

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO ĐỂ HỖ TRỢ XÂY DỰNG ĐỀ CƯƠNG HỌC PHẦN DỰA TRÊN CHUẨN ĐẦU RA

Đinh Tuấn Long¹, Trương Hữu Thành^{1*}

*Tác giả liên hệ, Email: thanhth@hou.edu.vn. ORCID: 0009-0006-3458-4595

Ngày tòa soạn nhận được bài báo: 15/06/2026

Ngày phản biện đánh giá: 22/06/2026

Ngày bài báo được duyệt đăng: 14/09/2026

DOI: 10.59266/houjs.2026.1424

Tóm tắt: Xây dựng đề cương học phần theo chuẩn đầu ra cần sự phối hợp giữa chuẩn đầu ra chương trình, học phần với nội dung, hoạt động dạy học và phương pháp đánh giá. Bài báo đề xuất một khung kiến trúc ứng dụng AI nhằm hỗ trợ giảng viên xây dựng và đánh giá đề cương học phần. Nghiên cứu sử dụng phương pháp Design Science Research (DSR), kết hợp nghiên cứu tài liệu, phân tích khái niệm và thiết kế artefact. Cơ sở sự phạm gồm: Constructive Alignment (tính đồng bộ kiến tạo), Backward Design (thiết kế ngược) và Bloom's Taxonomy (thang đo Bloom) Cơ sở công nghệ gồm mô hình ngôn ngữ lớn (LLM) và Retrieval-Augmented Generation (RAG). Khung đề xuất gồm bốn tầng: giao diện, ứng dụng, AI và nghiệp vụ cùng với dữ liệu và tri thức với sáu module từ phân tích chuẩn đầu ra của chương trình đào tạo (Program Learning Outcomes - PLO), tạo chuẩn đầu ra học phần (Course Learning Outcomes - CLO) đến đề xuất dạy học, đánh giá, kiểm tra chất lượng và tổng hợp đề cương. Cơ chế Human-in-the-Loop (mô hình làm việc kết hợp giữa trí tuệ nhân tạo (AI) và con người) quy định AI chỉ tạo đề xuất, còn giảng viên kiểm tra, chỉnh sửa và phê duyệt. Bài báo đề xuất thiết kế kiểm định dựa trên Bloom, PLO-CLO alignment, độ phủ nội dung và tính nhất quán, kết hợp độ đo tự động, đánh giá chuyên gia và khảo sát. Kết quả hiện tại là khung kiến trúc và thiết kế kiểm định.

Từ khóa: trí tuệ nhân tạo, chuẩn đầu ra, đề cương học phần, giáo dục đại học

I. Đặt vấn đề

Giáo dục đại học theo định hướng chuẩn đầu ra (Outcome-Based Education - OBE) yêu cầu thiết kế học phần thể hiện mối liên hệ giữa PLO, CLO, hoạt động dạy học và đánh giá. Đề cương học phần vừa định hướng dạy học, vừa là thành phần của hệ thống bảo đảm chất lượng.

Xây dựng đề cương đòi hỏi xử lý đồng thời quy định, PLO, CLO, nội dung, phương pháp dạy học, đánh giá và ma trận liên kết. Bài báo tập trung thiết kế một mô hình hỗ trợ có khả năng chuẩn hóa quy trình.

LLM và RAG tạo điều kiện xử lý tài liệu có cấu trúc và sinh nội dung theo ngữ cảnh. Tuy nhiên, LLM đơn thuần không

¹ Trường Đại học Mở Hà Nội, Hà Nội, Vietnam

bảo đảm nội dung phù hợp với quy định, PLO/CLO hoặc logic thiết kế học phần. Vấn đề nghiên cứu là tích hợp nguyên tắc thiết kế giáo dục với truy xuất tri thức và đánh giá tự động, đồng thời duy trì quyền kiểm soát học thuật của giảng viên.

Mục tiêu nghiên cứu là đề xuất khung kiến trúc AI hỗ trợ phân tích PLO, đề xuất nội dung, xây dựng CLO, đề xuất dạy học và đánh giá đến tổng hợp đề cương; đồng thời bổ sung kiểm tra chất lượng và Human-in-the-Loop. Phạm vi giới hạn ở xây dựng và đánh giá đề cương học phần.

II. Cơ sở lý thuyết

2.1. Nguyên lý Tương thích kiến tạo (Constructive Alignment) trong thiết kế đề cương

Constructive Alignment nhấn mạnh sự tương thích giữa kết quả học tập dự kiến, hoạt động dạy - học và đánh giá. Backward Design (thiết kế ngược) là một cách tổ chức quy trình thiết kế bắt đầu từ kết quả mong đợi, xác định bằng chứng đánh giá, sau đó mới lựa chọn hoạt động và nội dung. Hai khái niệm có quan hệ chặt chẽ nhưng không đồng nhất. Trong mô hình đề xuất, Backward Design được sử dụng để tổ chức trình tự thiết kế; Constructive Alignment được sử dụng như nguyên tắc kiểm tra sự nhất quán giữa các thành phần (Biggs và Tang, 2011; Wiggins và McTighe, 2005; Đoàn Thị Minh Trinh và cộng sự, 2012; Nguyễn Hữu Duyệt, 2017).

Trong bối cảnh đề cương học phần, trình tự có thể biểu diễn như: PLO → CLO → bằng chứng đánh giá → hoạt động dạy học và nội dung → kế hoạch triển khai.

Tuy nhiên, các bước không nhất thiết hoàn toàn tuyến tính; hệ thống cần cho phép quay lại điều chỉnh CLO hoặc nội dung khi kết quả kiểm tra alignment cho thấy vấn đề (Nguyễn Thanh Nhân, 2024).

2.2. Thang phân loại Bloom và ứng dụng trong viết chuẩn đầu ra

Thang phân loại Bloom (Bloom's Taxonomy) và phiên bản sửa đổi của Anderson và Krathwohl cung cấp khung phân loại các mức độ nhận thức: Nhớ, Hiểu, Vận dụng, Phân tích, Đánh giá và Sáng tạo. Trong thiết kế CLO, động từ hành động là tín hiệu quan trọng để xác định mức độ nhận thức và khả năng đo lường. Vì vậy, mô hình đề xuất sử dụng Bloom như một lớp kiểm tra hỗ trợ: phân tích động từ, nhận diện mức độ nhận thức và phát hiện CLO có động từ không rõ ràng hoặc chứa nhiều hành động cần được xem xét (Anderson và Krathwohl, 2001; Krathwohl, 2002).

2.3. Mô hình BERT trong phân loại mục tiêu học tập

Một trong những ứng dụng thành công nhất của AI trong giáo dục là việc sử dụng các mô hình ngôn ngữ như BERT để tự động phân tích và phân loại chuẩn đầu ra. Nghiên cứu của các học giả quốc tế đã chứng minh rằng việc tinh chỉnh mô hình BERT với các lớp mạng nơ-ron dày đặc cho phép máy tính nhận diện chính xác mức độ nhận thức của một câu mục tiêu học tập theo thang Bloom.

Trong một nghiên cứu so sánh quy mô lớn với 21.380 mục tiêu học tập từ 5.558 khóa học tại một đại học Úc, các bộ phân loại dựa trên BERT đã

đạt điểm F1 lên tới 0.95 và hệ số Kappa của Cohen là 0.93, vượt trội so với các phương pháp máy học truyền thống như SVM hay Random Forest (Li và cộng sự, 2022). Điều này cho phép xây dựng các hệ thống tự động quét hàng nghìn đề cương để vẽ ra bản đồ phân bố năng lực, phát hiện ngay lập tức những ngành học đang bị lệch về lý thuyết mà thiếu thực hành.

2.4. RAG

LLM có khả năng phân tích và sinh văn bản theo ngữ cảnh nhưng có thể tạo ra thông tin không phù hợp với nguồn tri thức hoặc quy định cụ thể. RAG khắc phục một phần hạn chế này bằng cách truy xuất các tài liệu liên quan từ kho tri thức rồi đưa chúng vào ngữ cảnh của LLM trước khi sinh. Trong nghiên cứu này, RAG được sử dụng như một cơ chế truy xuất và cung cấp thông tin làm căn cứ cho mô hình ngôn ngữ lớn LLM (Lewis và cộng sự, 2020). Hệ thống tìm kiếm các thông tin liên quan từ kho tri thức đã được xây dựng và kiểm soát, sau đó cung cấp những thông tin này cho LLM để hỗ trợ quá trình phân tích và sinh nội dung. RAG không làm thay đổi các tham số của LLM và không phải là quá trình huấn luyện mô hình.

Kho tri thức dự kiến gồm các văn bản quy định, chuẩn đầu ra, ma trận ánh xạ, danh mục động từ Bloom và các đề cương đã được chuẩn hóa và phê duyệt. Các tài liệu được tiền xử lý, chia đoạn, embedding và lập chỉ mục trong vector store. Khi giảng viên gửi yêu cầu, retriever tìm các đoạn phù hợp; context

được đưa vào prompt để LLM tạo đề xuất. GraphRAG/Knowledge Graph không được coi là thành phần bắt buộc của phiên bản mô hình này; nếu được triển khai ở giai đoạn sau, nó sẽ là một lớp truy xuất bổ sung cho các quan hệ PLO - CLO - nội dung - đánh giá.

2.5. Nghiên cứu liên quan và khoảng trống

Tổng quan các nghiên cứu liên quan cho thấy những công nghệ và phương pháp được sử dụng trong mô hình đã được nghiên cứu ở những khía cạnh khác nhau. Tuy nhiên, các nghiên cứu này chủ yếu tập trung vào từng bài toán riêng lẻ, chưa hình thành một khung thống nhất hỗ trợ xây dựng và kiểm tra đề cương học phần dựa trên chuẩn đầu ra. Trên cơ sở đó, nghiên cứu này kế thừa các kết quả đã có và đề xuất một hướng tiếp cận tích hợp các thành phần phù hợp trong một khung thống nhất: sử dụng BERT để tự động phân loại mục tiêu học tập theo các mức độ nhận thức; sử dụng GPT-4 để hỗ trợ xây dựng mục tiêu học tập; sử dụng hệ thống khái niệm và biểu diễn tri thức để mô tả mối quan hệ giữa chương trình đào tạo và chuẩn đầu ra; cùng với các công cụ trí tuệ nhân tạo hỗ trợ rà soát đề cương học phần ở một số cơ sở giáo dục quốc tế. Vì vậy, tính mới của nghiên cứu không nằm ở việc phát triển một thuật toán hoàn toàn mới cho từng thành phần riêng lẻ, mà ở việc tích hợp các thành phần này thành một khung thống nhất nhằm hỗ trợ xây dựng và kiểm tra đề cương học phần dựa trên chuẩn đầu ra (Li và cộng sự, 2022; Sridhar và cộng sự, 2023).

Bảng 1. So sánh các hướng nghiên cứu liên quan

Hướng nghiên cứu/giải pháp	Thành phần chính	Điểm liên quan	Khoảng trống đối với mô hình này
BERT/encoder + Bloom	Tự động phân loại mức nhận thức	Hỗ trợ xác định mức độ nhận thức của mục tiêu học tập	Chưa bao phủ toàn bộ quy trình xây dựng đề cương
LLM/GPT	Sinh CLO	Hỗ trợ xây dựng chuẩn đầu ra học phần	Chưa tích hợp đầy đủ kho tri thức và kiểm tra alignment
Hệ thống khái niệm và đồ thị tri thức	Biểu diễn các khái niệm và mối quan hệ trong chương trình đào tạo	Hỗ trợ mô tả và kiểm tra mối quan hệ giữa các thành phần của chương trình	Chưa tích hợp đầy đủ với mô hình ngôn ngữ lớn, cơ chế truy xuất thông tin và sự tham gia của giảng viên trong quá trình kiểm tra, hiệu chỉnh
AI hỗ trợ syllabus	Rà soát/sinh nội dung	Hỗ trợ một số nhiệm vụ trong xây dựng và rà soát đề cương	Khác bối cảnh quy định và kho tri thức tại Việt Nam
Mô hình đề xuất	DSR + CA + Bloom + RAG + LLM + validation	Tích hợp quá trình xây dựng nội dung và đánh giá mức độ phù hợp	Đóng góp ở cấp độ một khung tích hợp, có sự tham gia của giảng viên, hướng đến hỗ trợ xây dựng đề cương học phần trong bối cảnh đào tạo CNTT tại Trường Đại học Mở Hà Nội

Khoảng trống được xác định trong phạm vi tài liệu tổng hợp: thiếu khung tích hợp kết nối chuẩn đầu ra, thiết kế giáo dục, RAG/LLM, đánh giá định lượng và cơ chế giảng viên phê duyệt trong xây dựng đề cương tại một cơ sở giáo dục đại học Việt Nam.

III. Phương pháp nghiên cứu

Nghiên cứu sử dụng phương pháp Nghiên cứu khoa học thiết kế (Design

Science Research - DSR) làm phương pháp chủ đạo để xây dựng một khung kiến trúc và quy trình ứng dụng trí tuệ nhân tạo nhằm hỗ trợ xây dựng và đánh giá đề cương học phần. Quá trình nghiên cứu gồm các bước: xác định vấn đề, tổng hợp cơ sở lý luận và dữ liệu, thiết kế giải pháp, xây dựng phương án đánh giá và xác định hướng cải tiến.

Bảng 2. Các giai đoạn của quy trình Design Science Research

Giai đoạn	Nội dung	Đầu ra
1. Xác định vấn đề	Phân tích yêu cầu và điểm cần hỗ trợ	Yêu cầu sư phạm, nghiệp vụ
2. Cơ sở lý thuyết và dữ liệu	Tổng hợp Constructive Alignment, Backward Design, Bloom; quy định; dữ liệu dự kiến	Bộ tiêu chí và dữ liệu nguồn
3. Thiết kế artefact	Thiết kế kiến trúc 4 tầng, 6 module, RAG, validation và Human-in-the-Loop	Khung kiến trúc và workflow
4. Đánh giá	Thiết kế thử nghiệm 10 đề cương; chuyên gia; 15-20 giảng viên dự kiến	Bộ chỉ số và phản hồi
5. Cải tiến	Phân tích sai lệch và phản hồi để điều chỉnh mô hình	Phiên bản cải tiến

Về dữ liệu, thuyết minh dự kiến thu thập và chuẩn hóa 30-50 đề cương thuộc một chương trình Công nghệ thông tin, cùng PLO, CLO, quy định và tài liệu

chuẩn. Trong bài báo, đây là dữ liệu dự kiến phục vụ nguyên mẫu và kiểm định, chưa phải tập dữ liệu đã thực nghiệm.

IV. Kết quả nghiên cứu và thảo luận

4.1. Bối cảnh, phạm vi và yêu cầu hệ thống

Phạm vi nghiên cứu là Trường Đại học Mở Hà Nội, với thử nghiệm dự kiến tại một chương trình đào tạo thuộc Khoa Công nghệ Thông tin. Nội dung tập trung vào xây dựng và đánh giá đề cương dựa trên PLO/CLO, không mở rộng sang toàn bộ chương trình đào tạo.

Về pháp lý, bài báo sử dụng các văn bản về chuẩn chương trình đào tạo làm căn cứ xác định thành phần đề cương (Bộ Giáo dục và Đào tạo, 2021). Đối với kiểm định chương trình, bài báo tham chiếu Thông tư 04/2025/TT-BGDĐT; các yêu cầu chi tiết chỉ áp dụng khi có căn cứ từ quy định hoặc tiêu chuẩn tương ứng (Bộ Giáo dục và Đào tạo, 2025).

4.2. Khung kiến trúc đề xuất ứng dụng trí tuệ nhân tạo

Sản phẩm nghiên cứu là khung kiến trúc hệ thống AI và quy trình nghiệp vụ, gồm các thành phần công nghệ, dữ liệu, workflow và tiêu chí đánh giá; không chỉ là một mô hình lý thuyết.

4.2.3. Tầng trí tuệ nhân tạo và nghiệp vụ: sáu module

Bảng 3. Sáu module và nghiệp vụ của khung đề xuất

Module	Chức năng	Đầu ra
M1. Phân tích PLO	Phân tích năng lực cần đạt và các yêu cầu, ràng buộc của PLO	Hồ sơ chuẩn đầu ra chương trình
M2. Mô tả/gợi ý nội dung	Gợi ý mô tả học phần và phạm vi nội dung	Dự thảo mô tả và nội dung học phần
M3. Tạo CLO	Đề xuất CLO theo PLO, nội dung và Bloom	Dự thảo chuẩn đầu ra học phần
M4. Dạy học & đánh giá	Đề xuất dạy học và đánh giá theo CLO	Kế hoạch dạy học và đánh giá
M5. Kiểm tra chất lượng	Kiểm tra mức độ phù hợp với thang đo Bloom, sự thống nhất giữa các thành phần, mức độ bao phủ và tính nhất quán của đề cương	Cảnh báo và kết quả đánh giá
M6. Tổng hợp kế hoạch	Tổng hợp các thành phần thành đề cương	Dự thảo đề cương học phần



Hình 1. Kiến trúc tổng thể của hệ thống AI đề xuất

4.2.1. Tầng giao diện

Tầng giao diện cung cấp biểu mẫu nhập thông tin học phần, PLO liên quan và các ràng buộc; hiển thị dự thảo CLO, nội dung, hoạt động dạy học, đánh giá và kế hoạch; đồng thời cho phép giảng viên chỉnh sửa, xác nhận và lưu phiên bản. Mọi kết quả AI ở tầng này phải được đánh dấu là đề xuất cho đến khi được phê duyệt.

4.2.2. Tầng ứng dụng

Tầng ứng dụng quản lý workflow, phiên làm việc, quyền truy cập, versioning và điều phối API; không trực tiếp thay LLM suy luận mà chuyển yêu cầu đến các dịch vụ phù hợp.

Sáu module được thống nhất trong toàn bài. Module 4 kết hợp hoạt động dạy học và đánh giá vì hai thành phần được thiết kế đồng bộ theo CLO. Module 5 được bổ sung rõ: kiểm tra chất lượng và tạo cảnh báo để giảng viên rà soát.

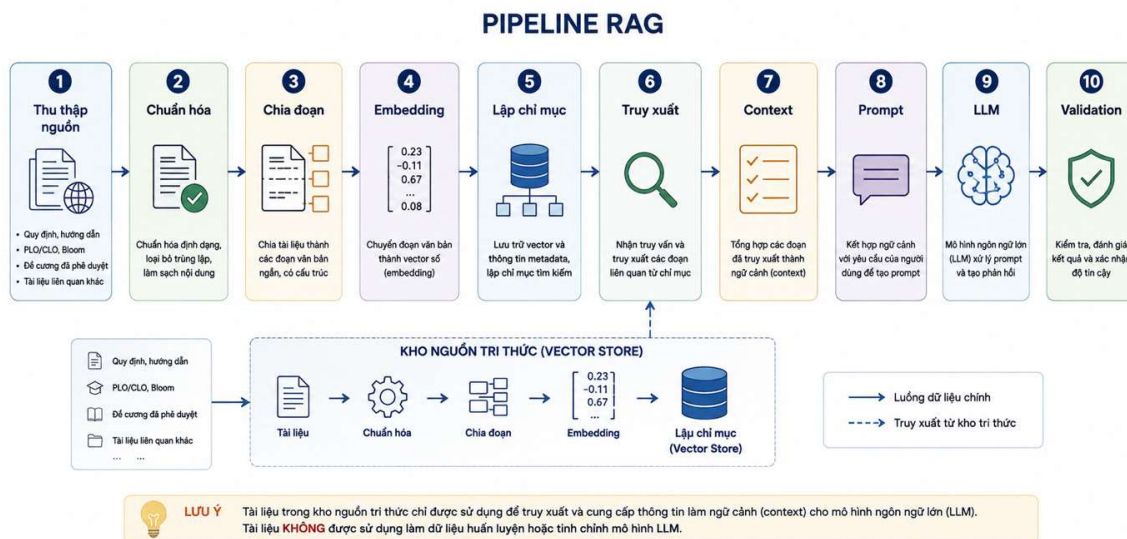
4.2.4. Tầng dữ liệu và tri thức

Tầng dữ liệu bao gồm kho các quy định, chuẩn đầu ra chương trình, chuẩn đầu ra học phần, thang đo Bloom, các đề cương học phần đã được phê duyệt, thông tin mô tả dữ liệu và kho dữ liệu phục vụ tìm kiếm. Các tài liệu được chuẩn hóa và lập chỉ mục để thuận tiện cho việc lưu trữ và truy xuất. Trong các giai đoạn phát triển tiếp theo, hệ thống có thể bổ sung

một đồ thị tri thức để biểu diễn các mối quan hệ giữa chuẩn đầu ra chương trình với chuẩn đầu ra học phần, giữa chuẩn đầu ra học phần với nội dung và đánh giá, cũng như giữa nội dung với hoạt động dạy học. Thành phần này hiện chỉ được đề xuất ở mức định hướng phát triển và chưa được xem là một chức năng đã triển khai trong nghiên cứu.

4.2.5. Pipeline RAG

Pipeline RAG gồm: thu thập nguồn → chuẩn hóa → chia đoạn → embedding → lập chỉ mục → truy xuất → context → prompt → LLM → validation. Tài liệu chỉ được sử dụng làm nguồn truy xuất, không mô tả như dữ liệu huấn luyện LLM.



Hình 2. Pipeline RAG đề xuất

4.3. Cơ chế đánh giá chất lượng và sự tham gia của giảng viên

Đối với Bloom, hệ thống không chỉ kiểm tra động từ mà còn đối chiếu mức độ nhận thức với bằng chứng đánh giá dự kiến: các CLO ở mức Vận dụng có thể gắn với bài tập/thực hành; mức Phân

tích với tình huống hoặc bài tập phân tích; mức Đánh giá với tiêu chí/rubric đánh giá lập luận; và mức Sáng tạo với sản phẩm hoặc dự án. Đây là quy tắc hỗ trợ để phát hiện sai lệch, không thay thế phán đoán của giảng viên (Anderson và Krathwohl, 2001).

4.3.1 Bộ tiêu chí đánh giá

Bảng 4. Bộ tiêu chí đánh giá chất lượng đề cương

Tiêu chí	Cách kiểm tra đề xuất	Vai trò
C1. Rõ ràng CLO	Động từ, đối tượng, khả năng đo lường	Phát hiện CLO mơ hồ
C2. Bloom	Phân loại mức nhận thức	Kiểm tra mức Bloom
C3. Mức độ phù hợp giữa chuẩn đầu ra chương trình và chuẩn đầu ra học phần	Sử dụng mức độ tương đồng giữa các nội dung, ma trận đối chiếu và kết quả rà soát của chuyên gia	Điểm đánh giá mức độ phù hợp
C4. Độ phủ nội dung	Đối chiếu nội dung với CLO	Kết quả đánh giá độ bao phủ
C5. Tính nhất quán	Kiểm tra nhất quán giữa CLO, nội dung, dạy học và đánh giá	Cảnh báo về những điểm chưa nhất quán

4.3.2. Điểm đánh giá mức độ phù hợp

Với mỗi cặp chuẩn CLO-PLO, hệ thống có thể tính mức độ tương đồng cosin giữa hai véc-tơ biểu diễn tương ứng. Chỉ số này được tính theo công thức:

$$[\text{sim}(\text{PLO}, \text{CLO}) = \frac{v_{\text{PLO}} \cdot v_{\text{CLO}}}{|v_{\text{PLO}}| |v_{\text{CLO}}|}]$$

Giá trị càng cao cho thấy mức độ tương đồng về ngữ nghĩa giữa hai chuẩn

đầu ra càng lớn. Tuy nhiên, chỉ số này không đồng nghĩa với việc hai chuẩn đầu ra đã hoàn toàn phù hợp về mặt sự phạm. Vì vậy, điểm đánh giá mức độ phù hợp chỉ được sử dụng như một chỉ báo hỗ trợ, kết hợp với ma trận đối chiếu đã được phê duyệt và kết quả đánh giá, rà soát của giảng viên.

4.3.3. Human-in-the-Loop



Hình 3. Quy trình Human-in-the-Loop

Đề cương là sản phẩm có trách nhiệm chuyên môn nên AI không tự động đưa CLO, đánh giá hoặc nội dung vào phiên bản chính thức. Kết quả AI ở trạng thái ‘đề xuất’; giảng viên chấp nhận, chỉnh sửa, từ chối hoặc yêu cầu sinh lại. Hệ thống lưu phiên bản, người phê duyệt, thời điểm và nguồn tri thức đã truy xuất.

4.4. Thiết kế kiểm định mô hình

Thuyết minh xác định thử nghiệm với ít nhất 10 đề cương và khảo sát khoảng 15-20 giảng viên. Trong bài báo, đây là thiết kế đánh giá dự kiến; chưa có kết quả định lượng được báo cáo.

Bảng 5. Thiết kế kiểm định mô hình

Đối tượng đánh giá	Thiết kế	Chỉ số / kết quả
Chất lượng văn bản sinh	So sánh bản dự thảo với đề cương chuẩn khi có tài liệu tham chiếu	Chỉ số BLEU, ROUGE và đánh giá của chuyên gia
Module kiểm tra	So sánh các cảnh báo hoặc kết quả phân loại của hệ thống với đánh giá của chuyên gia	Độ chính xác, độ bao phủ, F1
Mức độ phù hợp	So sánh điểm đánh giá của hệ thống với kết quả đánh giá hoặc ma trận đối chiếu của chuyên gia	Mức độ tương quan
Người dùng	Khảo sát mức hữu ích và tin cậy	Giá trị trung bình, độ lệch chuẩn
Case study	Ít nhất 10 đề cương dự kiến	Phân tích lỗi và tình huống sử dụng

Mục tiêu kỹ thuật là độ chính xác cấu trúc và nội dung tối thiểu 85% khi so sánh với đề cương chuẩn. Đây là chỉ tiêu mục tiêu, không phải kết quả hiện tại. Khi có dữ liệu cần báo cáo khoảng tin cậy, phân bố lỗi và các trường hợp AI khác chuyên gia.

4.5. Thảo luận

Điểm khác biệt quan trọng của khung đề xuất so với cách tiếp cận chỉ sử dụng LLM để sinh văn bản là hệ thống được tổ chức quanh logic thiết kế học phần. PLO là đầu vào định hướng; CLO được tạo và kiểm tra theo Bloom; nội dung, hoạt động dạy học và đánh giá được gắn với CLO; sau đó toàn bộ đề cương được kiểm tra trước khi giảng viên phê duyệt. Như vậy, AI được đặt trong một workflow có ràng buộc thay vì được xem như công cụ soạn thảo tự do.

Về mặt công nghệ, RAG cho phép tách kho tri thức khỏi mô hình tạo sinh. Điều này phù hợp với các quy định và biểu mẫu có thể thay đổi theo thời gian. Khi quy định hoặc đề cương chuẩn được cập nhật, kho tri thức có thể được cập nhật và lập chỉ mục lại mà không nhất thiết phải huấn luyện lại LLM. Tuy nhiên, chất lượng truy xuất thông tin phụ thuộc nhiều vào chất lượng nguồn dữ liệu, cách chia đoạn, cách biểu diễn và phương pháp tìm

kiếm. Việc truy xuất thông tin không phù hợp có thể dẫn đến kết quả sinh ra sai lệch.

Về mặt sư phạm, các độ đo tự động chỉ là bằng chứng hỗ trợ. Độ tương đồng cosin chỉ phản ánh mức độ tương đồng về ngữ nghĩa và không thể tự xác định sự phù hợp về mặt sư phạm giữa các chuẩn đầu ra. Tương tự, BLEU và ROUGE chủ yếu phản ánh mức độ giống nhau giữa nội dung được sinh ra và nội dung tham chiếu. Do đó, kết quả của hệ thống cần được kiểm tra và xác nhận bởi giảng viên hoặc chuyên gia.

Về tính mới, nghiên cứu không khẳng định BERT, LLM, RAG hay Bloom là các công nghệ mới. Đóng góp được xác định ở cấp độ tích hợp: kết hợp cơ sở thiết kế giáo dục, kho tri thức theo bối cảnh cơ sở đào tạo, sinh nội dung, đánh giá alignment và cơ chế phê duyệt trong một khung kiến trúc thống nhất. Phạm vi thử nghiệm tại một chương trình CNTT của Trường Đại học Mở Hà Nội cũng giúp xác định rõ bối cảnh ứng dụng, nhưng đồng thời giới hạn khả năng khái quát.

Các hạn chế gồm: chưa có kết quả thực nghiệm; dữ liệu dự kiến 30-50 đề cương; chưa kiểm chứng trên nhiều ngành; chưa có bằng chứng định lượng về tiết kiệm thời gian; knowledge graph/ GraphRAG chưa triển khai. Hướng

tiếp theo là hoàn thiện nguyên mẫu, thử nghiệm, đánh giá chuyên gia và so sánh chiến lược truy xuất.

V. Kết luận

Bài báo đề xuất một khung kiến trúc ứng dụng AI nhằm hỗ trợ xây dựng và đánh giá đề cương học phần dựa trên chuẩn đầu ra, theo phương pháp DSR. Khung đề xuất kết hợp nguyên tắc Constructive Alignment, Backward Design và thang đo Bloom với LLM cùng cơ chế truy xuất, cung cấp thông tin (RAG), được tổ chức thành bốn tầng với sáu mô-đun. Điểm nhấn của nghiên cứu là cơ chế kiểm tra chất lượng và sự tham gia của giảng viên, trong đó hệ thống trí tuệ nhân tạo đưa ra đề xuất và cảnh báo, còn giảng viên kiểm tra, hiệu chỉnh và phê duyệt kết quả.

Khung đề xuất làm cơ sở phát triển nguyên mẫu trong một chương trình Công nghệ thông tin tại Trường Đại học Mở Hà Nội. Tiêu chí kiểm định gồm độ rõ CLO, Bloom, PLO-CLO alignment, độ phủ nội dung và tính nhất quán; kế hoạch kết hợp độ đo tự động, đánh giá chuyên gia và khảo sát. Bài báo chưa có kết quả thực nghiệm nên không kết luận AI đã tiết kiệm thời gian, nâng cao chất lượng hoặc bảo đảm alignment. Các lợi ích này cần được kiểm chứng bằng dữ liệu thực tế.

Tài liệu tham khảo

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Biggs, J. B., & Tang, C. (2011). *Teaching for quality learning at university: What the student does* (4th ed.). McGraw-Hill/Society for Research into Higher Education/Open University Press.
- Bộ Giáo dục và Đào tạo. (2021). *Thông tư số 17/2021/TT-BGDĐT quy định về chuẩn chương trình đào tạo; xây dựng, thẩm định và ban hành chương trình đào tạo các trình độ của giáo dục đại học*.
- Bộ Giáo dục và Đào tạo. (2025). *Thông tư số 04/2025/TT-BGDĐT quy định về kiểm định chất lượng chương trình đào tạo các trình độ của giáo dục đại học*.
- Đoàn Thị Minh Trinh, Nguyễn Quốc Chính, Nguyễn Hữu Lộc, Phạm Công Bằng, Gray, P. J., và Hồ Tấn Nhựt. (2012). *Thiết kế và phát triển chương trình đào tạo đáp ứng chuẩn đầu ra*. Nhà xuất bản Đại học Quốc gia Thành phố Hồ Chí Minh.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212-218.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459-9474.
- Li, Y., Rakovic, M., Poh, B. X., Gašević, D., & Chen, G. (2022). Automatic classification of learning objectives based on Bloom's taxonomy. In A. Mitrovic & N. Bosch (Eds.), *Proceedings of the 15th International Conference on Educational Data Mining* (pp. 530-537). International Educational Data Mining Society. <https://doi.org/10.5281/zenodo.6853191>
- Nguyễn Hữu Duyệt. (2017). Thiết kế đề cương chi tiết học phần để tổ chức dạy học nhất quán với chuẩn đầu ra chương trình. *Tạp chí Khoa học Đại học Đồng Tháp*, 29, 49-52. <https://doi.org/10.52714/dthu.29.12.2017.529>
- Nguyễn Thanh Nhân. (2024). Phương pháp xây dựng đề cương học phần theo định hướng CDIO. *Tạp chí Thiết bị Giáo dục*, 311(2).

- Sridhar, P., Doyle, A., Agarwal, A., Bogart, C., Savelka, J., & Sakr, M. (2023). Harnessing LLMs in curricular design: Using GPT-4 to support authoring of learning objectives. In *Proceedings of the Workshop on Empowering Education with LLMs - the Next-Gen Interface and Content Generation at AIED 2023* (pp. 139-150). CEUR Workshop Proceedings.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (Expanded 2nd ed.). Association for Supervision and Curriculum Development.

AI-DRIVEN COURSE SYLLABUS DEVELOPMENT AN OUTCOME-BASED APPROACH

Dinh Tuan Long¹, Truong Huu Thanh¹

Abstract: *Outcome-based course syllabus development requires consistent coordination among PLO, CLO, content, teaching activities, and assessment methods. This paper proposes an artificial intelligence architecture to support lecturers in developing and evaluating course syllabi. The study adopts Design Science Research, combined with literature review, conceptual analysis, and artifact design. The pedagogical foundations include Constructive Alignment, Backward Design, and Bloom's Taxonomy, while the technological foundations include large language models and Retrieval-Augmented Generation. The proposed architecture consists of four layers, interface, application, AI and business services, and data and knowledge, with six functional modules covering PLO analysis, content suggestion, CLO generation, teaching and assessment recommendations, quality and alignment checking, and syllabus synthesis. A Human-in-the-Loop mechanism ensures that AI outputs remain proposals subject to lecturer review, revision, and approval. The paper also proposes an evaluation design based on Bloom's classification, PLO-CLO alignment, content coverage, and consistency, combining automated metrics, expert assessments, and lecturer surveys. The current contribution is the proposed architecture and evaluation design.*

Keywords: *artificial intelligence, course learning outcomes, course syllabi, higher education*

¹ Hanoi Open University, Hanoi, Vietnam