

TÓM TẮT BÁO CÁO KHOA HỌC

Hội thảo Quốc gia lần thứ 4 về Ngôn ngữ học Tính toán (VCL2026)

Tên đề tài: Nâng cao năng lực “Thẩm định lỗi từ AI” nhằm phát triển tư duy sự phạm độc lập cho sinh viên ngành Giáo dục Tiểu học

Tác giả: TS. Nguyễn Văn Bản, NGƯT. ThS. Tô Ngọc Sơn (Tác giả liên hệ)

Đơn vị: Khoa Giáo dục Tiểu học - Mầm non, Trường Đại học Đồng Tháp

1. ĐẶT VẤN ĐỀ VÀ TÍNH MỚI NỔI BẬT

Bối cảnh & Lỗ hổng nghiên cứu:

Sự bùng nổ của trí tuệ nhân tạo tạo sinh (Generative AI) và các mô hình ngôn ngữ lớn (LLMs) mang lại cơ hội tối ưu hóa học liệu, nhưng đồng thời gây ra nguy cơ suy giảm tư duy phản biện ngôn ngữ, dẫn đến hội chứng phụ thuộc và ủy thác tư duy cho thuật toán ở sinh viên sư phạm.

Bậc Giáo dục Tiểu học đòi hỏi tính chuẩn mực tuyệt đối về ngữ pháp, vốn từ, tâm lý lứa tuổi và giá trị văn hóa. Tuy nhiên, văn bản Tiếng Việt do AI tạo ra thường mắc lỗi hệ thống tinh vi: *ảo giác kiến thức (hallucination)*, *lạm dụng thuật ngữ trừu tượng đao to búa lớn*, và *lối tư duy dịch thuật lai căng (“Tây hóa”)*.

Tính mới nổi bật của đề tài:

Lần đầu tiên, năng lực "Thẩm định lỗi từ AI" (AI error auditing) được xác định là một cấu trúc năng lực nghề nghiệp cốt lõi, đóng vai trò "Màng lọc sư phạm" bảo vệ sự trong sáng của Tiếng Việt và nhận thức của học sinh tiểu học.

Chuyển dịch triết lý đào tạo: Từ vị thế tiêu thụ/sao chép công nghệ thụ động sang vị thế làm chủ, giám sát và phản biện AI có trách nhiệm xã hội và đạo đức học thuật.

2. PHƯƠNG PHÁP NGHIÊN CỨU VÀ PHÁT HIỆN THỰC TRẠNG

Thiết kế nghiên cứu: Khảo sát định lượng phối hợp phân tích định tính trên $N = 140$ sinh viên ngành Giáo dục Tiểu học (từ năm 1 đến năm 3).

Kết quả thực trạng nổi bật (Thực chứng định lượng):

Tần suất tương tác cao nhưng thiếu hoài nghi: 97.9% sinh viên tương tác với AI thường xuyên (40% dùng hằng ngày, 57.9% dùng vài lần/tuần). 82.1% hình thành thói quen dùng AI ngay từ năm thứ nhất.

Mục đích sử dụng: Chủ yếu để tìm ý tưởng trò chơi/khởi động (92.9%), giải thích kiến thức khó (65.7%), soạn Kế hoạch bài dạy (51.4%). 29.3% thừa nhận lười động não hơn từ khi dùng AI.

Tín hiệu tích cực về ý thức nghề nghiệp: 0% sinh viên sao chép thô nguyên bản (copy-paste) vào giáo án. 63.6% có ý thức biên tập, cắt ngắn câu cho phù hợp lứa tuổi; 20% tự viết lại hoàn toàn dựa trên ý tưởng.

Khoảng trống năng lực: 85.7% sinh viên đánh giá văn bản AI còn cảm tính, nghe "xuôi tai" là sử dụng, chỉ có 14.3% tự tin nhận diện và xử lý các lỗi ngôn ngữ/sur phạm tinh vi.

3. HỆ THỐNG GIẢI PHÁP ĐỘC ĐÁO VÀ MANG TÍNH ĐỘT PHÁ

Tác giả đề xuất hệ thống 3 giải pháp đồng bộ nhằm kích hoạt tư duy sư phạm độc lập:

* Giải pháp 1: Thiết lập và chuyển giao bộ tiêu chí "Màng lọc sư phạm 3 lớp"

Cụ thể hóa việc thẩm định văn bản AI thành quy trình kiểm định đa tầng định lượng:

1. **Lớp 1 - Khách quan & Chính xác:** Lỗi chính tả = 0%, Ảo giác kiến thức = 0%. Tra từ điển và đối chiếu trực tiếp với Sách giáo khoa hiện hành.
2. **Lớp 2 - Chuẩn mực sư phạm lứa tuổi:** Kiểm định độ dài câu (15–20 từ cho lớp 1–2; không quá 25 từ cho lớp 3–5), từ vựng phù hợp tâm lý học lứa tuổi, tương thích với "Yêu cầu cần đạt" của Chương trình GDPT 2018.
3. **Lớp 3 - Văn hóa & Thẩm mỹ Tiếng Việt:** Triệt tiêu các cấu trúc bị động "Tây hóa" (dịch máy), gọt giũa nhạc điệu, thanh lọc từ ngữ sáo rỗng/đao to búa lớn.

* Giải pháp 2: Thực hành "Bắt lỗi AI, gọt giũa ngôn từ" theo quy trình 4 bước

Đảo ngược thể chủ động giữa con người và thuật toán:

- Bước 1 (Thiết lập): Ra lệnh cho AI tạo văn bản thô (làm minh chứng gốc).
- Bước 2 (Mổ xẻ): Sinh viên đóng vai nhà thẩm định, bóc tách lỗi cú pháp, ngữ nghĩa, lỗi sư phạm.
- Bước 3 (Tái cấu trúc): Ngắt kết nối AI, tự tay biên tập, gọt giũa văn bản theo tư duy ngôn ngữ bản địa.
- Bước 4 (Phản hồi biện chứng): Báo cáo bảng so sánh song song và lập luận bảo vệ phương án ngôn từ tối ưu.

* Giải pháp 3: Đổi mới đánh giá nghiệp vụ sư phạm theo hướng "Kháng AI"

Chuyển trọng tâm đánh giá từ *sản phẩm cuối cùng* sang *tiến trình tư duy phản biện*:

- Đề bài theo ma trận phản biện: Cung cấp sẵn giáo án/học liệu do AI tạo ra có cài cắm sẵn lỗi tinh vi (kiến thức, sư phạm, văn phong dịch máy). Yêu cầu sinh viên viết bài luận phân tích vị trí lỗi, nguyên nhân và đề xuất bản chỉnh sửa.

- Đánh giá thời gian thực (Real-time): Yêu cầu sinh viên phát hiện lỗi và gọt giũa ngữ liệu AI trực tiếp tại lớp trong thời gian 5–7 phút không có thiết bị hỗ trợ internet.

4. MINH CHỨNG THỰC NGHIỆM VÀ KẾT LUẬN

Minh chứng thực nghiệm thuyết phục:

Qua phân tích ví dụ điển hình (ngữ liệu AI về chủ đề "Tình bạn" lớp 3 chứa thuật ngữ "phạm trù đạo đức", "đồng hành trung kiên", câu bị động dịch máy), sinh viên vận dụng mảng lọc 3 lớp để chuyển hóa thành đoạn văn trong sáng, chuẩn mực sư phạm lứa tuổi.

100% sinh viên trong nhóm thực nghiệm hình thành phản xạ hoài nghi khoa học, nâng cao năng lực ứng biến ngôn ngữ và lập luận biện chứng.

Kết luận & Khuyến nghị:

Việc nâng cao năng lực thẩm định lỗi từ AI cho sinh viên ngành Giáo dục Tiểu học không phải là một giải pháp tình thế mà là một chiến lược giáo dục lâu dài hướng tới việc đào tạo ra những nhà giáo tương lai biết làm chủ công nghệ một cách có trách nhiệm.

Nghiên cứu đóng góp một góc nhìn thực tiễn về ứng dụng và giới hạn ngữ nghĩa/ngữ cảnh của các mô hình ngôn ngữ lớn (LLMs) trong xử lý Tiếng Việt thuộc lĩnh vực giáo dục sư phạm.

Khuyến nghị tích hợp chuyên đề "Phản biện AI" vào khung chương trình đào tạo sư phạm và xây dựng **Kho dữ liệu số về các lỗi mẫu của AI** làm học liệu thực hành thường xuyên cho sinh viên.