VNU Journal of Foreign Studies

Journal homepage: https://jfs.ulis.vnu.edu.vn/

THE CURRENT STATE OF GENERATIVE AI USE IN PROJECT-BASED LEARNING IN BUSINESS CHINESE EDUCATION AND A PROPOSED ASSESSMENT CRITERIA FRAMEWORK

Nguyen Thi Ngoc Hien*, Nguyen Thi Hao, Tran Thi Bich Huong

Faculty of Chinese Language and Culture, VNU University of Languages and International Studies, No. 2 Pham Van Dong, Cau Giay, Hanoi, Vietnam

Received 03 June 2026
Revised 10 September 2026; Accepted 23 October 2026

Abstract: With the rapid advancement of Generative Artificial Intelligence (GenAI) in education, the teaching of the Chinese language, particularly Business Chinese, is undergoing profound transformations. This study focuses on students majoring in Chinese Language with an Economic orientation at the University of Languages and International Studies, Vietnam National University, Hanoi (ULIS-VNU). Utilizing the Project-Based Learning (PBL) framework, the research investigates the current application and impact of GenAI on students' project execution. The results from a combined survey and in-depth interview approach indicate that GenAI effectively supports task completion, enhances language proficiency, facilitates access to complex economic concepts, and streamlines business information retrieval. However, the study also identifies several challenges, including excessive AI dependency, limitations in critical thinking and in-depth feedback, a lack of information verification, and issues regarding transparency in AI usage. Based on these findings, the paper proposes the development of a PBL project evaluation framework that integrates AI literacy, critical thinking, and academic integrity requirements, thereby contributing to the refinement of assessment systems in Business Chinese pedagogy within the digital era.

Keywords: Project-Based Learning (PBL), Business Chinese teaching, Generative Artificial Intelligence (GenAI), AI literacy, assessment system

* Corresponding author.
Email address: hienntn0509@gmail.com
https://doi.org/10.63023/2525-2445/jfs.ulis.5887

THỰC TRẠNG ỨNG DỤNG AI TẠO SINH TRONG HỌC TẬP QUA DỰ ÁN TRONG HỌC PHẦN TIẾNG TRUNG KINH TẾ VÀ ĐỀ XUẤT KHUNG TIÊU CHÍ ĐÁNH GIÁ

Nguyễn Thị Ngọc Hiền, Nguyễn Thị Hảo, Trần Thị Bích Hường

Khoa Ngôn ngữ và Văn hóa Trung Quốc, Trường Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội, Số 2 Phạm Văn Đồng, Cầu Giấy, Hà Nội, Việt Nam

Nhận bài ngày 03 tháng 6 năm 2026
Chỉnh sửa ngày 10 tháng 9 năm 2026; Chấp nhận đăng ngày 23 tháng 10 năm 2026

Tóm tắt: Với sự phát triển nhanh chóng của trí tuệ nhân tạo tạo sinh (GenAI) trong giáo dục, việc giảng dạy tiếng Trung, đặc biệt là tiếng Trung kinh tế, đang trải qua những thay đổi sâu sắc. Nghiên cứu này tập trung vào sinh viên chuyên ngành Ngôn ngữ Trung Quốc định hướng kinh tế tại Trường Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội. Dựa trên khung phương pháp học tập qua dự án (PBL), nghiên cứu khảo sát thực trạng ứng dụng và tác động của GenAI trong quá trình thực hiện dự án của sinh viên. Kết quả điều tra kết hợp phỏng vấn sâu cho thấy GenAI hỗ trợ hiệu quả việc hoàn thành nhiệm vụ, nâng cao chất lượng diễn đạt ngôn ngữ, hỗ trợ tiếp cận các khái niệm kinh tế phức tạp và thu thập thông tin kinh doanh. Tuy nhiên, nghiên cứu cũng chỉ ra một số thách thức như: sự phụ thuộc quá mức vào GenAI, hạn chế trong tư duy phản biện và phản hồi chuyên sâu, thiếu kiểm chứng tính chính xác của thông tin, cũng như chưa minh bạch trong việc sử dụng GenAI. Trên cơ sở đó, bài viết đề xuất khung đánh giá sơ bộ đánh giá dự án PBL theo hướng tích hợp năng lực sử dụng AI, tư duy phản biện và yêu cầu về liêm chính học thuật, góp phần hoàn thiện hệ thống đánh giá trong giảng dạy tiếng Trung kinh tế trong bối cảnh kỷ nguyên số.

Từ khóa: Học tập qua dự án (PBL), giảng dạy tiếng Trung kinh tế, trí tuệ nhân tạo tạo sinh (GenAI), năng lực số AI, hệ thống đánh giá

1. Đặt vấn đề

Sự phát triển nhanh chóng của trí tuệ nhân tạo tạo sinh (GenAI) đang làm thay đổi sâu sắc cách dạy và học ngoại ngữ, đặc biệt trong các học phần mang tính định hướng chuyên ngành. Với ưu thế về xử lý ngôn ngữ, mô phỏng tình huống và phản hồi tức thì, GenAI mang lại nhiều hỗ trợ kỹ thuật mới cho giảng dạy và học tập trong các học phần tiếng Trung kinh tế. Trong bối cảnh đó, việc ứng dụng GenAI trong các dự án học tập đã mở ra những cơ hội mới cho việc nâng cao hiệu quả đào tạo.

Học tập qua dự án (PBL) là mô hình dạy học lấy người học làm trung tâm, nhấn mạnh nhiệm vụ thực tiễn và phát triển năng lực tổng hợp, được ứng dụng rộng rãi trong các học phần kinh tế và quản lý. Việc tích hợp GenAI vào các dự án PBL góp phần nâng cao hiệu quả triển khai thông qua việc cung cấp tài liệu học tập song ngữ phù hợp với trình độ và nhu cầu chuyên môn của sinh viên (SV), mô phỏng các tình huống kinh doanh thực tế để tăng cơ hội thực hành, đồng thời đưa ra phản hồi tức thì nhằm hỗ trợ cá nhân hóa việc dạy học và đánh giá.

Tuy nhiên, sự tiện lợi của GenAI cũng đặt ra thách thức mới. Việc phụ thuộc quá mức vào GenAI có thể làm suy giảm khả năng phân tích, suy luận, phản tư và tư duy phản biện của SV, ảnh hưởng đến giá trị cốt lõi của PBL. Bên cạnh đó, khi GenAI tham gia sâu vào quá trình tạo ra sản phẩm dự án, phương thức đánh giá truyền thống thiên về kết quả đầu ra trở nên kém phù hợp do khó phân biệt giữa nỗ lực của SV và đóng góp của GenAI. Vì vậy, giảng viên (GV)

cần điều chỉnh tiêu chuẩn đánh giá theo hướng coi trọng cả quá trình và kết quả học tập.

Hiện nay, các nghiên cứu về PBL trong giảng dạy tiếng Trung, đặc biệt trong các học phần tiếng Trung kinh tế, vẫn còn hạn chế; các nghiên cứu về thực hành và đánh giá PBL trong bối cảnh GenAI can thiệp sâu thậm chí mới ở giai đoạn khởi đầu. Các nghiên cứu hiện có chủ yếu tập trung vào vai trò hỗ trợ ngôn ngữ của GenAI, trong khi nghiên cứu về hành vi sử dụng GenAI của SV và tác động của nó đến hệ thống đánh giá trong PBL đối với các học phần tiếng Trung kinh tế vẫn chưa được quan tâm đầy đủ.

Trong phạm vi bài viết này, chúng tôi khảo sát phương thức SV sử dụng công cụ GenAI trong quá trình học tập qua dự án và phân tích tác động của nó đối với hiệu quả học tập trong khung giảng dạy PBL của các học phần tiếng Trung kinh tế. Trên cơ sở đó, bài viết đề xuất đưa hành vi sử dụng GenAI và yêu cầu về năng lực số AI (AI Literacy) vào hệ thống đánh giá kết quả dự án một cách hợp lý. Nghiên cứu không chỉ phản ánh thực tiễn ứng dụng GenAI trong học tập dự án tiếng Trung kinh tế mà còn cung cấp góc nhìn tham khảo cho việc tối ưu hóa tiêu chuẩn đánh giá trong đào tạo tiếng Trung trong bối cảnh GenAI phát triển nhanh chóng.

Bài viết tập trung trả lời các câu hỏi nghiên cứu sau:

(1) SV sử dụng GenAI trong các dự án PBL tiếng Trung kinh tế như thế nào?

(2) Việc sử dụng GenAI có mối liên hệ như thế nào với quá trình thực hiện dự án và hiệu quả học tập của SV?

(3) Trong bối cảnh GenAI can thiệp sâu, hệ thống đánh giá dự án PBL cần được điều chỉnh như thế nào?

2. Cơ sở lý luận

2.1. Đặc điểm của tiếng Trung kinh tế và sự phân định với giảng dạy kiến thức chuyên ngành

Dưới góc nhìn của lý thuyết Ngôn ngữ phục vụ mục đích chuyên biệt (Language for Specific Purposes – LSP) (Dudley-Evans & St John, 1998), tiếng Trung kinh tế có những đặc thù riêng so với tiếng Trung tổng hợp và giảng dạy kiến thức kinh tế. Nếu tiếng Trung tổng hợp chú trọng ngữ âm, từ vựng và ngữ pháp cơ bản, còn các học phần kinh tế tập trung vào mô hình, quy luật vận hành và quyết định chiến lược, thì tiếng Trung kinh tế lấy năng lực sử dụng ngôn ngữ chuyên ngành trong bối cảnh kinh doanh làm mục tiêu cốt lõi. Trong các dự án học tập, kiến thức kinh tế, chẳng hạn phân tích chiến lược marketing của doanh nghiệp, đóng vai trò là nội dung và ngữ cảnh để SV rèn luyện khả năng diễn đạt, lập luận và trình bày bằng tiếng Trung. Vì vậy, đánh giá cần chú trọng độ chính xác của thuật ngữ, tính logic trong diễn đạt và sự phù hợp của văn phong thương mại, thay vì chỉ đánh giá kiến thức kinh tế hay năng lực ngôn ngữ chung.

Theo phương pháp dạy học lồng ghép nội dung và ngôn ngữ (Content and Language Integrated Learning – CLIL) (Coyle và cộng sự, 2010), SV phải đồng thời tiếp thu kiến thức kinh tế phức tạp và diễn đạt chính xác bằng ngoại ngữ, tạo nên tải nhận thức kép. Đặc thù này khiến việc ứng dụng GenAI vào PBL tiếng Trung kinh tế khác với các dự án tiếng Trung tổng hợp. Nếu trong tiếng Trung tổng hợp, GenAI chủ yếu hỗ trợ tra cứu từ vựng, dịch thuật và chỉnh sửa diễn đạt, thì trong tiếng Trung kinh tế, công cụ còn có thể cung cấp dữ liệu thị trường, giải thích khái niệm quản trị và hỗ trợ chuyển hóa ý tưởng kinh doanh thành báo cáo thương mại chuẩn mực.

Tuy nhiên, GenAI tham gia sâu vào cả quá trình xử lý kiến thức chuyên môn và biểu đạt ngôn ngữ cũng làm gia tăng nguy cơ SV phụ thuộc vào AI, thay vì tự hình thành lập luận và làm chủ nội dung. Đây là cơ sở đặt ra yêu cầu xây dựng hệ thống tiêu chí đánh giá PBL chuyên biệt,

vừa xem xét chất lượng sản phẩm, vừa đánh giá năng lực tư duy độc lập, sử dụng và kiểm soát GenAI của người học.

2.2. Học tập qua dự án (PBL)

Trong nghiên cứu nền tảng năm 1918, Kilpatrick đã định nghĩa PBL là "hoạt động thực hành mang tính tích hợp do người học tự khởi xướng, nhằm đạt được mục tiêu cụ thể trong bối cảnh xã hội thực tiễn" (Kilpatrick, 1918, tr. 1).

Các trụ cột cốt lõi của mô hình này bao gồm: tính tự chủ của người học, định hướng mục tiêu, bối cảnh thực tế và sự tích hợp đa ngành. PBL được thiết kế qua các bước chính như: xác định mục tiêu học tập, xây dựng câu hỏi định hướng, lập kế hoạch dự án và gắn với bối cảnh thực tế. Đồng thời, cần trao quyền lựa chọn cho người học, kết hợp đánh giá - phản hồi - phản tư liên tục trong suốt quá trình, và hướng tới sản phẩm cuối cùng được trình bày công khai.

2.3. Khung năng lực số AI và liêm chính học thuật

Ng và cộng sự (2021) đề xuất bốn phương diện của năng lực số AI, gồm "know and understand, use, evaluate, and ethical issues" (tr. 504), tương ứng với năng lực nhận thức và hiểu biết, sử dụng, đánh giá và các vấn đề đạo đức trong sử dụng AI. Trong bối cảnh GenAI, nghiên cứu này kết hợp khung năng lực trên với sáu giá trị cốt lõi về liêm chính học thuật của ICAI bao gồm: trung thực, tin cậy, công bằng, tôn trọng, trách nhiệm và dũng cảm (ICAI, 2021, tr. 4). Sự kết hợp này nhằm xác định ranh giới cộng tác giữa người và máy, từ đó đảm bảo tính xác thực của quá trình khám phá trong PBL và xác định rõ trách nhiệm của người học.

2.4. Khung TPACK

Theo Mishra và Koehler (2006), TPACK là khung lý thuyết mô tả sự tích hợp giữa ba miền kiến thức cốt lõi gồm: kiến thức chuyên môn (Content Knowledge – CK), kiến thức sư phạm (Pedagogical Knowledge – PK) và kiến thức công nghệ (Technological Knowledge – TK), nhằm giải thích năng lực của GV trong việc tích hợp công nghệ vào dạy học (tr. 1028).

Mặc dù TPACK ban đầu được phát triển để mô tả hệ thống kiến thức của GV, nghiên cứu này cho rằng logic cốt lõi của khung - sự tích hợp giữa kiến thức chuyên môn, kiến thức về phương pháp và công nghệ - vẫn có giá trị tham chiếu khi xem xét năng lực của người học trong môi trường học tập có công nghệ. Trong PBL, SV cũng phải đồng thời huy động kiến thức chuyên môn, phương pháp thực hiện dự án và công nghệ để giải quyết nhiệm vụ, tạo ra sản phẩm và bảo vệ kết quả. Vì vậy, nghiên cứu không xem TPACK như một thang đo năng lực GV được chuyển nguyên dạng sang SV, mà vận dụng logic tích hợp của TPACK ở cấp độ người học.

Cụ thể, CK được hiểu là năng lực vận dụng kiến thức kinh tế và hệ thống thuật ngữ tiếng Trung kinh tế; TK là khả năng lựa chọn, sử dụng và điều chỉnh các công cụ GenAI nhằm phục vụ nhiệm vụ học tập; còn PK được diễn giải theo nghĩa mở rộng là kiến thức về phương pháp học tập và tổ chức dự án, bao gồm: lập kế hoạch, xử lý thông tin, phối hợp nhóm và triển khai nhiệm vụ. Trong bối cảnh nghiên cứu này, TPACK được sử dụng như khung kết nối về mặt lý luận giữa PBL và năng lực số AI.

2.5. Khung lí thuyết tích hợp và giả thuyết nghiên cứu

Sự kết hợp giữa ba khung lý thuyết PBL, năng lực số AI và TPACK tạo nên một hệ thống logic tương hỗ chặt chẽ: PBL thiết lập bối cảnh thực hành và nền tảng phương pháp luận, giúp người học phát huy tính tự chủ, khả năng khám phá và phản tư thông qua các đặc trưng thực tiễn và cộng tác. TPACK đóng vai trò là khung thiết kế tích hợp, giúp GV kết nối AI với giảng dạy chuyên ngành, đồng thời cung cấp khung tham chiếu để phân tích sự tích hợp cho

năng lực ứng dụng tổng hợp của SV. Trong tiến trình đó, khung năng lực số AI và liêm chính học thuật giữ vai trò quy chuẩn hóa, thiết lập các giới hạn đạo đức và nâng cao chiều sâu trong tương tác người – máy.

Trên cơ sở sự giao thoa của ba khung lý luận, nghiên cứu xác định năm phương diện sử dụng GenAI trong PBL, gồm: năng lực Prompt (PR), kiểm chứng thông tin (VE), tính minh bạch và trách nhiệm (IN), đánh giá và phản hồi (FB), và hỗ trợ nhận thức chuyên môn (SU). Các phương diện này được sử dụng để kiểm định mối quan hệ với hiệu quả học tập thực chất (EFF).Trên cơ sở đó, nghiên cứu đề xuất các giả thuyết sau nhằm kiểm định tác động của các nhân tố thuộc năng lực số AI đến hiệu quả học tập:

H_1 : Năng lực Prompt có mối quan hệ tích cực với hiệu quả học tập thực chất.

H_2 : Năng lực kiểm chứng thông tin có mối quan hệ tích cực với hiệu quả học tập thực chất.

H_3 : Tính minh bạch và trách nhiệm trong sử dụng GenAI có mối quan hệ tích cực với hiệu quả học tập thực chất.

H_4 : Việc sử dụng GenAI để đánh giá và phản hồi có mối quan hệ tích cực với hiệu quả học tập thực chất.

H_5 : Việc sử dụng GenAI để hỗ trợ nhận thức chuyên môn có mối quan hệ tích cực với hiệu quả học tập thực chất.

3. Phương pháp nghiên cứu

Nghiên cứu áp dụng thiết kế tuần tự giải thích (explanatory sequential design) thuộc phương pháp nghiên cứu hỗn hợp (mixed methods), kết hợp dữ liệu định lượng từ bảng hỏi và dữ liệu định tính từ phỏng vấn bán cấu trúc. Giai đoạn định lượng nhằm khảo sát thực trạng sử dụng GenAI, mức độ sử dụng và thái độ của SV trong quá trình học tập theo dự án (PBL). Kết quả định lượng được sử dụng làm cơ sở lựa chọn đối tượng và nội dung cho giai đoạn định tính, qua đó giải thích sâu hơn trải nghiệm, cách thức sử dụng và những khó khăn của SV khi tương tác với GenAI.

Nghiên cứu lựa chọn dự án "Phân tích chiến lược marketing của một doanh nghiệp" trong học phần Quản trị doanh nghiệp làm trường hợp nghiên cứu. Dự án được triển khai trong chương trình đào tạo cử nhân Ngôn ngữ Trung Quốc định hướng kinh tế tại Trường Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội, trong đó, nội dung chiến lược marketing tập trung vào bốn phương diện: sản phẩm (Product), giá (Price), phân phối (Place) và xúc tiến (Promotion). Khảo sát được tiến hành vào cuối học phần, sau khi SV hoàn thành báo cáo dự án.

3.1. Phương pháp điều tra bằng bảng hỏi và đo lường biến số

3.1.1. Thiết kế công cụ thu thập dữ liệu

Bảng hỏi định lượng sử dụng thang Likert 5 mức, gồm ba phần: (1) thông tin cá nhân và thói quen sử dụng GenAI; (2) năng lực số AI, kế thừa và điều chỉnh theo Ng và cộng sự (2021), gồm: nhận thức, ứng dụng, đánh giá và đạo đức; (3) hiệu quả sử dụng GenAI trong dự án, xây dựng dựa trên TCK và TPK thuộc khung TPACK của Mishra và Koehler (2006), có điều chỉnh theo bối cảnh tiếng Trung kinh tế. Các biến quan sát được kế thừa, điều chỉnh từ thang đo công bố và bổ sung theo mục tiêu nghiên cứu, với cơ sở lý luận từ PBL, TPACK và năng lực số AI.

Phỏng vấn bán cấu trúc dành cho SV và GV tập trung vào quá trình thực hiện dự án, cách ứng dụng GenAI, khó khăn khi tương tác và đề xuất cải tiến đánh giá dự án.

3.1.2. Giá trị nội dung và thử nghiệm công cụ

Bảng hỏi được thử nghiệm (pilot test) với một nhóm SV trước khi khảo sát chính thức nhằm kiểm tra mức độ rõ ràng của câu hỏi và sự phù hợp của các thuật ngữ chuyên ngành. Bản dự thảo sau đó được hai GV có chuyên môn về phương pháp giảng dạy tiếng Trung và PBL góp ý trước khi hoàn thiện.

3.1.3. Phương pháp chọn mẫu và mẫu nghiên cứu

Nghiên cứu sử dụng phương pháp chọn mẫu thuận tiện kết hợp chọn mẫu chủ đích. Tiêu chí lựa chọn gồm: (1) là SV khóa QH2022 hoặc QH2023 thuộc chương trình Ngôn ngữ Trung Quốc định hướng kinh tế; (2) trực tiếp tham gia dự án "Phân tích chiến lược marketing 4P" trong học phần Quản trị doanh nghiệp.

Bảng hỏi được gửi tới 153 SV, thu về 90 phản hồi, trong đó, 85 phiếu hợp lệ sau khi loại bỏ các trường hợp thiếu thông tin, thời gian hoàn thành quá ngắn hoặc có biểu hiện trả lời thiếu phân biệt. Tỷ lệ phản hồi hợp lệ đạt 55,6%. Mẫu gồm 80 nữ và 5 nam, 100% SV đạt trình độ tiếng Trung tương đương bậc 5 theo Khung năng lực ngoại ngữ 6 bậc dùng cho Việt Nam (VSTEP) trở lên.

3.1.4. Phân loại và đo lường biến số

Bảng hỏi gồm 22 câu, trong đó, câu 1–6 là các câu hỏi tìm hiểu thực trạng sử dụng GenAI và câu 7–22 sử dụng thang đo Likert 5 mức. Các biến được phân loại như sau:

Bảng 1

Phân loại nội dung đo lường trong bảng hỏi

Nhóm	Mã biến	Nội dung đo lường	Câu hỏi
Biến độc lập	PR	Năng lực thiết kế và điều chỉnh Prompt	Q7–Q8
	VE	Năng lực kiểm chứng thông tin từ GenAI	Q9–Q10
	IN	Trách nhiệm và minh bạch trong sử dụng GenAI	Q11–Q12
	FB	Sử dụng GenAI cho đánh giá và phản hồi	Q13–Q15
	SU	Sử dụng GenAI hỗ trợ nhận thức chuyên môn	Q16–Q17
Biến phụ thuộc	EFF	Hiệu quả học tập và tiến bộ thực chất	Q18–Q20
Biến mô tả bổ sung	ASSESS	Quan điểm về đưa AI vào đánh giá quá trình	Q21–Q22

ASSESS được sử dụng để mô tả xu hướng của SV đối với việc đưa tiêu chí sử dụng AI vào đánh giá và không được đưa vào mô hình kiểm định giả thuyết.

3.2. Phương pháp phỏng vấn

3.2.1. Phỏng vấn SV

Sau khi hoàn thành phân tích định lượng, nghiên cứu lựa chọn 7 SV từ nhóm khảo sát để phỏng vấn bán cấu trúc. Việc lựa chọn dựa trên mức độ sử dụng GenAI khác nhau và sự đa dạng trong kết quả khảo sát, nhằm bổ sung các góc nhìn khác nhau về quá trình sử dụng GenAI.

Các cuộc phỏng vấn được thực hiện theo hình thức một - một qua Zoom hoặc trực tiếp, có ghi âm với sự đồng ý của người tham gia, mỗi cuộc kéo dài khoảng 45 phút. Nội dung tập trung vào trải nghiệm thực tế, cách thức tương tác với GenAI, những khó khăn và cách SV đánh giá hiệu quả của đầu ra do GenAI hỗ trợ.

3.2.2. Phỏng vấn GV

Hai GV có chuyên môn về kinh tế và kinh nghiệm triển khai PBL được mời tham gia phỏng vấn bán cấu trúc. Nội dung tập trung vào thực trạng ứng dụng GenAI, những vấn đề trong tổ chức và đánh giá dự án, cũng như quan điểm của GV về các tiêu chí cần xem xét khi đánh giá dự án PBL có sử dụng GenAI.

3.3. Phương pháp phân tích và tích hợp dữ liệu

3.3.1. Phân tích dữ liệu định lượng

Dữ liệu định lượng được mã hóa, làm sạch và xử lý bằng Microsoft Excel. Các phương pháp phân tích gồm: (1) *Xử lý thống kê mô tả (Descriptive Statistics)*: Dữ liệu được tính toán dựa trên tần suất và tỷ lệ phần trăm (%) để xây dựng dữ liệu tổng thể về mức độ đồng thuận của người học đối với mô hình PBL và thực trạng ứng dụng GenAI; (2) *Kiểm định độ tin cậy thang đo (Cronbach's Alpha):* Đánh giá tính nhất quán nội tại của các biến số trong bảng hỏi; (3) *Phân tích tương quan Pearson:* xác định mối quan hệ giữa các biến độc lập và EFF; (4) *Hồi quy tuyến tính đa biến (Multiple Linear Regression):* kiểm định mối quan hệ giữa năm phương diện sử dụng GenAI (PR, VE, IN, FB, SU) và EFF, qua đó kiểm định các giả thuyết H1–H5.

3.3.2. Phân tích dữ liệu định tính

Dữ liệu phỏng vấn được xử lý theo quy trình mã hóa theo chủ đề (thematic coding). Trước hết, các bản ghi phỏng vấn được đọc nhiều lần để xác định những ý nghĩa và nội dung liên quan đến câu hỏi nghiên cứu, sau đó tiến hành mã hóa ban đầu và nhóm các mã có nội dung tương đồng thành các chủ đề. Các chủ đề được đối chiếu với dữ liệu khảo sát định lượng nhằm tìm kiếm sự tương đồng, bổ sung hoặc khác biệt giữa hai nguồn dữ liệu. Cuối cùng, các mã và chủ đề được rà soát lại theo toàn bộ dữ liệu để bảo đảm tính nhất quán trước khi sử dụng cho phân tích và diễn giải kết quả.

3.3.3. Tích hợp và đối chiếu dữ liệu

Theo thiết kế tuần tự giải thích, kết quả định tính được sử dụng để bổ sung và giải thích các kết quả định lượng. Việc đối chiếu hai nguồn dữ liệu giúp làm rõ cách thức sử dụng GenAI, những khác biệt giữa nhận thức và thực hành của SV, đồng thời cung cấp cơ sở thực tiễn cho việc đề xuất khung tiêu chí đánh giá dự án PBL có ứng dụng GenAI.

4. Phân tích dữ liệu và kết quả nghiên cứu

4.1. Thực trạng sử dụng GenAI trong dự án của SV

4.1.1. Khái quát tình hình sử dụng GenAI trong dự án

Q1: *Trong dự án PBL lần này, bạn có sử dụng GenAI không?*

Q2: *Bạn đã sử dụng GenAI vào những giai đoạn nào của dự án?*

Kết quả khảo sát cho thấy 100% SV được hỏi đã sử dụng GenAI trong quá trình thực hiện dự án, trong đó, gần 70% thuộc nhóm sử dụng với tần suất cao. GenAI được sử dụng chủ yếu cho phân tích marketing, xây dựng khung báo cáo và trau chuốt nội dung; 34% SV sử dụng công cụ này trong xác định đề tài và thu thập thông tin thị trường, trong khi 39% sử dụng để soạn thảo báo cáo. Có thể thấy, GenAI đã được sử dụng ở nhiều giai đoạn của dự án, tuy nhiên, mức độ sử dụng khác nhau giữa các nhiệm vụ, trong đó hoạt động phân tích và xử lý nội dung được sử dụng phổ biến hơn.

Q3: *Những công cụ GenAI nào bạn thường xuyên sử dụng?*

Q4: *Bạn có thể dễ dàng xác định loại công cụ GenAI cần thiết cho từng giai đoạn của dự án marketing không?*

Về lựa chọn công cụ, ChatGPT được sử dụng nhiều nhất (94%), tiếp theo là Gemini (74%) và DeepSeek (58%); các công cụ khác có tỷ lệ sử dụng thấp hơn. Phỏng vấn cho thấy SV có xu hướng ưu tiên các công cụ miễn phí, cho thấy chi phí là một trong những yếu tố được cân nhắc khi lựa chọn công cụ. Đáng chú ý, chỉ 29,4% SV cho biết có thể xác định công cụ phù hợp với từng giai đoạn của dự án, trong khi 70,6% chưa có tiêu chí lựa chọn rõ ràng. Kết quả cho thấy SV đã có mức độ tiếp cận GenAI cao nhưng khả năng lựa chọn công cụ theo đặc điểm nhiệm vụ còn hạn chế.

Q5: *Khi viết báo cáo phân tích chiến lược marketing doanh nghiệp, nội dung do GenAI tạo ra chiếm khoảng bao nhiêu ?*

Q6: *Sau khi GenAI đưa ra kết quả, bạn thường sẽ làm gì?*

Mức độ sử dụng GenAI trong soạn thảo báo cáo cũng tương đối cao: 9,4% SV cho biết trên 80% nội dung do GenAI tạo ra, 35,3% cho biết tỷ lệ này khoảng 60–70%, trong khi chỉ 14,1% kiểm soát tỷ lệ nội dung AI ở mức dưới 30%. Về hậu xử lý, 32,9% SV chỉ chỉnh sửa khoảng 20–30% nội dung do AI cung cấp, 37,6% chỉnh sửa khoảng 50%, và tỷ lệ chỉnh sửa ở mức 60–70% chưa đạt một nửa.

Các kết quả trên cho thấy mức độ tham gia của GenAI vào quá trình tạo sản phẩm khá lớn, đồng thời cho thấy sự khác biệt về mức độ chủ động xử lý đầu ra AI giữa các SV. Đây là cơ sở để tiếp tục xem xét năng lực sử dụng, kiểm chứng và minh bạch hóa quá trình tương tác với GenAI ở các phân tích tiếp theo.

4.1.2. Thực trạng năng lực số AI trong dự án

4.1.2.1. Năng lực ứng dụng câu lệnh (Prompt)

Bảng 2

Kết quả khảo sát năng lực ứng dụng câu lệnh (Prompt) của SV

	Hoàn toàn không đồng ý		Không đồng ý		Trung lập		Đồng ý		Hoàn toàn đồng ý	
Q7	4	4,7%	4	4,7%	25	29,4%	45	52,9%	7	8,3%
Q8	2	2,4%	3	3,5%	11	12,9%	55	64,7%	14	16,5%

**** Q7:*** *Tôi biết cách soạn thảo các câu lệnh (prompts) chi tiết và có cấu trúc rõ ràng để GenAI có thể hiểu chính xác bối cảnh của doanh nghiệp mà tôi đang nghiên cứu.*

Q8: *Khi GenAI đưa ra câu trả lời không thỏa đáng, tôi biết cách điều chỉnh câu lệnh hoặc cung cấp thêm dữ liệu để nhận được kết quả tốt hơn.*

Dữ liệu cho thấy SV rất tự tin về khả năng sử dụng câu lệnh cũng như điều chỉnh câu lệnh để đạt kết quả mong muốn, song phỏng vấn chuyên sâu lại phản ánh một thực tế trái ngược. Đa số SV được phỏng vấn phản ánh cách thức truy vấn tương đối đơn giản, ví dụ: chưa biết cung cấp nguồn tài liệu tham khảo đầu vào cho GenAI, cùng với những câu lệnh chung chung, chưa sát với nội dung lý thuyết của môn học. Chẳng hạn, một ý kiến tiêu biểu phản ánh thói quen đặt câu lệnh khá sơ sài khi phân tích chiến lược doanh nghiệp: *"Em hỏi ChatGPT ví dụ chiến lược về giá của Vinamilk là gì? Nếu nó cho kết quả em thấy có vẻ chưa đúng thì sẽ tiếp tục hỏi, như là chiến lược giá xâm nhập là gì?"* (SV2).

4.1.2.2. Năng lực kiểm chứng thông tin AI

Bảng 3

Kết quả khảo sát về chiều sâu và phương thức kiểm chứng thông tin AI của SV

	Hoàn toàn không đồng ý		Không đồng ý		Trung lập		Đồng ý		Hoàn toàn đồng ý	
Q9	6	7,1%	0	0%	9	10,6%	42	49,4%	28	32,9%
Q10	4	4,7%	1	1,2%	9	10,6%	43	50,6%	28	32,9%

**** Q9:** Tôi không trực tiếp chấp nhận kết quả do GenAI đưa ra mà luôn kiểm tra tính chính xác của chúng.*

***Q10:** Đối với các thông tin chuyên sâu từ GenAI, tôi luôn đối chiếu với các kênh chính thống khác để xác thực.*

Đa số SV cho biết có thực hiện kiểm chứng thông tin do GenAI cung cấp. Tỷ lệ đồng ý hoặc hoàn toàn đồng ý đạt 82,3% đối với Q9 và 83,5% đối với Q10. Tuy nhiên, dữ liệu phỏng vấn cho thấy việc kiểm chứng ở một bộ phận SV còn mang tính hình thức; một số SV chưa có tiêu chí rõ ràng để phân biệt nguồn thông tin phổ biến với nguồn có tính chuyên môn và thẩm quyền.

4.1.2.3. Trách nhiệm học thuật và tính minh bạch trong sử dụng AI

Bảng 4

Kết quả khảo sát về trách nhiệm học thuật và tính minh bạch trong sử dụng GenAI của SV

	Hoàn toàn không đồng ý		Không đồng ý		Trung lập		Đồng ý		Hoàn toàn đồng ý	
Q11	1	1,2%	6	7,1%	38	44,7%	35	41,1%	5	5,9%
Q12	3	3,5%	11	12,9%	26	30,6%	33	38,8%	12	14,2%

** **Q11:** Tôi tự tin chịu trách nhiệm 100% về tính chính xác của các thông tin trong báo cáo, ngay cả khi nội dung đó do GenAI gợi ý.*

***Q12:** Tôi sẽ đánh dấu/ghi chú rõ ràng những nội dung nào trong dự án được tạo ra hoặc được hoàn thiện với sự hỗ trợ của GenAI.*

Về tính minh bạch, khảo sát cho thấy một nghịch lý đáng chú ý: dù đa số SV sẵn sàng chịu trách nhiệm về tính xác thực của nội dung, nhiều người vẫn dè dặt trong việc công khai mức độ can thiệp của GenAI. Điều này cho thấy một bộ phận SV chưa phân biệt rõ ranh giới giữa hỗ trợ, cải biên và tạo nội dung trực tiếp từ GenAI. Kết quả phản ánh nhu cầu cấp thiết phải làm rõ quy chuẩn học thuật và tích hợp tiêu chí minh bạch trong sử dụng GenAI vào hệ thống đánh giá nhằm đảm bảo liêm chính học thuật trong kỷ nguyên GenAI.

4.1.3. Hiệu quả của GenAI trong dự án

4.1.3.1. Sử dụng GenAI cho đánh giá và phản hồi

Bảng 5

Kết quả khảo sát thực trạng ứng dụng GenAI như một công cụ đánh giá và phản hồi

	Hoàn toàn không đồng ý		Không đồng ý		Trung lập		Đồng ý		Hoàn toàn đồng ý	
Q13	6	7%	6	7%	25	29,5%	40	47,1%	8	9,4%
Q14	2	2,3%	7	8,2%	23	27,2%	42	49,4%	11	12,9%
Q15	3	3,5%	8	9,4%	25	29,5%	42	49,4%	7	8,2%

** **Q13:** Tôi thường yêu cầu GenAI tranh luận/mô phỏng để phát hiện các điểm yếu trong chiến lược marketing của mình.*

***Q14:** Tôi sử dụng GenAI để nhận phản hồi và thường xuyên chỉnh sửa nội dung dự án dựa trên những phản hồi đó.*

***Q15:** Tôi sử dụng GenAI để hỗ trợ đưa ra các đánh giá mang tính xây dựng cho kết quả của các thành viên trong nhóm hoặc các nhóm khác.*

Kết quả cho thấy tỷ lệ SV đồng ý hoặc hoàn toàn đồng ý với Q13, Q14 và Q15 lần lượt là 56,5%, 62,3% và 57,6%. Như vậy, GenAI đã được một bộ phận đáng kể SV sử dụng cho tranh luận, phản hồi và đánh giá trong quá trình thực hiện dự án. Phỏng vấn cho thấy hoạt động này chủ yếu tập trung vào việc phát hiện lỗi diễn đạt, chuẩn hóa thuật ngữ và điều chỉnh nội dung theo yêu cầu của nhiệm vụ; mức độ sử dụng GenAI để phản biện sâu về lập luận chuyên môn còn hạn chế. Điển hình như chia sẻ của một SV về thói quen sử dụng GenAI để nghiệm thu nội dung ở mức độ bề mặt: *"Em chưa bao giờ dùng AI để nhận xét bài của các bạn khác, em chủ yếu dùng để tìm kiếm thông tin phần mình phụ trách. Sau khi viết xong em sẽ gửi lại bảo nó nhận xét lần cuối xem đã ok (ổn) chưa?"* (SV5).

4.1.3.2. Sử dụng GenAI hỗ trợ ngôn ngữ và nhận thức chuyên môn

Bảng 6

Kết quả khảo sát vai trò hỗ trợ của GenAI trong trau chuốt ngôn ngữ và nhận thức chuyên môn

	Hoàn toàn không đồng ý		Không đồng ý		Trung lập		Đồng ý		Hoàn toàn đồng ý	
Q16	2	2,3%	1	1,2%	18	21,2%	44	51,8%	20	23,5%
Q17	2	2,3%	3	3,5%	14	16,5%	41	48,2%	25	29,4%

** **Q16:** Tôi thường xuyên yêu cầu GenAI chỉnh sửa văn phong hoặc cách diễn đạt ngôn ngữ để báo cáo dự án mang tính chuyên nghiệp hơn.*

***Q17:** GenAI giúp tôi hiểu các lý thuyết kinh tế phức tạp hoặc các thuật ngữ chuyên ngành.*

Tỷ lệ SV đồng ý hoặc hoàn toàn đồng ý đạt 75,3% đối với Q16 và 77,6% đối với Q17. Kết quả cho thấy GenAI được sử dụng phổ biến cho cả việc trau chuốt ngôn ngữ chuyên ngành và hỗ trợ hiểu nội dung kinh tế. Phỏng vấn cũng cho thấy đây là một trong những mục đích sử dụng GenAI nổi bật trong quá trình thực hiện dự án.

4.1.4. Nhận thức về vai trò hỗ trợ của GenAI đối với dự án

Bảng 7

Kết quả khảo sát về niềm tin năng lực AI và lợi thế nghề nghiệp

	Hoàn toàn không đồng ý		Không đồng ý		Trung lập		Đồng ý		Hoàn toàn đồng ý	
Q18	4	4,7%	1	1,2%	15	17,6%	44	51,8%	21	24,7%
Q19	3	3,5%	2	2,4%	23	27,1%	48	56,4%	9	10,6%
Q20	2	2,4%	0	0%	13	15,3%	54	63,5%	16	18,8%

**** Q18:** GenAI giúp tôi tiết kiệm thời gian, đồng thời giúp tôi hiểu sâu hơn nội dung bài học.*

***Q19:** Thông qua dự án có sự hỗ trợ của GenAI, tôi cảm thấy năng lực tư duy và diễn đạt tiếng Trung kinh tế của mình có sự tiến bộ thực chất.*

***Q20:** Tôi tin rằng những kỹ năng sử dụng GenAI học được trong dự án này sẽ trở thành lợi thế lớn khi tôi làm việc tại doanh nghiệp sau khi tốt nghiệp.*

Kết quả cho thấy tỷ lệ SV đồng ý hoặc hoàn toàn đồng ý với Q18, Q19 và Q20 lần lượt là 76,5%, 67,0% và 82,3%. Nhìn chung, SV đánh giá tích cực vai trò của GenAI đối với hiệu quả học tập và lợi thế nghề nghiệp. Tuy nhiên, phỏng vấn cho thấy một số SV lo ngại việc phụ thuộc quá mức vào GenAI có thể làm giảm cơ hội tự xây dựng ý tưởng và rèn luyện khả năng diễn đạt.

4.1.5. Quan điểm về tiêu chí đánh giá dự án

Bảng 8

Kết quả khảo sát mức độ đồng thuận đối với đánh giá quá trình và các chiều đánh giá AI

	Hoàn toàn không đồng ý		Không đồng ý		Trung lập		Đồng ý		Hoàn toàn đồng ý	
Q21	3	3,5%	3	3,5%	18	21,1%	47	55,4%	14	16,5%
Q22	4	4,7%	1	1,2%	19	22,3%	49	57,7%	12	14,1%

**** Q21:** Khi đánh giá dự án, GV nên chú trọng hơn vào quá trình thực hiện của SV thay vì chỉ tập trung vào kết quả cuối cùng.*

***Q22:** Tôi cho rằng tiêu chuẩn đánh giá PBL nên bao gồm các chiều đánh giá liên quan đến việc sử dụng GenAI.*

Kết quả cho thấy 71,9% SV đồng ý hoặc hoàn toàn đồng ý rằng GV nên chú trọng đánh giá quá trình thực hiện dự án. Đối với việc đưa các tiêu chí liên quan đến sử dụng GenAI vào đánh giá, tỷ lệ đồng ý hoặc hoàn toàn đồng ý đạt 71,8%. Kết quả này cho thấy SV nhìn chung ủng hộ việc mở rộng đánh giá từ sản phẩm cuối cùng sang quá trình thực hiện và cách thức sử dụng GenAI.

4.2. Kiểm định mô hình và các giả thuyết nghiên cứu

Sau khi phân tích thực trạng, nghiên cứu tiến hành kiểm định độ tin cậy của thang đo, phân tích tương quan Pearson và hồi quy tuyến tính đa biến nhằm kiểm định mối quan hệ giữa năm phương diện sử dụng GenAI và hiệu quả học tập thực chất (EFF).

4.2.1. Kiểm định độ tin cậy thang đo (Cronbach's Alpha)

Nghiên cứu sử dụng hệ số Cronbach's Alpha để kiểm định tính nhất quán nội tại của các biến quan sát. Theo tiêu chuẩn thống kê, thang đo được đánh giá là đạt yêu cầu khi hệ số Cronbach's Alpha từ 0,7 trở lên.

Bảng 9

Kết quả kiểm định độ tin cậy thang đo Cronbach's Alpha

Ký hiệu	Tên nhân tố	Số biến quan sát	Hệ số Cronbach's Alpha
PR	Năng lực Prompt	2	0,841
VE	Năng lực kiểm chứng thông tin	2	0,912
IN	Trách nhiệm và minh bạch học thuật	2	0,718
FB	Sử dụng AI để đánh giá và phản hồi	3	0,75
SU	Mức độ tận dụng AI hỗ trợ nhận thức	2	0,729
EFF	Hiệu quả học tập và tiến bộ thực chất	3	0,824

Kết quả tại Bảng 9 cho thấy hệ số Cronbach’s Alpha của sáu thang đo dao động từ 0,718 đến 0,912, đều vượt ngưỡng chấp nhận 0,70, cho thấy các thang đo có mức độ nhất quán nội tại đạt yêu cầu và đủ điều kiện để sử dụng trong các phân tích tiếp theo.

4.2.2. Phân tích tương quan Pearson

Phân tích tương quan Pearson được sử dụng để kiểm tra mối liên hệ tuyến tính giữa các biến nghiên cứu. Đồng thời, hiện tượng đa cộng tuyến trong mô hình hồi quy được đánh giá thông qua hệ số VIF.

Bảng 10

Ma trận tương quan Pearson

Nhân tố	Mô_tả	Mean	SD	PR	VE	IN	FB	SU	EFF
PR	Năng lực Prompt	3,724	0,789	1	0,712***	0,505***	0,658***	0,569***	0,642***
VE	Năng lực kiểm chứng thông tin	4,071	0,973	0,712***	1	0,473***	0,460***	0,603***	0,614***
IN	Trách nhiệm & minh bạch học thuật	3,453	0,789	0,505***	0,473***	1	0,569***	0,429***	0,471***
FB	Sử dụng AI đánh giá và phản hồi	3,522	0,767	0,658***	0,460***	0,569***	1	0,516***	0,614***
SU	Tận dụng AI hỗ trợ nhận thức	3,959	0,776	0,569***	0,603***	0,429***	0,516***	1	0,712***
EFF	Hiệu quả học tập và tiến bộ thực chất	3,851	0,728	0,642***	0,614***	0,471***	0,614***	0,712***	1

** Ghi chú: *** tương đương mức ý nghĩa p < 0.001*

Kết quả phân tích (Bảng 10) chỉ ra rằng cả 5 biến độc lập thuộc nhóm năng lực số AI đều có mối tương quan thuận và có ý nghĩa thống kê rất cao ($p < 0.001$) đối với biến phụ thuộc Hiệu quả học tập (EFF). Đáng chú ý, Mức độ tận dụng AI hỗ trợ nhận thức (SU) có hệ số tương quan mạnh nhất ($r = 0,712$), tiếp đó là Năng lực Prompt (PR, $r = 0,642$). Hệ số tương quan giữa các biến độc lập đều nhỏ hơn 0,8, cho thấy không có dấu hiệu nghiêm trọng của hiện tượng đa cộng tuyến. Mối liên hệ chặt chẽ này là tiền đề phù hợp để tiến hành phân tích hồi quy đa biến.

4.2.3. Phân tích hồi quy tuyến tính và kiểm định giả thuyết

Mô hình hồi quy tuyến tính bội được thực hiện với phương pháp Enter nhằm kiểm định tác động đồng thời của 5 yếu tố độc lập đến Hiệu quả học tập thực chất (EFF).

Bảng 11

Trọng số hồi quy và kết quả kiểm định giả thuyết

Biến độc lập	Hệ số chưa chuẩn hóa (B)	Sai số chuẩn (SE)	Hệ số chuẩn hóa (Beta)	Giá trị t	Mức ý nghĩa (p-value)	Đa cộng tuyến (VIF)
(Hằng số)	0,543	0,298	-	1,818	0,072	-
PR	0,127	0,107	0,138	1,195	0,235	2,844
VE	0,111	0,079	0,149	1,405	0,163	2,393
IN	0,021	0,08	0,022	0,26	0,794	1,622
FB	0,215	0,095	0,226	2,25	0,027*	2,144
SU	0,39	0,086	0,416	4,526	0,000***	1,793

** Thông số mô hình: $R^2 = 0{,}627$; R^2 hiệu chỉnh $= 0{,}603$; $F = 26{,}565$ ($p < 0{,}001$). (Ghi chú: * tương đương p < 0,05, *** tương đương p < 0,001)*

Đánh giá độ phù hợp của mô hình: Hệ số R-bình phương hiệu chỉnh (Adjusted R-squared) đạt 0,603, chứng tỏ 5 biến độc lập trong mô hình có thể giải thích được 60,3% sự biến thiên của hiệu quả học tập thực chất. Kiểm định ANOVA cho thấy mô hình hồi quy tổng thể có ý nghĩa thống kê (F = 26,565, p < 0,001), cho thấy tập hợp các biến độc lập có khả năng giải thích có ý nghĩa đối với biến phụ thuộc EFF. Hệ số phóng đại phương sai (VIF) của các biến dao động từ 1,622 đến 2,844 (đều nhỏ hơn 3), xác nhận mô hình không xảy ra hiện tượng đa cộng tuyến.

Kết quả kiểm định từng giả thuyết cho thấy FB và SU có mối quan hệ tích cực và có ý nghĩa thống kê với EFF. Cụ thể, FB có hệ số hồi quy chuẩn hóa β = 0,226 (p = 0,027), trong khi SU có β = 0,416 (p < 0,001). Do đó, H4 và H5 được ủng hộ. Trong đó, SU có mức đóng góp tương đối lớn nhất trong mô hình.

Ngược lại, PR (β = 0,138, p = 0,235), VE (β = 0,149, p = 0,163) và IN (β = 0,022, p = 0,794) không có mối quan hệ độc lập có ý nghĩa thống kê với EFF khi đồng thời kiểm soát các biến còn lại. Vì vậy, H1, H2 và H3 không được ủng hộ.

Đáng chú ý, kết quả tương quan Pearson cho thấy cả năm yếu tố đều có mối liên hệ thuận với EFF, trong khi kết quả hồi quy chỉ xác nhận đóng góp độc lập có ý nghĩa thống kê của FB và SU. Điều này cho thấy mối liên hệ giữa năng lực sử dụng GenAI và hiệu quả học tập không hoàn toàn đồng nhất giữa các phương diện. Khả năng sử dụng GenAI để hỗ trợ nhận thức chuyên môn và để thực hiện đánh giá, phản hồi có vai trò nổi bật hơn so với năng lực thao tác công cụ, kiểm chứng thông tin hoặc trách nhiệm và minh bạch học thuật khi các yếu tố được xem xét đồng thời. Kết quả này cũng gợi ý rằng việc có kỹ năng sử dụng GenAI hoặc nhận thức được yêu cầu kiểm chứng thông tin chưa nhất thiết tạo ra sự gia tăng trực tiếp về hiệu quả học tập. Giá trị giáo dục của GenAI có thể phụ thuộc nhiều hơn vào cách người học tích hợp công cụ này vào quá trình suy luận, phản hồi, điều chỉnh và kiến tạo kiến thức chuyên môn.

4.3. Thực trạng ứng dụng GenAI của GV trong dự án

Qua phỏng vấn, nghiên cứu cho thấy trong giai đoạn khởi động dự án, GV hầu như chưa sử dụng GenAI để thiết kế mục tiêu, nhiệm vụ, tiêu chí đánh giá hay kế hoạch triển khai, dù nhận thức rõ SV sử dụng GenAI rộng rãi, đặc biệt ở nhóm năng lực yếu. Một số SV thậm chí phụ thuộc vào nội dung GenAI tạo sẵn, dẫn đến hiểu biết hời hợt và giảm khả năng tư duy sâu mà PBL hướng tới.

Trong giai đoạn triển khai, GenAI chủ yếu hỗ trợ GV trong giảng giải nội dung phức tạp, tối ưu bài giảng và đưa ra gợi ý phản hồi khi đánh giá tiến độ; tuy nhiên, quyết định đánh giá cuối cùng vẫn do GV thực hiện. Ở giai đoạn báo cáo kết quả và phản tư sau dự án, GenAI ít được sử dụng vì GV chú trọng đánh giá khả năng trình bày trực tiếp, mức độ hiểu sâu và ứng biến của SV, đồng thời chưa có cơ chế phản tư hệ thống về việc dùng GenAI.

Sự thiếu hụt quy chuẩn và công cụ đánh giá GenAI ở phía GV, trong khi GenAI đã thâm nhập sâu vào quá trình học tập của SV, tạo ra khoảng trống trong quản lý chất lượng dạy học dự án. Đây cũng là cơ sở để nghiên cứu đề xuất xây dựng hệ thống tiêu chí đánh giá PBL mới nhằm hỗ trợ GV định hướng và kiểm soát hiệu quả việc sử dụng GenAI.

5. Thảo luận và đề xuất giảng dạy

5.1. Thực trạng sử dụng GenAI trong PBL

Kết quả nghiên cứu cho thấy tồn tại khoảng cách giữa mức độ sử dụng GenAI của SV

và mức độ tích hợp công nghệ vào tổ chức dạy học PBL. Mặc dù GenAI đã được sử dụng phổ biến, chỉ 29,4% SV có khả năng lựa chọn công cụ phù hợp với từng giai đoạn dự án (Q4), phản ánh năng lực Kiến thức – Nội dung – Công nghệ (TCK) còn hạn chế và việc khai thác công cụ chưa thực sự phù hợp với từng nhiệm vụ.

Về phía GV, phỏng vấn cho thấy GenAI chủ yếu được dùng để giải thích nội dung, chuẩn bị học liệu và gợi ý phản hồi, chưa được tích hợp có hệ thống vào thiết kế và quản lý dự án theo định hướng Kiến thức – Sư phạm – Công nghệ (TPK). Điều này cho thấy khoảng cách giữa việc sử dụng GenAI trong học tập và tích hợp công nghệ có chủ đích vào toàn bộ chu trình PBL vẫn còn đáng kể.

Bên cạnh đó, một bộ phận SV cho biết tỷ lệ nội dung có sự hỗ trợ của AI trong báo cáo ở mức cao (Q5). Phỏng vấn GV cũng ghi nhận những sản phẩm có ngôn ngữ tương đối hoàn chỉnh nhưng chưa phản ánh đầy đủ mức độ hiểu biết của người học. Vì vậy, cần phân biệt giữa việc sử dụng GenAI để hỗ trợ học tập và việc dựa vào AI để thay thế đáng kể hoạt động tư duy của SV.

5.2. Những giá trị tích cực của GenAI đối với hoạt động PBL

Kết quả nghiên cứu cho thấy SV đánh giá tương đối tích cực vai trò của GenAI trong PBL:

- *Thúc đẩy phản hồi và điều chỉnh sản phẩm (Q13, Q14):* 56,5% SV sử dụng GenAI để trao đổi, tranh luận và tìm điểm cần cải thiện; 62,3% sử dụng công cụ để tiếp nhận phản hồi và điều chỉnh sản phẩm. Tuy nhiên, phỏng vấn cho thấy phản hồi vẫn chủ yếu tập trung vào ngôn ngữ và cách trình bày, trong khi phản biện lập luận kinh tế và chiến lược marketing còn hạn chế.

- *Hỗ trợ kiến thức chuyên môn và năng lực ngôn ngữ (Q16, Q17, Q19):* 75,3% SV nhận thấy GenAI giúp trau chuốt ngôn ngữ, 77,6% cho rằng công cụ giúp hiểu khái niệm và thuật ngữ kinh tế, còn 67,0% ghi nhận sự tiến bộ về tư duy và diễn đạt. Đáng chú ý, hỗ trợ nhận thức chuyên môn có mối liên hệ riêng nổi bật nhất với EFF, cho thấy tiềm năng hỗ trợ SV kết hợp kiến thức kinh tế với năng lực biểu đạt bằng tiếng Trung.

- *Hỗ trợ quản lý thời gian và tăng tính chủ động trong học tập (Q7, Q8, Q11, Q18):* 76,5% SV cho rằng GenAI giúp tiết kiệm thời gian và hiểu nội dung tốt hơn; 61,2% có thể viết câu lệnh tương đối chi tiết, 81,2% biết điều chỉnh câu lệnh khi kết quả chưa phù hợp. Tuy nhiên, chất lượng câu lệnh còn chưa đồng đều, nhất là về dữ liệu đầu vào, bối cảnh chuyên môn và chuỗi câu lệnh nhiều bước.

- *Hỗ trợ hình thành năng lực hợp tác và nhận thức nghề nghiệp (Q15, Q20, Q22):* 57,6% SV cho rằng GenAI hỗ trợ đánh giá và trao đổi nhóm; 82,3% nhận thấy kỹ năng sử dụng AI có thể tạo lợi thế nghề nghiệp, và 71,8% ủng hộ đưa tiêu chí sử dụng AI vào đánh giá PBL. Điều này cho thấy SV bắt đầu gắn năng lực sử dụng AI với yêu cầu nghề nghiệp và đánh giá học tập.

5.3. Những vấn đề tồn tại

Bên cạnh những giá trị tích cực, kết quả nghiên cứu cũng cho thấy một số vấn đề cần được lưu ý khi tích hợp GenAI vào PBL.

- *Chiến lược sử dụng công cụ còn đơn điệu (Q4):* Chỉ 29,4% SV có khả năng lựa chọn công cụ phù hợp với từng giai đoạn dự án. Phỏng vấn cho thấy nhiều SV sử dụng câu lệnh chung, cung cấp ít dữ liệu và chưa biết tương tác nhiều bước. Vì vậy, cần phát triển năng lực lựa chọn công cụ theo nhiệm vụ và mục đích cụ thể.

- *Khoảng cách giữa sử dụng GenAI và phản tư thực chất (Q13, Q14):* Dù 56,5% SV

dùng GenAI để tìm điểm yếu của sản phẩm và 62,3% sử dụng công cụ để phản hồi, điều chỉnh, hoạt động phản tư vẫn chưa đi sâu vào lập luận chuyên môn. Kiến thức kinh tế chưa vững có thể khiến SV khó đánh giá đầu ra, làm tăng nguy cơ phụ thuộc AI thay vì sử dụng công cụ như một giàn giáo nhận thức.

- *Khoảng cách giữa tự đánh giá và thực hành kiểm chứng thông tin (Q9, Q10):* Mặc dù 82,3% SV cho biết thường kiểm tra thông tin và 83,5% đối chiếu thông tin chuyên sâu với nguồn khác, phỏng vấn cho thấy một số SV chưa phân biệt được nguồn phổ biến với nguồn có thẩm quyền, cũng như chưa có quy trình kiểm chứng hệ thống.

- *Vấn đề về trách nhiệm và minh bạch trong sử dụng AI (Q11, Q12):* Chỉ 47% SV đồng ý hoặc hoàn toàn đồng ý rằng họ sẵn sàng chịu trách nhiệm về độ chính xác của sản phẩm; tỷ lệ này là 53% đối với việc minh bạch mức độ sử dụng GenAI. Điều này cho thấy cần làm rõ ranh giới giữa AI hỗ trợ và AI thay thế thông qua hướng dẫn và tiêu chí đánh giá cụ thể.

Nhìn chung, SV đã tiếp cận GenAI khá phổ biến nhưng còn hạn chế trong lựa chọn công cụ, phản tư chuyên môn, kiểm chứng thông tin và chịu trách nhiệm về sản phẩm. Vì vậy, cần chuyển trọng tâm từ khả năng sử dụng sang năng lực đánh giá và kiểm soát AI trong giải quyết nhiệm vụ chuyên môn; đây cũng là cơ sở để đề xuất khung tiêu chí đánh giá PBL ở phần tiếp theo.

5.4. Đề xuất giảng dạy

Từ kết quả nghiên cứu, bài viết đề xuất định hướng giảng dạy và đánh giá theo nguyên tắc: GenAI cần hỗ trợ quá trình hiểu vấn đề, phản biện, tiếp nhận phản hồi và điều chỉnh sản phẩm, thay vì chỉ tạo hoặc hoàn thiện nội dung. Năng lực đặt câu lệnh, kiểm chứng thông tin và minh bạch trong sử dụng AI cần được duy trì như những điều kiện bảo đảm chất lượng:

- *Thứ nhất, tái cấu trúc quy trình dự án theo nguyên tắc tư duy độc lập trước - AI hỗ trợ sau:* SV cần trình bày ý tưởng, sơ đồ tư duy hoặc giả thuyết ban đầu trước khi sử dụng GenAI để mở rộng, phản biện và hoàn thiện. Cách làm này duy trì vai trò chủ thể của người học, đồng thời giúp GV đánh giá quá trình tư duy thay vì chỉ dựa vào sản phẩm cuối cùng.

- *Thứ hai, hướng dẫn SV lựa chọn công cụ theo chức năng và giai đoạn dự án:* GV cần giúp SV sử dụng công cụ phù hợp với mục tiêu: ChatGPT hoặc Claude để phát triển ý tưởng; Perplexity hoặc Gemini để tìm kiếm, tổng hợp thông tin; GenAI để hỗ trợ phân tích, diễn đạt và chuẩn hóa thuật ngữ; cuối cùng, đối chiếu thông tin với nguồn gốc và nguồn chuyên ngành đáng tin cậy như: Caixin, Sina Finance. Trọng tâm là lựa chọn đúng công cụ cho đúng nhiệm vụ, không phải sử dụng càng nhiều công cụ càng tốt.

- *Thứ ba, đổi mới đánh giá theo hướng coi trọng năng lực nhận thức và tương tác với AI:* Bên cạnh prompt log và nhật ký kiểm chứng, cần ghi nhận cách SV sử dụng GenAI để giải thích khái niệm, so sánh phương án, phát hiện điểm yếu và điều chỉnh sản phẩm. Đánh giá vấn đáp sau dự án giúp xác định mức độ hiểu biết thực chất và khả năng giải thích các quyết định chuyên môn của SV.

- *Thứ tư, tăng cường năng lực TPACK của GV trong tổ chức PBL có GenAI:* GV cần kết hợp kiến thức chuyên môn, phương pháp PBL và công nghệ trong thiết kế nhiệm vụ, đồng thời xác định khi nào nên sử dụng GenAI và khi nào cần duy trì tư duy độc lập. Nên ưu tiên các nhiệm vụ phân tích dữ liệu thực tế, tình huống doanh nghiệp, giải thích quyết định và phản biện kết quả.

- *Thứ năm, xây dựng quy chuẩn về tính minh bạch và liêm chính trong sử dụng GenAI:* Mỗi dự án cần quy định rõ hình thức sử dụng AI được phép, yêu cầu khai báo và cách ghi nhận mức độ hỗ trợ. SV có thể nộp AI usage log, ghi lại mục đích sử dụng, câu lệnh chính, nội dung

đầu ra được sử dụng và cách kiểm chứng, chỉnh sửa, qua đó nâng cao trách nhiệm học thuật và ý thức sử dụng AI có đạo đức.

6. Đề xuất khung tiêu chí đánh giá dự án PBL có sử dụng GenAI

6.1. Cơ sở và quá trình xây dựng rubric

Rubric được xây dựng trên cơ sở kế thừa, điều chỉnh các tiêu chí từ PBL/PBLL, năng lực số AI và đánh giá AI, với TPACK làm khung định hướng sự phù hợp giữa công nghệ, hoạt động học tập và nội dung chuyên ngành.

Các tiêu chí về kiến thức chuyên ngành, năng lực ngôn ngữ và thuyết trình – hỏi đáp được điều chỉnh theo đặc thù tiếng Trung kinh tế; các tiêu chí về sử dụng GenAI, tư duy phản biện và thẩm định thông tin tập trung vào lựa chọn công cụ, thiết kế prompt, minh bạch, kiểm chứng và điều chỉnh đầu ra AI.

Quá trình xây dựng gồm ba bước: (1) tổng hợp tiêu chí từ các khung liên quan; (2) loại bỏ trùng lặp và điều chỉnh theo đặc thù học phần; (3) xây dựng chỉ báo cho từng mức độ thực hiện.

6.2. Các tiêu chí đánh giá

Rubric gồm năm tiêu chí: (1) kiến thức chuyên ngành kinh tế; (2) năng lực ngôn ngữ chuyên ngành; (3) sử dụng và tương tác với GenAI; (4) tư duy phản biện và thẩm định thông tin; (5) thuyết trình và hỏi đáp. Tiêu chí 3 yêu cầu AI log ghi nhận quá trình lựa chọn công cụ, thiết kế, điều chỉnh prompt và đóng góp của AI/người học; tiêu chí 4 đánh giá khả năng kiểm chứng, đánh giá và điều chỉnh đầu ra AI. Hai tiêu chí này giúp chuyển trọng tâm đánh giá từ sản phẩm cuối cùng sang cả quá trình người học tương tác và xử lý đầu ra của AI.

6.3. Trọng số và khung đánh giá

Trọng số được đề xuất dựa trên chuẩn đầu ra và đặc thù của học phần, ưu tiên kiến thức chuyên ngành và năng lực ngôn ngữ, đồng thời bổ sung tiêu chí về AI và tư duy phản biện. Đây là trọng số sơ bộ, chưa được chuẩn hóa thống kê, cần tiếp tục thẩm định qua ý kiến chuyên gia và thử nghiệm thực tế.

Bảng 22

Bảng tiêu chí đánh giá dự án PBL tiếng Trung kinh tế tích hợp năng lực sử dụng AI

Tiêu chí đánh giá	Trọng số	Xuất sắc (9-10 điểm)	Khá (7-8 điểm)	Trung bình (5-6 điểm)	Không đạt (dưới 5 điểm)
1. Nội dung chuyên ngành kinh tế (CK - TPACK - PBL)	20%	Vận dụng chính xác và linh hoạt các lý thuyết kinh tế; sử dụng dữ liệu cập nhật, đáng tin cậy; lập luận chặt chẽ và gắn với bối cảnh doanh nghiệp.	Nội dung đầy đủ, cấu trúc logic; vận dụng đúng các kiến thức chuyên ngành nhưng phân tích chưa thật sâu hoặc liên hệ thực tiễn còn hạn chế.	Nội dung còn rời rạc; vận dụng lý thuyết chưa chính xác hoặc còn sử dụng dữ liệu cũ; phân tích chủ yếu mang tính mô tả.	Lạc đề; kiến thức chuyên môn sai lệch nghiêm trọng; dữ liệu thiếu tin cậy hoặc sao chép nguyên văn từ nguồn khác.
2. Năng lực ngôn ngữ chuyên ngành (CK - TPACK- PBL)	20%	Sử dụng chính xác thuật ngữ tiếng Trung kinh tế, phù hợp lý thuyết bài học; văn phong học thuật, phù hợp ngữ cảnh giao tiếp thương mại; diễn	Sử dụng đúng phần lớn thuật ngữ chuyên ngành; còn một số lỗi nhỏ về từ vựng hoặc ngữ pháp nhưng không ảnh hưởng nội dung; một số thuật ngữ	Thuật ngữ chuyên ngành chưa thống nhất; còn nhiều lỗi diễn đạt hoặc lạm dụng dịch máy nhưng vẫn hiểu được ý chính.	Sai nhiều về ngữ pháp và thuật ngữ; diễn đạt khó hiểu; lạm dụng dịch máy khiến nội dung sai lệch hoặc không thể đọc hiểu.

		đạt tự nhiên, không có lỗi ngôn ngữ.	kinh tế chưa phù hợp lý thuyết bài học		
3. Năng lực sử dụng và tương tác với GenAI (Yêu cầu nộp AI log)	10%	Lựa chọn công cụ AI phù hợp với từng nhiệm vụ; Prompt có cấu trúc rõ ràng, đa vòng và được tối ưu qua nhiều lần tương tác; AI log đầy đủ, minh bạch; biết sử dụng GenAI để đặt câu hỏi, so sánh, phản biện, nhận phản hồi nhằm cải thiện sản phẩm	Có lựa chọn công cụ AI phù hợp ở một số giai đoạn; Prompt tương đối rõ nhưng chủ yếu là câu lệnh đơn; AI log đầy đủ nhưng giải trình còn đơn giản; có sử dụng AI để nhận phản hồi hoặc điều chỉnh sản phẩm nhưng chưa thường xuyên	Chỉ sử dụng một công cụ AI cho mọi nhiệm vụ; Prompt đơn giản; AI log sơ sài; phụ thuộc nhiều vào nội dung AI tạo ra; chủ yếu sử dụng AI để tạo hoặc chỉnh sửa nội dung, ít thông tin phản biện và phản hồi	Vi phạm liêm chính: Không nộp AI log, cố tình che giấu việc sử dụng AI, nộp AI log giả mạo; hoặc để AI tạo nội dung 100% không qua xử lý; tiếp nhận phần lớn đầu ra của AI mà không có tương tác/ điều chỉnh đáng kể
4. Tư duy phản biện và thẩm định thông tin (PBL)	10%	Kiểm chứng thông tin AI bằng nhiều nguồn đáng tin cậy; đánh giá, lựa chọn và điều chỉnh các đề xuất của AI; phát hiện và hiệu chỉnh các lỗi/ hiện tượng ảo giác của AI.	Có kiểm chứng và điều chỉnh một phần nội dung do AI tạo ra; biết lựa chọn đề xuất phù hợp nhưng việc giải trình còn chưa đầy đủ hoặc chưa sâu sắc.	Chỉ kiểm chứng một phần; tiếp nhận phần lớn đề xuất AI; ít giải thích được lý do lựa chọn; bỏ sót một số sai lệch của AI.	Tiếp nhận hoàn toàn đầu ra của AI; không kiểm chứng; không phát hiện các lỗi hoặc hiện tượng ảo giác; lập luận thiếu căn cứ hoặc mâu thuẫn.
5.Thuyết trình và hỏi đáp trực tiếp	40%	Trình bày tự tin, logic; bảo vệ thuyết phục các quyết định chuyên môn và quá trình sử dụng AI; trả lời đầy đủ các câu hỏi; chứng minh sự thấu hiểu và khả năng vận dụng kiến thức.	Trình bày rõ ràng; có thể trả lời các câu hỏi cơ bản nhưng lúng túng hoặc chưa đủ sâu sắc khi bị hội đồng/nhóm khác truy vấn thêm.	Chủ yếu đọc slide; trả lời được ít câu hỏi; giải thích chưa rõ quá trình hình thành ý tưởng hoặc việc sử dụng AI; mức độ hiểu biết còn bề mặt.	Không trình bày được hoặc không trả lời được các câu hỏi chính; không giải thích được quá trình thực hiện; thể hiện phụ thuộc hoàn toàn vào AI.

7. Kết luận

7.1. Kết luận chính

Nghiên cứu cho thấy GenAI được sử dụng khá phổ biến trong PBL tiếng Trung kinh tế, nhưng còn hạn chế về lựa chọn công cụ, kiểm chứng thông tin và minh bạch. Pearson cho thấy PR, VE, IN, FB và SU đều tương quan thuận với EFF; tuy nhiên, hồi quy chỉ xác định FB (β = 0,226; p = 0,027) và SU (β = 0,416; p < 0,001) có mối liên hệ riêng có ý nghĩa thống kê với EFF. Điều này nhấn mạnh vai trò của GenAI trong hỗ trợ nhận thức chuyên môn, phản hồi và điều chỉnh sản phẩm. Trên cơ sở đó, nghiên cứu đề xuất khung đánh giá PBL kết hợp chất lượng sản phẩm, quá trình sử dụng GenAI, năng lực kiểm chứng và khả năng giải trình của SV.

7.2. Đóng góp của nghiên cứu

Về mặt lý luận, nghiên cứu góp phần mở rộng cách tiếp cận đánh giá PBL trong bối cảnh GenAI bằng cách kết nối các định hướng của PBL, năng lực số AI và TPACK. Thay vì xem năng lực AI như một kỹ năng công nghệ tách biệt, nghiên cứu đặt năng lực sử dụng và kiểm chứng AI trong mối quan hệ với kiến thức chuyên ngành, năng lực ngôn ngữ và hoạt động học tập qua dự án. Kết quả hồi quy đồng thời cung cấp thêm một góc nhìn về sự khác biệt giữa mối tương quan chung và mối liên hệ riêng của các phương diện sử dụng GenAI đối với hiệu quả học tập.

Về mặt thực tiễn, nghiên cứu cung cấp bằng chứng ban đầu cho việc tổ chức PBL có sử dụng GenAI trong dạy học tiếng Trung kinh tế. Kết quả cho thấy GV cần chuyển từ việc chỉ hướng dẫn SV sử dụng công cụ sang hướng dẫn sử dụng AI có mục đích, kiểm chứng thông tin, phản tư và giải trình quá trình sử dụng. Khung tiêu chí được đề xuất có thể được sử dụng như một cơ sở tham khảo để thiết kế đánh giá dự án theo hướng kết hợp giữa sản phẩm, quá trình sử dụng AI và khả năng bảo vệ kết quả học tập.

7.3. Hạn chế và hướng nghiên cứu tiếp theo

Nghiên cứu có ba hạn chế: (1) mẫu tại một cơ sở đào tạo, chủ yếu dựa trên dữ liệu tự báo cáo, hạn chế khả năng khái quát hóa; (2) thiết kế cắt ngang chưa xác định được quan hệ nhân quả giữa sử dụng GenAI và hiệu quả học tập; (3) rubric mới ở mức đề xuất sơ bộ, chưa được kiểm định đầy đủ về độ giá trị và độ tin cậy. Nghiên cứu tiếp theo cần mở rộng mẫu, áp dụng thiết kế thực nghiệm/bán thực nghiệm và thử nghiệm, thẩm định rubric trong thực tế.

Tài liệu tham khảo

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press. https://doi.org/10.1017/9781009024549

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press. https://www.cambridge.org/9780521596756

International Center for Academic Integrity. (2021). *The fundamental values of academic integrity* (3rd ed.). https://academicintegrity.org/aws/ICAI/asset_manager/get_file/911282?ver=1

Kilpatrick, W. H. (1918). *The project method: The use of the purposeful act in the educative process*. Teachers College, Columbia University. https://archive.org/details/projectmethodus00kilpgoog

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record: The Voice of Scholarship in Education, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). AI literacy: Definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology*, *58*(1), 504–509. https://doi.org/10.1002/pra2.487

PHỤ LỤC

Thông tin khảo sát xin truy cập mã QR dưới đây: