ĐẠO ĐỨC AI TRONG GIÁO DỤC NGHỀ NGHIỆP: THÁCH THỨC VÀ GIẢI PHÁP

Trần Phan Phú Sơn
Trường Trung cấp Thanh Bình

Tóm tắt: *Sự phát triển nhanh chóng của trí tuệ nhân tạo (AI) đang tạo ra những thay đổi sâu rộng trong giáo dục nghề nghiệp (GDNN), từ thiết kế bài giảng, hỗ trợ dạy học, đánh giá người học đến tư vấn nghề nghiệp và quản lý đào tạo. Bên cạnh các lợi ích rõ rệt, việc ứng dụng AI trong GDNN cũng đặt ra nhiều vấn đề đạo đức như bảo mật dữ liệu cá nhân, thiên kiến thuật toán, minh bạch trong ra quyết định, trách nhiệm pháp lý khi AI hỗ trợ đánh giá, nguy cơ lệ thuộc AI và bất bình đẳng trong tiếp cận. Trên cơ sở nghiên cứu lý luận, bài báo phân tích khái niệm đạo đức AI trong GDNN, xác định những thách thức cốt lõi và đề xuất một số giải pháp nhằm bảo đảm việc ứng dụng AI diễn ra an toàn, công bằng và nhân văn. Kết quả nghiên cứu cho thấy đạo đức AI trong GDNN không chỉ là yêu cầu kỹ thuật mà còn là vấn đề quản trị giáo dục, đòi hỏi sự phối hợp giữa nhà trường, giáo viên, người học và cơ quan quản lý. Từ đó, bài báo đề xuất một số định hướng về xây dựng quy định nội bộ, năng lực đạo đức AI cho giáo viên, tăng cường kiểm chứng đầu ra AI, bảo vệ dữ liệu và bảo đảm quyền chủ động của người học trong môi trường đào tạo nghề.*

Từ khóa: *Trí tuệ nhân tạo, đạo đức AI, giáo dục nghề nghiệp, quản trị AI, AI có trách nhiệm.*

AI ETHICS IN VOCATIONAL EDUCATION: CHALLENGES AND SOLUTIONS

Abstract: *The rapid development of artificial intelligence (AI) is driving profound changes in vocational education and training (VET), spanning areas such as lesson design, instructional support, learner assessment, career counseling, and training management. Alongside its clear benefits, the application of AI in VET raises significant ethical issues, including personal data privacy, algorithmic bias, decision-making transparency, legal liability in AI-assisted assessment, the risk of AI dependency, and inequalities in access. Drawing on theoretical research, this paper analyzes the concept of AI ethics within VET, identifies core challenges, and proposes solutions to ensure that AI implementation is safe, equitable, and human-centric. The findings indicate that AI ethics in VET is not merely a technical requirement but also a matter of educational governance, necessitating collaboration among institutions, educators, learners, and regulatory bodies. Consequently, the paper outlines directions for establishing internal regulations, enhancing teachers' AI ethics competencies, strengthening the validation of AI outputs, ensuring data protection, and safeguarding learner agency within the vocational training environment.*

Keywords: *Artificial intelligence, AI ethics, vocational education, AI governance, responsible AI.*

Nhận bài: 12/06/2026 Phản biện: 12/07/2026 Duyệt đăng: 15/07/2026

I. ĐẶT VẤN ĐỀ

Trí tuệ nhân tạo (AI) đang làm thay đổi hoạt động dạy học, đánh giá và quản trị giáo dục, tạo điều kiện cá nhân hóa học tập và thúc đẩy thực hiện Mục tiêu Phát triển Bền vững (UNESCO, 2021a). Sự phát triển của AI tạo sinh mở rộng khả năng hỗ trợ thiết kế học liệu, đánh giá và nghiên cứu, đồng thời cũng đặt ra nhiều yêu cầu về minh bạch, trách nhiệm và quản trị AI trong giáo dục (European Commission, 2019; UNESCO, 2025).

Trong giáo dục nghề nghiệp (GDNN), AI đang được ứng dụng ngày càng rộng rãi nhằm hỗ trợ cá nhân hóa học tập, hỗ trợ phát triển kỹ năng nghề, nâng cao hiệu quả đào tạo. Tuy nhiên, hiệu quả triển khai phụ thuộc vào bối cảnh của cơ sở đào tạo, năng lực của người sử dụng và cơ chế quản trị phù hợp. Đồng thời làm nảy sinh các vấn đề về quyền riêng tư, thiên kiến thuật toán, minh bạch và trách nhiệm giải trình (Deutscher et al., 2026). Tại Việt Nam, các cơ sở GDNN ngày càng quan tâm đến việc ứng dụng AI nhằm nâng cao chất lượng đào tạo và hiệu quả quản lý. Đồng thời, việc Quốc hội ban hành Luật Trí tuệ nhân tạo (QUỐC HỘI, 2025) đã tạo cơ sở pháp lý ban đầu cho phát triển và quản trị AI có trách nhiệm.

Mặc dù, nhiều nghiên cứu đã phân tích các vấn đề đạo đức AI trong giáo dục, phần lớn các nghiên cứu chủ yếu tập trung vào giáo dục phổ thông hoặc giáo dục đại học (Gouseti et al., 2025). Đối với GDNN, các nghiên cứu hiện có chủ yếu đề cập đến ứng dụng AI trong đào tạo và chuyển đổi số nhưng chưa phân tích một cách hệ thống các thách thức đạo đức gắn với đặc thù đào tạo kỹ năng thực hành, đánh giá năng lực nghề và hợp tác với doanh nghiệp (Deutscher et al., 2026; Wiese et al., 2025). Tại Việt Nam, Huyền (Huyền & Huyen, 2025) đã chỉ ra cơ hội và thách thức của AI trong quản lý đào tạo nghề, song chưa tiếp cận dưới góc độ đạo đức AI. Vì vậy, vẫn thiếu một khung phân tích tổng hợp về các thách thức đạo đức AI và các giải pháp quản trị phù hợp với bối cảnh GDNN.

Xuất phát từ khoảng trống đó, bài báo nhằm hệ thống hóa cơ sở lý luận về đạo đức AI trong

GDNN, phân tích các nhóm thách thức đạo đức chủ yếu và đề xuất các giải pháp thúc đẩy ứng dụng AI theo hướng an toàn, minh bạch, có trách nhiệm và lấy con người làm trung tâm.

II. PHƯƠNG PHÁP NGHIÊN CỨU

Bài báo được thực hiện theo hướng nghiên cứu lý luận, sử dụng phương pháp tổng quan tài liệu kết hợp phân tích nội dung nhằm hệ thống hóa các vấn đề liên quan đến đạo đức AI trong GDNN. Nguồn tài liệu gồm các công trình khoa học quốc tế và văn kiện của các tổ chức uy tín như UNESCO, OECD và European Commission, được lựa chọn dựa trên mức độ liên quan đến đạo đức AI, giáo dục và GDNN, chủ yếu công bố trong giai đoạn 2019–2026. Trên cơ sở thu thập, phân tích và tổng hợp tài liệu, nghiên cứu xác định các nhóm thách thức đạo đức AI trong GDNN và đề xuất các giải pháp thúc đẩy ứng dụng AI có trách nhiệm. Do sử dụng phương pháp tổng quan tài liệu, nghiên cứu chưa thực hiện khảo sát thực nghiệm tại các cơ sở GDNN. Vì vậy, các kết luận chủ yếu mang tính khái quát và định hướng.

III. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Khái niệm đạo đức AI trong giáo dục nghề nghiệp

Đạo đức AI là hệ thống các giá trị và nguyên tắc định hướng việc phát triển, triển khai và sử dụng AI theo hướng tôn trọng quyền con người và mang lại lợi ích cho xã hội. UNESCO (2021b) và European Commission (2019) nhấn mạnh các nguyên tắc cốt lõi gồm minh bạch, công bằng, bảo vệ quyền riêng tư, trách nhiệm giải trình và giám sát của con người. Waelen (2022) cho rằng các nguyên tắc này cần được xem xét trên cơ sở tác động của AI đối với con người và xã hội.

Trong giáo dục nghề nghiệp, AI được ứng dụng trong cá nhân hóa học tập, mô phỏng kỹ năng, đánh giá năng lực và hỗ trợ đào tạo. Vì vậy, đạo đức AI được hiểu là hệ thống các giá trị và nguyên tắc quản trị nhằm bảo đảm AI được sử dụng theo hướng lấy con người làm trung tâm, minh bạch, công bằng và có trách nhiệm, trong đó AI chỉ đóng vai trò hỗ trợ, không thay thế giáo viên. Trên cơ sở đó, bài báo xây dựng khung phân tích gồm năm nhóm thách thức đạo đức AI trong giáo dục nghề nghiệp.

3.2. Các nhóm thách thức đạo đức AI trong giáo dục nghề nghiệp

3.2.1. Thách thức về dữ liệu và quyền riêng tư

Một trong những thách thức lớn nhất của AI trong GDNN là bảo vệ và quản lý dữ liệu người học. Các hệ thống AI thường sử dụng dữ liệu về kết quả học tập, hành vi học tập, sản phẩm thực hành và các thông tin cá nhân để hỗ trợ cá nhân hóa học tập. Quyền riêng tư và bảo vệ dữ liệu là nguyên tắc cốt lõi của đạo đức AI, yêu cầu việc thu thập và xử lý dữ liệu phải minh bạch, đúng mục đích và được sự đồng thuận của chủ thể dữ liệu (UNESCO, 2021b). Tương tự, European Commission (European Commission, 2019) nhấn mạnh dữ liệu phải được quản lý hợp pháp, minh bạch và bảo đảm quyền kiểm soát của người dùng. Kết quả khảo sát của Marín và cộng sự (Marín et al., 2025) cũng cho thấy hơn một nửa giảng viên và gần một nửa sinh viên lo ngại về quyền riêng tư và an toàn dữ liệu khi sử dụng AI trong giáo dục.

Trong GDNN, dữ liệu người học còn được chia sẻ giữa nhà trường và doanh nghiệp trong quá trình thực tập, đánh giá kỹ năng hoặc tuyển dụng. Vì vậy, cần xác định rõ quyền truy cập, mục đích sử dụng và sự đồng thuận của người học nhằm tránh lạm dụng dữ liệu, bảo vệ quyền lợi của người học và duy trì niềm tin đối với các hệ thống AI.

3.2.2. Thách thức về thiên kiến và công bằng

AI học từ dữ liệu nên có thể kế thừa và khuếch đại các thiên kiến sẵn có. Trong GDNN, điều này có thể dẫn đến thiếu công bằng khi AI được sử dụng để tư vấn nghề nghiệp, phân luồng, gợi ý kỹ năng hoặc hỗ trợ đánh giá người học. Nếu dữ liệu huấn luyện không đại diện đầy đủ, AI có thể tạo ra các khuyến nghị thiên lệch đối với một số nhóm người học, làm giảm hiệu quả của quá trình cá thể hóa đào tạo.

UNESCO (UNESCO, 2021b) khẳng định AI cần được thiết kế theo nguyên tắc công bằng, không phân biệt đối xử và giảm thiểu thiên kiến dữ liệu. Tương tự, European Commission (European Commission, 2019) nhấn mạnh các hệ thống AI phải bảo đảm tính đa dạng và hạn chế thiên lệch trong toàn bộ quá trình phát triển và sử dụng. Đồng thời, Khung năng lực AI cho học sinh (UNESCO, 2024a) cũng yêu cầu người học có khả năng nhận diện các thiên kiến liên quan đến giới tính, dân tộc, văn hóa và các hình thức phân biệt đối xử trong các công cụ AI.

Trong GDNN, thiên kiến AI không chỉ ảnh hưởng đến kết quả học tập mà còn tác động đến tư vấn nghề nghiệp, đánh giá kỹ năng và cơ hội việc làm. Nếu các hệ thống AI được huấn luyện dựa trên dữ liệu lịch sử chưa đa dạng, chúng có thể củng cố định kiến giới, vùng miền hoặc thành tích học tập, làm hạn chế cơ hội tiếp cận của các nhóm yếu thế. Điều này đi ngược lại mục tiêu bảo đảm công bằng và phát triển nguồn nhân lực chất lượng cao trong GDNN.

3.2.3. Thách thức về minh bạch và trách nhiệm

Một vấn đề đạo đức quan trọng khác là tính minh bạch và trách nhiệm của AI. Trong nhiều trường hợp, người dùng chỉ nhận được kết quả từ hệ thống mà không hiểu rõ vì sao AI đưa ra khuyến nghị hoặc đánh giá như vậy. Điều này tạo ra nguy cơ "hộp đen" trong giáo dục, nhất là khi AI tham gia vào chấm điểm, phân tích năng lực hay tư vấn nghề nghiệp.

UNESCO (UNESCO, 2021b) cho rằng các hệ thống AI cần bảo đảm tính minh bạch, khả năng giải thích và trách nhiệm giải trình trong toàn bộ vòng đời phát triển và sử dụng. Tương tự, European Commission (European Commission, 2019) xác định tính minh bạch, khả năng giải thích và truy xuất nguồn gốc là những yêu cầu cốt lõi của AI đáng tin cậy. Thêm vào đó, nghiên cứu của Marín và cộng sự (Marín et al., 2025) cho thấy nhiều giảng viên và sinh viên còn lo lắng về tính minh bạch trong cách AI vận hành và ra quyết định, từ đó nhấn mạnh cần tăng cường cơ chế kiểm tra và trách nhiệm giải trình. UNUESCO (UNESCO, 2024b) cũng khẳng định rằng trách nhiệm đối với các quyết định có sự hỗ trợ của AI luôn thuộc về con người.

Trong GDNN, nếu AI gợi ý một học viên không phù hợp với ngành học hoặc chấm điểm thực hành chưa chính xác, cần phải xác định rõ ai chịu trách nhiệm: người lập trình, đơn vị cung cấp nền tảng, giáo viên hay nhà trường. Vấn đề này cho thấy đạo đức AI phải đi cùng cơ chế trách nhiệm giải trình. Không thể để AI trở thành một chủ thể "vô hình" nhưng lại ảnh hưởng lớn đến đời sống học tập và tương lai nghề nghiệp của người học.

3.2.4. Thách thức về lệ thuộc AI và vai trò trung tâm của con người

Việc AI ngày càng được sử dụng trong thiết kế học liệu, đánh giá và hỗ trợ học tập đã làm gia tăng nguy cơ lệ thuộc vào AI nếu giáo viên và người học sử dụng AI như một chủ thể thay vì một công cụ hỗ trợ. UNESCO (UNESCO, 2024b) nhấn mạnh rằng AI cần được phát triển theo định hướng lấy con người làm trung tâm, trong đó AI đóng vai trò hỗ trợ thay vì thay thế năng lực nhận thức, quyền tự quyết và trách nhiệm của con người. Các nghiên cứu thực nghiệm cũng cho thấy việc phụ thuộc quá mức vào AI làm suy giảm quyền tự chủ học tập và năng lực tư duy phản biện của người học (Wiese et al., 2025).

Trong GDNN, AI hỗ trợ cá nhân hóa học tập, mô phỏng kỹ năng và nâng cao hiệu quả đào tạo. Tuy nhiên, Deutscher và cộng sự (Deutscher et al., 2026) cho rằng chưa có đủ bằng chứng để khẳng định AI có thể thay thế giáo viên trong hướng dẫn thực hành, đánh giá kỹ năng và hình thành năng lực nghề nghiệp. Vì vậy, AI chỉ nên được sử dụng như công cụ hỗ trợ, trong khi quyền quyết định và trách nhiệm đào tạo vẫn thuộc về giáo viên.

3.2.5. Thách thức về bình đẳng trong tiếp cận

Bình đẳng trong tiếp cận AI được xem là một trong những nguyên tắc cốt lõi của đạo đức AI trong giáo dục. Khi AI trở thành một thành phần của quá trình dạy học, sự khác biệt về khả năng tiếp cận hạ tầng số, công cụ AI và năng lực số có thể làm gia tăng khoảng cách giáo dục giữa các nhóm người học và giữa các cơ sở đào tạo.

UNESCO (UNESCO, 2024a) nhấn mạnh rằng các hệ thống AI trong giáo dục phải được thiết kế theo hướng bao trùm, bảo đảm khả năng tiếp cận đối với người học có hoàn cảnh khó khăn, người khuyết tật và các nhóm yếu thế, đồng thời phù hợp với bối cảnh văn hóa và điều kiện địa phương. Nghiên cứu của Metreveli và cộng sự (Metreveli et al., 2025) cho thấy việc triển khai AI trong GDNN có nguy cơ làm gia tăng khoảng cách giữa các nhóm người học nếu thiếu các chính sách hỗ trợ phù hợp về năng lực số và điều kiện tiếp cận.

Đối với GDNN, bất bình đẳng trong tiếp cận AI không chỉ xuất phát từ hạ tầng số mà còn liên quan đến sự khác biệt về ngành nghề đào tạo, năng lực số của giáo viên và điều kiện kinh tế của người học. Sự chênh lệch này có thể dẫn đến khoảng cách về chất lượng đào tạo và cơ hội nghề nghiệp của người học, đi ngược lại mục tiêu công bằng trong GDNN.

Bảng 1: Tổng hợp các nhóm thách thức và định hướng giải pháp trong GDNN

Nhóm thách thức	Biểu hiện trong GDNN	Định hướng giải pháp
Dữ liệu và quyền riêng tư	Thu thập, lưu trữ, chia sẻ dữ liệu người học	Xây dựng quy định nội bộ; bảo vệ dữ liệu
Thiên kiến và công bằng	Thiên lệch trong tư vấn nghề, đánh giá kỹ năng	Kiểm soát dữ liệu; nâng cao nhận thức đạo đức AI
Minh bạch và trách nhiệm	Khó giải thích kết quả AI, khó xác định trách nhiệm	Kiểm chứng đầu ra; giám sát của giáo viên
Lệ thuộc AI và vai trò con người	Phụ thuộc AI trong dạy học và đánh giá	AI hỗ trợ, con người quyết định; bồi dưỡng giáo viên
Bình đẳng trong tiếp cận	Chênh lệch hạ tầng, thiết bị, năng lực số	Bảo đảm tiếp cận công bằng; hỗ trợ nhóm yếu thế

IV. GIẢI PHÁP BẢO ĐẢM ĐẠO ĐỨC AI TRONG GIÁO DỤC NGHỀ NGHIỆP

4.1. Xây dựng quy định nội bộ về sử dụng AI

Các cơ sở GDNN cần xây dựng quy định nội bộ nhằm bảo đảm AI được triển khai theo hướng an toàn, đáng tin cậy và lấy con người làm trung tâm. Theo OECD (OECD, 2019) và UNESCO (UNESCO, 2024a) khuyến nghị quản trị AI dựa trên các nguyên tắc minh bạch, trách nhiệm giải trình, công bằng, giám sát của con người và quản lý rủi ro. Trên cơ sở đó, bài báo đề xuất bốn nhóm quy định nội bộ cho các cơ sở GDNN: (1) quy định nguyên tắc sử dụng AI theo hướng lấy con người làm trung tâm; (2) quy trình kiểm chứng và giám sát đầu ra của AI; (3) xác định trách nhiệm của giáo viên trong giảng dạy, đánh giá và hướng dẫn người học; và (4) quản lý, lưu trữ và bảo vệ dữ liệu người học.

4.2. Nâng cao năng lực đạo đức AI cho giáo viên

Giáo viên cần phát triển đồng thời năng lực công nghệ và năng lực đạo đức AI nhằm sử dụng AI một cách minh bạch, công bằng và có trách nhiệm (UNESCO, 2024b). Đồng thời, các cơ sở GDNN cần xây dựng chương trình bồi dưỡng gắn với đặc thù đào tạo nghề, kết hợp tập huấn, cộng đồng học tập nghề nghiệp và đánh giá năng lực định kỳ. Bảng 2 trình bày khung năng lực đạo đức AI dành cho giáo viên GDNN.

Bảng 2: Đề xuất khung năng lực đạo đức AI dành cho giáo viên GDNN (tổng hợp của tác giả từ UNESCO, Chee, Wiese và Mouta).

Nhóm năng lực	Nội dung cốt lõi	Minh họa trong GDNN
Nhận thức đạo đức AI	Minh bạch, công bằng, trách nhiệm	Lựa chọn công cụ AI phù hợp
Quản trị dữ liệu	Quyền riêng tư, bảo mật	Không đưa dữ liệu người học lên AI công cộng
Đánh giá đầu ra AI	Kiểm chứng, phát hiện thiên kiến	Đối chiếu kết quả AI với nguồn chuyên môn
Tư duy phản biện	Không phụ thuộc AI	Giáo viên quyết định cuối cùng
Thiết kế học tập	Tích hợp AI có trách nhiệm	Thiết kế bài thực hành có AI
Phát triển nghề nghiệp	Học tập liên tục	Tham gia tập huấn và cộng đồng chuyên môn

4.3. Kiểm chứng và giám sát đầu ra của AI

Các cơ sở GDNN cần thiết lập cơ chế kiểm chứng đầu ra của AI nhằm ba03 đảm tính chính xác, minh bạch và trách nhiệm giải trình trong giảng dạy và đánh giá (UNESCO, 2021b). Theo nguyên tắc "AI hỗ trợ, con người quyết định", giáo viên là người thẩm định cuối cùng. Nội dung do AI tạo ra cần được đối chiếu với giáo trình, tiêu chuẩn nghề hoặc nhiều nguồn tin cậy nhằm hạn chế hiện tượng "ảo giác AI" (UNESCO, 2025). Bài báo đề xuất quy trình kiểm chứng đầu ra AI gồm ba bước: (1) xác định và gắn nhãn những nội dung do AI tạo ra; (2) đối chiếu kết quả AI với giáo trình, tiêu chuẩn nghề và các nguồn thông tin đáng tin cậy nhằm kiểm tra tính chính xác và phù hợp; (3) giáo viên xem xét, điều chỉnh và phê duyệt trước khi sử dụng chính thức trong hoạt động dạy học hoặc đánh giá. Quy trình này góp

phần tăng cường tính minh bạch, trách nhiệm giải trình và bảo đảm AI được sử dụng như công cụ hỗ trợ, trong khi quyền quyết định cuối cùng vẫn thuộc về con người.

4.4. Bảo vệ dữ liệu người học

Các cơ sở GDNN cần xây dựng chính sách quản trị dữ liệu theo các nguyên tắc minh bạch, tối thiểu hóa dữ liệu, xác định rõ mục đích sử dụng và bảo đảm sự đồng thuận của người học trong toàn bộ quá trình thu thập, xử lý và chia sẻ dữ liệu (UNESCO, 2021b). Trên cơ sở đó, bài báo đề xuất các cơ sở GDNN xây dựng khung quản trị dữ liệu người học gồm bốn thành phần: (1) chính sách bảo vệ dữ liệu; (2) quy trình thu thập, lưu trữ và chia sẻ dữ liệu; (3) hướng dẫn sử dụng AI an toàn cho giáo viên và người học; và (4) cơ chế giám sát, kiểm tra định kỳ việc tuân thủ các quy định về bảo vệ dữ liệu.

4.5. Đảm bảo tiếp cận công bằng

Các cơ sở GDNN cần xây dựng chính sách bảo đảm mọi người học đều có cơ hội tiếp cận công nghệ AI phù hợp với điều kiện học tập và đặc điểm nghề đào tạo. UNESCO (UNESCO, 2021a, 2021b) khuyến nghị việc triển khai AI phải bảo đảm tính bao trùm, ưu tiên các nhóm người học ở vùng khó khăn, người khuyết tật và các đối tượng dễ bị tổn thương. Trên cơ sở các phân tích trên, bài báo đề xuất các cơ sở GDNN ưu tiên đầu tư hạ tầng số, xây dựng nguồn tài nguyên AI dùng chung, bồi dưỡng năng lực AI cho giáo viên và người học, đồng thời chuẩn bị các phương án dạy học thay thế đối với những đơn vị chưa đủ điều kiện triển khai AI. Việc triển khai AI cần phù hợp với điều kiện địa phương và đặc thù từng ngành nghề nhằm bảo đảm mọi người học đều có cơ hội tiếp cận công nghệ, góp phần thu hẹp khoảng cách số và thúc đẩy GDNN theo hướng công bằng, bao trùm và bền vững.

V. KẾT LUẬN

Bài báo hệ thống hóa cơ sở lý luận về đạo đức AI trong giáo dục nghề nghiệp (GDNN) và phân tích các vấn đề đạo đức phát sinh khi AI được ứng dụng trong dạy học, đánh giá và quản trị đào tạo. Kết quả cho thấy AI chỉ phát huy hiệu quả khi bảo đảm các nguyên tắc minh bạch, công bằng, bảo vệ quyền riêng tư, trách nhiệm giải trình và lấy con người làm trung tâm.

Nghiên cứu đề xuất khung phân tích đạo đức AI và các định hướng quản trị nhằm hỗ trợ các cơ sở GDNN triển khai AI an toàn, minh bạch và có trách nhiệm, đồng thời cung cấp cơ sở tham khảo cho việc xây dựng chính sách và phát triển năng lực đạo đức AI tại Việt Nam.

Do sử dụng phương pháp nghiên cứu lý luận, bài báo chưa phản ánh thực tiễn triển khai AI trong các cơ sở GDNN. Các nghiên cứu tiếp theo cần khảo sát thực nghiệm để đánh giá năng lực đạo đức AI của giáo viên, người học và kiểm chứng hiệu quả của các giải pháp đề xuất.

TÀI LIỆU THAM KHẢO

Deutscher, V., Thomann, H., Zlatkin-Troitschanskaia, O., Weyland, U., Abele, S., Danek, A. H., Greiff, S., Rausch, A., Seeber, S., Seifried, J., & Winther, E. (2026). Artificial intelligence in vocational education and training: A systematic review of educational purposes, theoretical conceptualizations, and empirical effectiveness. Computers and Education: Artificial Intelligence, 11, 100628. https://doi.org/10.1016/j.caeai.2026.100628

European Commission. (2019). Ethics Guidelines for Trustworthy AI. European Commission. https://digital-strategy.ec.europa.eu

Gouseti, A., James, F., Fallin, L., & Burden, K. (2025). The ethics of using AI in K-12 education: A systematic literature review. Technology, Pedagogy and Education, 34(2), 161–182. https://doi.org/10.1080/1475939X.2024.2428601

Huyền P. T. N., & Huyen P. T. N. (2025). QUẢN LÝ ĐÀO TẠO NGHỀ TRONG BỐI CẢNH CHUYỂN ĐỔI SỐ VÀ ỨNG DỤNG AI. https://tamlygiaoduc.com.vn/wp-content/uploads/2025/09/55.Quan-ly-dao-tao-nghe-trong-boi-canh-chuyen-doi-so-va-ung-dung-AI.pdf

Marín, Y. R., Caro, O. C., Rituay, A. M. C., Llanos, K. A. G., Perez, D. T., Bardales, E. S., Tuesta, J. N. A., & Santos, R. C. (2025). Ethical Challenges Associated with the Use of Artificial Intelligence in University Education. Journal of Academic Ethics, 23(4), 2443–2467. https://doi.org/10.1007/s10805-025-09660-w

Metreveli, A., Chen, X., Hedman, A., & Sergeeva, A. (2025). "Who will be left behind?": A Swedish case of learning AI in vocational education. International Journal of Educational Research, 133, 102697. https://doi.org/10.1016/j.ijer.2025.102697

OECD. (2019). Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO) (OECD Digital Economy Papers No. 291; OECD Digital Economy Papers, Vol. 291). https://doi.org/10.1787/d62f618a-en

QUỐC HỘI. (2025). luat134_TRI TUE NHAN TAO. https://datafiles.chinhphu.vn/cpp/files/vbpq/2026/01/luat134.signed.pdf

UNESCO. (2021a). AI and education: Guidance for policy-makers. UNESCO. https://doi.org/10.54675/PCSP7350

UNESCO. (2021b). Recommendation on the Ethics of Artificial Intelligence. UNESCO. https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence

UNESCO. (2024a). AI competency framework for students. UNESCO. https://doi.org/10.54675/JKJB9835

UNESCO. (2024b). AI competency framework for teachers. UNESCO. https://doi.org/10.54675/ZJTE2084

UNESCO. (2025). Hướng dẫn về AI tạo sinh trong giáo dục và nghiên cứu. UNESCO. https://doi.org/10.54675/KMQF4065

Waelen, R. (2022). Why AI Ethics Is a Critical Theory. Philosophy & Technology, 35(1), 9. https://doi.org/10.1007/s13347-022-00507-5

Wiese, L. J., Patil, I., Schiff, D. S., & Magana, A. J. (2025). AI ethics education: A systematic literature review. Computers and Education: Artificial Intelligence, 8, 100405. https://doi.org/10.1016/j.caeai.2025.100405

Zhu, Y., Zuo, H., & Chen, Y. (2024). Digital Transformation in Vocational Education: Challenges, Strategies, and an Experimental Proposal. In F. Huang, Z. Zhan, I. A. Khan, & M. C. Birkök (Eds.), Proceedings of the 2023 4th International Conference on Artificial Intelligence and Education (ICAIE 2023) (Vol. 15, pp. 643–650). Atlantis Press International BV. https://doi.org/10.2991/978-94-6463-242-2_79