

ỨNG DỤNG AGENT AI TRONG PHÁT TRIỂN NHÀ TRƯỜNG THÔNG MINH: TỔNG QUAN NGHIÊN CỨU VÀ ĐỊNH HƯỚNG TRIỂN KHAI TẠI VIỆT NAM

AGENT AI FOR SMART SCHOOL DEVELOPMENT: A STRUCTURED LITERATURE REVIEW
AND IMPLEMENTATION DIRECTIONS FOR VIETNAM

Nguyễn Hữu Trí¹,
Nguyễn Phương Thảo^{1,+},
Lê Ngọc Phượng¹,
Trương Văn Thạo²,
Trần Trung^{3,4}

¹Trường Đại học An Giang - Đại học Quốc gia Thành phố Hồ Chí Minh;

²Viện Chiến lược và Chính sách Giáo dục;

³Trường Đại học Giáo dục - Đại học Quốc gia Hà Nội;

⁴Bộ Dân tộc và Tôn giáo

+Tác giả liên hệ • Email: npthao@agu.edu.vn

Article history

Received: 10/7/2026

Accepted: 12/8/2026

Published: 05/9/2026

Keywords

Agent AI, agentic AI, multi-agent system, smart school, digital transformation in education

ABSTRACT

The rapid development of Generative Artificial Intelligence (Generative AI) and Agent AI has created new opportunities for digital transformation and the development of smart schools. However, existing studies mainly focus on isolated applications of Agent AI in teaching, learning, or educational management, while comprehensive reviews and integrated architectural models suitable for the Vietnamese educational context remain limited. This study employs a structured literature review method, combined with qualitative content analysis, to synthesize, analyze, and systematize research on AI agents in smart schools. Relevant publications were collected from Scopus, Web of Science, IEEE Xplore, SpringerLink, ScienceDirect, and Google Scholar, with the inclusion criteria covering peer-reviewed publications from 2020 to 2025 directly related to the research topic. The findings indicate that Agent AI is evolving from task-oriented AI assistants toward multi-agent ecosystems capable of autonomous planning, collaboration, and orchestration across educational activities. Based on the synthesized evidence, this paper proposes a conceptual Agent AI ecosystem model for Vietnamese smart schools, consisting of five functional layers and specialized agents supporting teaching, learning, assessment, research, educational management, and digital services. The proposed model provides a theoretical foundation and practical orientation for future research and implementation of Agent AI in educational institutions, contributing to the digital transformation of education in Vietnam.

1. Mở đầu

Trong những năm gần đây, sự phát triển nhanh chóng của trí tuệ nhân tạo (AI), đặc biệt là các mô hình ngôn ngữ lớn (LLMs), đã tạo ra bước ngoặt quan trọng trong nhiều lĩnh vực, trong đó có giáo dục. Một xu hướng nổi bật là sự xuất hiện của Agent AI - các hệ thống có khả năng tự động hóa, lập kế hoạch, ra quyết định và tương tác trong môi trường phức tạp. Agent AI không chỉ thực hiện các tác vụ đơn lẻ mà còn có thể phối hợp đa tác nhân để giải quyết các nhiệm vụ phức hợp, từ đó mở rộng phạm vi ứng dụng của AI trong giáo dục (Xi và cộng sự, 2025).

Trong nghiên cứu này, thuật ngữ Generative AI được sử dụng để chỉ các hệ thống AI có khả năng tạo sinh nội dung như văn bản, hình ảnh hoặc mã nguồn dựa trên dữ liệu đã được huấn luyện, điển hình là ChatGPT, Gemini và Claude. Trong khi đó, Agent AI (hay Agentic AI) là các tác nhân thông minh có khả năng tự xác định mục tiêu, lập kế hoạch, lựa chọn và sử dụng công cụ, ghi nhớ ngữ cảnh, thực hiện chuỗi nhiệm vụ và phối hợp với các tác nhân khác dưới sự giám sát của con người. Khi nhiều Agent AI phối hợp thực hiện các nhiệm vụ thông qua cơ chế điều phối chung sẽ hình thành hệ đa tác nhân (multi-agent system). Như vậy, Agent AI không chỉ tạo sinh nội dung mà còn có khả năng chủ động thực hiện và điều phối các hoạt động nhằm đạt được mục tiêu đã xác định.

Song song với sự phát triển của Agent AI, khái niệm nhà trường thông minh đã trở thành một hướng nghiên cứu và triển khai quan trọng. Nhà trường thông minh được hiểu là một hệ sinh thái giáo dục tích hợp công nghệ số, dữ liệu lớn, Internet vạn vật (IoT) và AI nhằm nâng cao chất lượng dạy học và quản lý (Rojas và Chiappe, 2024). Nhà

trường thông minh không chỉ là việc ứng dụng công nghệ mà còn là sự chuyển đổi toàn diện trong cách thức tổ chức, vận hành và ra quyết định trong giáo dục (Terzieva và cộng sự, 2022). Một trong những động lực chính thúc đẩy sự phát triển của nhà trường thông minh là nhu cầu cá nhân hóa học tập. Các nghiên cứu cho thấy rằng người học có nhu cầu và phong cách học tập khác nhau, do đó việc áp dụng các phương pháp giảng dạy truyền thống không còn phù hợp trong bối cảnh hiện đại. Agent AI, với khả năng phân tích dữ liệu và thích ứng theo ngữ cảnh, có thể hỗ trợ xây dựng các lộ trình học tập cá nhân hóa, giúp nâng cao hiệu quả học tập (Sharma và cộng sự, 2025). Ngoài ra, sự gia tăng của dữ liệu giáo dục cũng là một yếu tố quan trọng. Dữ liệu từ các hệ thống quản lý học tập (LMS), bài kiểm tra trực tuyến, và các nền tảng học tập số tạo ra một lượng lớn thông tin có thể khai thác để cải thiện chất lượng giáo dục. Tuy nhiên, việc xử lý và phân tích dữ liệu này đòi hỏi các công cụ thông minh, trong đó Agent AI đóng vai trò quan trọng (Possaghi và cộng sự, 2025).

Trong phạm vi bài báo này, thuật ngữ nhà trường thông minh được hiểu là hệ sinh thái giáo dục số tích hợp dữ liệu, hạ tầng số, các mô hình AI nền tảng và hệ thống Agent AI nhằm hỗ trợ dạy học, học tập, đánh giá, quản trị và cung cấp dịch vụ giáo dục. Vì vậy, bài báo không nhằm tổng quan toàn bộ các ứng dụng AI trong giáo dục mà tập trung vào vai trò của Agent AI như một thành phần trung tâm trong kiến trúc nhà trường thông minh.

Bên cạnh những cơ hội, việc triển khai Agent AI trong nhà trường thông minh cũng đối mặt với nhiều thách thức, việc tích hợp công nghệ vào giáo dục gặp phải các rào cản như thiếu hạ tầng, thiếu kỹ năng của GV và sự kháng cự từ người dùng (Alfoudari và cộng sự, 2021), các vấn đề về đạo đức, bảo mật và quyền riêng tư cũng cần được xem xét kỹ lưỡng (Díaz-Rodríguez và cộng sự, 2023). Các nghiên cứu tổng quan trước đây chủ yếu tập trung vào chatbot giáo dục, Generative AI, Intelligent Tutoring Systems hoặc Learning Analytics, trong khi rất ít nghiên cứu tổng hợp chuyên sâu về Agent AI theo hướng hệ đa tác nhân cho nhà trường thông minh. Đặc biệt, chưa có nghiên cứu nào đề xuất mô hình Agent AI tích hợp phù hợp với bối cảnh chuyển đổi số giáo dục tại Việt Nam. Khoảng trống này là cơ sở để bài báo tiến hành tổng hợp tài liệu và đề xuất mô hình Agent AI cho nhà trường thông minh Việt Nam.

2. Phương pháp nghiên cứu

2.1. Thiết kế nghiên cứu

Bài báo sử dụng phương pháp tổng quan tài liệu có cấu trúc, kết hợp với phân tích nội dung định tính nhằm tổng hợp, phân tích và hệ thống hóa các nghiên cứu về Agent AI trong nhà trường thông minh. Phương pháp này được lựa chọn nhằm nhận diện các xu hướng nghiên cứu, lĩnh vực ứng dụng, kiến trúc công nghệ và những vấn đề đặt ra trong quá trình triển khai Agent AI trong giáo dục, đồng thời làm cơ sở đề xuất mô hình Agent AI phù hợp với bối cảnh chuyển đổi số giáo dục tại Việt Nam. Quá trình thực hiện nghiên cứu tham khảo một số nguyên tắc báo cáo của PRISMA 2020 nhằm tăng tính minh bạch của quy trình tìm kiếm và lựa chọn tài liệu, tuy nhiên nghiên cứu không áp dụng đầy đủ quy trình systematic review theo chuẩn PRISMA. Việc tìm kiếm được thực hiện trên các cơ sở dữ liệu Scopus, Web of Science, IEEE Xplore, SpringerLink, ScienceDirect và Google Scholar nhằm bảo đảm phạm vi bao quát các nghiên cứu liên quan. Chiến lược tìm kiếm được xây dựng dựa trên các nhóm từ khóa chính gồm: (“Agent AI” OR “Agentic AI” OR “AI Agent” OR “Multi-agent System”) AND (“Smart School” OR “Smart Education” OR Education OR Learning). Các từ khóa được kết hợp bằng toán tử Boolean (AND, OR) nhằm mở rộng phạm vi tìm kiếm nhưng vẫn bảo đảm tính liên quan đến chủ đề nghiên cứu.

Quá trình lựa chọn tài liệu được thực hiện theo bốn bước: (1) Tìm kiếm tài liệu từ các cơ sở dữ liệu; (2) Loại bỏ các tài liệu trùng lặp; (3) Sàng lọc tiêu đề và tóm tắt theo tiêu chí lựa chọn; (4) Đọc toàn văn và lựa chọn các tài liệu phù hợp để đưa vào phân tích. Các nghiên cứu được lựa chọn là các công trình tập trung vào Agent AI, hệ đa tác nhân, ứng dụng AI trong giáo dục và nhà trường thông minh, đồng thời có nội dung liên quan trực tiếp đến mục tiêu nghiên cứu của bài báo.

2.2. Phương pháp phân tích dữ liệu

Các tài liệu được phân tích bằng phương pháp phân tích nội dung định tính kết hợp với phương pháp tổng hợp tài liệu. Trước hết, các nghiên cứu được đọc toàn văn nhằm xác định mục tiêu, phương pháp, đối tượng nghiên cứu và các kết quả chính. Tiếp theo, các công trình được mã hóa và phân loại theo các chủ đề nghiên cứu nhằm nhận diện xu hướng phát triển của Agent AI trong giáo dục. Quá trình phân tích tập trung vào năm nhóm nội dung chính: (1) Xu hướng nghiên cứu về Agent AI trong giáo dục; (2) Kiến trúc và mô hình hệ đa tác nhân trong nhà trường thông minh; (3) Các lĩnh vực ứng dụng của Agent AI trong dạy học, học tập, đánh giá và quản lý giáo dục; (4) Những thách thức và yêu cầu đặt ra khi triển khai Agent AI trong giáo dục; (5) Khoảng trống nghiên cứu và các định hướng phát triển trong tương lai. Trên cơ sở kết quả phân tích, các nghiên cứu được tổng hợp theo từng chủ đề nhằm so sánh, đối chiếu và nhận diện những xu hướng nổi bật cũng như những vấn đề còn chưa được giải quyết. Từ đó, bài

báo đề xuất một khung định hướng triển khai Agent AI trong nhà trường thông minh phù hợp với bối cảnh chuyển đổi số giáo dục tại Việt Nam. Các công trình công bố trong giai đoạn 2020-2025 được sử dụng làm nguồn dữ liệu chính để phân tích xu hướng nghiên cứu và thực tiễn ứng dụng Agent AI trong giáo dục. Bên cạnh đó, một số tài liệu công bố trước năm 2020 hoặc sau năm 2025 được sử dụng với vai trò tài liệu nền, tài liệu phương pháp hoặc cơ sở lý luận nhằm làm rõ các khái niệm, bối cảnh nghiên cứu và sự phát triển của lĩnh vực. Các tài liệu này không thuộc tập dữ liệu phân tích chính của nghiên cứu.

3. Kết quả nghiên cứu

3.1. Xu hướng nghiên cứu về Agent AI trong nhà trường thông minh

Sự phát triển của các mô hình ngôn ngữ lớn (LLMs) đã thúc đẩy quá trình chuyển dịch từ các ứng dụng AI hỗ trợ từng tác vụ sang các hệ thống Agent AI có khả năng tự chủ, lập kế hoạch, sử dụng công cụ và phối hợp thực hiện chuỗi nhiệm vụ. Khác với chatbot truyền thống chỉ phản hồi theo yêu cầu của người dùng, Agent AI có thể xác định mục tiêu, phân rã nhiệm vụ, phối hợp với các tác nhân khác và tự động thực hiện các bước cần thiết để hoàn thành công việc, qua đó mở rộng đáng kể phạm vi ứng dụng của AI trong giáo dục.

Tổng hợp các nghiên cứu cho thấy sự phát triển của Agent AI trong giáo dục có thể khái quát thành ba giai đoạn. Giai đoạn thứ nhất, AI chủ yếu hỗ trợ các tác vụ đơn lẻ như trả lời câu hỏi, tạo học liệu, dịch thuật hoặc chấm bài thông qua các công cụ AI tạo sinh. Giai đoạn thứ hai, Agent AI đơn tác nhân xuất hiện với khả năng lập kế hoạch, gọi công cụ và thực hiện chuỗi nhiệm vụ tương đối độc lập như thiết kế bài giảng, xây dựng ngân hàng câu hỏi, hỗ trợ đánh giá hoặc nghiên cứu khoa học. Giai đoạn thứ ba, các nghiên cứu chuyển sang phát triển hệ đa tác nhân, trong đó nhiều Agent chuyên biệt phối hợp dưới sự điều phối của một tác nhân trung tâm hoặc nền tảng điều phối (như CrewAI, AutoGen, LangGraph) để giải quyết các nhiệm vụ giáo dục phức hợp.

Mặc dù đạt được nhiều tiến bộ, phần lớn các nghiên cứu hiện nay vẫn tập trung vào từng ứng dụng hoặc từng tác nhân riêng lẻ. Các mô hình tích hợp toàn diện kết nối dạy học, học tập, đánh giá, quản trị, nghiên cứu khoa học và dịch vụ giáo dục trong một hệ sinh thái Agent AI còn hạn chế. Khoảng trống này là cơ sở để bài báo đề xuất mô hình Agent AI cho nhà trường thông minh Việt Nam ở các mục tiếp theo.

3.2. Kiến trúc Agent AI trong nhà trường thông minh

Kiến trúc Agent AI là nền tảng quyết định khả năng vận hành, mở rộng và tích hợp của hệ sinh thái nhà trường thông minh. Tổng hợp các nghiên cứu cho thấy Agent AI không được triển khai dưới dạng một tác nhân đơn lẻ mà chủ yếu theo kiến trúc hệ đa tác nhân (Multi-agent System), trong đó các tác nhân chuyên biệt phối hợp với nhau thông qua một cơ chế điều phối trung tâm nhằm thực hiện các nhiệm vụ dạy học, học tập, đánh giá và quản trị (Xi và cộng sự, 2025). Kiến trúc này cho phép hệ thống xử lý đồng thời nhiều tác vụ, tăng tính linh hoạt và khả năng thích ứng trong môi trường giáo dục số.

Bảng 1. Các thành phần chính trong kiến trúc Agent AI của nhà trường thông minh

Thành phần	Vai trò chính	Kết quả nghiên cứu tiêu biểu	Hàm ý đối với nhà trường thông minh
Lớp dữ liệu (Data Layer)	Tích hợp dữ liệu học tập, quản trị, IoT và học liệu số	Chất lượng và khả năng liên thông dữ liệu quyết định hiệu quả của Agent AI (Figueroa và cộng sự, 2026).	Xây dựng cơ sở dữ liệu thống nhất và chuẩn hóa dữ liệu giáo dục.
Lớp tác nhân (Agent Layer)	Bao gồm các Agent chuyên biệt hỗ trợ dạy học, học tập, đánh giá, quản trị và nghiên cứu	Phân chia chức năng theo Agent giúp tăng khả năng mở rộng và xử lý song song (Aguilar và cộng sự, 2016).	Hình thành hệ sinh thái đa tác nhân phục vụ nhiều hoạt động của nhà trường.
Lớp điều phối (Agent Orchestrator)	Phân tích yêu cầu, phân rã nhiệm vụ, điều phối và tổng hợp kết quả	Các LLM đóng vai trò điều phối, lập kế hoạch và phối hợp các Agent (Xi và cộng sự, 2025).	Bảo đảm sự phối hợp hiệu quả giữa các Agent và tối ưu hóa quy trình xử lý.
Lớp ứng dụng (Application Layer)	Cung cấp giao diện và dịch vụ cho GV, HS, CBQL và phụ huynh	Giao diện thân thiện góp phần nâng cao trải nghiệm và mức độ chấp nhận công nghệ (Cetinkaya và Krämer, 2025).	Tăng khả năng khai thác Agent AI trong các hoạt động giáo dục hằng ngày.
Các thành phần hỗ trợ	RAG, bảo mật, phân quyền, quản trị dữ liệu	RAG giúp giảm hiện tượng “ảo giác”; bảo mật và quản trị dữ liệu là điều kiện triển khai AI trong giáo dục (Li và cộng sự, 2025; Díaz-Rodríguez và cộng sự, 2023).	Bảo đảm độ tin cậy, an toàn và khả năng mở rộng của hệ thống.

Nguồn: Tổng hợp của nhóm tác giả.

Các nghiên cứu cho thấy kiến trúc Agent AI đang phát triển theo hướng hệ đa tác nhân có điều phối, kết hợp LLM và RAG nhằm nâng cao khả năng truy xuất tri thức và hỗ trợ ra quyết định. Tuy nhiên, việc triển khai vẫn phụ

thuộc vào chất lượng dữ liệu, bảo mật và cơ chế quản trị. Bên cạnh những lợi ích về khả năng mở rộng và tự động hóa, các nghiên cứu cũng chỉ ra một số thách thức trong triển khai Agent AI, bao gồm chất lượng và khả năng liên thông dữ liệu, bảo mật và quyền riêng tư, chi phí triển khai, cũng như việc thiếu các tiêu chuẩn chung trong thiết kế và đánh giá hệ thống (Acharya và cộng sự, 2025). Vì vậy, các kiến trúc Agent AI cần được phát triển theo hướng mở, bảo đảm cơ chế quản trị dữ liệu, giám sát của con người (Human-in-the-loop) và khả năng tích hợp với các nền tảng số hiện có của nhà trường.

3.3. Các lĩnh vực ứng dụng của Agent AI

Bảng 2. Tổng hợp các nghiên cứu tiêu biểu về Agent AI trong nhà trường thông minh

Tác giả	Đối tượng Lĩnh vực nghiên cứu	Đóng góp chính	Hạn chế	Hàm ý cho nghiên cứu này
Smutny và Schreiberova (2020)	Chatbot trong giáo dục	Tổng quan các chatbot giáo dục, phân tích vai trò trong hỗ trợ học tập trực tuyến và tương tác người học	Chủ yếu tập trung vào chatbot, chưa đề cập Agent AI đa tác nhân	Là cơ sở để phân tích sự chuyển dịch từ chatbot sang Agent AI
Okonkwo và Ade-Ibijola (2021)	Chatbot giáo dục	Đánh giá hiệu quả chatbot trong hỗ trợ học tập và phản hồi tức thời	Chưa xem xét khả năng lập kế hoạch và phối hợp tác vụ	Khẳng định tiềm năng của Agent AI trong hỗ trợ học tập
Aguilar và cộng sự (2018)	AI hỗ trợ GV	AI hỗ trợ thiết kế bài giảng, đánh giá và phân tích kết quả học tập	Chưa tích hợp trong hệ sinh thái nhà trường thông minh	Gợi ý vai trò của Teaching Agent
Liao và cộng sự (2024)	Giáo dục STEM	AI hỗ trợ phát triển tư duy tính toán và giải quyết vấn đề	Chủ yếu nghiên cứu trong giáo dục STEM	Là cơ sở phát triển Teaching Agent và Learning Agent
Lu và cộng sự (2024)	Thiết kế hoạt động học tập	AI hỗ trợ GV thiết kế hoạt động học tập sáng tạo	Chưa đề xuất mô hình tích hợp	Bổ sung cơ sở cho Teaching Agent
Ng và cộng sự (2024)	ChatGPT trong giáo dục	Phân tích khả năng hỗ trợ viết, giải quyết vấn đề và tư duy phản biện	Tập trung vào Generative AI, chưa đề cập Agent AI	Làm rõ sự khác biệt giữa Generative AI và Agent AI
Levine và cộng sự (2025)	AI trong học tập đại học	AI góp phần nâng cao động lực học tập và năng lực tự học	Chưa đánh giá khả năng phối hợp nhiều Agent	Khẳng định vai trò của Learning Agent
Sharma và cộng sự (2025)	Cá nhân hóa học tập	AI xây dựng lộ trình học tập cá nhân và giảm tỷ lệ bỏ học	Chưa xem xét tích hợp với quản trị nhà trường	Cơ sở phát triển Learning Agent
Shi và cộng sự (2026)	Adaptive Learning	AI điều chỉnh nội dung học theo phản hồi của người học	Chủ yếu nghiên cứu môi trường học tập trực tuyến	Gợi ý phát triển Agent học tập thích ứng
Xi và cộng sự (2025)	Agentic AI	Làm rõ đặc điểm Agent AI, khả năng lập kế hoạch, sử dụng công cụ và phối hợp nhiều tác nhân	Chưa đề xuất mô hình cho giáo dục	Cơ sở lý luận trực tiếp cho mô hình Agent AI
Các nghiên cứu về Smart School	Nhà trường thông minh	Phân tích kiến trúc trường học thông minh dựa trên IoT, dữ liệu và AI	Chưa tích hợp hệ sinh thái Agent AI	Khoảng trống để đề xuất mô hình của bài báo

Nguồn: Tổng hợp của nhóm tác giả.

Các nghiên cứu hiện nay chủ yếu tập trung vào từng khía cạnh riêng lẻ của AI trong giáo dục như chatbot, trợ lý học tập, học tập thích ứng, giáo dục STEM hoặc ứng dụng mô hình ngôn ngữ lớn. Một số nghiên cứu gần đây đã bắt đầu đề cập đến Agent AI và hệ đa tác nhân, tuy nhiên phần lớn vẫn dừng ở việc phân tích các tác nhân độc lập hoặc các tình huống ứng dụng cụ thể. Đến nay, còn thiếu các nghiên cứu tổng hợp đề xuất một kiến trúc Agent AI tích hợp theo hướng hệ sinh thái đa tác nhân phục vụ đồng thời các hoạt động dạy học, học tập, kiểm tra, đánh giá, quản trị, nghiên cứu khoa học và dịch vụ giáo dục trong nhà trường thông minh. Khoảng trống này là cơ sở để bài báo đề xuất mô hình Agent AI năm lớp chức năng phù hợp với bối cảnh chuyển đổi số giáo dục tại Việt Nam.

Bảng 3. Các lĩnh vực ứng dụng Agent AI trong nhà trường thông minh

Lĩnh vực	Chức năng chính của Agent AI	Giá trị đối với nhà trường thông minh	Yêu cầu/lưu ý
Dạy học và học tập	Hỗ trợ giải đáp và phản hồi; thiết kế bài giảng, học liệu; cá nhân hóa nội dung và tốc độ học; hỗ trợ STEM và học tập thích ứng	Tăng khả năng hỗ trợ học tập cá nhân hóa; nâng cao khả năng tự học; hỗ trợ GV đổi mới dạy học và quản lý lớp học	Cần bảo đảm chất lượng nội dung và duy trì vai trò chuyên môn của GV

Lĩnh vực	Chức năng chính của Agent AI	Giá trị đối với nhà trường thông minh	Yêu cầu/lưu ý
Kiểm tra, đánh giá	Sinh câu hỏi; chấm điểm và phản hồi; phân tích minh chứng học tập; theo dõi tiến bộ	Hỗ trợ đánh giá thường xuyên, rút ngắn thời gian xử lý và phát hiện khó khăn học tập	GV cần giám sát để bảo đảm công bằng, minh bạch
Quản trị nhà trường	Hỗ trợ quản lý đào tạo, HS, nhân sự; phân tích dữ liệu; dashboard; tự động hóa dịch vụ hành chính	Tối ưu nguồn lực, hỗ trợ ra quyết định dựa trên dữ liệu và cải thiện dịch vụ giáo dục	Phụ thuộc vào chất lượng dữ liệu, bảo mật, quyền riêng tư và khả năng tích hợp hệ thống
Nghiên cứu khoa học	Tìm kiếm tài liệu; tổng hợp tri thức; hỗ trợ phân tích dữ liệu và viết	Nâng cao hiệu quả tổng quan, phân tích và công bố khoa học	Cần kiểm chứng nguồn, kiểm soát sai lệch và tuân thủ đạo đức nghiên cứu

Nguồn: Tổng hợp của nhóm tác giả.

Kết quả tổng hợp cho thấy Agent AI có khả năng hỗ trợ hầu hết các hoạt động cốt lõi của nhà trường, từ dạy học, học tập và đánh giá đến quản trị và nghiên cứu khoa học. Tuy nhiên, các nghiên cứu hiện nay chủ yếu tiếp cận từng chức năng hoặc từng nhóm hoạt động riêng biệt; việc kết nối các Agent chuyên biệt thành một hệ sinh thái thống nhất còn hạn chế. Đồng thời, hiệu quả triển khai phụ thuộc vào chất lượng dữ liệu, cơ chế giám sát của con người, bảo mật, quyền riêng tư và các yêu cầu đạo đức.

3.4. Những thách thức và vấn đề đặt ra

Mặc dù Agent AI mở ra nhiều cơ hội đổi mới dạy học, học tập và quản trị nhà trường, tổng hợp các nghiên cứu cho thấy việc triển khai công nghệ này cũng đặt ra nhiều thách thức về kỹ thuật, quản trị, đạo đức và năng lực triển khai. Nếu không có cơ chế quản trị phù hợp, các rủi ro này có thể ảnh hưởng đến chất lượng giáo dục, tính công bằng và mức độ tin cậy của hệ thống AI trong nhà trường (Phạm và cộng sự, 2025).

Bảng 4. Những thách thức trong triển khai Agent AI tại nhà trường thông minh

Nhóm thách thức	Biểu hiện	Tài liệu tiêu biểu	Hàm ý đối với triển khai	Nhóm thách thức
Độ tin cậy của AI	AI có thể tạo thông tin không chính xác (hallucination), trích dẫn sai hoặc tạo tài liệu giả	Huang và cộng sự (2025); Walters và Wilder (2023)	Cần cơ chế kiểm chứng thông tin và giám sát của con người	Độ tin cậy của AI
Thiên lệch thuật toán	Quyết định của AI có thể chịu ảnh hưởng từ dữ liệu huấn luyện, gây mất công bằng trong đánh giá và hỗ trợ học tập	Tao và cộng sự (2024)	Xây dựng dữ liệu chất lượng và đánh giá tính công bằng của mô hình	Thiên lệch thuật toán
Bảo mật và quyền riêng tư	Nguy cơ rò rỉ dữ liệu, tấn công mạng và vi phạm quyền riêng tư	Liu và Khalil (2023)	Áp dụng cơ chế bảo mật, phân quyền và quản trị dữ liệu	Bảo mật và quyền riêng tư
Năng lực triển khai	Hạn chế về hạ tầng số, chi phí và năng lực AI của GV	Alfoudari và cộng sự (2021)	Đầu tư hạ tầng và bồi dưỡng năng lực số cho đội ngũ	Năng lực triển khai
Tính minh bạch và trách nhiệm	Khó giải thích quyết định của AI và xác định trách nhiệm khi xảy ra sai sót	Tao và cộng sự (2024)	Phát triển AI có khả năng giải thích (Explainable AI) và cơ chế trách nhiệm rõ ràng	Tính minh bạch và trách nhiệm
Đạo đức và phụ thuộc công nghệ	Nguy cơ phụ thuộc AI, giảm tư duy độc lập và phát sinh các vấn đề đạo đức	Tian và Zhang (2025); Cetinkaya và Krämer (2025)	Xây dựng quy tắc sử dụng AI có trách nhiệm và tăng cường năng lực AI cho người học	Đạo đức và phụ thuộc công nghệ

Nguồn: Tổng hợp của nhóm tác giả.

Các nghiên cứu đều thống nhất rằng việc triển khai Agent AI trong giáo dục cần được thực hiện theo nguyên tắc AI lấy con người làm trung tâm (Human-in-the-loop), trong đó GV và CBQL giữ vai trò giám sát, kiểm chứng và đưa ra quyết định cuối cùng. Đồng thời, việc xây dựng khung AI Governance với các thành phần như quản trị dữ liệu, bảo mật, quyền riêng tư, minh bạch, khả năng giải thích và trách nhiệm giải trình là điều kiện quan trọng để bảo đảm việc ứng dụng Agent AI an toàn, hiệu quả và bền vững trong nhà trường thông minh (McCormack và Bendeckache, 2025).

Tổng hợp các nghiên cứu cũng cho thấy các thách thức nêu trên không phải là rào cản đối với việc triển khai Agent AI, mà là những yêu cầu cần được xem xét ngay từ giai đoạn thiết kế kiến trúc hệ thống. Đây cũng là cơ sở để bài báo đề xuất mô hình Agent AI theo hướng hệ sinh thái đa tác nhân, tích hợp cơ chế Human-in-the-loop, quản trị dữ liệu và điều phối tập trung nhằm nâng cao tính khả thi và độ tin cậy khi triển khai trong các cơ sở giáo dục Việt Nam.

3.5. Khoảng trống nghiên cứu và định hướng triển khai Agent AI trong nhà trường thông minh

Mặc dù Agent AI đang phát triển nhanh và được ứng dụng trong nhiều lĩnh vực của giáo dục, tổng hợp các nghiên cứu cho thấy việc triển khai Agent AI trong nhà trường thông minh vẫn đối mặt với nhiều hạn chế về công nghệ, quản trị và đánh giá hiệu quả. Phần lớn các nghiên cứu hiện nay tập trung vào phát triển các tác nhân hoặc ứng dụng riêng lẻ, trong khi các mô hình tích hợp theo hướng hệ sinh thái đa tác nhân còn tương đối hạn chế. Những khoảng trống này là cơ sở quan trọng để định hướng các nghiên cứu và triển khai Agent AI trong thời gian tới.

Bảng 5. Khoảng trống nghiên cứu và định hướng phát triển Agent AI trong nhà trường thông minh

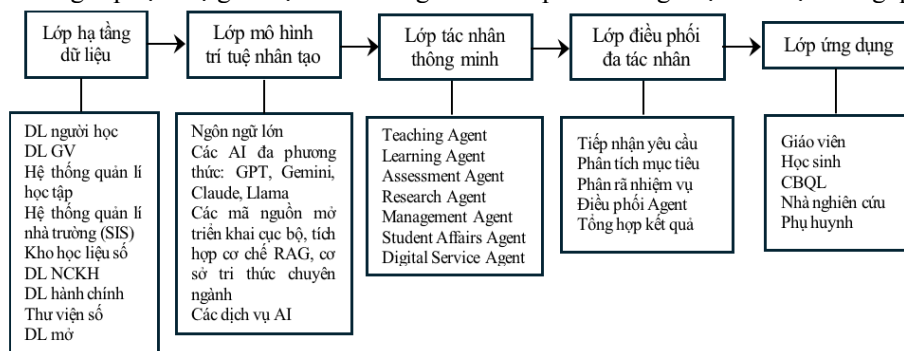
Khoảng trống nghiên cứu	Minh chứng từ các nghiên cứu	Định hướng triển khai
Thiếu mô hình Agent AI tích hợp cho nhà trường thông minh	Phần lớn nghiên cứu tập trung vào từng Agent hoặc từng lĩnh vực ứng dụng riêng lẻ	Phát triển hệ sinh thái đa tác nhân tích hợp dạy học, học tập, đánh giá, quản trị và nghiên cứu
Thiếu khung đánh giá toàn diện	Chủ yếu đánh giá hiệu quả kỹ thuật, chưa xem xét tác động sư phạm, độ tin cậy và tính bền vững (McCormack và Bendechache, 2025)	Xây dựng bộ tiêu chí đánh giá đa chiều cho Agent AI trong giáo dục
Thiếu nghiên cứu thực nghiệm dài hạn	Các nghiên cứu chủ yếu thực hiện ở quy mô nhỏ hoặc thời gian ngắn (Deng và cộng sự, 2025)	Triển khai thí điểm tại các cơ sở giáo dục và đánh giá tác động dài hạn
Hạn chế trong tích hợp dữ liệu	Khó kết hợp dữ liệu định lượng, định tính và dữ liệu thời gian thực (Hauk và Soujon, 2026)	Phát triển nền tảng dữ liệu tích hợp và cơ chế chia sẻ dữ liệu giáo dục

Nguồn: Tổng hợp của nhóm tác giả.

Các nghiên cứu cho thấy xu hướng phát triển Agent AI đang chuyển sang hệ đa tác nhân tích hợp. Tuy nhiên, vẫn thiếu kiến trúc tổng thể, khung đánh giá và nghiên cứu thực nghiệm dài hạn. Đây là khoảng trống để bài báo đề xuất mô hình Agent AI cho nhà trường thông minh Việt Nam. Tuy nhiên, các nghiên cứu hiện nay vẫn thiếu một kiến trúc tổng thể có khả năng tích hợp đồng thời các thành phần dữ liệu, mô hình AI nền tảng, các Agent chuyên biệt, cơ chế điều phối và lớp ứng dụng trong một hệ sinh thái thống nhất. Bên cạnh đó, các vấn đề về quản trị dữ liệu, bảo mật, quyền riêng tư, trách nhiệm giải trình và đánh giá hiệu quả triển khai vẫn chưa được nghiên cứu đầy đủ. Đây chính là khoảng trống nghiên cứu mà bài báo hướng tới giải quyết thông qua việc đề xuất mô hình Agent AI năm lớp chức năng cho nhà trường thông minh Việt Nam.

3.6. Đề xuất mô hình Agent AI cho nhà trường thông minh Việt Nam

Xuất phát từ khoảng trống nghiên cứu đó, bài báo đề xuất mô hình Agent AI cho nhà trường thông minh Việt Nam theo hướng hệ sinh thái đa tác nhân, được thiết kế dựa trên nguyên tắc phân tách chức năng theo từng nhóm tác nhân chuyên biệt, sử dụng chung nguồn dữ liệu, khai thác năng lực của các mô hình ngôn ngữ lớn và được điều phối thông qua một tác nhân trung tâm nhằm bảo đảm sự phối hợp thống nhất giữa các hoạt động dạy học, quản trị, nghiên cứu khoa học và cung cấp dịch vụ giáo dục. Mô hình gồm năm lớp chức năng được thể hiện thông qua sơ đồ 1:



Sơ đồ 1. Mô hình Agent AI cho nhà trường thông minh Việt Nam (nguồn: tác giả đề xuất)

GV và CBQL giữ vai trò Human-in-the-loop để giám sát và phê duyệt các quyết định của Agent AI.

4. Kết luận và bình luận

Nghiên cứu đã tổng hợp và phân tích có hệ thống các công trình tiêu biểu về Agent AI trong nhà trường thông minh, qua đó làm rõ xu hướng phát triển, các lĩnh vực ứng dụng, kiến trúc công nghệ và những vấn đề đặt ra trong quá trình triển khai. Kết quả tổng quan cho thấy Agent AI đang phát triển từ các công cụ hỗ trợ đơn lẻ sang các hệ sinh thái đa tác nhân có khả năng phối hợp, tự động hóa và hỗ trợ ra quyết định, mở ra nhiều tiềm năng trong đổi mới hoạt động dạy học, quản trị nhà trường, nghiên cứu khoa học và cung cấp dịch vụ giáo dục.

Trên cơ sở tổng hợp các kết quả nghiên cứu trong và ngoài nước, bài báo đã đề xuất mô hình Agent AI cho nhà trường thông minh Việt Nam theo kiến trúc năm lớp chức năng theo hướng hệ sinh thái đa tác nhân, bảo đảm khả

năng tích hợp với các nền tảng số hiện có, hỗ trợ phối hợp giữa các tác nhân chuyên biệt và đáp ứng yêu cầu chuyển đổi số trong giáo dục. Kết quả nghiên cứu góp phần bổ sung cơ sở lý luận về ứng dụng Agent AI trong giáo dục, đồng thời cung cấp một khung tham chiếu có thể hỗ trợ các cơ sở giáo dục trong quá trình nghiên cứu, thiết kế và triển khai hệ sinh thái Agent AI phù hợp với điều kiện thực tiễn của Việt Nam.

Nghiên cứu sử dụng phương pháp tổng quan tài liệu kết hợp với đề xuất mô hình khái niệm, do đó mô hình Agent AI mới được xây dựng trên cơ sở tổng hợp các nghiên cứu đã công bố và chưa được kiểm định thông qua đánh giá của chuyên gia hoặc triển khai thực nghiệm tại các cơ sở giáo dục. Ngoài ra, phạm vi phân tích chủ yếu tập trung vào các công trình công bố trong giai đoạn 2020-2025, nên có thể chưa phản ánh đầy đủ các nghiên cứu mới nhất về Agent AI đang phát triển rất nhanh trong thời gian gần đây.

Trong thời gian tới, nhóm tác giả sẽ tiếp tục hoàn thiện mô hình thông qua đánh giá của các chuyên gia trong lĩnh vực giáo dục và trí tuệ nhân tạo, phát triển nguyên mẫu (prototype) cho một số tác nhân AI chuyên biệt, triển khai thử nghiệm tại các cơ sở giáo dục để đánh giá tính khả thi và hiệu quả của mô hình, đồng thời nghiên cứu các vấn đề liên quan đến quản trị dữ liệu, an toàn thông tin, đạo đức AI và các rủi ro phát sinh trong quá trình ứng dụng Agent AI trong nhà trường thông minh.

Tuyên bố về vai trò của các tác giả: Trần Trung: Giám sát, hướng dẫn, định hướng và rà soát quá trình nghiên cứu. Nguyễn Hữu Trí, Nguyễn Phương Thảo: lên ý tưởng, xác định phương pháp nghiên cứu, phân tích dữ liệu, viết bản thảo. Lê Ngọc Phượng, Trương Văn Thọ: rà soát, chỉnh sửa bản thảo.

Tuyên bố về GenAI và Quyền tác giả: Trong quá trình chuẩn bị bản thảo này, các tác giả có sử dụng một số công cụ GenAI như ChatGPT, Gemini để hỗ trợ tìm từ khóa, tóm tắt tài liệu và chỉnh sửa văn phong học thuật. Các công cụ này chỉ được sử dụng như phương tiện hỗ trợ và không thay thế việc phân tích xử lý dữ liệu hay diễn giải kết quả nghiên cứu.

Tuyên bố về xung đột lợi ích: Các tác giả tuyên bố không có xung đột lợi ích.

Lời cảm ơn: Nhóm tác giả cảm ơn sự tài trợ của quỹ NAFOSTED qua đề tài với mã số: NCUD.10-2025.03.

Tài liệu tham khảo

- Acharya, D. B., Kuppan, K., & Divya, B. (2025). Agentic AI: Autonomous Intelligence for Complex Goals-A Comprehensive Survey. *IEEE Access*, 13, 18912-18936. <https://doi.org/10.1109/ACCESS.2025.3532853>
- Aguilar, J., Chamba-Eras, L., & Cordero, J. (2016). *Specification of a Smart Classroom Based on Agent Communities* (pp. 1003-1012). https://doi.org/10.1007/978-3-319-31232-3_95
- Aguilar, J., Sánchez, M., Cordero, J., Valdiviezo-Díaz, P., Barba-Guamán, L., & Chamba-Eras, L. (2018). Learning analytics tasks as services in smart classrooms. *Universal Access in the Information Society*, 17(4), 693-709. <https://doi.org/10.1007/s10209-017-0525-0>
- Alfoudari, A. M., Durugbo, C. M., & Aldhmour, F. M. (2021). Understanding socio-technological challenges of smart classrooms using a systematic review. *Computers & Education*, 173, 104282. <https://doi.org/10.1016/j.compedu.2021.104282>
- Cetinkaya, N. E., & Krämer, N. (2025). Between transparency and trust: identifying key factors in AI system perception. *Behaviour & Information Technology*, 1-15. <https://doi.org/10.1080/0144929X.2025.2533358>
- Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224. <https://doi.org/10.1016/j.compedu.2024.105224>
- Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, 99, 101896. <https://doi.org/10.1016/j.inffus.2023.101896>
- Figueroa, E., Batista, E., Molero-Aranda, T., Ferre, M., & Martínez-Ballesté, A. (2026). Smart classroom ontology: enhancing interoperability in learning environments. *Knowledge and Information Systems*, 68(1), 75. <https://doi.org/10.1007/s10115-026-02695-4>
- Hauk, D., & Soujon, N. (2026). How reliable are large language models in analyzing the quality of written lesson plans? A mixed-methods study from a teacher internship program. *Computers and Education: Artificial Intelligence*, 10, 100538. <https://doi.org/10.1016/j.caeai.2025.100538>

- Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. *ACM Transactions on Information Systems*, 43(2), 1-55. <https://doi.org/10.1145/3703155>
- Levine, S., Beck, S. W., Mah, C., Phalen, L., & Pittman, J. (2025). How do students use ChatGPT as a writing support? *Journal of Adolescent & Adult Literacy*, 68(5), 445-457. <https://doi.org/10.1002/jaal.1373>
- Liao, J., Zhong, L., Zhe, L., Xu, H., Liu, M., & Xie, T. (2024). Scaffolding Computational Thinking With ChatGPT. *IEEE Transactions on Learning Technologies*, 17, 1628-1642. <https://doi.org/10.1109/TLT.2024.3392896>
- Liu, Q., & Khalil, M. (2023). Understanding privacy and data protection issues in learning analytics using a systematic review. *British Journal of Educational Technology*, 54(6), 1715-1747. <https://doi.org/10.1111/bjet.13388>
- Li, Z., Wang, Z., Wang, W., Hung, K., Xie, H., & Wang, F. L. (2025). Retrieval-augmented generation for educational application: A systematic survey. *Computers and Education: Artificial Intelligence*, 8, 100417. <https://doi.org/10.1016/j.caeai.2025.100417>
- Lu, J., Zheng, R., Gong, Z., & Xu, H. (2024). Supporting Teachers' Professional Development With Generative AI: The Effects on Higher Order Thinking and Self-Efficacy. *IEEE Transactions on Learning Technologies*, 17, 1267-1277. <https://doi.org/10.1109/TLT.2024.3369690>
- McCormack, L., & Bendeche, M. (2025). A comprehensive survey and classification of evaluation criteria for trustworthy artificial intelligence. *AI and Ethics*, 5(3), 1973-1994. <https://doi.org/10.1007/s43681-024-00590-8>
- Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, 55(4), 1328-1353. <https://doi.org/10.1111/bjet.13454>
- Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- Pham, N., Pham Ngoc, H., & Nguyen-Duc, A. (2025). Fairness for machine learning software in education: A systematic mapping study. *Journal of Systems and Software*, 219, 112244. <https://doi.org/10.1016/j.jss.2024.112244>
- Possaghi, I., Vesin, B., Zhang, F., Sharma, K., Knudsen, C., Bjørkum, H., & Papavlasopoulou, S. (2025). Integrating multi-modal learning analytics dashboard in K-12 education: insights for enhancing orchestration and teacher decision-making. *Smart Learning Environments*, 12(1), 53. <https://doi.org/10.1186/s40561-025-00410-4>
- Rojas, M. P., & Chiappe, A. (2024). Artificial Intelligence and Digital Ecosystems in Education: A Review. *Technology, Knowledge and Learning*, 29(4), 2153-2170. <https://doi.org/10.1007/s10758-024-09732-7>
- Sharma, S., Mittal, P., Kumar, M., & Bhardwaj, V. (2025). The role of large language models in personalized learning: a systematic review of educational impact. *Discover Sustainability*, 6(1), 243. <https://doi.org/10.1007/s43621-025-01094-z>
- Shi, Y., Yu, K., Dong, Y., & Chen, F. (2026). Large language models in education: a systematic review of empirical applications, benefits, and challenges. *Computers and Education: Artificial Intelligence*, 10, 100529. <https://doi.org/10.1016/j.caeai.2025.100529>
- Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, 151, 103862. <https://doi.org/10.1016/j.compedu.2020.103862>
- Tao, Y., Viberg, O., Baker, R. S., & Kizilcec, R. F. (2024). Cultural bias and cultural alignment of large language models. *PNAS Nexus*, 3(9). <https://doi.org/10.1093/pnasnexus/pgae346>
- Terzieva, V., Ilchev, S., & Todorova, K. (2022). The Role of Internet of Things in Smart Education. *IFAC-PapersOnLine*, 55(11), 108-113. <https://doi.org/10.1016/j.ifacol.2022.08.057>
- Tian, J., & Zhang, R. (2025). Learners' AI dependence and critical thinking: The psychological mechanism of fatigue and the social buffering role of AI literacy. *Acta Psychologica*, 260, 105725. <https://doi.org/10.1016/j.actpsy.2025.105725>
- Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports*, 13(1), 14045. <https://doi.org/10.1038/s41598-023-41032-5>
- Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., Zheng, R., Fan, X., Wang, X., Xiong, L., Zhou, Y., Wang, W., Jiang, C., Zou, Y., Liu, X., ... Gui, T. (2025). The rise and potential of large language model based agents: a survey. *Science China Information Sciences*, 68(2), 121101. <https://doi.org/10.1007/s11432-024-4222-0>