

# PHÁT TRIỂN HỆ THỐNG HỎI ĐÁP ĐA PHƯƠNG THỨC DỰA TRÊN RAG TRONG LĨNH VỰC GIAO THÔNG - VĂN HÓA - DU LỊCH

**Thành viên thực hiện:**

Ngô Anh Tuấn (521H0379@student.tdtu.edu.vn)

Trần Ngọc Phát (521H0129@student.tdtu.edu.vn)

**Giảng viên hướng dẫn:**

TS. Trần Thanh Phước (tranthanhphuoc@tdtu.edu.vn)

Trường Đại học Tôn Đức Thắng, Quận 7, TP.HCM

**Tháng 09, 2025**

## I. GIỚI THIỆU

### 1. Bối cảnh thực hiện (Tập trung vào Giao thông, Văn hóa và Du lịch)

Trong bối cảnh hội nhập quốc tế và sự bùng nổ của du lịch thông minh (Smart Tourism), nhu cầu truy xuất thông tin **Giao thông, Văn hóa, và Du lịch** một cách tổng hợp và chính xác là rất lớn. Người dùng hiện đại cần một công cụ có thể trả lời các câu hỏi phức tạp, liên ngành, ví dụ: *"Làm thế nào để đi từ sân bay A đến Di sản B bằng phương tiện công cộng, và những quy tắc văn hóa nào cần lưu ý khi tham quan?"*

Lĩnh vực này tồn tại nhiều thách thức khiến các hệ thống hỏi đáp thông thường gặp khó khăn:

- **Tính Đa Phương Thức (Multimodality):** Dữ liệu không chỉ là văn bản (mô tả địa điểm, quy định giao thông) mà còn là **hình ảnh** (bản đồ, biển báo, hình ảnh điểm đến). Các hệ thống truyền thống chỉ xử lý văn bản không thể trả lời các câu hỏi dựa trên nhận dạng thị giác (ví dụ: *"Đây là biển báo gì và nó có ý nghĩa gì trong giao thông?"*).

- **Tính Liên Ngành và Quan hệ Phức tạp:** Thông tin Giao thông (biển báo) có mối quan hệ chặt chẽ với Văn hóa (lễ hội, phong tục, ứng xử) và Du lịch (địa điểm, dịch vụ). Việc truy vấn đòi hỏi khả năng **kết nối logic** giữa các thực thể (ví dụ: Địa điểm A **có thể đến bằng** Tuyến Giao thông B **trong thời điểm** Lễ hội C).
- **Hạn chế của Chatbot Tổng quát (Hallucination):** Các LLM tổng quát thường thiếu dữ liệu chuyên sâu, chính xác về **luật giao thông địa phương, biển báo, hoặc các phong tục văn hóa cụ thể** của Việt Nam, dẫn đến việc đưa ra các câu trả lời sai lệch (*hallucination*), đặc biệt là các thông tin có tính cập nhật cao.

→ Giải pháp Đề xuất: **“RAG Đa Phương Thức kết hợp Knowledge Graph”**

Để phát triển một hệ thống đáp ứng được tính phức tạp và đa dạng của lĩnh vực này, đề tài này đề xuất kiến trúc tiên tiến: **RAG Đa Phương Thức (Multimodal RAG)** kết hợp với **Đồ thị Tri thức (Knowledge Graph – KG)**.

- **RAG Đa Phương Thức:** Sử dụng các **mô hình nhúng đa phương thức** để xử lý và lưu trữ đồng thời dữ liệu **Văn bản và Hình ảnh** dưới dạng vector. Điều này cho phép hệ thống truy xuất không chỉ dựa trên ngữ nghĩa văn bản mà còn dựa trên nội dung hình ảnh.
- **Knowledge Graph (KG):** Là thành phần cốt lõi để quản lý các **mối quan hệ logic** giữa các thực thể Giao thông (biển báo), Văn hóa (sự kiện, quy tắc) và Du lịch (địa điểm). KG giúp hệ thống thực hiện các **lập luận đa bước** và **truy vấn cấu trúc** (ví dụ: tìm tất cả các địa điểm du lịch nổi tiếng về lịch sử ở Tp.HCM).

### Ý nghĩa thực tiễn

Hệ thống được kỳ vọng sẽ:

- **Tăng độ chính xác và tính toàn diện** cho các câu hỏi liên ngành bằng cách tổng hợp thông tin từ văn bản, hình ảnh, và mối quan hệ cấu trúc.
- **Hỗ trợ du khách và người dân** lập kế hoạch di chuyển, tìm hiểu văn hóa và khám phá các địa điểm du lịch một cách hiệu quả.
- **Tạo ra một nền tảng tri thức có cấu trúc** cho lĩnh vực Giao thông Văn hoá và Du lịch Việt Nam.

## 2. Mục tiêu dự án (Tối ưu hóa cho Đa phương thức)

**Mục tiêu chính:**

- Phát triển một Hệ thống Chatbot HỏiĐáp Đa Phương Thức Tiếng Việt về Giao thông – Văn hoá – Du lịch, dựa trên kiến trúc Multimodal RAG tích hợp Knowledge Graph.
- Xây dựng kho dữ liệu Đa Phương Thức và Đồ thị Tri thức cho lĩnh vực này.

**Mục tiêu cụ thể:**

- Xử lý và Chuẩn hóa dữ liệu Đa Phương Thức (Văn bản và Hình ảnh) từ các nguồn về Giao thông, Văn hóa và Du lịch.
- Trích xuất Thực thể – Quan hệ để xây dựng Đồ thị Tri thức Giao thông Văn hoá Du lịch.
- Phát triển mô hình hỏi đáp dựa trên kiến trúc Multimodal RAG kết hợp Knowledge Graph (ví dụ: sử dụng Fusion Retrieval để kết hợp truy xuất vector và truy vấn graph).
- Triển khai giao diện demo có khả năng chấp nhận đầu vào đa phương thức (văn bản, ảnh) và trả lời bằng văn bản, kèm theo các hình ảnh/bản đồ minh họa liên quan.

**Giới hạn phạm vi đồ án (Giữ nguyên, bổ sung chi tiết):**

- Dữ liệu: Tập trung vào các quy tắc Biển báo Giao thông, các quy tắc văn hóa, và các điểm Văn hoá Du lịch trọng điểm.
- Dữ liệu đầu vào: bao gồm cả Văn bản và Hình ảnh.
- Ngôn ngữ: Tiếng Việt.

**II. Kiến thức nền tảng****1. Bối cảnh**

Với tính chất phức tạp của lĩnh vực Giao thông Văn hóa Du lịch, các hệ thống HỏiĐáp (QA) cần vượt qua giới hạn của truy vấn văn bản đơn thuần. Hai xu hướng tiên tiến được đề tài áp dụng là:

- Multimodal RAG: Mở rộng RAG từ xử lý văn bản sang xử lý Văn bản và Hình ảnh (hoặc các phương thức khác), cho phép mô hình hiểu được ngữ cảnh thị giác (biển báo giao thông, hình ảnh địa điểm) để trả lời.
- KG-RAG: Tích hợp Đồ thị Tri thức để hỗ trợ các truy vấn yêu cầu lập luận logic, tìm kiếm lộ trình và kết nối quan hệ giữa các lĩnh vực, điều mà Vector Search truyền thống gặp nhiều khó khăn.

**2. Các công trình nổi bật**

Nhóm Công trình	Mục đích trong Đồ án	Các Công trình Tiêu biểu (Đã sắp xếp)
Multimodal RAG & Multimodal Embedding	Cung cấp cơ sở lý thuyết để xử lý dữ liệu hình ảnh và kết hợp truy xuất vector đa phương thức vào RAG. (Mặc dù bạn không liệt kê chi tiết, cần ngầm hiểu sẽ dùng các công trình về Multimodal Embedding và Multimodal QA).	Các công trình nền tảng về Multimodal Embedding (CLIP, ViLT, v.v.) và các kiến trúc Multimodal RAG (ví dụ: VRAG).
KGRAG (Tích hợp RAG với Đồ thị Tri thức)	Cung cấp phương pháp để kết hợp truy xuất ngữ nghĩa (vector) và truy vấn cấu trúc (graph) cho các câu hỏi về lộ trình/quan hệ.	Yunfan Gao et al. (2023) [1]; Sanmartin et al. (2024) KGRAG: Chain of Explorations [2]; Zhu et al. (2025) KG2RAG: Graphguided Chunk Expansion [3]; Mavromatis Karypis (2024) GNNRAG [13].
Knowledge Graph Fundamentals (Xây dựng và Truy vấn KG)	Cung cấp nền tảng để mô hình hóa các mối quan hệ Giao thông, Văn hóa, Du lịch (ví dụ: tuyến đường, sự kiện liên quan, địa điểm kề cận).	Hogan et al. (2021) Khảo sát KG [4]; Angles et al. (2017) Ngôn ngữ truy vấn đồ thị [7]; Hur et al. (2021) Xây dựng KG tự động [8].
Ứng dụng RAG Chuyên ngành & Trong nước	Tham khảo kinh nghiệm triển khai RAG/KGRAG trong các lĩnh vực cần độ chính xác và chuyên môn cao.	Đại học Bách Khoa TP.HCM (2024) KG cho giáo dục [19]; Dự án GraphRAG (HuyLD) QA Lịch sử Việt [20].