

REDESIGNING LEARNER ASSESSMENT IN THE CONTEXT OF AI-INTEGRATED HIGHER EDUCATION

Huynh Van Tung¹, Huynh Huu Tho²,
Nguyen Le Hoa Tuyet^{*3}

* Corresponding author
Email: nlhtuyet@ctu.edu.vn

¹ Email: hvtung@ctu.edu.vn
Can Tho University of Technology
256 Nguyen Van Cu street, Cai Khe ward,
Can Tho City, Vietnam

² Email: hhtho@ctu.edu.vn
School of Economics, Can Tho University
3/2 road, Ninh Kieu ward, Can Tho City, Vietnam

³ Can Tho University of Technology
256 Nguyen Van Cu street, Cai Khe ward,
Can Tho City, Vietnam

Received: 19/5/2026
Revised: 28/7/2026
Accepted: 10/8/2026
Published: 20/10/2026

Abstract: The rapid rise of AI has severely compromised the validity of product-based assessment, creating an urgent need for a new theoretical framework to monitor learners' knowledge construction. Applying the integrative literature review method, this study proposes the Research - Operation - Evaluation (R-O-E) framework, shifting the assessment paradigm from product-based to process-based evaluation. The core contribution lies in establishing a comprehensive assessment ecosystem through the synchronous integration of the R-O-E framework with the ETHICAL principles and the SARPS reconstruction strategies. This system mandates transparency in students' knowledge construction by leveraging digital footprints, reflective logs, oral defense, and adversarial validation of AI outputs. Consequently, the study provides a scientific and feasible roadmap to help higher education institutions safeguard academic integrity and enhance substantive educational quality in the AI era.

Keywords: *Product-based assessment, process-oriented assessment, AI, R-O-E, students.*

TÁI THIẾT PHƯƠNG PHÁP ĐÁNH GIÁ NGƯỜI HỌC DƯỚI BỐI CẢNH GIÁO DỤC ĐẠI HỌC TÍCH HỢP AI

Huỳnh Văn Tùng¹, Huỳnh Hữu Thọ²,
Nguyễn Lê Hoa Tuyết^{*3}

* Tác giả liên hệ
Email: nlhtuyet@ctu.edu.vn

¹ Email: hvtung@ctu.edu.vn
Trường Đại học Kỹ thuật - Công nghệ Cần Thơ
256 Nguyễn Văn Cừ, phường Cái Khế,
Thành phố Cần Thơ, Việt Nam

² Email: hhtho@ctu.edu.vn
Trường Kinh tế, Đại học Cần Thơ
Khu II, Đường 3/2, phường Ninh Kiều,
Thành phố Cần Thơ, Việt Nam

³ Trường Đại học Kỹ thuật - Công nghệ Cần Thơ
256 Nguyễn Văn Cừ, phường Cái Khế,
Thành phố Cần Thơ, Việt Nam

Nhận bài: 19/5/2026
Chỉnh sửa xong: 28/7/2026
Chấp nhận đăng: 10/8/2026
Xuất bản: 20/10/2026

Tóm tắt: Sự bùng nổ của trí tuệ nhân tạo đã làm suy giảm nghiêm trọng hiệu lực phương pháp đánh giá sản phẩm, đặt ra yêu cầu cấp thiết về một khung lý thuyết mới giám sát tiến trình kiến tạo tri thức của người học. Vận dụng phương pháp tổng quan tài liệu tích hợp, nghiên cứu này đề xuất khung "Nghiên cứu - Thực thi - Đánh giá" (gọi tắt là R-O-E) dịch chuyển trọng tâm đánh giá từ sản phẩm sang quá trình. Đóng góp cốt lõi của nghiên cứu là thiết lập được một hệ sinh thái đánh giá toàn diện dựa trên tích hợp đồng bộ khung R-O-E với nguyên tắc đạo đức ETHICAL và chiến lược tái thiết SARPS. Mô hình này buộc người học minh bạch tiến trình kiến tạo tri thức qua vết tích số, nhật ký phản tư, vấn đáp và thẩm định ngược trí tuệ nhân tạo. Qua đó, nghiên cứu cung cấp một lộ trình khoa học, khả thi giúp các cơ sở giáo dục đại học bảo vệ liêm chính học thuật và nâng cao chất lượng đào tạo thực chất trong kỷ nguyên trí tuệ nhân tạo.

Từ khóa: *Đánh giá sản phẩm, đánh giá quá trình, trí tuệ nhân tạo, R-O-E, sinh viên.*

1. Đặt vấn đề

Sự bùng nổ của trí tuệ nhân tạo (AI) đang thách thức trực tiếp hệ thống giáo dục đại học, làm suy giảm hiệu lực của phương pháp đánh giá sản phẩm đối với loại hình bài luận, chuyên đề, đồ án hay nghiên cứu khoa học. Các mô hình ngôn ngữ lớn, đặc biệt là công cụ nhân bản hóa văn bản AI đã tạo ra

sự đứt gãy nghiêm trọng giữa chất lượng sản phẩm nộp với năng lực thực của sinh viên (AIBlooshi, 2026; Chan & Colloton, 2024; Freeman, 2025). Hệ lụy của việc lạm dụng AI không chỉ đe dọa trực tiếp đến tính liêm chính học thuật mà còn tạo ra bất bình đẳng trong môi trường giáo dục. Nghiêm trọng hơn, nó có thể dẫn đến việc hình thành một thế hệ người học

giới ảo và lệ thuộc vào công nghệ (AlBlooshi, 2026; Cotton và cộng sự, 2024; Chan & Colloton, 2024; Freeman, 2025). Dù nhiều nghiên cứu đi trước đã nỗ lực đề xuất giải pháp ứng phó, song đa số vẫn tập trung vào công cụ phát hiện AI (Kofinas và cộng sự, 2025; Lodge và cộng sự, 2023; Weber-Wulff và cộng sự, 2023). Thực tế, các công cụ kiểm soát đầu ra này thường thiếu độ tin cậy, dễ bị qua mặt bởi phần mềm nhân bản hóa văn bản AI (Weber-Wulff và cộng sự, 2023); khiến người chấm khó nhận biết được sự can thiệp của thuật toán (Kofinas và cộng sự, 2025). Nghiêm trọng hơn, các công cụ phát hiện AI còn tồn tại thiên kiến hệ thống đối với người dùng không sử dụng tiếng Anh bản ngữ (Liang và cộng sự, 2023), làm tổn hại sự công bằng và xói mòn niềm tin sư phạm (Lodge và cộng sự, 2023). Bối cảnh này đẩy các trường đại học vào cuộc đua công nghệ không hồi kết, đặt ra yêu cầu cấp thiết về việc chuyển dịch trọng tâm từ dò tìm gian lận sản phẩm cuối khóa sang giám sát quá trình kiến tạo tri thức (Perkins và cộng sự, 2024; Roe và cộng sự, 2026). Tuy nhiên, hiện vẫn thiếu vắng một khung lý thuyết hệ thống để hiện thực hóa sự dịch chuyển này, đặc biệt trong bối cảnh giáo dục đại học tại Việt Nam.

Nghiên cứu này lấp đầy khoảng trống đó bằng việc đề xuất hệ sinh thái đánh giá quá trình tích hợp ba chiều: Khung R-O-E (Research - Operation - Evaluation) định hướng hành vi; nguyên tắc đạo đức ETHICAL (Adamakis & Rachiotis, 2025) kiểm soát liên chính số và các chiến lược SARPS (Chan & Colloton, 2024) tái thiết phương pháp kiểm tra. Cụ thể, bài viết giải quyết ba câu hỏi nghiên cứu gồm: 1) Sự bùng nổ AI đã làm suy giảm hiệu lực phương pháp đánh giá sản phẩm truyền thống như thế nào? 2) Hệ sinh thái đánh giá quá trình đã đề xuất để thay thế phương pháp đánh giá sản phẩm (gồm khung R-O-E, nguyên tắc ETHICAL và chiến lược SARPS) cần được cấu trúc ra sao? 3) Những thách thức và điều kiện thực tế nào mà các trường đại học cần giải quyết để triển khai hiệu quả hệ sinh thái đánh giá quá trình trên? Đóng góp khoa học cốt lõi của nghiên cứu này là không đề xuất các giải pháp rời rạc mà kiến tạo một hệ thống đồng bộ chuyển dịch trọng tâm đánh giá từ sản phẩm tĩnh sang quá trình động. Sự tích hợp này phân định rõ ranh giới giữa hỗ trợ và lạm dụng công nghệ, cung cấp lộ trình khả thi bảo chứng tư duy nguyên bản của người học trong kỷ nguyên AI.

2. Phương pháp nghiên cứu

Nghiên cứu vận dụng phương pháp tổng quan

tài liệu tích hợp (Integrative Literature Review) để xây dựng khung R-O-E (Torraco, 2005; Whittemore & Knafl, 2005) qua năm giai đoạn. Ở giai đoạn đầu, nghiên cứu làm rõ sự suy giảm hiệu lực của phương pháp đánh giá sản phẩm và sự thiếu hụt các mô hình kiểm soát quá trình kiến tạo tri thức. Ở giai đoạn 2, dữ liệu thứ cấp từ 2023 - 2026 (tức giai đoạn AI bùng nổ mạnh mẽ) được truy xuất từ các cơ sở dữ liệu uy tín (Google Scholar, Web of Science, ScienceDirect) và báo cáo thực chứng của các tổ chức lớn như Viện chính sách giáo dục đại học (Freeman, 2025), Hội đồng giáo dục kỹ thuật số (Digital Education Council, 2025), Tổ chức Giáo dục, Khoa học và Văn hóa Liên Hợp Quốc (UNESCO, 2025), tổ chức tư vấn chiến lược giáo dục Tyton Partners (Shaw và cộng sự, 2025). Ở giai đoạn 3, các tài liệu này được sàng lọc nghiêm ngặt dựa trên uy tín tạp chí, chất lượng bài viết và mức độ đóng góp học thuật. Kết quả có 22 tài liệu đạt chuẩn được đưa vào phân tích sâu. Trong quá trình này, NotebookLM được sử dụng để hỗ trợ trích xuất thông tin thô và kiểm chứng chéo với văn bản gốc. Tuy nhiên, toàn bộ phân tích, lập luận và phát triển mô hình đều là tư duy nghiên cứu độc lập của tác giả. Ở giai đoạn 4, dữ liệu được phân tích phê bình chuyên sâu để chỉ ra giới hạn của các công cụ phát hiện AI. Cuối cùng, quy trình khép lại bằng việc đề xuất một hệ sinh thái đánh giá quá trình tích hợp đồng bộ khung lý thuyết mới R-O-E với nguyên tắc đạo đức ETHICAL và các chiến lược tái thiết SARPS.

3. Kết quả nghiên cứu

3.1. Thực trạng người học sử dụng AI trong giáo dục đại học

AI giúp tối ưu hóa lộ trình học tập cá nhân, giúp tiết kiệm thời gian và nâng cao hiệu suất nghiên cứu khoa học (Adamakis & Rachiotis, 2025; Chan & Colloton, 2024; Jacques và cộng sự, 2024). Do đó, nó đã nhanh chóng phủ sóng trong cộng đồng sinh viên. Nghiên cứu của Viện chính sách giáo dục đại học (Freeman, 2025) cho thấy giai đoạn 2024-2025, tỉ lệ sinh viên sử dụng AI đã tăng vọt từ 66% lên 92%; tỉ lệ dùng AI làm bài tập đánh giá năng lực tăng từ 53% lên 88%. Tuy nhiên, đáng chú ý nhất vẫn là xu hướng dùng AI để tạo lập văn bản hoàn chỉnh tăng hơn gấp đôi, từ mức 30% lên mức 64%.

Về động cơ, mục tiêu hàng đầu khiến sinh viên sử dụng AI là tối ưu hiệu suất học tập (AlBlooshi, 2026; Adamakis & Rachiotis, 2025; Chan & Colloton, 2024; Freeman, 2025). Cụ thể, 51% dùng AI vì muốn tiết kiệm thời gian, 50% kì vọng AI giúp cải thiện chất lượng bài làm và 40% tin AI giúp đạt điểm số cao

Bảng 1: Mức độ sử dụng AI của sinh viên

Tiêu chí	Năm 2024	Năm 2025
Đã sử dụng AI dưới mọi hình thức	66%	92%
Sử dụng AI làm bài tập đánh giá năng lực	53%	88%
Sử dụng AI để tạo văn bản hoàn chỉnh	30%	64%

(Nguồn: Freeman, 2025)

hơn. Mặc dù có tới 28% người học muốn cải thiện kỹ năng công nghệ và 20% cảm thấy học được nhiều hơn khi có sự hỗ trợ của AI nhưng tỉ lệ các trường đại học khuyến khích sử dụng chỉ ở mức 13%. Điều này cho thấy một khoảng cách chính sách rõ rệt giữa nhu cầu thực tế của sinh viên và định hướng của các trường đại học (xem Bảng 2).

Mặc dù việc dùng AI khá phổ biến nhưng vẫn tồn tại khoảng cách rõ rệt giữa các sinh viên có nền tảng kinh tế khác nhau (Freeman, 2025). Nhóm sinh viên khá giả có mức độ sử dụng cao hơn, thể hiện qua tỉ lệ dùng AI cấu trúc luồng suy nghĩ chiếm đến 50%, cao hơn mức 42% của nhóm sinh viên có điều kiện kinh tế thấp. Đặc biệt, tỉ lệ nhóm khá giả dùng

Bảng 2: Lí do sinh viên sử dụng AI

Lí do sinh viên sử dụng AI	Tỉ lệ (%)
Tiết kiệm thời gian	51
Cải thiện chất lượng bài làm	50
Tin AI giúp đạt điểm cao hơn	40
Tin trường đại học không thể phát hiện dùng AI	20
Nhận được hỗ trợ ngay tức thì	40
Nhận được hỗ trợ mang tính cá nhân hóa	32
Nhận được hỗ trợ ngoài giờ học truyền thống	29
Cải thiện kỹ năng công nghệ	28
Khi dùng AI học được nhiều hơn	20
Vì các sinh viên khác cũng sử dụng	15
Trường đại học khuyến khích dùng	13

(Nguồn: Freeman, 2025)

AI tạo lập và chỉnh sửa văn bản hoàn chỉnh chiếm 20%, cao gấp đôi nhóm có điều kiện thấp (10%) (xem Bảng 3).

Bảng 3: Mục đích sử dụng AI phân theo nền tảng kinh tế - xã hội của sinh viên

Mục đích sử dụng	Nhóm sinh viên điều kiện kinh tế khá giả	Nhóm sinh viên điều kiện kinh tế thấp
Sử dụng AI tóm tắt bài viết	44%	35%
Sử dụng AI cấu trúc luồng suy nghĩ	50%	42%
Dùng AI tạo văn bản và chỉnh sửa lỗi	20%	10%
Không sử dụng AI	13%	14%

(Nguồn: Freeman, 2025)

Việc sử dụng AI cũng tiềm ẩn nhiều rủi ro; trong đó, nghiêm trọng nhất là lệ thuộc vào thuật toán làm suy giảm trí nhớ, thoái hóa khả năng tư duy phản biện và sáng tạo độc lập. AI cũng cung cấp thông tin ảo ở mức khá thường xuyên chiếm 30% và ở mức rất thường xuyên chiếm 11% (Freeman, 2025). Mặt khác, thiên kiến từ thuật toán và nguy cơ rò rỉ dữ liệu cũng là những hệ lụy khó tránh. Sự phổ cập của AI còn kéo theo thay đổi lệch lạc trong tư duy về liêm chính học thuật. Năm 2024, chỉ có 17% sinh viên cho rằng, nộp văn bản do AI tạo ra là được chấp nhận nhưng đến năm 2025, tỉ lệ này tăng lên 25%; trong đó, 18% thừa nhận đã thật sự đưa nội dung từ AI vào bài làm (Freeman, 2025).

Thống kê của Digital Education Council (2025) cho thấy có sự đồng thuận cao giữa sinh viên (59%) và giảng viên (61%) trong chủ động tiếp nhận AI. Tuy nhiên, chỉ có khoảng một nửa sinh viên (52%) lo ngại về việc lệ thuộc công nghệ trong khi có tới 82% giảng viên bày tỏ quan ngại đối với khả năng tự học, tư duy độc lập của sinh viên trước sự bùng nổ AI. Mặt khác, dù AI mang đến sự hỗ trợ rất lớn song cả giảng viên (64%) và sinh viên (84%) vẫn ưu tiên lựa chọn tương tác với con người (từ bạn bè, thầy cô) khi gặp khó khăn (Shaw và cộng sự, 2025). Kết quả này đã chỉ ra một nghịch lí thú vị rằng, công nghệ dù phát triển mạnh mẽ đến đâu vẫn không thể thay thế được cảm xúc và niềm tin trong mối quan hệ thầy trò.

3.2. Sự suy giảm hiệu quả phương pháp đánh giá sản phẩm dưới tác động của AI

Sự phát triển của AI làm vô hiệu hóa phương pháp đánh giá sản phẩm bởi nó có thể tạo ra bài luận chỉ với vài câu lệnh đơn giản (AlBlooshi, 2026; Adamakis & Rachiotis, 2025; Chan & Colloton, 2024; Freeman, 2025; Jacques và cộng sự, 2024; Koh & Doroudi, 2023). Thống kê của Freeman (2025) chỉ ra 50% sinh viên cho rằng dùng AI sẽ giúp cải thiện chất lượng bài làm và có 40% nhận định dùng AI mới đạt điểm số cao. Trong đó, 20% tin rằng, trường đại học không thể phát hiện được việc họ dùng AI. Ngoài ra, cùng với sự phát triển của các phần mềm nhận diện văn bản AI như Turnitin, GPTZero, Copyleaks, Winston AI... là sự gia tăng các công cụ nhân bản hóa văn bản AI như Undetectable AI, QuyllBot, Smodin, StealthGPT... (Koh & Doroudi, 2023; Hodges & Kirschner, 2024). Tuy nhiên, các công cụ quét AI vẫn thường xuyên mắc lỗi, nhận diện nhầm văn bản do con người viết là sản phẩm của AI và ngược lại, dẫn đến một số trường hợp cáo buộc oan người học, làm rạn nứt nghiêm trọng niềm tin giữa người học và tổ chức giáo dục trực thuộc (Koh & Doroudi, 2023; AlBlooshi, 2026). Điều đó cho thấy nỗ lực phân biệt rạch ròi giữa văn bản do người hay máy thực hiện bằng phần mềm không đạt hiệu quả cao như kì vọng của người dùng.

Mặt khác, việc thẩm định các sản phẩm từ AI còn gây lãng phí nguồn lực xã hội, gia tăng chi phí cơ hội và gây quá tải cho giảng viên trong việc nỗ lực phát hiện gian lận. Theo Shaw và cộng sự (2025), việc đối phó với gian lận AI khiến 71% giảng viên tăng giờ làm và 61% tốn thời gian thiết kế đề kháng AI. Ngoài ra, nếu sử dụng công cụ phát hiện AI không minh bạch thuật toán sẽ gây rủi ro lớn về độ tin cậy. Trước thực trạng này, các trường đại học cần cấp thiết tái cấu trúc lại phương pháp đánh giá người học nhằm đảm bảo AI chỉ hỗ trợ nhưng không thay thế tư duy và năng lực nội sinh của người học (Adamakis & Rachiotis, 2025; Chan & Colloton, 2024; Khlaif và cộng sự, 2025). Điều này tương thích với kết quả khảo sát của Digital Education Council (2025) khi 54% giảng viên yêu cầu cải tổ phương pháp đánh giá người học và 50% đồng thuận thiết kế lại đề bài kháng AI để phù hợp với bối cảnh giáo dục đại học hiện đại.

3.3. Tái thiết phương pháp đánh giá người học phù hợp với sự phát triển của AI

Khác với đánh giá sản phẩm, đánh giá quá trình buộc người học trực tiếp tham gia vào hành trình

kiến tạo tri thức qua các yếu tố AI khó mô phỏng như tư duy phản biện, ra quyết định, tuân thủ đạo đức học thuật và sự phản tư (Adamakis & Rachiotis, 2025; AlBlooshi, 2026; Chan & Colloton, 2024; Jacques và cộng sự, 2024; Khlaif và cộng sự, 2025). AI không thể thay người học tương tác trực tiếp, thảo luận và chỉnh sửa dài hạn với giảng viên. Do đó, kiểm soát tiến trình qua nhật kí thực hiện giúp ghi nhận sự tiến bộ và cách người học kết nối tri thức đã tiếp thu với thực tiễn (Chan & Colloton, 2024; Hodges & Kirschner, 2024). Để cụ thể hóa tiến trình học tập thành các hoạt động đo lường được, nghiên cứu đề xuất khung R-O-E (Research - Operation - Evaluation), phát triển dựa trên “Thang nhận thức Bloom” (Anderson & Krathwohl, 2001) và triết lí “Đánh giá như một quá trình học tập” (Earl, 2003). Khung này vận hành nhất quán dựa trên cơ chế kiểm soát đa điểm chạm và tam giác hóa minh chứng học tập qua ba giai đoạn:

(1) *R - Research (Nghiên cứu)*: Sinh viên được dùng AI để phát triển ý tưởng và tìm kiếm thông tin, nhưng phải bảo vệ bản đồ tư duy, đề cương và danh mục tài liệu trực tiếp tại lớp (không dùng AI) nhằm định hình tư duy logic độc lập.

(2) *O - Operation (Thực thi)*: Sinh viên cụ thể hóa đề cương thành bản thảo dưới sự hỗ trợ của AI nhưng phải lưu vết tích số (qua Google Docs Version History hoặc Word Track Changes). Nhằm tuân thủ Luật Bảo vệ dữ liệu cá nhân 2025, việc thu thập phải có sự đồng thuận trước từ sinh viên, cam kết dữ liệu chỉ dùng đánh giá học phần và xóa ngay sau khi kết thúc. Để giảm rủi ro lộ dữ liệu cá nhân, giảng viên nên hướng dẫn sinh viên xuất file PDF kèm phụ lục lịch sử chỉnh sửa (hoặc video minh chứng) đối với Google Docs, hoặc file Word bật Track Changes ở chế độ “All markup”. Nhằm bảo đảm liêm chính học thuật, bước “Thực thi” cần tích hợp nguyên tắc đạo đức ETHICAL (Adamakis & Rachiotis, 2025). Thay vì cấm đoán, ETHICAL cho phép sinh viên dùng AI nhưng yêu cầu phân biệt rõ giữa hỗ trợ với lạm dụng thông qua việc đính kèm phụ lục khai báo bắt buộc. Cơ chế này tương thích với khảo sát của Digital Education Council (2025), khi 57% giảng viên sẵn sàng cho phép sử dụng AI nếu sinh viên minh bạch và tuân thủ hướng dẫn. Nhờ đó, bộ nguyên tắc tái khẳng định mục tiêu cốt lõi của mọi bài đánh giá không nằm không nằm ở sản phẩm nộp lại mà ở quá trình lĩnh hội tri thức (xem Bảng 4).

(3) *E - Evaluation (Đánh giá)*: Sinh viên nộp bài kết hợp bảo vệ trực tiếp trước giảng viên hoặc hội

Bảng 4: Khung nguyên tắc đạo đức ETHICAL khi sinh viên làm bài cùng AI

Nguyên tắc	Mô tả thực hiện
Nâng cao nhận thức	Phát triển nhận thức về khả năng và những giới hạn của AI.
Minh bạch	Minh bạch và báo cáo chính xác việc sử dụng các công cụ AI hỗ trợ.
Tôn vinh sáng tạo tự thân	Nêu bật sự sáng tạo tự thân trong quá trình giải quyết vấn đề cùng AI.
Luôn luôn liêm chính	Luôn tuân thủ liêm chính học thuật, trích dẫn nguồn minh bạch rõ ràng.
Trau dồi sự hiểu biết	Trau dồi sự hiểu biết, nâng cao mức độ linh hoạt thay vì lệ thuộc AI.
Khai báo sử dụng AI	Đính kèm phụ lục dùng AI, giải trình về cách thức và lí do sử dụng.
Phân định ranh giới	Phân biệt và xác định rõ ràng ranh giới giữa việc sử dụng AI hợp lí với lạm dụng, gian lận.

(Nguồn: Adamakis & Rachiotis, 2025)

đồng. Trường hợp không tổ chức báo cáo, giảng viên nên vấn đáp ngắn tại thời điểm nộp bài (Hodges & Kirschner, 2024) để kiểm tra chiều sâu tư duy khi không có AI. Đồng thời, giảng viên có thể yêu cầu sinh viên nộp nhật kí phản tư (Chan & Colloton, 2024; Hodges & Kirschner, 2024) phân tích sự tiến bộ và bài học kinh nghiệm, những trải nghiệm cá nhân gắn liền bối cảnh thực tế mà AI khó mô phỏng. Để giảm tải áp lực chấm thi, giảng viên có thể tổ chức phản biện chéo giữa các sinh viên (Hodges & Kirschner, 2024). Toàn bộ tiến trình này nên được tái thiết dựa trên sáu chiến lược SARPS của Chan & Colloton (2024) (xem Bảng 5).

Khung SARPS tái thiết phương thức kiểm tra qua sáu chiến lược (Bảng 5). Bốn chiến lược đầu đáp ứng lo ngại của giảng viên về năng lực phản biện (83%) và sự lệ thuộc AI của sinh viên (82%) (Digital Education Council, 2025). Chiến lược thứ năm (được 81% giảng viên ủng hộ) ưu tiên đánh giá các kĩ năng thuật toán khó mô phỏng như trí tuệ cảm xúc, sự thấu cảm, làm việc nhóm và lãnh đạo (Digital Education Council, 2025). Chiến lược thứ sáu hướng đến xây dựng văn hóa phản hồi chuyên sâu và bồi dưỡng tư duy học tập suốt đời. Sáu chiến lược SARPS được cụ thể hóa thành chín loại hình đánh giá (xem Bảng 6), kết hợp thể mạnh xử lí dữ liệu của công nghệ với nhận thức

Bảng 5: Chiến lược tái thiết SARPS đánh giá người học trong kỉ nguyên AI

TT	Chiến lược	Hành động thực thi
1	Tích hợp đa phương pháp	Đa dạng hóa nguồn đánh giá để loại bỏ thiên kiến: 1) AI kiểm tra nhanh; 2) Phản biện chéo đồng cấp; 3) Giảng viên đánh giá.
2	Đánh giá xác thực	Chuyển trọng tâm sang giải quyết vấn đề/tình huống thực tiễn.
3	Liêm chính và chân thực	Đào tạo năng lực AI; quy định ranh giới giữa hỗ trợ và gian lận; đánh giá đồng cấp để tăng minh bạch và tôn vinh tư duy nguyên bản.
4	Chấp nhận AI làm đối tác học tập	Khuyến khích hợp tác với AI để tìm lỗi, phản biện ngược kết quả thuật toán.
5	Đánh giá kĩ năng mềm AI khó sao chép	Tập trung đánh giá yếu tố AI chưa thể vượt qua (tư duy phản biện, đạo đức, thấu cảm, kĩ năng làm việc nhóm và lãnh đạo).
6	Ưu tiên phản hồi hơn điểm số	Chuyển từ chấm điểm tĩnh sang phản hồi theo thời gian thực từ AI và phản hồi chuyên sâu từ giảng viên nhằm giảm áp lực điểm số; xây dựng tư duy học tập suốt đời.

(Nguồn: Chan & Colloton, 2024)

Bảng 6: Các loại hình đánh giá người học có tích hợp AI

TT	Loại hình đánh giá	Yêu cầu ra đề
1	Đánh giá dựa trên hiệu suất	Yêu cầu tạo sản phẩm thực, có dùng AI thu thập dữ liệu thô nhưng người học tự giải thích số liệu và ra quyết định.
2	Đánh giá được cá nhân hóa theo bối cảnh	Thiết kế đề buộc liên hệ lí thuyết với trải nghiệm cá nhân hoặc thực tiễn địa phương.
3	Đánh giá năng lực lấy con người làm trung tâm	Ra đề mô phỏng tình huống làm việc nhóm/giải quyết khủng hoảng để đo lường kĩ năng mềm (thấu cảm, giao tiếp, khả năng chịu áp lực).
4	Đánh giá quan hệ hợp tác giữa người và AI	Ra đề mở bắt buộc dùng AI; nộp kèm câu lệnh (prompt), kết quả thô và phần phản biện, tinh chỉnh tự thân.
5	Đánh giá dựa trên dự án hoặc tình huống	Xây dựng các đề án phức tạp kéo dài nhiều tuần, buộc sinh viên phải thu thập dữ liệu sơ cấp.
6	Đánh giá thích ứng do AI tạo ra có giới hạn về thời gian	Dùng AI tạo ngân hàng câu hỏi phân hóa độ khó; thi vấn đáp hoặc làm bài thời gian siêu ngắn để hạn chế tra cứu AI.
7	Đánh giá siêu nhận thức	Đánh giá qua nhật kí phản tư thay vì đáp án cuối cùng.
8	Đánh giá tác động đạo đức & xã hội	Cung cấp sản phẩm do AI tạo sẵn; yêu cầu phản biện ngược và chỉ ra lỗi của AI.
9	Đánh giá hồ sơ học tập suốt đời	Xây dựng hồ sơ lưu trữ bản nháp và lịch sử chỉnh sửa để chứng minh sự tiến bộ xuyên suốt học phần.

(Nguồn: Chan & Colloton, 2024)

đặc thù của người học (Chan & Colloton, 2024). Hệ thống này không nhằm đa dạng hóa các hình thức chấm điểm sản phẩm cuối khóa, mà là công cụ đo lường chuỗi hành vi kiến tạo tri thức dài hạn, giúp bộc lộ năng lực thực chất và khắc phục hạn chế của việc phân loại cơ học dựa trên điểm số tĩnh (xem Bảng 6).

Khi AI dễ dàng thay thế các bậc nhận thức thấp trong thang Bloom (ghi nhớ, hiểu, vận dụng), sự

chuyển tiếp giữa ba giai đoạn của khung R-O-E lại buộc sinh viên phải vận dụng tư duy bậc cao (phân tích, đánh giá, sáng tạo). Đồng thời, cơ chế phản tư liên tục khi tương tác với AI trong khung R-O-E giúp phát triển “siêu nhận thức” (Metacognition) - tức năng lực thấu cảm và kiểm soát luồng tư duy mà thuật toán chưa thể vượt qua. Khung R-O-E giải quyết triệt để ba thách thức cốt lõi của giáo dục đại học trong kỉ nguyên AI. Thứ nhất, xóa bỏ “Hộp đen tri thức” (Knowledge black box) bằng cách chuyển

Bảng 7: Ma trận tích hợp khung R-O-E, nguyên tắc ETHICAL và chiến lược SARPS

Khung R-O-E	Nguyên tắc đạo đức ETHICAL	Chiến lược tái thiết SARPS
R - Research (Nghiên cứu)	- Tôn vinh sáng tạo tự thân - Trau dồi sự hiểu biết	- Tích hợp đa phương pháp - Chấp nhận AI làm đối tác học tập
O - Operation (Thực thi)	- Minh bạch - Luôn luôn liên chính - Khai báo sử dụng AI	- Liêm chính và chân thực
E - Evaluation (Đánh giá)	- Nâng cao nhận thức - Phân định ranh giới	- Đánh giá xác thực - Đánh giá kĩ năng mềm AI khó sao chép - Ưu tiên phản hồi hơn điểm số

(Nguồn: Nhóm tác giả tổng hợp và đề xuất, 2026)

từ kiểm duyệt sản phẩm vốn dễ bị thao túng bởi công nghệ sang kiểm soát quá trình, buộc người học minh bạch hóa dòng tư duy và tương tác thực chất. Thứ hai, khắc phục sự bất công trong đánh giá bằng cách thay thế công cụ quét AI thụ động, thiếu chính xác bằng các điểm chạm trực tiếp như vấn đáp nhanh và nhật ký phản tư. Thứ ba, giảm tải áp lực cho giảng viên nhờ kết hợp phản biện chéo đồng cấp và tự động hóa theo dõi tiến độ qua công nghệ nhúng (Embedded AI) (xem Bảng 7).

Khác với những nghiên cứu trước đây vốn tập trung vào các giải pháp kỹ thuật rời rạc, kết quả nghiên cứu này có tính mới ở việc cấu trúc hóa toàn bộ hành vi học tập thành một hệ sinh thái kiểm soát quá trình khép kín. Hệ sinh thái này tích hợp đồng bộ ba chiều: 1) Khung R-O-E định hướng hành vi; 2) Nguyên tắc ETHICAL định hình thái độ; 3) Chiến lược SARPS tái thiết phương pháp kiểm tra. Sự kết hợp này xóa bỏ khoảng trống vận hành, đảm bảo mỗi bước đi của người học đều có hành lang đạo đức giám sát và chiến lược sư phạm hỗ trợ. Nhờ đó, khung R-O-E tích hợp chuyển trọng tâm từ dò tìm gian lận sản phẩm sang chủ động kiến tạo môi trường kháng AI tự thân trong suốt quá trình học tập.

4. Thảo luận

So với các mô hình khác, khung R-O-E vừa có tính kế thừa vừa sở hữu những đóng góp khoa học riêng biệt. R-O-E chia sẻ tầm nhìn chung với “Thang đánh giá AI” (AI Assessment Scale - AIAS) của Perkins và cộng sự (2024) ở việc từ bỏ tư duy cấm đoán để thiết lập ranh giới sử dụng công nghệ minh bạch. Tuy nhiên, nếu AIAS chỉ gán nhãn mức độ cho phép sử dụng AI đối với từng nhiệm vụ riêng lẻ, R-O-E lại đề xuất một quy trình đồng bộ xuyên suốt tiến trình học tập, giúp kiểm soát sự phát triển tư duy thực chất. Đồng thời, R-O-E khắc phục hạn chế về nguồn lực của mô hình “Đánh giá song song” (Assessment Twins) của Roe và cộng sự (2026). Thay vì thiết kế hai nhiệm vụ độc lập để đối chiếu chéo như mô hình “Đánh giá song song” sẽ tiêu tốn nhiều thời gian, R-O-E tích hợp trực tiếp hoạt động phản tư và dấu vết kỹ thuật số vào duy nhất một bài đánh giá, giúp tối ưu hóa chi phí vận hành.

Mặc dù nhiều ưu điểm song thách thức lớn nhất khi áp dụng khung R-O-E là gánh nặng thời gian và khối lượng công việc cho giảng viên khi theo dõi vết tích số và vấn đáp trực tiếp, đặc biệt ở các lớp quy mô lớn. Thực tế, 40% giảng viên ngại dùng AI vì thiếu thời gian và 38% cho rằng, AI gia tăng áp lực công việc (Shaw và cộng sự, 2025). Đáng chú ý, 46%

giảng viên vẫn chấm điểm thủ công và 71% trong số đó khẳng định việc dò tìm gian lận làm tăng mạnh tải trọng công việc (Shaw và cộng sự, 2025). Để khắc phục, đánh giá đồng cấp từ sinh viên (thuộc nhóm chiến lược SARPS) nên được ứng dụng nhằm chia sẻ áp lực chấm điểm cho giảng viên; đồng thời, ứng dụng Embedded AI tích hợp vào hệ thống quản lý học tập LMS để tự động phân tích dữ liệu vi mô về tiến độ bản thảo. Giải pháp này giúp nâng mức độ hài lòng của giảng viên lên 27 điểm - vượt trội so với 2 điểm ở nhóm đối chứng (Shaw và cộng sự, 2025). Ngoài ra, các nền tảng Low-code/No-code (chiếm 13% công cụ AI tự phát triển) cho phép người dạy tự thiết kế công cụ theo dõi tự động phù hợp với đặc thù học phần mà không cần kỹ năng lập trình cao (UNESCO, 2025).

Bên cạnh đó, việc triển khai khung R-O-E có nguy cơ nới rộng khoảng cách số giữa các nhóm sinh viên có điều kiện kinh tế đối lập. Nếu thiếu chính sách hỗ trợ hạ tầng công nghệ cho người học yếu thế, bất bình đẳng vô hình sẽ phát sinh. Đồng thời, hạn chế về năng lực số của giảng viên cũng là rào cản lớn, bởi chỉ khi thấu hiểu giới hạn thuật toán, người dạy mới có thể thiết kế được các đề bài kháng AI hiệu quả. Do đó, các trường đại học cần tổ chức tập huấn nâng cao năng lực số cho cả giảng viên và sinh viên. Giảng viên cần được trang bị trang bị kỹ năng thiết kế đề kháng AI và đánh giá quá trình; trong khi sinh viên cần được đào tạo về kỹ thuật thiết lập câu lệnh (Prompt) chuẩn gắn với đạo đức học thuật số.

Mặt khác, sự thiếu vắng quy chế quản lý AI cũng tạo ra áp lực thể chế rất lớn cho người dạy. Theo Digital Education Council (2025), 80% giảng viên phản ánh nhà trường thiếu hướng dẫn toàn diện, 78% cho rằng tài nguyên đào tạo AI yếu kém, và chỉ 4% nhận được hướng dẫn cụ thể. Đáng chú ý, chỉ 28% trường đại học đã ban hành chính sách chính thức về AI (Shaw và cộng sự, 2025). Thiếu khung pháp lý vĩ mô, việc chuyển đổi sang đánh giá quá trình sẽ chỉ mang tính tự phát. Vì vậy, các trường đại học cần cấp thiết thể chế hóa quy định sử dụng AI, trang bị công cụ và cung cấp hướng dẫn thực tiễn nhằm trao quyền cho giảng viên tự tin thực thi khung R-O-E.

5. Kết luận và khuyến nghị

5.1. Kết luận

Để ứng phó với sự suy giảm hiệu lực của phương pháp đánh giá sản phẩm trong kỷ nguyên AI, nghiên cứu đề xuất dịch chuyển sang đánh giá quá trình theo khung R-O-E, tích hợp đồng bộ với nguyên tắc

đạo đức ETHICAL và chiến lược tái thiết SARPS. Việc buộc sinh viên minh bạch quá trình kiến tạo tri thức thông qua vết tích số, nhật ký phản tư, phụ lục khai báo AI và vấn đáp mang lại giá trị kép. Về mặt lí luận, khung R-O-E tạo ra sự giao thoa giữa thang nhận thức Bloom và triết lí “Đánh giá như một quá trình học tập”. Về mặt thực tiễn, việc chuyển đổi này không chỉ cung cấp cho các trường đại học một lộ trình khả thi để bảo vệ liêm chính học thuật số mà còn giúp bồi dưỡng năng lực siêu nhận thức, tư duy phản biện và năng lực làm chủ công nghệ cho thế hệ người học tương lai. Dù có nhiều đóng góp thiết thực, nghiên cứu hiện tại vẫn chỉ dừng ở phương pháp tổng quan tài liệu với minh chứng từ dữ liệu thứ cấp mà chưa kiểm chứng bằng các mô hình định lượng với dữ liệu khảo sát. Do đó, các nghiên cứu thực nghiệm trong tương lai có thể thu thập dữ liệu sơ cấp và ứng dụng các mô hình định lượng để đo lường hiệu quả của đánh giá quá trình theo khung R-O-E hay đánh giá mức độ tương thích và hiệu quả giảm tải khối lượng công việc cho giảng viên khi tích hợp Embedded AI vào hệ thống quản lí học tập.

5.2. Khuyến nghị

Để hiện thực hóa khung đánh giá quá trình R-O-E cần sự phối hợp đồng bộ từ cấp quản lí vĩ mô đến các chủ thể vi mô gồm:

- *Đối với cơ quan quản lí Nhà nước (QLNN)*: 1) Thiết lập kho lưu trữ vết tích số học thuật quốc gia: Nên yêu cầu các trường đại học lưu vết tích số quá trình tạo ra sản phẩm khoa học để phục vụ kiểm định và hậu kiểm; 2) Hoàn thiện hành lang pháp lí về quyền sở hữu trí tuệ đối với sản phẩm có yếu tố AI: Lấy Luật Trí tuệ nhân tạo 2025 làm khung quản lí và bổ sung Luật Sở hữu trí tuệ hiện hành thêm các hướng dẫn chi tiết về quyền tác giả và định mức đóng góp của AI nhằm tạo cơ sở cho người dùng sử dụng AI hợp pháp và có trách nhiệm.

- *Đối với các trường đại học*: 1) Ban hành bộ quy tắc đạo đức học thuật ứng dụng AI theo nguyên tắc ETHICAL nhằm quy định minh bạch ranh giới giữa sự hỗ trợ công nghệ hợp lệ với hành vi gian lận, lạm dụng; 2) Thể chế hóa phương pháp đánh giá quá trình theo khung R-O-E và thực thi chiến lược tái thiết SARPS, đưa chín loại hình đánh giá tích hợp AI

vào quy chế đào tạo đối với các học phần như đồ án, chuyên đề, khóa luận, báo cáo thực tập tốt nghiệp...; 3) Nâng cấp hạ tầng công nghệ để triển khai theo dõi vết tích số một cách đồng bộ mà không gây quá tải; 4) Chuẩn hóa quy trình bảo mật dữ liệu, tự động hóa việc xóa dữ liệu sau khi kết thúc học phần nhằm tuân thủ tuyệt đối nguyên tắc “lưu trữ trong thời hạn cần thiết” của Luật Bảo vệ dữ liệu cá nhân 2025; 5) Phát triển năng lực số cho giảng viên và sinh viên. Trọng tâm tập huấn đối với giảng viên là kĩ năng thiết kế đề kháng AI trong khi đối với sinh viên là kĩ thuật điều vấn AI (tức kĩ thuật thiết lập câu lệnh AI).

Theo Chan & Colloton (2024), một câu lệnh chuẩn phải đủ bảy thành phần gồm: 1) Chỉ thị: ra yêu cầu rõ ràng; 2) Ngữ cảnh: cung cấp bối cảnh; 3) Vai trò: gán cho AI một vai trò cụ thể; 4) Súc tích: văn phong ngắn gọn, ngôn từ không gây nhiễu hay hiểu lầm; 5) Ví dụ cụ thể: đưa ra mẫu tham chiếu cho AI hiểu sâu sắc vấn đề; 6) Định dạng: yêu cầu cách trình bày cụ thể; 7) Sự điều chỉnh: tinh chỉnh dần câu lệnh dựa trên các câu trả lời trước đó.

- *Đối với giảng viên*: 1) Thay đổi tư duy cấm đoán dùng AI sang dẫn dắt và cộng tác; đồng thời, chủ động nâng cao năng lực số của bản thân và tự nguyện ứng dụng linh hoạt khung đánh giá quá trình R-O-E nhằm thiết kế hiệu quả các đề bài kháng AI; 2) Đa dạng hóa phương thức đánh giá để loại bỏ thiên kiến; trong đó, ưu tiên kết hợp phản biện chéo đồng cấp giữa các sinh viên cùng nhóm hướng dẫn.

- *Đối với sinh viên*: 1) Đảm bảo liêm chính học thuật số theo khung nguyên tắc ETHICAL, tức nhận thức rõ AI chỉ là công cụ hỗ trợ không thể thay thế quá trình tư duy nguyên bản và thực hiện phụ lục khai báo sử dụng AI; 2) Chủ động thích nghi với khung đánh giá quá trình R-O-E, lưu trữ vết tích số; đồng thời, rèn luyện kĩ năng viết nhật ký phản tư và năng lực vấn đáp.

Tuyên bố xung đột lợi ích: Nhóm tác giả chịu trách nhiệm hoàn toàn về mọi nhận định, đánh giá trong bài báo và cam kết đây là nghiên cứu độc lập, không có bất kì xung đột lợi ích hay vi phạm quyền lợi hợp pháp nào của các bên liên quan.

Tài liệu tham khảo

Adamakis, M. & Rachiotis, T. (2025). Artificial intelligence in higher education: A state-of-the-art overview of pedagogical integrity, artificial

intelligence literacy, and policy integration. *Encyclopedia*, 5(4), pp.180-205. <https://doi.org/10.3390/encyclopedia5040180>

- AlBlooshi, S. (2026). Artificial intelligence in higher education, opportunities, and challenges: a review. *Frontiers in Education*, 10: 1683968. <https://doi.org/10.3389/feduc.2025.1683968>
- Anderson, L. W. & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition*. New York: Longman.
- Chan, C. K. Y. & Colloton, T. (2024). *Generative AI in higher education: The ChatGPT effect*. Routledge. <https://doi.org/10.4324/9781003459026>
- 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), pp.228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- Digital Education Council - DEC. (2025). *Global AI faculty survey 2025*. DEC.
- Earl, L. (2003). *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning*. Thousand Oaks, CA: Corwin Press.
- Freeman, J. (2025). Student generative AI survey 2025. *Higher Education Policy Institute: London, UK*. <https://www.hepi.ac.uk/wp-content/uploads/2025/02/HEPI-Policy-Note-61-2.pdf>
- Hodges, C. B. & Kirschner, P. A. (2024). Innovation of instructional design and assessment in the age of generative artificial intelligence. *TechTrends*, 68(1), pp.195-199. <https://doi.org/10.1007/s11528-023-00926-x>
- Jacques, P. H., Moss, H. K. & Garger, J. (2024). A synthesis of AI in higher education: Shaping the future. *Journal of Behavioral and Applied Management*, 24(2), pp.103-111. <https://doi.org/10.21818/001c.122146>
- Khlaif, Z. N., Alkhouk, W. A., Salama, N. & Abu Eideh, B. (2025). Redesigning assessments for AI-enhanced learning: A framework for educators in the generative AI era. *Education Sciences*, 15(2), pp.174-201. <https://doi.org/10.3390/educsci15020174>
- Kofinas, A. K., Tsay, C. H. H. & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, 56(6), pp.2522-2549. <https://doi.org/10.1111/bjet.13585>
- Koh, E. & Doroudi, S. (2023). Learning, teaching, and assessment with generative artificial intelligence: towards a plateau of productivity. *Learning: Research and Practice*, 9(2), pp.109-116. <https://doi.org/10.1080/23735082.2023.226408>
- Liang, W., Yuksekogonul, M., Mao, Y., Wu, E. & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), 100779. <https://doi.org/10.1016/j.patter.2023.100779>
- Lodge, J., Howard, S., Bearman, M. L., Dawson, P. & Associates (2023). *Assessment Reform for The Age of Artificial Intelligence*. Tertiary Education Quality and Standards Agency. <https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf>
- Perkins, M., Furze, L., Roe, J. & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, 21(6), pp.49-66. <https://doi.org/10.53761/q3azde36>
- Roe, J., Perkins, M. & Gray, L. (2026). Assessment twins: An approach for strengthening assessment validity in the age of generative AI. *Journal of Applied Learning & Teaching*, 9(2), pp.1-15. <https://doi.org/10.37074/jalt.2026.9.2.3>
- Shaw, C., Dorn, H., Martin, S., Hay, L., Janson, N. & Bryant, G. (2025, June). *Time for Class - 2025*. Tyton Partners.
- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human resource development review*, 4(3), pp.356-367. <https://doi.org/10.1177/1534484305278283>
- United Nations Educational, Scientific and Cultural Organization - UNESCO. (2025). *AI and the future of education: Disruptions, dilemmas and directions*. <https://doi.org/10.54675/KECK1261>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., ... & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19(1), pp.26-65. <https://doi.org/10.1007/s40979-023-00146-z>
- Whittemore, R. & Knafl, K. (2005). The integrative review: updated methodology. *Journal of advanced nursing*, 52(5), pp.546-553. <https://doi.org/10.1111/j.1365-2648.2005.03621.x>