SỬ DỤNG TRÍ TUỆ NHÂN TẠO CỦA SINH VIÊN ĐẠI HỌC QUỐC GIA HÀ NỘI: THỰC TRẠNG VÀ MỘT SỐ VẤN ĐỀ ĐẠO ĐỨC ĐẶT RA(*)

PHẠM THU TRANG
Trường Đại học Khoa học Xã hội và Nhân văn, Đại học Quốc gia Hà Nội

Nhận bài ngày 05/11/2025. Sửa chữa xong 20/12/2025. Duyệt đăng 28/12/2025.

Abstract

The rapid development of artificial intelligence (AI), particularly generative AI (GenAI) models, has brought about profound changes in the learning and research activities of university students. At Vietnam National University, Hanoi (VNU), students represent one of the groups that have early and extensively adopted AI in their academic practices. This study aims to assess the current state of AI use among VNU students and identify the ethical issues arising from such usage. Data were collected from a survey of 746 students across various academic disciplines, focusing on levels of AI use, purposes of use, perceptions of AI, and behaviors that may pose risks of violating academic integrity. Based on the findings, the study proposes several orientations to promote ethical AI use within higher education settings.

Keywords: *Artificial intelligence, ethics, students, Vietnam National University Hanoi.*

1. Đặt vấn đề

Sự xuất hiện và phổ biến nhanh chóng của trí tuệ nhân tạo tạo sinh (Generative Artificial Intelligence - GenAI) như ChatGPT, Gemini, Claude hay Copilot đang định hình lại toàn bộ hoạt động học thuật trong các trường đại học trên toàn thế giới. AI không chỉ mang đến những cơ hội mới cho học tập, nghiên cứu mà còn đặt ra hàng loạt thách thức mang tính nền tảng đối với đạo đức học thuật, chuẩn mực nghề nghiệp và năng lực tư duy của người học. Trong bối cảnh ấy, việc nghiên cứu cách sinh viên (SV) sử dụng AI và những hệ quả đạo đức phát sinh trở thành một yêu cầu cấp thiết. Tại Việt Nam, AI đã nhanh chóng được SV tiếp nhận như một công cụ hỗ trợ học tập hiệu quả, đặc biệt trong các hoạt động như viết bài tập, tóm tắt tài liệu, xử lý số liệu hay chuẩn bị bài thuyết trình. Tuy nhiên, chính sự phổ biến rộng rãi này lại làm dấy lên lo ngại rằng SV có thể đang sử dụng AI mà chưa được trang bị đầy đủ về kỹ năng đánh giá độ tin cậy, nguyên tắc an toàn dữ liệu cũng như các chuẩn mực đạo đức cần thiết trong bối cảnh học thuật. Vì vậy, việc nắm bắt đúng thực trạng sử dụng AI của SV không chỉ giúp nhận diện các vấn đề đạo đức nổi bật mà còn là cơ sở để xây dựng các chính sách phù hợp ở cấp trường và cấp hệ thống. Đại học Quốc gia Hà Nội (ĐHQG Hà Nội) - đại học đa ngành, đa lĩnh vực hàng đầu cả nước là môi trường mà SV có mức độ tiếp xúc công nghệ sớm, đa dạng về ngành học và năng động trong việc ứng dụng các công cụ số. Trong bối cảnh đó, nghiên cứu này được thực hiện dựa trên khảo sát 746 SV của các ngành khác nhau trong ĐHQG Hà Nội nhằm phân tích mức độ, mục đích, hành vi và nhận thức liên quan đến việc sử dụng AI; đồng thời xác định các rủi ro đạo đức học thuật nảy sinh. Các kết quả thu được sẽ góp phần cung cấp dữ liệu thực chứng cho việc xây dựng môi trường giáo dục số lành mạnh, phát triển năng lực sử dụng AI có trách nhiệm cho SV trong thời đại mới.

(*) Bài viết thuộc đề tài khoa học cấp cơ sở mã số CS. 2025.23 Trường Đại học Khoa học Xã hội và Nhân văn, Đại học Quốc gia Hà Nội.
Email: trangtriet84.ussh@gmail.com

2. Nội dung nghiên cứu

2.1. Tổng quan lý luận về trí tuệ nhân tạo (AI)

Trí tuệ nhân tạo (AI) thường được hiểu như khả năng của máy móc mô phỏng những hoạt động tư duy của con người, bao gồm học hỏi, lập luận, giải quyết vấn đề và ra quyết định. Theo Stuart Russell và Peter Norvig, vấn đề đặt ra là AI được xem xét đồng nhất với hành vi thông minh hay quá trình tư duy nội tại hay không. Từ đó, hai tác giả đã khái quát bốn cách tiếp cận cơ bản trong nghiên cứu và phát triển AI đó là: 1) *AI hành động như con người*, nhấn mạnh việc mô phỏng hành vi con người nhưng hiện nay ít được ưu tiên; 2) *AI tư duy như con người*, tập trung mô hình hóa các quá trình nhận thức của con người; 3) *AI tư duy hợp lý*, dựa trên logic và các hệ thống ký hiệu, song gặp thách thức khi phải xử lý sự bất định; 4) *AI hành động hợp lý*, chú trọng tối ưu hóa hành vi hợp lý mà không cần sao chép tư duy hay hành vi con người. Trong bốn cách tiếp cận này, Russell và Norvig coi *AI hành động hợp lý* là cách tiếp cận tiêu chuẩn và có ảnh hưởng sâu rộng nhất trong lịch sử phát triển AI [7, tr. 31-38].

Theo Luckin và các tác giả, AI được định nghĩa là các hệ thống máy tính được thiết kế để tương tác với thế giới thông qua các khả năng và hành vi thông minh mà chúng ta thường coi là đặc trưng của con người. Các khả năng này bao gồm *năng lực nhận thức* (như nhận dạng thị giác và nhận dạng giọng nói) và *hành vi thông minh* (Chẳng hạn như việc đánh giá thông tin sẵn có và sau đó thực hiện hành động hợp lý nhất để đạt được mục tiêu đã đề ra). Các tác giả cho rằng AI giúp tạo ra các môi trường học tập thích ứng, linh hoạt và được cá nhân hóa cao đồng thời như một trợ lý đa năng giúp hỗ trợ tối đa cho người dùng [5, tr. 14]. Các nghiên cứu của UNESCO (2022) và OECD (2023) cũng nhấn mạnh vai trò của AI trong việc tạo ra môi trường học tập linh hoạt, thông minh và định hướng kỹ năng tương lai, đồng thời cảnh báo rủi ro liên quan đến đạo đức học thuật, bảo mật dữ liệu và sự phụ thuộc công nghệ. Nghiên cứu của Alenezi cho thấy AI là yếu tố trung tâm trong chuyển đổi số giáo dục đại học (GDĐH) nhưng cần được tích hợp đi kèm chính sách kiểm soát phù hợp với văn hóa học thuật từng quốc gia [1, tr. 1-13].

Tại Việt Nam, các nghiên cứu học thuật về việc SV sử dụng AI tuy đã bắt đầu xuất hiện nhưng còn rời rạc và thiên về mô tả nhận thức chung hơn là phân tích hành vi thực tế dựa trên dữ liệu quy mô lớn. Một số khảo sát gần đây cho thấy SV rất chủ động tìm hiểu và tiếp cận các công cụ AI phục vụ học tập, tuy nhiên, mức độ hiểu biết về rủi ro đạo đức, quyền riêng tư dữ liệu hay cách sử dụng có trách nhiệm còn hạn chế. Đồng thời, nhiều cơ sở GDĐH chưa xây dựng được quy định hay hướng dẫn chính thức giúp SV phân biệt giữa hỗ trợ học tập hợp lệ và hành vi gian lận khi sử dụng các công cụ AI tạo sinh. Điều này khiến AI vừa được nhìn nhận như một cơ hội thúc đẩy học tập, vừa gây lo ngại về những hệ lụy tiềm ẩn. Để tiếp cận những vấn đề này một cách toàn diện, nghiên cứu lựa chọn phương pháp hỗn hợp (mixed methods), kết hợp định lượng và định tính. Trọng tâm của nghiên cứu là phân tích dữ liệu thu thập từ 746 phiếu khảo sát được phát ra cho SV thuộc nhiều trường đại học thành viên của ĐHQG Hà Nội. Bên cạnh dữ liệu định lượng, nghiên cứu thực hiện 20 cuộc phỏng vấn sâu với SV thuộc nhiều ngành học khác nhau nhằm đi sâu vào trải nghiệm, động cơ, lo ngại, kỳ vọng và quan điểm cá nhân của người học đối với AI trong học tập. Việc kết hợp hai nguồn dữ liệu giúp bổ sung chiều sâu cho các phân tích định lượng, đồng thời phản ánh rõ hơn những sắc thái nhận thức và hành vi mà khảo sát định lượng khó nắm bắt được.

2.2. Kết quả nghiên cứu về thực trạng sử dụng AI của sinh viên Đại học Quốc gia Hà Nội và một số vấn đề đạo đức đặt ra

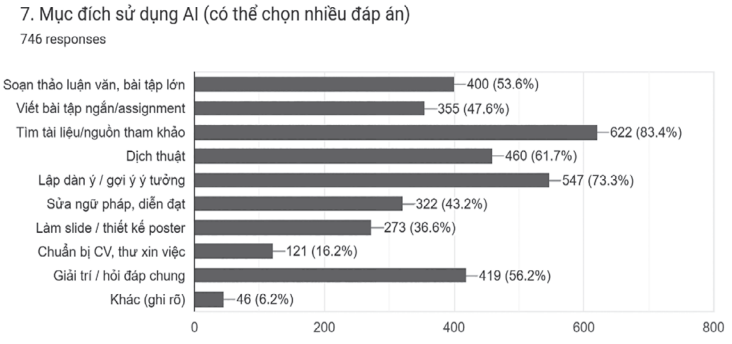
Bảng trả lời khảo sát nhận được từ 746 SV các khóa khác nhau của các trường trực thuộc ĐHQG Hà Nội, trải rộng các nhóm ngành khác nhau như: Trường Đại học Khoa học Xã hội và Nhân văn, Trường Đại học Khoa học Tự nhiên, Trường Đại học Luật, Trường Đại học Kinh tế, Trường Khoa học Liên ngành và Nghệ thuật, Trường Đại học Y Dược, Trường Đại học Giáo dục, Trường Quốc tế, Trường Đại học Công nghệ… Kết quả phản ánh một bức tranh tương đối toàn diện về thực trạng sử dụng AI trong học tập của SV ĐHQG Hà Nội hiện nay.

Kết quả cho thấy gần như toàn bộ SV đều đã từng sử dụng AI phục vụ cho mục đích học tập, trong đó tỷ lệ SV sử dụng thường xuyên và thỉnh thoảng chiếm đa số. SV hiếm khi sử dụng hoặc chưa từng sử

dụng AI chỉ chiếm một tỷ lệ rất nhỏ và không đáng kể. Điều này phản ánh AI đã thâm nhập sâu vào đời sống học thuật của SV, trở thành công cụ quen thuộc phục vụ nhiều mục đích học tập khác nhau. Đối với câu hỏi lựa chọn nhiều đáp án về *"Bạn đã từng sử dụng công cụ AI nào sau đây để hỗ trợ học tập?"*, kết quả cho thấy ChatGPT là công cụ AI được sử dụng ở mức độ thường xuyên nhất, áp đảo so với các công cụ còn lại. Sau ChatGPT là Gemini (Google AI), tiếp đến là nhóm AI dịch thuật (Google Translate) và các AI tạo hình ảnh, video. Ngoài ra, AI kiểm tra ngữ pháp như Grammarly, Quillbot cũng được dùng với tần suất tương đối, đặc biệt trong các môn đòi hỏi viết nhiều. Điều này cho thấy SV ĐHQG Hà Nội không chỉ giới hạn trong một công cụ duy nhất mà có xu hướng kết hợp đa nền tảng để tối ưu hóa hiệu quả học tập. Đa số SV chỉ mới bắt đầu sử dụng AI trong hai năm trở lại đây, đặc biệt từ năm 2023 và 2024, trùng với thời điểm các mô hình tạo sinh như ChatGPT bùng nổ trên quy mô toàn cầu. Việc phần lớn SV tiếp cận AI rất nhanh và gần như đồng loạt cho thấy sự phổ biến của AI trong GDĐH Việt Nam diễn ra theo hình thức lan truyền mạnh mẽ, chủ yếu thông qua tác động xã hội và môi trường số.

Về mục đích sử dụng AI và cách thức ứng dụng trong học tập, kết quả khảo sát cho thấy, SV sử dụng AI vào nhiều hoạt động, trong đó nổi bật nhất là tìm kiếm tài liệu, nguồn tham khảo chiếm 83,4 %; tiếp theo là 73,3% sử dụng AI để lập dàn ý, gợi ý ý tưởng viết bài thể hiện vai trò AI như "công cụ khơi gợi ý tưởng"; với mục đích dịch thuật có khoảng 61,7%; sau đó là 56,2% sử dụng AI để giải trí hoặc hỏi đáp chung, điều này cho thấy mức độ hòa nhập của AI vào đời sống cá nhân cũng khá rõ nét. Các SV còn sử dụng AI để soạn bài tập lớn hoặc làm luận văn, viết bài tập ngắn cũng lần lượt chiếm khoảng 53,6 % và 47,6% cho thấy mức độ can thiệp sâu của AI vào sản phẩm học thuật. Bên cạnh đó, AI còn đóng vai trò như một trợ lý viết với việc sửa lỗi diễn đạt và ngữ pháp chiếm 43,2%, ngoài ra còn giúp SV chuẩn bị CV, viết thư xin việc… Kết quả này cho thấy phần lớn SV coi AI như công cụ hỗ trợ học tập đa chức năng, mang lại lợi ích về tốc độ xử lý thông tin, hiệu quả tìm kiếm và trợ giúp trong tư duy.

Biểu đồ 1: Mục đích sử dụng AI



Nguồn: Kết quả khảo sát tháng 11/2025

Đánh giá về mức độ hiểu biết về đạo đức AI thì đa số SV đánh giá mức độ hiểu biết của mình ở mức trung bình hoặc tốt, trong khi nhóm tự nhận "rất tốt", kém hoặc rất kém chiếm tỷ lệ thấp. Điều này cho thấy SV có sự tự tin nhất định vào khả năng đánh giá đúng sai trong bối cảnh AI nhưng niềm tin này có thể không phản ánh đúng thực tế. Liên quan đến hành vi để AI viết bài thay, kết quả khảo sát cho thấy 26,4% SV thừa nhận đã từng nộp bài mà phần lớn nội dung do AI viết, trong khi gần 23% "không chắc chắn" liệu mình có làm vậy hay không. Điều này phản ánh SV không thực sự phân biệt được thế nào là "AI hỗ trợ" và "AI làm thay"; Họ sử dụng AI như một phần của quá trình viết đến mức khó tách bạch tỷ lệ đóng góp; Hành vi gian lận vô thức có thể diễn ra thường xuyên mà SV không nhận ra.

Về mức độ minh bạch trong trích dẫn AI cho thấy hành vi minh bạch còn rất hạn chế: 32,8% chưa bao giờ trích dẫn, 16,9% hiếm khi trích dẫn, 35,9% thỉnh thoảng trích dẫn, chỉ 14,3% luôn trích dẫn. Điều này cho thấy đa số SV chưa xem việc trích dẫn AI là yêu cầu học thuật. Đây là vấn đề nghiêm trọng

vì minh bạch thông tin chính là tiêu chí cơ bản của liêm chính học thuật. Thói quen kiểm chứng thông tin AI và nhận thức về rủi ro, SV thể hiện thái độ tương đối thận trọng: 77% SV cho biết luôn kiểm tra lại thông tin AI cung cấp, 17% chỉ kiểm tra trong một số tình huống, 6% không kiểm tra. Việc đa số kiểm chứng lại thông tin cho thấy SV không hoàn toàn tin tưởng AI nhưng nhóm "không kiểm tra" vẫn đủ lớn để tạo ra rủi ro lan truyền thông tin sai trong môi trường học thuật.

Nhận thức về quyền sở hữu trí tuệ khi dùng AI kết quả cho thấy 53,2% SV chỉ "biết một phần", 31,1% biết rõ, 15,7% không biết gì. Có thể thấy đa số SV thiếu hiểu biết đầy đủ về vấn đề bản quyền khi đưa dữ liệu, tài liệu, văn bản hoặc ý tưởng vào hệ thống AI. Đây là khoảng trống kiến thức nguy hiểm, bởi AI có thể lưu trữ, tái sử dụng hoặc biến đổi nội dung do người dùng nhập vào mà họ không hề biết. Khi được hỏi về việc tìm hiểu chính sách quyền riêng tư của các nền tảng AI: Chỉ 15,3% SV đọc đầy đủ, gần 50% hiếm khi hoặc thỉnh thoảng tìm hiểu; 23,8% chưa từng đọc. Xu hướng này cho thấy SV đang vô tình chia sẻ dữ liệu cá nhân cho các mô hình AI mà không ý thức được hệ quả. Đây là rủi ro lớn trong bối cảnh AI thu thập dữ liệu mờ và khó kiểm soát.

Quan niệm của SV về trách nhiệm đạo đức, quan điểm của SV là nhà trường và giảng viên (GV) được xem là đối tượng chịu trách nhiệm chính, SV cũng được thừa nhận có trách nhiệm nhưng không phải yếu tố hàng đầu. Điều này phản ánh tâm lý phổ biến: SV mong muốn có khung hướng dẫn chính thức thay vì phải tự định hình chuẩn mực. Điều này hợp lý trong giai đoạn công nghệ đang thay đổi nhanh và chưa có chuẩn mực chung về việc sử dụng AI trong học thuật

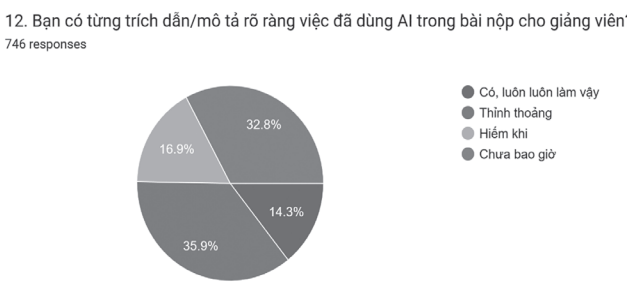
Kết quả phỏng vấn sâu 20 SV cho thấy nhiều điểm nổi bật: Về lợi ích: SV cho rằng lợi ích lớn nhất của AI là: tiết kiệm thời gian, hỗ trợ tìm tài liệu hiệu quả hơn, gợi ý ý tưởng, hỗ trợ lập luận; bổ sung khoảng trống kiến thức, giúp xây dựng khung nội dung ban đầu, tăng tính sáng tạo khi có công cụ hỗ trợ. Tất cả đều đồng thuận rằng AI đem lại hiệu quả vượt trội nếu biết dùng đúng cách. Tuy nhiên, phần lớn các em đều lo ngại rằng, rủi ro lớn nhất của việc sử dụng AI là lệ thuộc và mất khả năng tư duy độc lập. Bên cạnh đó, các em còn nêu ra một số lo ngại khác như: thông tin sai lệch, bóp méo kiến thức, quên cách viết; giảm động lực học tập và tự tin ảo khi dựa dẫm quá nhiều vào AI. Từ đó, các SV được phỏng vấn đã đề xuất một số giải pháp như: nhà trường có thể đưa AI vào giảng dạy bài bản; tổ chức hội thảo, seminar về AI; hướng dẫn cách sử dụng AI đúng mục đích; không nên cấm AI nhưng phải có khung đạo đức rõ ràng. Xu hướng chung của SV là tiếp nhận AI tích cực nhưng không muốn bị lệ thuộc.

Như vậy, kết quả nghiên cứu về tình hình sử dụng AI của SV ĐHQG Hà Nội cho thấy một bức tranh nhiều chiều, trong đó hiện hữu cả những cơ hội to lớn lẫn thách thức đáng kể đối với GDĐH. Hiện nay AI đã trở thành công cụ học tập chính của SV - một sự chuyển đổi nhanh ngoài dự báo. Điều này cho thấy GDĐH đang bước vào giai đoạn mà AI không còn là "công cụ hỗ trợ tùy chọn" mà đã trở thành hạ tầng học tập mới. Hiện tượng này phù hợp với xu hướng toàn cầu, khi các mô hình tạo sinh (LLM) như GPT-4/5, Gemini, Claude hay DeeSeek làm thay đổi cách con người tìm kiếm thông tin, viết, lập luận và sáng tạo nội dung. Tuy nhiên, điều đáng chú ý là sự lan rộng của AI trong cộng đồng SV xảy ra cực nhanh, gần như không qua giai đoạn chuyển tiếp. Phần lớn SV mới tiếp cận AI trong vòng 1-2 năm trở lại đây nhưng tốc độ thích nghi lại rất cao. Điều này phản ánh hai điểm quan trọng: 1. SV vốn đã quen với môi trường công nghệ số nên khả năng tiếp nhận và sử dụng công nghệ mới diễn ra tự nhiên; 2. Sự phổ biến của AI chủ yếu là kết quả của truyền thông xã hội, xu thế công nghệ toàn cầu và nhu cầu thực tế trong học tập chứ không phải do định hướng chính thức từ nhà trường. Điều này tạo ra một nghịch lý: sử dụng nhiều nhưng thiếu hướng dẫn, dẫn đến sử dụng tự phát và rủi ro đạo đức - học thuật. Một mặt, AI là công cụ hỗ trợ mạnh mẽ nhưng lại tiềm ẩn nguy cơ suy giảm tư duy độc lập. Kết quả khảo sát và phỏng vấn cho thấy SV đánh giá rất cao các lợi ích của AI, trong đó lợi ích nổi bật nhất là "tiết kiệm thời gian". Ngoài ra, AI giúp tìm tài liệu nhanh hơn, hỗ trợ viết lách, tạo dàn ý, phân tích nội dung và gợi ý ý tưởng. SV coi AI giống một "trợ lý học thuật 24/7", có khả năng phản hồi tức thời và chi phí thấp. Tuy nhiên, chính vì AI hỗ trợ quá nhiều, SV cũng nhận thức rõ nguy cơ bị phụ thuộc, dẫn đến: giảm khả năng tư duy phản biện, mất dần kỹ năng tạo lập văn bản, suy giảm năng lực tự tìm kiếm và xử lý thông tin, khó phân biệt đâu là ý tưởng của bản thân và đâu là sản phẩm do AI tạo sinh. Hầu hết

người được phỏng vấn đều nhắc đến rủi ro "lạm dụng và lệ thuộc", thậm chí xem đó là "nguy cơ lớn nhất". Các kết quả nghiên cứu cũng cho thấy một bộ phận SV không kiểm chứng thông tin từ AI, điều này có thể dẫn đến việc lan truyền thông tin sai hoặc sai lệch nhận thức. Thêm nữa, có nghịch lý đáng chú ý: SV đánh giá mình hiểu đạo đức AI ở mức trung bình hoặc tốt nhưng lại tồn tại hành vi không phù hợp như: để AI viết bài thay nhưng nghĩ đó không phải gian lận, không trích dẫn AI, sử dụng AI để hoàn thành sản phẩm học tập mà không kiểm chứng nội dung, không phân biệt được giữa hỗ trợ tư duy và thay thế tư duy. Điều này cho thấy SV đang rơi vào trạng thái "ảo tưởng về năng lực đạo đức" (ethical overconfidence). Họ tin rằng mình có thể kiểm soát AI nhưng thực tế đang dùng AI theo cách vượt khỏi vùng an toàn học thuật.

Bên cạnh vấn đề chưa thực sự hiểu đạo đức AI thì *vấn đề liêm chính học thuật* chính là một khoảng trống trong quá trình sử dụng AI của SV. Khi được hỏi *"Bạn có từng trích dẫn hoặc mô tả rõ ràng việc đã dùng AI trong bài nộp cho GV?"*, kết quả khảo sát cho thấy có tới 245 SV (32,8%) chưa bao giờ khai báo, 126 SV (16,9%) hiếm khi thực hiện, 268 SV (35,9%) chỉ thỉnh thoảng mới khai báo, trong khi nhóm luôn trích dẫn đầy đủ chỉ chiếm 14,3% (107 người). Như vậy, phần lớn SV vẫn chưa có thói quen minh bạch hóa việc sử dụng AI - một trong những yêu cầu tối thiểu để đảm bảo tính liêm chính trong học thuật. Điều này cũng đồng nghĩa rằng quy định về trích dẫn AI hiện chưa rõ ràng hoặc chưa được phổ biến rộng rãi trong nhà trường, khiến SV lúng túng trong việc xác định khi nào và bằng cách nào cần khai báo.

Biểu đồ 2: Về việc khai báo đã sử dụng AI để nộp bài cho GV



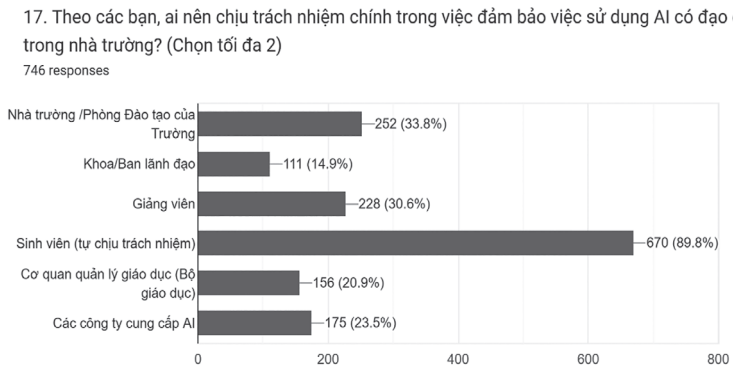
Nguồn: Kết quả khảo sát tháng 11/2025

Về mức độ hiểu biết của SV đối với các *vấn đề pháp lý và đạo đức khi sử dụng AI*. Kết quả khảo sát cho thấy, chỉ có 232 SV (31,1%) cho biết họ hiểu rõ các quy định liên quan đến bản quyền khi đưa dữ liệu vào các công cụ AI; trong khi đó, 397 SV (53,2%) chỉ có hiểu biết ở mức một phần và 117 SV (15,7%) hoàn toàn không có kiến thức về vấn đề này. Như vậy, gần 70% SV chưa được trang bị đầy đủ kiến thức về quyền tác giả và sở hữu trí tuệ trong bối cảnh sử dụng AI, đặc biệt trong khi nhiều công cụ AI hiện nay có điều khoản cho phép thu thập, lưu trữ và tái sử dụng dữ liệu đầu vào. Sự thiếu hụt hiểu biết này tiềm ẩn nguy cơ rò rỉ dữ liệu, mất quyền tác giả hoặc sao chép ngoài ý muốn, nhất là trong môi trường học thuật. Đối với chính sách quyền riêng tư của các công cụ AI, kết quả cho thấy thói quen đọc và hiểu các điều khoản sử dụng của SV còn rất hạn chế. Gần một nửa số SV (48%) cho biết chỉ thỉnh thoảng hoặc hiếm khi đọc các chính sách này, 178 SV (23,8%) thừa nhận chưa bao giờ đọc, trong khi chỉ có 114 SV (15,3%) luôn đọc đầy đủ trước khi sử dụng. Điều này phản ánh một xu hướng phổ biến là chấp nhận điều khoản sử dụng một cách thụ động, thiếu cân nhắc, từ đó làm gia tăng rủi ro liên quan đến việc khai thác và sử dụng dữ liệu cá nhân, đặc biệt trong bối cảnh các hệ thống AI ngày càng được tích hợp sâu vào hệ sinh thái số và các nền tảng học tập trực tuyến.

Liên quan đến quan niệm *về trách nhiệm đạo đức trong sử dụng AI*, kết quả khảo sát cho thấy sự phân tán trong nhận thức của SV về chủ thể chịu trách nhiệm chính. Nổi bật nhất là phương án "SV (tự chịu trách nhiệm)" với 670 lượt lựa chọn, chiếm 89,8%, cao vượt trội so với tất cả các nhóm còn lại. Điều này cho thấy đa số SV đã ý thức khá rõ vai trò chủ thể của bản thân trong việc sử dụng AI một cách có trách nhiệm, thay vì hoàn toàn lệ thuộc vào các cơ chế kiểm soát từ bên ngoài. Kết quả này phản ánh sự

chuyển biến tích cực trong nhận thức đạo đức học thuật, khi người học ngày càng xem việc sử dụng AI là một phần của năng lực cá nhân và đạo đức nghề nghiệp tương lai. Tuy nhiên, bên cạnh trách nhiệm cá nhân, SV vẫn đặt kỳ vọng đáng kể vào các thiết chế giáo dục. Cụ thể, nhà trường/Phòng Đào tạo được 252 SV lựa chọn (33,8%) và GV nhận được 228 lượt chọn (30,6%). Điều này cho thấy SV không xem trách nhiệm đạo đức trong sử dụng AI là vấn đề cá nhân đơn lẻ mà cần được đặt trong một khuôn khổ thể chế, bao gồm quy định, hướng dẫn, chuẩn mực đánh giá và sự đồng hành của đội ngũ cán bộ giảng dạy. Vai trò của GV ở đây không chỉ là người truyền đạt tri thức mà còn là chủ thể định hướng cách thức sử dụng AI phù hợp với chuẩn mực học thuật và mục tiêu đào tạo. Ngoài ra, cơ quan quản lý giáo dục (20,9%) và các công ty cung cấp AI (23,5%) cũng được SV ghi nhận như những chủ thể có trách nhiệm nhất định. Điều này phản ánh nhận thức ngày càng rõ rằng đạo đức AI không chỉ là vấn đề nội bộ của nhà trường mà còn liên quan đến chính sách vĩ mô, khung pháp lý và trách nhiệm xã hội của các doanh nghiệp công nghệ trong việc thiết kế, minh bạch hóa và kiểm soát các công cụ AI. Ngược lại, Khoa/Ban lãnh đạo có tỷ lệ lựa chọn thấp nhất (14,9%), cho thấy SV có xu hướng nhìn nhận trách nhiệm đạo đức AI nghiêng nhiều về cấp quản trị trung tâm (nhà trường), người trực tiếp giảng dạy (GV) và bản thân người học, hơn là cấp trung gian trong cấu trúc tổ chức đại học. Như vậy, kết quả khảo sát cho thấy SV đang hình thành một cách tiếp cận trách nhiệm đa tầng đối với việc sử dụng AI có đạo đức: trong đó trách nhiệm cá nhân của SV là trung tâm, được hỗ trợ bởi khuôn khổ thể chế của nhà trường, sự định hướng của GV cùng với vai trò điều tiết của cơ quan quản lý nhà nước và doanh nghiệp công nghệ. Đây là nền tảng quan trọng để xây dựng các chính sách quản trị AI trong giáo dục theo hướng kết hợp giữa tự chủ - trách nhiệm - và kiểm soát có định hướng.

Biểu đồ 3: Trách nhiệm chính về đảm bảo việc sử dụng AI có đạo đức trong nhà trường



Nguồn: Kết quả khảo sát tháng 11/2025

Khi được hỏi về các giải pháp quản lý AI trong GDĐH thì SV cho rằng xu hướng chung là giáo dục và bồi dưỡng nhận thức cho SV được xếp ở vị trí quan trọng nhất, với tỷ lệ lựa chọn mức "1 - quan trọng nhất" cao hơn rõ rệt so với các giải pháp khác. Điều này khẳng định rằng SV xem nâng cao nhận thức và năng lực sử dụng AI là nền tảng cốt lõi để bảo đảm đạo đức học thuật. Tiếp theo, quy định và chế tài của nhà trường được đánh giá ở mức độ quan trọng cao, phản ánh kỳ vọng của SV vào vai trò thiết lập khung pháp lý và chuẩn mực hành vi của cơ sở giáo dục. Vai trò định hướng chuyên môn của GV cũng được xếp ở nhóm giải pháp quan trọng, cho thấy SV coi GV là cầu nối trực tiếp giữa quy định thể chế và thực tiễn sử dụng AI trong học tập và nghiên cứu. Ngược lại, các giải pháp mang tính kỹ thuật như công cụ phát hiện nội dung do AI tạo ra và hợp tác với nhà cung cấp AI để tăng tính minh bạch có xu hướng được xếp ở mức ít ưu tiên hơn. Điều này không đồng nghĩa với việc SV phủ nhận vai trò của công nghệ mà cho thấy họ xem các giải pháp kỹ thuật là biện pháp hỗ trợ, chỉ thực sự hiệu quả khi được đặt trong một nền tảng giáo dục, quy định và định hướng đạo đức rõ ràng.

3. Kết luận và một số kiến nghị

Nghiên cứu về thực trạng sử dụng AI của SV ĐHQG Hà Nội cho thấy AI, đặc biệt là các công cụ AI tạo sinh đã nhanh chóng trở thành một bộ phận không thể thiếu trong đời sống học thuật của người học. AI mang lại nhiều lợi ích rõ rệt, tuy nhiên, song hành với những lợi ích đó là hàng loạt thách thức mang tính nền tảng đối với đạo đức học thuật, liêm chính khoa học và năng lực tư duy độc lập của SV. Kết quả nghiên cứu cho thấy tồn tại một nghịch lý đáng chú ý: SV đánh giá tương đối cao mức độ hiểu biết của bản thân về đạo đức AI nhưng trên thực tế lại xuất hiện nhiều hành vi tiềm ẩn nguy cơ vi phạm, đồng thời, việc SV tiếp cận AI chủ yếu theo hướng tự phát, thiếu hướng dẫn chính thức từ nhà trường càng làm gia tăng các rủi ro đạo đức và pháp lý trong môi trường học thuật.

Từ những kết quả trên, nghiên cứu cho rằng việc quản trị AI trong GDĐH không nên tiếp cận theo hướng cấm đoán mà cần xây dựng một mô hình quản trị cân bằng giữa đổi mới, trách nhiệm và kiểm soát có định hướng. Trên cơ sở đó, nghiên cứu đề xuất một số kiến nghị sau: 1. Các cơ sở GDĐH cần sớm ban hành khung hướng dẫn chính thức về sử dụng AI trong học tập và nghiên cứu, trong đó làm rõ các khái niệm cốt lõi như "AI hỗ trợ" và "AI làm thay", quy định cụ thể về trích dẫn và khai báo AI cũng như các hành vi được phép và không được phép trong bối cảnh học thuật; 2. Cần tăng cường giáo dục và bồi dưỡng nhận thức về đạo đức AI cho SV, thông qua việc tích hợp các nội dung về AI, đạo đức công nghệ, quyền sở hữu trí tuệ và bảo vệ dữ liệu cá nhân vào chương trình đào tạo, đặc biệt là các học phần đại cương. Đây là giải pháp nền tảng giúp SV hình thành năng lực sử dụng AI một cách có trách nhiệm và bền vững; 3. Vai trò của GV cần được phát huy như một chủ thể định hướng và đồng hành, không chỉ trong việc kiểm soát mà còn trong việc hướng dẫn SV khai thác AI để nâng cao tư duy phản biện, sáng tạo và năng lực học tập suốt đời; 4. Các công cụ kỹ thuật như phần mềm phát hiện nội dung do AI tạo ra nên được sử dụng như biện pháp hỗ trợ, đi kèm với nguyên tắc minh bạch và công bằng, tránh lạm dụng công nghệ kiểm soát thay cho giáo dục đạo đức; 5. Xây dựng văn hóa sử dụng AI có đạo đức trong GDĐH cần sự phối hợp đa tầng giữa SV, nhà trường, cơ quan quản lý giáo dục và các doanh nghiệp công nghệ. Chỉ khi AI được đặt trong một khuôn khổ giáo dục nhân văn và có trách nhiệm, công nghệ này mới thực sự trở thành động lực thúc đẩy chất lượng đào tạo và phát triển con người trong kỷ nguyên số.

Tài liệu tham khảo

[1] Alenezi, M. (2021). *Deep Dive into Digital Transformation in Higher Education Institutions*. Education Sciences, 11(12), 770. https://doi.org/10.3390/educsci11120770.

[2] Bùi Minh Thuận (2025). *Đạo đức của người sử dụng trí tuệ nhân tạo trong nền giáo dục 5.0.* Tạp chí Khoa học Giáo dục kỹ thuật, Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh, JTE, tập 20, số đặc biệt 01, tháng 02, tr. 34-43.

[3] Bùi Trọng Tài, Nguyễn Minh Tuấn (2024). *Nghiên cứu ảnh hưởng của trí tuệ nhân tạo trong giáo dục tới hoạt động học tập của sinh viên*. Tạp chí Giáo dục, 24(10), 6-11.

[4] Lê Anh Vinh, Trần Mỹ Ngọc (2024). *Tác động của trí tuệ nhân tạo (AI) đối với hệ thống giáo dục toàn cầu và giáo dục Việt Nam*. Tạp chí Khoa học Giáo dục Việt Nam, 20 (5), 1-11.

[5] Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

[6] OECD. (2023). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. OECD Publishing. https://www.oecd.org/education/ai-in-education/.

[7] Russel, Stuart and Norvig, Peter (eds.). (2021). *Artificial Intelligence: A Modern Approach*. Fourth edition. Hoboken: Pearson.

[8] UNESCO. (2022). *AI and Education: Guidance for Policy-makers*. United Nations Educational, Scientific and Cultural Organization. https://unesco.org/ark:/48223/pf0000376709.