026), 26(16), 107-114 ISSN: 2354-0753

ÁP DỤNG MÔ HÌNH KAP ĐỂ KHẢO SÁT VIỆC SỬ DỤNG AI CÓ TRÁCH NHIỆM CỦA SINH VIÊN: NGHIÊN CỨU TẠI TRƯỜNG ĐẠI HỌC KHOA HỌC XÃ HỘI VÀ NHÂN VĂN - ĐẠI HỌC QUỐC GIA HÀ NỘI

APPLYING KAP MODEL TO EXAMINE STUDENTS' RESPONSIBLE AI USE: A STUDY AT THE UNIVERSITY OF SOCIAL SCIENCES AND HUMANITIES, VIETNAM NATIONAL UNIVERSITY, HANOI

Đinh Gia Hân+,
Nguyễn Thị Lan Anh,
Nguyễn Thị Minh Ánh,
Hà Thu Hằng,
Phan Thị Minh Hòa,
Mai Thị Thanh Hoa

Trường Đại học Khoa học Xã hội và Nhân văn - Đại học Quốc gia Hà Nội
+Tác giả liên hệ ● Email: giahan190906@gmail.com

Article history
Received: 30/5/2026
Accepted: 29/6/2026
Published: 20/8/2026

Keywords
Responsible AI use, KAP model, higher education, Generative AI, academic integrity

ABSTRACT
The rapid development of artificial intelligence (AI) is creating many changes in higher education and raising issues related to academic ethics and students' dependence on technology. This study investigates the knowledge, attitudes, and practices regarding responsible AI use among students at the University of Social Sciences and Humanities, Vietnam National University, Hanoi, based on the KAP model and an explanatory sequential mixed-methods design, including a survey of 360 students and semi-structured in-depth interviews. The research findings show that students have relatively positive knowledge and attitudes toward responsible AI use; however, their actual AI usage practices remain inconsistent and are strongly influenced by the learning context. Qualitative data indicate that many students have not clearly identified the boundary between learning support and AI dependence in academic activities. The findings emphasize the role of universities in guiding responsible AI use. Accordingly, the study proposes strengthening academic ethics education and implementing digital competency frameworks for students. At the same time, further studies are needed to more fully explain AI usage behaviors among students.

1. Mở đầu

Sự phát triển nhanh chóng của trí tuệ nhân tạo (Artificial Intelligence - AI) đang làm thay đổi mạnh mẽ giáo dục đại học (GDĐH), đặc biệt trong cách sinh viên (SV) tiếp cận tri thức, xử lí thông tin và thực hiện các nhiệm vụ học thuật. Các nghiên cứu về AI trong GDĐH cho thấy công nghệ này đang ngày càng được tích hợp vào các hoạt động dạy và học, góp phần hỗ trợ người học trong việc tìm kiếm thông tin, phân tích tài liệu và phát triển ý tưởng học thuật (Zawacki-Richter và cộng sự, 2019). Bên cạnh những lợi ích tiềm năng, sự phát triển của AI cũng đặt ra nhiều thách thức liên quan đến vai trò của người học, nhất là nguy cơ thay đổi cách thức học tập truyền thống và phụ thuộc vào công nghệ. Trong bối cảnh đó, Tổ chức Giáo dục, Khoa học và Văn hóa của Liên Hợp Quốc (UNESCO) nhấn mạnh rằng việc ứng dụng AI trong giáo dục cần được triển khai theo hướng lấy con người làm trung tâm, trong đó AI chỉ đóng vai trò hỗ trợ chứ không thay thế năng lực tư duy của người học (UNESCO, 2021a). Tuy nhiên, các nghiên cứu cũng chỉ ra rằng nếu thiếu định hướng phù hợp, việc sử dụng AI có thể làm thay đổi cách người học tham gia vào quá trình học tập và đặt ra yêu cầu cao hơn đối với việc đảm bảo vai trò chủ động và tư duy độc lập của SV (Zawacki-Richter và cộng sự, 2019). Các nghiên cứu gần đây cho thấy AI tạo sinh vừa mang lại lợi ích học tập vừa tạo ra rủi ro học thuật. Abbas và cộng sự (2024) chỉ ra rằng việc sử dụng ChatGPT thường xuyên có thể làm gia tăng trì hoãn học tập, giảm ghi nhớ và ảnh hưởng đến kết quả học tập. Đồng thời, áp lực học tập và thời gian thúc đẩy SV sử dụng AI nhiều hơn, từ đó làm tăng nguy cơ phụ thuộc và giảm mức độ tư duy phản biện. Tại Việt Nam, SV đã nhanh chóng tiếp cận AI tạo sinh trong học tập, chủ yếu để tìm kiếm thông tin, hỗ trợ viết bài và phát triển ý tưởng (Đặng Văn Em và cộng sự, 2024). Tuy nhiên, mức độ hiểu biết về chuẩn mực học thuật và kĩ năng đánh giá nội dung do

AI tạo ra vẫn còn hạn chế, đặt ra yêu cầu về năng lực kiểm chứng thông tin và tư duy phản biện (Kiều Phương Thùy và cộng sự, 2025).

Mặc dù có nhiều nghiên cứu về AI trong giáo dục nhưng phần lớn tập trung vào chấp nhận công nghệ thông qua các mô hình như: Mô hình chấp nhận công nghệ (Technology Acceptance Model), Mô hình chấp nhận và sử dụng công nghệ (Unified Theory of Acceptance and Use of Technology) và Lý thuyết hành vi có kế hoạch (Theory of Planned Behavior) (Sanchez và cộng sự, 2025). Các mô hình này chủ yếu giải thích ý định sử dụng nhưng chưa phản ánh đầy đủ hành vi sử dụng AI trong bối cảnh học thuật và yêu cầu liêm chính học tập. Trong khi đó, mô hình Knowledge - Attitude - Practice (KAP) cho phép phân tích đồng thời nhận thức, thái độ và hành vi thực tế của người học (World Health Organization - WHO, 2008). Trong lĩnh vực giáo dục, Sanchez và cộng sự (2025) đã vận dụng mô hình này để nghiên cứu việc sử dụng AI trong giáo dục và cho thấy KAP là công cụ phù hợp để đánh giá mức độ hiểu biết, thái độ và hành vi thực tế của người học đối với AI. Vì vậy, nghiên cứu lựa chọn mô hình KAP làm khung tiếp cận nhằm phân tích thực trạng sử dụng AI có trách nhiệm của SV. Trong phạm vi nghiên cứu này, mô hình KAP được sử dụng như một khung mô tả nhằm phản ánh đồng thời ba thành tố nói trên, không nhằm mục đích kiểm định quan hệ nhân quả hoặc vai trò trung gian giữa các thành tố.

Từ "khoảng trống" trên, nghiên cứu này được thực hiện nhằm khảo sát thực trạng nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm của SV Trường Đại học Khoa học Xã hội và Nhân văn - Đại học Quốc gia Hà Nội. Trên cơ sở đó, nghiên cứu hướng tới phân tích chênh lệch quan sát được giữa nhận thức - thái độ - hành vi của SV trong quá trình sử dụng AI, đồng thời đề xuất các hàm ý thực tiễn nhằm định hướng sử dụng AI theo hướng có trách nhiệm, bảo đảm liêm chính học thuật và phù hợp với mục tiêu thúc đẩy giáo dục hướng tới phát triển bền vững. Để đạt được mục tiêu trên, nghiên cứu tập trung trả lời các câu hỏi sau: (1) Mức độ nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm của SV Trường Đại học Khoa học Xã hội và Nhân văn hiện nay như thế nào? (2) Có tồn tại chênh lệch giữa nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm của SV hay không? (3) Những yếu tố bối cảnh nào góp phần lí giải chênh lệch giữa nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm của SV?

2. Phương pháp nghiên cứu

Nghiên cứu sử dụng phương pháp hỗn hợp (Mixed Methods), kết hợp khảo sát định lượng và phỏng vấn định tính nhằm phản ánh thực trạng nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm của SV. Thiết kế nghiên cứu được triển khai theo mô hình hỗn hợp tuần tự giải thích (Explanatory Sequential Mixed Methods Design), trong đó dữ liệu định lượng được thu thập trước để xác định xu hướng chung của hiện tượng nghiên cứu, sau đó dữ liệu định tính được thu thập nhằm giải thích và làm rõ sâu hơn các kết quả định lượng. Thiết kế này nhấn mạnh quy trình hai giai đoạn liên tiếp và cách thức tích hợp kết quả giữa dữ liệu định lượng và định tính nhằm cung cấp sự hiểu biết toàn diện hơn về vấn đề nghiên cứu (Ivankova và cộng sự, 2006).

Bảng 1. Đặc điểm nhân khẩu học của SV tham gia khảo sát

Đặc điểm	Phân loại	Tần số (n)	Tỉ lệ (%)
Giới tính	Nam	108	30
	Nữ	252	70
Năm học	Năm 1	58	16,1
	Năm 2	164	45,6
	Năm 3	100	27,8
	Năm 4	38	10,6
Ngành học	Khoa học quản lí	37	10,3
	Báo chí	28	7,8
	Đông Nam Á học	27	7,5
	Hàn Quốc học	18	5
	Quản trị dịch vụ du lịch và lữ hành	17	4,7
	Quản lí thông tin	14	3,9
	Khác	219	60,8

Như mô tả ở bảng 1, mẫu nghiên cứu gồm 360 SV Trường Đại học Khoa học Xã hội và Nhân văn, được lựa chọn theo phương pháp lấy mẫu thuận tiện. Cơ cấu mẫu gồm tỉ lệ giới tính (70% nữ và 30% nam), năm học (năm 2: 45,6%; năm 3: 27,8%; năm 1: 16,1%; năm 4: 10,6%) và 28 ngành học (100% trong đó Khoa học quản lí chiếm 10,3%; Báo chí 7,8%; Đông Nam Á học 7,5%...). Khảo sát được thực hiện bằng hình thức trực tuyến thông qua Google Forms

(https://forms.gle/YqS1DvjN1kuqDHj18) từ ngày 16/3/2026 đến ngày 28/3/2026. Đường dẫn khảo sát được gửi tới SV thông qua các nhóm lớp học phần, đăng tải trên các nhóm Facebook của SV Trường Đại học Khoa học Xã hội và Nhân văn và tiếp cận trực tiếp SV tại khuôn viên trường để mời tham gia khảo sát. Tiêu chí lựa chọn người tham gia gồm: (1) Đang theo học tại Trường Đại học Khoa học Xã hội và Nhân văn; (2) Đã từng sử dụng ít nhất một công cụ AI tạo sinh phục vụ học tập hoặc nghiên cứu. Sau khi rà soát dữ liệu và loại bỏ các phiếu không hợp lệ, nghiên cứu thu được 360 phiếu khảo sát hợp lệ đưa vào phân tích. Mẫu phỏng vấn định tính gồm 06 SV được lựa chọn theo phương pháp chọn mẫu có chủ đích từ kết quả khảo sát nhằm đảm bảo đa dạng về mức độ hành vi sử dụng AI. Các đối tượng phỏng vấn được lựa chọn từ những SV có mức độ sử dụng AI khác nhau, đảm bảo sự đa dạng về năm học và ngành đào tạo nhằm khai thác nhiều góc nhìn khác nhau về việc sử dụng AI có trách nhiệm.

Công cụ khảo sát được xây dựng dựa trên mô hình KAP của WHO (2008). Tuy nhiên, nghiên cứu không sử dụng nguyên mẫu một bảng hỏi có sẵn từ WHO (2008) mà vận dụng khung tiếp cận KAP như cơ sở lí thuyết để xây dựng các biến quan sát phù hợp với bối cảnh sử dụng AI trong GDĐH. Quá trình xây dựng thang đo được thực hiện thông qua tổng hợp các nguyên tắc sử dụng AI có trách nhiệm được đề cập trong các tài liệu của UNESCO, Tổ chức Hợp tác và Phát triển Kinh tế (OECD) và các nghiên cứu gần đây về sử dụng AI trong giáo dục. Bảng hỏi gồm 17 biến quan sát theo thang đo Likert 5 mức độ, chia thành các nhóm nội dung: (1) Nhận thức về bản chất và tính minh bạch của AI, quyền riêng tư và bảo mật dữ liệu, đạo đức và liêm chính học thuật; (2) Thái độ đối với định hướng sử dụng AI của nhà trường; (3) Hành vi sử dụng AI có trách nhiệm. Bảng hỏi được khảo sát thử (pilot) với 50 SV trước khi triển khai chính thức nhằm kiểm tra và điều chỉnh độ rõ ràng của các biến quan sát.

Dữ liệu định lượng được xử lí bằng SPSS 26.0 với các kĩ thuật thống kê mô tả gồm giá trị trung bình (Mean), độ lệch chuẩn (SD), kiểm định độ tin cậy Cronbach's Alpha và phân tích tương quan Pearson nhằm đánh giá độ tin cậy của thang đo và mối quan hệ giữa các thành phần trong mô hình KAP. Phần định tính sử dụng phỏng vấn bán cấu trúc kéo dài từ 15-20 phút với 06 SV. Dữ liệu phỏng vấn được phân tích theo phương pháp phân tích chủ đề (Thematic Analysis). Sau khi phiên mã toàn bộ nội dung phỏng vấn, nhóm nghiên cứu tiến hành mã hóa mở và tổng hợp thành các chủ đề chính gồm: (1) Thiếu hướng dẫn chính thức từ nhà trường; (2) Sử dụng AI do áp lực học tập và thời gian; (3) Khó khăn trong kiểm chứng thông tin; (4) Nhận thức về đạo đức và liêm chính học thuật. Các chủ đề này được đối chiếu với kết quả khảo sát định lượng nhằm giải thích những khác biệt giữa nhận thức và hành vi sử dụng AI của SV.

3. Kết quả nghiên cứu

3.1. Khung lí thuyết về sử dụng AI có trách nhiệm

AI được hiểu là các hệ thống được thiết kế hoặc dựa trên máy móc, có khả năng tạo ra các kết quả đầu ra như dự đoán, khuyến nghị hoặc quyết định ảnh hưởng đến môi trường thực hoặc ảo, dựa trên một tập hợp các mục tiêu do con người xác định (National Institute of Standards and Technology, 2023). Theo UNESCO (2021b), AI là các hệ thống công nghệ có khả năng xử lí dữ liệu và thông tin theo cách tương tự hành vi thông minh của con người, thường bao gồm các năng lực như học tập, suy luận, nhận thức, dự đoán, lập kế hoạch và kiểm soát. OECD (2024) cũng cho rằng AI là hệ thống dựa trên máy được thiết kế để suy luận từ dữ liệu đầu vào nhằm tạo ra các đầu ra như dự đoán, nội dung, khuyến nghị hoặc quyết định, từ đó có thể tác động đến môi trường vật lí hoặc môi trường ảo. Từ các cách tiếp cận trên, nghiên cứu thống nhất hiểu AI là các hệ thống dựa trên máy có khả năng xử lí và suy luận từ dữ liệu để tạo ra các đầu ra như dự đoán, nội dung, khuyến nghị hoặc quyết định, qua đó hỗ trợ con người thực hiện các nhiệm vụ trong nhiều lĩnh vực. Trong nghiên cứu này, AI được hiểu là các công cụ trí tuệ nhân tạo được SV sử dụng phục vụ hoạt động học tập và là đối tượng xem xét dưới góc độ sử dụng AI có trách nhiệm. Tuy nhiên, sự phát triển nhanh chóng của AI cũng đặt ra nhiều vấn đề liên quan đến đạo đức, quyền riêng tư, tính minh bạch của dữ liệu và nguy cơ phụ thuộc công nghệ. Vì vậy, khái niệm "sử dụng AI có trách nhiệm" ngày càng được nhấn mạnh trong các nghiên cứu và chính sách giáo dục quốc tế. Theo Khuyến nghị về đạo đức của trí tuệ nhân tạo của UNESCO (2021b), sử dụng AI có trách nhiệm là việc phát triển và ứng dụng AI dựa trên các nguyên tắc minh bạch, công bằng, an toàn, trách nhiệm giải trình và tôn trọng quyền con người. OECD (2019) cũng cho rằng AI cần được sử dụng theo hướng lấy con người làm trung tâm, bảo đảm công nghệ phục vụ lợi ích xã hội và phát triển bền vững.

Trong nghiên cứu giáo dục, năng lực AI (AI Literacy) là một hướng tiếp cận quan trọng. Lintner (2024) cho rằng năng lực AI bao gồm khả năng hiểu, sử dụng và đánh giá hệ thống AI, đồng thời nhận thức các vấn đề đạo đức phát sinh. UNESCO (2024) cũng nhấn mạnh năng lực AI cần kết hợp giữa hiểu biết công nghệ, tư duy lấy con người làm trung tâm và trách nhiệm đạo đức. Song song, việc sử dụng AI có trách nhiệm gắn chặt với yêu cầu liêm chính học thuật. Cotton và cộng sự (2024), Bittle và El-Gayar (2025) chỉ ra rằng AI tạo sinh vừa hỗ trợ học tập vừa làm gia

tăng nguy cơ đạo văn, giảm tính xác thực và làm mờ ranh giới giữa hỗ trợ học tập và gian lận học thuật. Trên cơ sở đó, nghiên cứu tiếp cận sử dụng AI có trách nhiệm như một hành vi học tập trong môi trường đại học, bao gồm việc khai thác AI để hỗ trợ học tập nhưng vẫn bảo đảm liêm chính học thuật, tư duy độc lập, kiểm chứng thông tin, bảo vệ dữ liệu cá nhân và tuân thủ chuẩn mực đạo đức.

Trong nghiên cứu này, thành phần nhận thức được phân tách thành ba nhóm nội dung gồm: (1) Nhận thức về bản chất và tính minh bạch của AI; (2) Nhận thức về quyền riêng tư và bảo mật dữ liệu; (3) Nhận thức về đạo đức và liêm chính học thuật. Việc phân tách này được thực hiện trên cơ sở tổng hợp các nguyên tắc sử dụng AI có trách nhiệm của UNESCO (2021b), OECD (2019), khung năng lực AI của UNESCO (2024), các nghiên cứu về năng lực AI (Lintner, 2024) và các nghiên cứu về liêm chính học thuật trong bối cảnh AI tạo sinh (Cotton và cộng sự, 2024; Bittle và El-Gayar, 2025). Cụ thể, nhận thức về bản chất và tính minh bạch của AI phản ánh hiểu biết của SV về cách thức hoạt động, giới hạn và khả năng tạo ra sai lệch thông tin của AI; nhận thức về quyền riêng tư và bảo mật dữ liệu phản ánh hiểu biết của SV đối với các rủi ro liên quan đến dữ liệu cá nhân khi sử dụng AI; trong khi nhận thức về đạo đức và liêm chính học thuật phản ánh hiểu biết của SV về các chuẩn mực học thuật, trách nhiệm cá nhân và việc sử dụng AI một cách minh bạch trong học tập. Ba nhóm nhận thức này được sử dụng làm cơ sở xây dựng các biến quan sát trong bảng 2 và phản ánh những thành tố cốt lõi của việc sử dụng AI có trách nhiệm trong môi trường GDĐH.

3.2. Thực trạng sử dụng AI có trách nhiệm của sinh viên Trường Đại học Khoa học Xã hội và Nhân văn - Đại học Quốc gia Hà Nội

Kết quả thống kê mô tả các biến quan sát được trình bày trong Bảng 2 bao gồm giá trị trung bình (Mean) và độ lệch chuẩn (Standard Deviation), qua đó phản ánh mức độ nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm của sinh viên Trường Đại học Khoa học Xã hội và Nhân văn, Đại học Quốc gia Hà Nội.

Bảng 2. Bảng mô tả nhận thức - thái độ - hành vi sử dụng AI có trách nhiệm của SV

Nhóm biến	Mã biến	Nội dung biến quan sát	Mean	Độ lệch chuẩn
Nhóm nhận thức về bản chất và tính minh bạch của AI	**B1**	SV hiểu rõ cách thức các mô hình AI (như ChatGPT, Gemini,...) thu thập và xử lí dữ liệu để đưa ra câu trả lời.	3.4556	1.04133
	B2	SV nhận thức được rằng các câu trả lời của AI có thể chứa đựng "định kiến thuật toán" (sai lệch về giới tính, văn hóa, chính trị,...).	3.5111	0.98450
	B3	SV biết rằng AI có hiện tượng tự tạo ra thông tin giả, nguồn trích dẫn không có thật.	3.4639	1.00074
Nhóm nhận thức về quyền riêng tư và bảo mật dữ liệu	**C1**	SV biết rõ rằng những dữ liệu, thông tin cá nhân tôi cung cấp cho AI có thể được dùng để huấn luyện máy móc hoặc có nguy cơ bị rò rỉ.	3.4472	1.09315
	C2	SV hiểu rằng việc chia sẻ các dữ liệu nghiên cứu nội bộ hoặc thông tin chưa công bố lên các nền tảng AI công cộng là hành vi thiếu an toàn.	3.5167	0.98724
Nhóm nhận thức về đạo đức và liêm chính học thuật	**D1**	SV hiểu rõ ranh giới giữa việc dùng AI để "gợi ý ý tưởng" và việc dùng AI để "viết thay/đạo văn".	3.3972	1.10984
	D2	SV biết các quy định của Trường Đại học Khoa học Xã hội và Nhân văn (hoặc Đại học Quốc gia Hà Nội) về việc sử dụng các công cụ hỗ trợ trong học tập và nghiên cứu.	3.4278	1.12718
	D3	SV nhận thức được tầm quan trọng của việc phải trích dẫn nguồn rõ ràng khi sử dụng nội dung do AI tạo ra trong bài làm.	3.4944	0.99580
	D4	SV hiểu rằng việc lạm dụng AI có thể làm giảm kĩ năng tư duy phản biện và khả năng sáng tạo của bản thân.	3.5500	1.08561
Thái độ của SV đối với giải pháp và định hướng của nhà trường	**G1**	SV đồng ý rằng nhà trường cần hướng dẫn cách sử dụng AI hiệu quả để hỗ trợ học tập và hạn chế việc lạm dụng AI.	3.7639	0.93661
	G2	SV cho rằng nhà trường cần trang bị cho SV năng lực sử dụng AI có trách nhiệm thông qua các hoạt động đào tạo và hỗ trợ học thuật phù hợp.	3.7722	0.93123
	G3	SV cho rằng giảng viên cần hướng dẫn rõ ràng về việc sử dụng AI trong học tập và nghiên cứu.	3.7611	0.93739

	G4	SV sẵn sàng điều chỉnh cách sử dụng AI nếu có định hướng từ nhà trường.	3.7833	0.94574
Hành vi sử dụng AI có trách nhiệm	**E1**	SV chủ động đối chiếu và kiểm chứng thông tin do AI cung cấp bằng các nguồn học thuật đáng tin cậy trước khi sử dụng trong bài học hoặc nghiên cứu của mình.	3.4944	0.97602
	E2	SV cân nhắc kĩ lưỡng các yếu tố bảo mật và quyền riêng tư khi cung cấp dữ liệu cho các nền tảng AI.	3.6222	0.96256
	E3	SV sử dụng AI với vai trò hỗ trợ học thuật, đồng thời đảm bảo duy trì tính độc lập và liêm chính trong sản phẩm học tập của mình.	3.5861	0.99488
	E4	SV khai thác AI như một công cụ thúc đẩy năng lực tự học và phát triển học tập lâu dài.	3.6250	1.03474

Kết quả khảo sát 360 SV cho thấy nhận thức của SV về AI đạt mức trung bình, SV đã có hiểu biết nhất định về AI nhưng mức độ nhận thức tương đối đồng đều giữa ba nhóm nội dung. Cụ thể, nhóm B (nhận thức về bản chất và tính minh bạch của AI) dao động trong khoảng Mean=3.46-3.51, nhóm C (nhận thức về quyền riêng tư và bảo mật dữ liệu) dao động trong khoảng Mean=3.45-3.52 và nhóm D (nhận thức về đạo đức và liêm chính học thuật) dao động trong khoảng Mean=3.40-3.55. Đáng chú ý, biến D1 (ranh giới giữa việc dùng AI để "gợi ý ý tưởng" và việc dùng AI để "viết thay/đạo văn") có điểm trung bình thấp nhất trong toàn bộ nhóm nhận thức (Mean=3.40), cho thấy rằng đây là nội dung SV còn lúng túng nhất và cần được tăng cường hướng dẫn trong thực tiễn đào tạo. Ngược lại, biến D4 (nhận thức về nguy cơ suy giảm tư duy phản biện khi lạm dụng AI) đạt điểm cao nhất trong nhóm nhận thức (Mean=3.55), cho thấy SV có ý thức khá rõ về rủi ro này dù chưa chuyển hóa thành ranh giới hành vi cụ thể như phản ánh ở biến D1.

Đối với thái độ của SV đối với giải pháp và định hướng của nhà trường (nhóm G), các biến đều có điểm trung bình cao (Mean=3.76-3.78), điều này phản ánh sự đồng thuận tương đối cao trong nhu cầu sinh viên cần được nhà trường hướng dẫn cách sử dụng AI hiệu quả, trang bị năng lực sử dụng AI có trách nhiệm, được giảng viên hướng dẫn rõ ràng về việc sử dụng AI trong học tập và nghiên cứu, đồng thời sẵn sàng điều chỉnh nếu có định hướng từ nhà trường. Trong khi đó, các biến thuộc nhóm hành vi sử dụng AI có trách nhiệm (nhóm E) có điểm trung bình thấp hơn nhóm thái độ (Mean=3.49-3.62). Kết quả này cho thấy, mặc dù SV có thái độ tích cực đối với tiếp nhận định hướng từ nhà trường trong việc sử dụng AI có trách nhiệm, mức độ tự báo cáo về việc thực hiện hành vi tương ứng còn ở mức khiêm tốn hơn. Tuy nhiên, việc kiểm định thống kê suy luận nhằm xác nhận ý nghĩa thống kê của chênh lệch này nằm ngoài phạm vi của nghiên cứu hiện tại và được đề xuất như một hướng kiểm chứng cho các nghiên cứu tiếp theo.

Bảng 3. Kết quả kiểm định độ tin cậy Cronbach's Alpha của các thang đo

Thang đo	Nội dung	Số biến quan sát	Cronbach's Alpha
B	Nhận thức về bản chất và tính minh bạch của AI	3	0.830
C	Nhận thức về quyền riêng tư và bảo mật dữ liệu	2	0.732
D	Nhận thức về đạo đức và liêm chính học thuật	4	0.900
G	Thái độ của SV đối với giải pháp và định hướng của nhà trường	4	0.834
E	Hành vi sử dụng AI có trách nhiệm	4	0.777

Độ tin cậy của các thang đo được đánh giá thông qua hệ số Cronbach's Alpha và hệ số tương quan biến - tổng hiệu chỉnh (Corrected Item-Total Correlation). Theo Hair và cộng sự (2019), các biến quan sát được giữ lại khi có hệ số Corrected Item-Total Correlation lớn hơn 0,3 và thang đo đạt hệ số Cronbach's Alpha từ 0,6 trở lên đối với nghiên cứu khám phá. Kết quả kiểm định cho thấy tất cả các biến quan sát đều đáp ứng yêu cầu và được giữ lại. Các thang đo có hệ số Cronbach's Alpha dao động từ 0,732 đến 0,900 - vượt ngưỡng chấp nhận và bảo đảm tính nhất quán nội tại.

Kết quả phân tích tương quan Pearson cho thấy các thành phần của mô hình KAP có mối tương quan thuận với nhau và đều có ý nghĩa thống kê. Cụ thể, nhận thức (K) có tương quan với thái độ (A) ở mức $r = 0,226$ ($p < 0,001$); nhận thức (K) tương quan với hành vi (P) ở mức $r = 0,139$ ($p = 0,008$); và thái độ (A) tương quan với hành vi (P) ở mức $r = 0,157$ ($p = 0,003$). Các hệ số tương quan đều mang dấu dương, cho thấy nhận thức và thái độ tích cực hơn có xu hướng đi kèm với hành vi sử dụng AI có trách nhiệm cao hơn. Tuy nhiên, mức độ tương quan còn tương đối thấp, cho thấy hành vi sử dụng AI có trách nhiệm của SV có thể chịu tác động bởi nhiều yếu tố khác ngoài nhận thức và thái độ. Do mục tiêu của

nghiên cứu là mô tả thực trạng nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm của SV, nghiên cứu tập trung vào thống kê mô tả, kiểm định độ tin cậy và phân tích tương quan nhằm phản ánh mối liên hệ giữa các thành phần của mô hình KAP, thay vì xây dựng mô hình dự báo hành vi.

Bảng 4. Ma trận tương quan Pearson giữa các thành phần của mô hình KAP

Biến	K	A	P
K (Knowledge)	1	0.226**	0.139**
A (Attitude)	0.226**	1	0.157**
P (Practice)	0.139**	0.157**	1

Ghi chú: N = 360; p < 0,01 (kiểm định hai phía).

Kết quả phỏng vấn sâu cho thấy hành vi sử dụng AI của SV chịu ảnh hưởng mạnh từ môi trường học tập và cách giảng viên định hướng trong từng môn học. Nhiều SV cho biết việc sử dụng AI hiện nay chủ yếu dựa trên "kinh nghiệm tự học" thay vì được hướng dẫn chính thức từ nhà trường. Một SV ngành Quản trị dịch vụ du lịch và lữ hành chia sẻ: "Mình dùng AI khá nhiều để tìm ý tưởng hoặc viết dàn ý nhưng hầu như chưa có môn nào hướng dẫn cụ thể là được dùng đến mức nào hay phải trích dẫn ra sao".

Các kết quả này cho thấy sự thiếu thống nhất trong định hướng học thuật có thể là một trong những nguyên nhân khiến hành vi sử dụng AI của SV mang tính tự phát và thiếu nhất quán. Dữ liệu định tính đồng thời góp phần lí giải chênh lệch quan sát được giữa nhận thức và hành vi trong khảo sát định lượng. Mặc dù nhiều SV nhận thức được các rủi ro liên quan đến AI như sai lệch thông tin hoặc phụ thuộc công nghệ, trên thực tế họ vẫn sử dụng AI thường xuyên do áp lực thời gian, khối lượng bài tập lớn và sự tiện lợi của công nghệ. Một SV cho biết: "Em biết AI có thể đưa thông tin sai hoặc làm mình bị phụ thuộc nhưng khi deadline nhiều thì vẫn phải dùng vì nhanh hơn rất nhiều". Điều này cho thấy hành vi sử dụng AI của SV không chỉ chịu ảnh hưởng bởi nhận thức cá nhân mà còn gắn chặt với bối cảnh học tập thực tế trong môi trường đại học.

Một điểm đáng chú ý trong kết quả nghiên cứu là sự khác biệt giữa dữ liệu định lượng và dữ liệu định tính liên quan đến nhận thức của SV về quy định sử dụng AI của nhà trường. Trong khảo sát định lượng, biến D2 ("SV biết các quy định của Trường Đại học Khoa học Xã hội và Nhân văn hoặc Đại học Quốc gia Hà Nội về việc sử dụng AI trong học tập và nghiên cứu") đạt mức trung bình 3,4278 - tương đương mức trung bình khá. Tuy nhiên, trong phỏng vấn sâu, nhiều SV lại cho biết chưa được hướng dẫn cụ thể và "không biết giới hạn ở đâu" khi sử dụng AI. Sự khác biệt này có thể liên quan đến hiện tượng "social desirability bias" (thiên lệch đáp ứng kì vọng xã hội), tức người tham gia khảo sát có xu hướng tự đưa ra câu trả lời phù hợp với những chuẩn mực hoặc hành vi được xã hội đánh giá là tích cực và đáng mong đợi (Randall và Fernandes, 1991). Ngoài ra, một số SV có thể hiểu "biết quy định" theo nghĩa đã từng nghe hoặc tiếp cận thông tin chung về AI nhưng chưa thực sự nắm rõ các hướng dẫn cụ thể trong bối cảnh học thuật. Kết quả phỏng vấn sâu cho thấy phần lớn SV chưa phân biệt rõ ràng giữa việc sử dụng AI như công cụ hỗ trợ học tập với hành vi phụ thuộc AI hoặc vi phạm liêm chính học thuật. Những băn khoăn này cho thấy ranh giới giữa việc AI hỗ trợ và thay thế hoạt động học thuật còn chưa rõ ràng. Một số SV cũng gặp khó khăn trong việc kiểm chứng thông tin do AI cung cấp vì thiếu thời gian và kỹ năng đối chiếu nguồn, dẫn đến xu hướng sử dụng kết quả AI một cách nhanh chóng. Điều này cho thấy nhận thức của SV về AI hiện nay vẫn mang tính khái quát và chưa chuyển hóa đầy đủ thành hành vi thực tế.

3.3. Giải pháp nhằm nâng cao việc sử dụng AI có trách nhiệm của sinh viên để thúc đẩy giáo dục hướng tới phát triển bền vững

Từ kết quả nghiên cứu, có thể thấy việc sử dụng AI có trách nhiệm trong môi trường đại học không chỉ liên quan đến nhận thức và thái độ của SV mà còn chịu ảnh hưởng từ bối cảnh học tập và sự định hướng trong môi trường giáo dục. Kết quả khảo sát cho thấy mặc dù SV có nhận thức và thái độ tương đối tích cực đối với việc sử dụng AI có trách nhiệm nhưng hành vi thực tế chưa hoàn toàn tương ứng với mức độ nhận thức. Dữ liệu định tính cũng cho thấy nhiều SV hiện nay chủ yếu sử dụng AI dựa trên kinh nghiệm cá nhân và chưa được tiếp cận các hướng dẫn học thuật một cách thống nhất. Trên cơ sở đó, nghiên cứu đưa ra một số hàm ý như sau:

Thứ nhất, kết quả nghiên cứu cho thấy SV có nhu cầu được định hướng rõ ràng hơn về việc sử dụng AI trong học tập và nghiên cứu. Nhiều SV tham gia phỏng vấn cho biết chưa thực sự hiểu rõ giới hạn của việc sử dụng AI trong các nhiệm vụ học thuật. Điều này cho thấy nhà trường có thể xem xét tăng cường hoạt động truyền thông, phổ biến và hướng dẫn về sử dụng AI trong môi trường học thuật nhằm giúp SV hiểu rõ hơn về trách nhiệm và chuẩn mực khi sử dụng công nghệ này; *Thứ hai,* kết quả khảo sát và phỏng vấn đều cho thấy một bộ phận SV chưa thường xuyên thực hiện các hành vi kiểm chứng thông tin hoặc đánh giá độ tin cậy của nội dung do AI tạo ra. Điều này gợi ý rằng việc phát triển năng lực đánh giá thông tin, tư duy phản biện và kĩ năng học thuật tiếp tục là những nội dung cần được quan tâm trong bối cảnh AI ngày càng được sử dụng phổ biến trong GDĐH; *Thứ ba,* kết quả nghiên cứu

cho thấy việc sử dụng AI trong giáo dục cần được tiếp cận theo hướng hỗ trợ quá trình học tập thay vì thay thế tư duy của người học. Đối với SV khối Khoa học Xã hội và Nhân văn, các năng lực như tư duy phản biện, lập luận học thuật, phân tích và xây dựng quan điểm cá nhân vẫn giữ vai trò cốt lõi. Do đó, AI nên được xem là công cụ hỗ trợ học tập, trong khi người học vẫn giữ vai trò trung tâm trong quá trình tiếp nhận, đánh giá và tạo lập tri thức; *Thứ tư,* việc nâng cao nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm được xem là một yêu cầu quan trọng trong bối cảnh GDĐH đang đẩy mạnh chuyển đổi số. Điều này không chỉ góp phần nâng cao hiệu quả học tập của SV mà còn hỗ trợ hình thành năng lực thích ứng với môi trường học tập và làm việc trong thời đại số. Tuy nhiên, các hàm ý nêu trên được rút ra từ kết quả khảo sát và phỏng vấn trong phạm vi nghiên cứu này nên những đề xuất cụ thể về chính sách quản lí hoặc chương trình đào tạo liên quan đến AI cần tiếp tục được kiểm chứng thông qua các nghiên cứu sâu hơn trong tương lai.

3.4. Thảo luận

Kết quả nghiên cứu cho thấy SV Trường Đại học Khoa học Xã hội và Nhân văn sử dụng AI khá phổ biến trong các hoạt động học tập như tìm kiếm thông tin, xây dựng ý tưởng và hỗ trợ thực hiện các nhiệm vụ học thuật. Đối với khối ngành Khoa học Xã hội và Nhân văn, nơi các nhiệm vụ học tập đòi hỏi khả năng phân tích, lập luận và tư duy phản biện, AI chủ yếu được xem là công cụ hỗ trợ hơn là thay thế hoàn toàn quá trình học tập. Kết quả này phù hợp với nghiên cứu của Qu và cộng sự (2024), cho thấy mức độ và cách thức sử dụng AI có sự khác biệt giữa các lĩnh vực đào tạo, đồng thời gợi ý rằng đặc thù ngành học là một trong những yếu tố ảnh hưởng đến hành vi sử dụng AI của SV.

Kết quả định tính cho thấy nhiều SV mong muốn có thêm hướng dẫn và định hướng cụ thể về việc sử dụng AI trong học tập và nghiên cứu. Điều này góp phần giải thích khoảng cách giữa nhận thức, thái độ và hành vi sử dụng AI có trách nhiệm, đặc biệt là kiểm chứng thông tin và đánh giá độ tin cậy của kết quả do AI tạo ra. Từ đó có thể thấy việc hình thành hành vi sử dụng AI có trách nhiệm không chỉ phụ thuộc vào nhận thức cá nhân mà còn chịu ảnh hưởng của môi trường học tập, sự hướng dẫn của giảng viên và các quy định của cơ sở GDĐH.

4. Kết luận và bình luận

Nghiên cứu cho thấy SV Trường Đại học Khoa học Xã hội và Nhân văn đã bước đầu hình thành nhận thức và thái độ tích cực đối với việc sử dụng AI trong học tập, song việc thực hành sử dụng AI có trách nhiệm vẫn chưa thực sự đồng bộ. Khoảng cách giữa nhận thức, thái độ và hành vi thể hiện rõ ở các hành vi như kiểm chứng thông tin, trích dẫn nguồn và bảo đảm liêm chính học thuật khi sử dụng AI. Về đóng góp, nghiên cứu vận dụng mô hình KAP để khảo sát việc sử dụng AI có trách nhiệm của SV Trường Đại học Khoa học Xã hội và Nhân văn, qua đó cung cấp thêm một góc nhìn về nhận thức, thái độ và hành vi của SV trong bối cảnh GDĐH. Từ kết quả nghiên cứu, bài báo gợi mở một số nội dung có thể được cân nhắc trong quá trình xây dựng hướng dẫn sử dụng AI, phát triển năng lực số và giáo dục liêm chính học thuật cho SV.

Bên cạnh đó, nghiên cứu vẫn tồn tại một số hạn chế nhất định: (1) Mẫu nghiên cứu được lựa chọn theo phương pháp thuận tiện và tập trung tại một cơ sở đào tạo nên mức độ khái quát hóa kết quả còn hạn chế; (2) Số lượng mẫu định tính chưa lớn và chủ yếu phản ánh trải nghiệm của SV trong bối cảnh Trường Đại học Khoa học Xã hội và Nhân văn; (3) nghiên cứu tiếp cận theo hướng khảo sát thực trạng nên chưa phân tích được sự thay đổi hành vi sử dụng AI của SV theo thời gian. Ngoài ra, nghiên cứu chưa xem xét đầy đủ các yếu tố bối cảnh như chính sách quản lí học thuật, mức độ kiểm soát của giảng viên hoặc áp lực học tập trong việc hình thành hành vi sử dụng AI của SV. Đây có thể là hướng nghiên cứu tiếp theo nhằm mở rộng khả năng giải thích hành vi sử dụng AI trong GDĐH. Các nghiên cứu tiếp theo có thể mở rộng phạm vi khảo sát sang nhiều cơ sở GDĐH và nhóm ngành đào tạo khác nhau nhằm tăng khả năng so sánh và khái quát hóa kết quả. Cùng với đó, cần xem xét thêm các yếu tố như năng lực AI, năng lực số, chính sách của nhà trường hoặc liêm chính học thuật để có cái nhìn toàn diện hơn về các yếu tố liên quan đến hành vi sử dụng AI có trách nhiệm của SV.

Tuyên bố về vai trò của các tác giả: *Đinh Gia Hân: Xây dựng ý tưởng nghiên cứu, Thiết kế phương pháp nghiên cứu, Phân tích dữ liệu, Thực hiện nghiên cứu và thu thập dữ liệu, Cung cấp tài nguyên, Chỉnh sửa và phản biện bản thảo, Trực quan hóa dữ liệu, Quản lí và điều phối. Nguyễn Thị Lan Anh: Giám sát và hướng dẫn, Thẩm định và kiểm chứng kết quả, Chỉnh sửa và phản biện bản thảo. Nguyễn Thị Minh Ánh: Xây dựng ý tưởng nghiên cứu, Thực hiện nghiên cứu và thu thập dữ liệu, Cung cấp tài nguyên, Viết bản thảo ban đầu. Hà Thu Hằng: Xây dựng ý tưởng nghiên cứu, Thực hiện nghiên cứu và thu thập dữ liệu, Cung cấp tài nguyên, Viết bản thảo ban đầu. Mai Thị Thanh Hoa: Thiết kế phương pháp nghiên cứu, Phân tích dữ liệu, Thực hiện nghiên cứu và thu thập dữ liệu, Cung cấp tài nguyên. Phan Thị Minh Hòa: Thực hiện nghiên cứu và thu thập dữ liệu, Cung cấp tài nguyên, Viết bản thảo ban đầu.*

Tuyên bố về GenAI và Quyền tác giả: *Trong quá trình chuẩn bị bản thảo này, các tác giả đã sử dụng ChatGPT, phiên bản Free cho việc chỉnh sửa văn phong để tăng tính mạch lạc cho nội dung.*
Tuyên bố về xung đột lợi ích: *Các tác giả tuyên bố không có xung đột lợi ích.*
Thông tin tài trợ: *Nghiên cứu này được tài trợ bởi Trường Đại học Khoa học Xã hội và Nhân văn theo mã số: SV.2026.12.*
Lời cảm ơn: *Nhóm tác giả cảm ơn sự tài trợ của Trường Đại học Khoa học Xã hội và Nhân văn qua đề tài với mã số: SV.2026.12.*

Tài liệu tham khảo

Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education, 21*(1), 10. https://doi.org/10.1186/s41239-024-00444-7

Bittle, K., & El-Gayar, O. (2025). Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda. *Information, 16*(4), 296. https://doi.org/10.3390/info16040296

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

Đặng Văn Em, Nguyễn Đình Loan Phương, Nguyễn Thị Hảo (2024). Thực trạng ứng dụng ChatGPT trong việc học tập, nghiên cứu của sinh viên Đại học Quốc gia Thành phố Hồ Chí Minh. *Tạp chí Giáo dục, 24*(1), 36-41. https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/1212

Hair, J. F., Babin, B. J., Anderson, R. E., & Black, W. C. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods, 18*(1), 3-20. https://doi.org/10.1177/1525822X05282260

Kiều Phương Thùy, Nguyễn Khắc Ân, Vũ Thái Giang (2025). Một số phương pháp ứng dụng tư duy phản biện khi sử dụng AI tạo sinh trong giáo dục phổ thông. *Tạp chí Khoa học, Trường Đại học Sư phạm Hà Nội, 70*(3), 161-172. https://doi.org/10.18173/2354-1075.2025-0060

Lintner, T. (2024). A systematic review of AI literacy scales. *npj Science of Learning, 9*, 50. https://doi.org/10.1038/s41539-024-00264-4

National Institute of Standards and Technology (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

OECD (2019). *Artificial Intelligence in Society*. https://doi.org/10.1787/eedfee77-en

OECD (2024). *Explanatory memorandum on the updated OECD definition of an AI system: OECD Artificial Intelligence Papers, No. 8*. https://doi.org/10.1787/623da898-en

Qu, Y., Tan, M. X. Y., & Wang, J. (2024). Disciplinary differences in undergraduate students' engagement with generative artificial intelligence. *Smart Learning Environments, 11*, 51. https://doi.org/10.1186/s40561-024-00341-6

Randall, D. M., & Fernandes, M. F. (1991). The social desirability response bias in ethics research. *Journal of Business Ethics*, 10(11), 805-817. https://link.springer.com/article/10.1007/BF00383696

Sanchez, J. M. P., Sumalinog, G. G., Mananay, J. A., Goles, C. E., Alejandro, I. M. V., & Fernandez, C. B. (2025). Scale development and validation of knowledge-attitudes-practices (KAP) on artificial intelligence (AI) productivity tools in education. *STEM Education, 5*(6), 1022-1057. https://doi.org/10.3934/steme.2025045

UNESCO (2021a). *Artificial intelligence in education*. https://www.unesco.org/en/digital-education/artificial-intelligence

UNESCO (2021b). *Recommendation on the Ethics of Artificial Intelligence.* https://unesdoc.unesco.org/ark:/48223/pf0000381137

UNESCO (2024). *AI competency framework for students*. https://unesdoc.unesco.org/ark:/48223/pf0000391105

World Health Organization (2008). *Advocacy, communication and social mobilization for TB control: A guide to developing knowledge, attitude and practice surveys*. https://iris.who.int/server/api/core/bitstreams/a08347d2-942d-42ae-9ff3-b1569895a923/content

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - where are the educators? *International Journal of Educational Technology in Higher Education, 16*, 39. https://link.springer.com/article/10.1186/s41239-019-0171-0