

**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**  
**KHOA KHOA HỌC MÁY TÍNH**

---



**ĐỒ ÁN MÔN HỌC**  
**TRUY VẤN THÔNG TIN ĐA PHƯƠNG TIỆN**  
**HỆ THỐNG HỎI ĐÁP DỊCH VỤ CÔNG**

**Giảng viên hướng dẫn : ThS. Đỗ Văn Tiến**  
**Sinh viên thực hiện 1 : Nguyễn Thị Thùy**  
**Mã sinh viên 1 : 21521514**  
**Sinh viên thực hiện 2 : Đào Duy Thịnh**  
**Mã sinh viên 2 : 21520463**

Thành phố Hồ Chí Minh, tháng 1 năm 2024

# MỤC LỤC

<b>CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI.....</b>	<b>2</b>
1. Lý do chọn đề tài.....	2
2. Mục tiêu đồ án.....	2
3. Các ứng dụng hiện có.....	3
<b>CHƯƠNG 2. PHƯƠNG PHÁP THỰC HIỆN.....</b>	<b>5</b>
1. Kỹ thuật Retrieval-Augmented Generation - RAG.....	5
1.1. Giới thiệu về RAG.....	5
1.2. Kiến trúc của RAG.....	5
2. Áp dụng RAG trong Hệ thống hỏi đáp Dịch vụ công.....	6
3. Xây dựng các tính năng cho ứng dụng.....	9
3.1. Bộ phân loại nhị phân.....	9
3.2. Bộ thêm dấu.....	10
3.3. Lưu bộ nhớ đệm.....	11
4. Cấu trúc hệ thống.....	11
<b>CHƯƠNG 3. ỨNG DỤNG DEMO.....</b>	<b>13</b>
<b>CHƯƠNG 4. TỔNG KẾT.....</b>	<b>15</b>
1. Kết quả đạt được.....	15
2. Ưu điểm, hạn chế và hướng mở rộng.....	15
2.1. Ưu điểm.....	15
2.2. Điểm hạn chế.....	15
2.3. Hướng mở rộng.....	15
<b>TÀI LIỆU THAM KHẢO.....</b>	<b>17</b>

## CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI

### 1. Lý do chọn đề tài

Xuất phát từ thực tế, nhóm thấy rằng cộng đồng đang phải đối mặt với những vấn đề liên quan đến việc tìm hiểu và thực hiện các thủ tục hành chính.

Các câu hỏi thường gặp của người dân về các dịch vụ công cần một phương tiện hiệu quả để giải đáp. Tuy nhiên, người dân thường gặp khó khăn khi tìm kiếm thông tin chính xác và chi tiết từ trang web và các nguồn thông tin khác. Mất nhiều thời gian và công sức để đạt đến thông tin mong muốn.

Xuất phát từ thực tế này, nhóm muốn xây dựng Hệ thống hỏi đáp dịch vụ công để hỗ trợ người dân trong việc giải quyết các thủ tục hành chính. Ứng dụng này có nhiều tiềm năng mang lại lợi ích cho cả người dân và cơ quan hành chính nhà nước.

#### - Một số lợi ích cho người dân:

+ Tiết kiệm thời gian và công sức: ứng dụng có thể cung cấp thông tin về các thủ tục hành chính một cách nhanh chóng và thuận tiện, giúp người dân đặc biệt là những người cao tuổi giảm thiểu thời gian và công sức tìm hiểu, chuẩn bị hồ sơ.

+ Tăng cường tính minh bạch và tiếp cận của dịch vụ công: ứng dụng có thể cung cấp thông tin chính xác và đầy đủ về các thủ tục hành chính, giúp người dân hiểu rõ hơn về quyền lợi và nghĩa vụ của mình.

#### - Một số lợi ích cho cơ quan hành chính nhà nước:

+ Giảm tải công việc cho cán bộ công chức: công dân chủ động tìm hiểu và có kiến thức về các quy trình, hồ sơ, thủ tục hành chính giúp giảm thiểu lượng công việc ở khâu hướng dẫn cho các cán bộ.

+ Tiết kiệm chi phí: tiết kiệm chi phí đào tạo và tuyển dụng nhân lực, giúp cơ quan hành chính nhà nước tiết kiệm chi phí vận hành.

+ Hỗ trợ 24/7: Luôn luôn có mặt để giải đáp thắc mắc cũng như hỗ trợ người dân trong quá trình làm dịch vụ công.

+ Thúc đẩy số hoá: Giúp chính phủ nâng cao chương trình số hóa, thành phố thông minh.

## 2. Mục tiêu đề án

Mục tiêu chính của ứng dụng này là tạo điều kiện thuận lợi cho người dân để tìm kiếm và sử dụng các dịch vụ công một cách dễ dàng và hiệu quả. Các mục tiêu đó được chia thành các ý như sau:

- Khả năng dễ dàng truy cập thông tin để đảm bảo rằng người dùng có thể dễ dàng truy cập thông tin về các dịch vụ công một cách nhanh chóng và thuận tiện.
- Tích hợp thông tin đa nguồn để cung cấp thông tin từ nhiều nguồn khác nhau, bao gồm cả các cơ quan và tổ chức chính phủ cũng như các dịch vụ công tư để dữ liệu mang đến người dùng đa dạng mà đúng trọng tâm câu hỏi mà người dùng cần giải đáp.
- Ứng dụng cần cung cấp thông tin chi tiết về mỗi dịch vụ công, bao gồm quy trình, yêu cầu và hồ sơ cần thiết.
- Cần đảm bảo tính minh bạch cũng như sự uy tín của dữ liệu được dùng để trả lời câu hỏi của người dùng.

## 3. Các ứng dụng hiện có

Theo như nhóm tìm hiểu, đã có nhiều chủ trương, kế hoạch bài báo nói về việc cần thiết của các phần mềm tư vấn thủ tục hành chính cũng như việc ứng dụng trí tuệ nhân tạo trong lĩnh vực hành chính.

Theo khảo sát của nhóm nghiên cứu về ứng dụng AI trên lĩnh vực hành chính công, hiện tại vẫn chưa có chương trình hay hệ thống hỏi đáp thông minh được áp dụng để giải đáp thắc mắc của người dân một cách tự động. Hiện nay, để truy vấn và hỏi đáp các thông tin liên quan đến dịch vụ hành chính công thì hầu như người dùng phải tìm kiếm và đưa câu hỏi để các nhân viên trả lời thông qua trang web. Hình 1 mô

**CỔNG DỊCH VỤ CÔNG QUỐC GIA**  
Kết nối, cung cấp thông tin và dịch vụ công mọi lúc, mọi nơi

Đăng ký Đăng nhập

Giới thiệu Thông tin và dịch vụ Thanh toán trực tuyến Phản ánh kiến nghị Thủ tục hành chính Hỗ trợ

Cổng dân Doanh nghiệp Dịch vụ công trực tuyến Dịch vụ công nổi bật Tra cứu hồ sơ Tòa án nhân dân

Câu hỏi thường gặp

Trang chủ > Câu hỏi thường gặp

Số đó -- Chọn Bộ ngành -- Tìm kiếm

**Tìm thấy 14 kết quả**

- Hồ sơ cấp lại sổ đỏ cho hộ gia đình, cá nhân trong trường hợp bị mất sổ đỏ?
- Năm 2010, tôi đứng chủ quyền nên biết thự trong khu quy hoạch Đô thị mới (có Sổ Đỏ). Năm 2011, tôi đi du học và sau đó kết hôn ở lại Đức. Cuối tháng 11 tôi bắt đầu tiến hành xây dựng nhà, do tôi sống ở Đức nên tôi nhờ mẹ tôi giám sát và quản lý. Đồng thời, tôi ký Giấy ủy quyền cho mẹ được phép vay ngân hàng, được thế chấp, xin phép xây dựng, làm đăng hồ nước, xin sổ nhà, hoàn công.v.v.Xin hỏi 1. Năm 2013 tôi đã có quyết định cấp sổ nhà, nhưng đến nay vẫn chưa hoàn công do nhà xây chưa xong. Hiện tôi đang xin nhập quốc tịch Đức, yêu cầu ở phía Đức là tôi phải từ bỏ quốc tịch VN. Như vậy tôi có được giữ nhà đất của tôi ở VN không? Mẹ tôi hiện giờ đang mắc bệnh hiểm nghèo, nếu tôi về VN ủy quyền lại cho mẹ tôi toàn bộ, là không may mẹ tôi qua đời thì như thế nào? Hiện nay tôi vẫn chưa có kế hoạch về VN sinh sống. Vậy theo pháp luật của VN tôi nên làm gì?
- Thời gian giải quyết cấp sổ đỏ cho hộ gia đình, cá nhân lần đầu là bao lâu?
- Thời gian thực hiện thủ tục cấp lại sổ đỏ cho hộ gia đình, cá nhân trong trường hợp bị mất sổ đỏ?
- Thủ tục sang tên sổ đỏ được tiến hành như thế nào
- Trình tự cấp lại sổ đỏ cho hộ gia đình, cá nhân trong trường hợp bị mất?
- Trình tự, thủ tục cấp sổ đỏ cho hộ gia đình, cá nhân?
- Trình tự, thủ tục cấp sổ đỏ cho người nhận chuyển nhượng quyền sử dụng đất, mua nhà ở, công trình xây dựng trong các dự án phát triển nhà ở?
- Tôi làm Giấy ủy quyền cho người nhà làm sổ đỏ mảnh đất ở của gia đình. Đại sứ quán có thể chứng thực chữ ký của tôi được không?
- Tôi muốn làm thủ tục đăng ký sổ đỏ, tôi cần những thủ tục gì? (Đất đã tranh chấp được cơ quan Tòa án tuyên giao quyền sử dụng đất, cơ quan THADS thực hiện giao đất). Tôi muốn làm thủ tục sang tên đăng ký đối với tài sản mua từ cơ quan tài chính (tài sản được tuyên tịch thu sung quỹ nhà nước do Bẩn án của Tòa án tuyên), tôi cần làm những thủ tục gì?

Hiện thị 10 trên 14 bản ghi

tả giao diện trên trang dịch vụ trang dichvucong.gov.vn có đơn vị chủ quản là Văn phòng Chính phủ.

*Hình 1. Mô tả giao diện tìm kiếm trên website dichvucong.gov*

Mặc dù trang dịch vụ công dichvucong.gov.vn của Văn phòng Chính phủ cung cấp nguồn thông tin uy tín, và đầy đủ. Tuy nhiên, người dùng sẽ gặp khó khăn trong việc tìm kiếm câu trả lời cho câu hỏi mong muốn, đôi khi phải mất rất nhiều thời gian để tìm kiếm thông tin từ các bài viết có liên quan.

Dựa trên những phân tích và khảo sát những vấn đề trên, nhóm nghiên cứu đã quyết định ứng dụng các kỹ thuật Xử lý ngôn ngữ Tự nhiên kết hợp với Máy học để xây dựng một Hệ thống hỏi đáp dịch vụ công. Nguồn dữ liệu để phản hồi cho người dùng được lấy trực tiếp từ trang dichvucong.gov.vn của chính phủ. Điều này đảm bảo các thông tin trả lời của ứng dụng chính xác dựa trên dữ liệu thực tế.

## CHƯƠNG 2. PHƯƠNG PHÁP THỰC HIỆN

Kỹ thuật chính mà nhóm dùng để xây dựng chatbot này là kỹ thuật Retrieval-Augmented Generation (RAG) cùng với một số kỹ thuật khác để xây dựng các tính năng của chatbot.

### 1. Kỹ thuật Retrieval-Augmented Generation - RAG

#### 1.1. Giới thiệu về RAG

- RAG là một kỹ thuật giúp cải thiện độ chính xác và tin cậy của generative AI cách cho phép mô hình ngôn ngữ lớn (LLM) sử dụng thêm các dữ liệu bổ sung mà không cần đào tạo lại

- RAG xuất hiện lần đầu tiên trong bài báo "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," một bài do Patrick Lewis và một nhóm tại Facebook AI Research công bố năm 2020.

- Một số lợi ích khi sử dụng RAG để xây dựng hệ thống hỏi đáp:

+ Trả lời được các thông tin chuyên sâu về chủ đề được quan tâm (trong đồ án này là chủ đề về dịch vụ công).

+ Cho phép Xác định nguồn thông tin của các trả lời được sinh ra từ các mô hình ngôn ngữ lớn.

+ Có khả năng cập nhật và mở rộng linh hoạt.

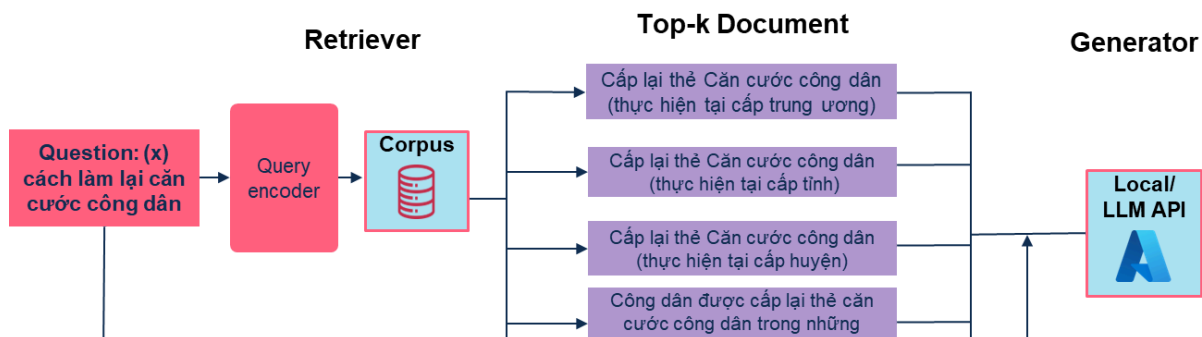
#### 1.2. Kiến trúc của RAG

Trong mô hình RAG sẽ có 3 thành phần:

- Corpus: Kho dữ liệu bên ngoài để cung cấp thêm thông tin cho bộ phận generator.

- Retriever: Từ một câu truy vấn  $x$ , trước tiên, mô hình RAG sẽ truy xuất một top-k các tài liệu hoặc đoạn văn có liên quan đến  $X$  từ một corpus bên ngoài. Điều này được thực hiện bằng cách sử dụng cơ chế truy xuất, thường dựa trên các biểu diễn vector của tài liệu và câu truy vấn.

- Generator: Sau khi các đoạn văn hoặc tài liệu có liên quan được truy xuất, chúng sẽ kết hợp cùng với câu truy vấn ban đầu để đưa vào các mô hình tạo sinh văn



bản. Sau đó, mô hình này tạo ra phản hồi, tận dụng cả kiến thức và thông tin được đào tạo trước từ các đoạn văn được truy xuất.

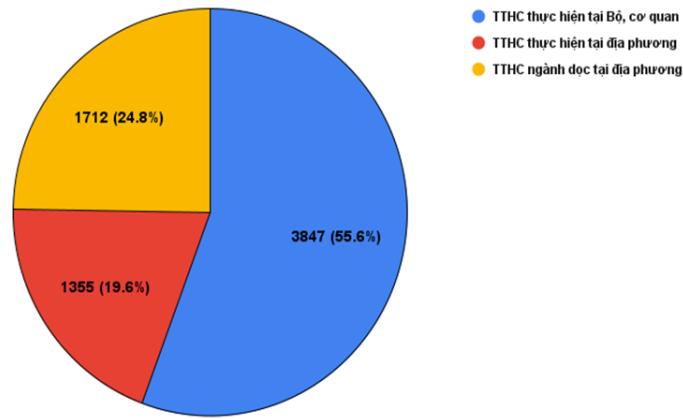
*Hình 2. Sơ đồ các bước thực hiện của RAG*

- Question: Đây là câu truy vấn mà người dùng muốn được trả lời
- Retriever: Hệ thống mã hóa câu truy vấn đầu vào để tìm kiếm các tài liệu hoặc đoạn văn bản có trong kho dữ liệu.
- Top-k documents: Sau bước truy vấn ta thu được top-k tài liệu liên quan đến câu truy vấn ban đầu nhất.
- Kết hợp Question + Top-k documents: Kết hợp câu truy vấn ban đầu và tài liệu truy xuất được lại với nhau để tạo thành prompt cho bộ phận generator. Sự kết hợp này cung cấp ngữ cảnh phong phú mà hệ thống sẽ sử dụng để tạo ra câu trả lời cuối cùng.
- Generator: Sử dụng prompt ở trên vào các mô hình ngôn ngữ lớn để tạo ra các câu trả lời mạch lạc, tự nhiên. Nó sẽ cung cấp phản hồi dựa trên khối lượng dữ liệu khổng lồ mà đã đã được học cùng với thông tin ngữ cảnh mà ta đã bổ sung vào.
- Output (answer): Đây là câu trả lời cuối cùng mà hệ thống tạo ra để trả lời truy vấn đầu vào ban đầu.

## **2. Áp dụng RAG trong Hệ thống hỏi đáp Dịch vụ công**

Hệ thống hỏi đáp Dịch vụ công ứng dụng kỹ thuật RAG với 3 thành phần Corpus, Retriever, Generator được thiết lập như sau:

- \* Corpus: sử dụng dữ liệu được thu thập trên web dichvucong.gov gồm:
  - Thủ tục hành chính có tổng cộng 6914 mẫu (*Hình 3*).
  - + TTHC thực hiện tại Bộ, cơ quan có 3847 mẫu.
  - + TTHC thực hiện tại địa phương có 1355 mẫu.
  - + TTHC ngành dọc tại địa phương có 1712 mẫu.



Hình 3. Chi tiết số liệu về các TTHC

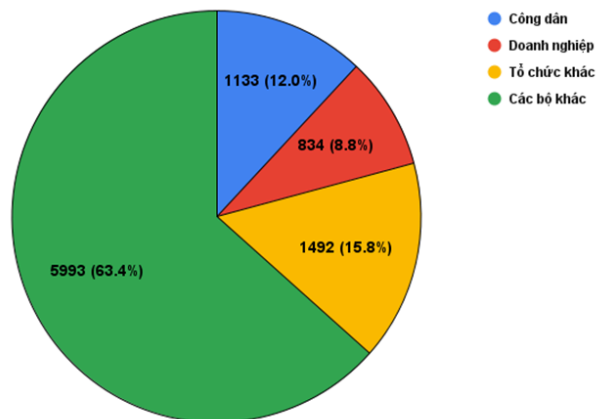
- Câu hỏi liên quan có tổng cộng 9452 mẫu (Hình 4).

+ Công dân: 1133 mẫu

+ Doanh nghiệp: 834 mẫu

+ Tổ chức khác: 1492 mẫu

+ Các bộ khác: 5993 mẫu



Hình 4. Chi tiết số liệu về các câu hỏi liên quan

### \* Retriever

Trong ứng dụng này, nhóm đề xuất một phương pháp xây dựng retriever sử dụng kỹ thuật retrieval và rerank. Phương pháp này bao gồm hai giai đoạn:

- Giai đoạn retrieval: nhóm sử dụng kết hợp hai kỹ thuật là BM25 và SimCSE. Đối với BM25, một mô hình truy vấn thông tin truyền thống, tính sự tương đồng giữa câu truy vấn và các văn bản trong tập dữ liệu dựa trên xác suất. Một cải tiến đáng chú ý khác là việc tích hợp SimCSE, một mô hình Self-supervised Learning dựa trên việc tối ưu hóa sự tương đồng giữa các biểu diễn câu. Nhóm sử dụng mô hình huấn luyện trước VoVanPhuc/sup-SimCSE-VietNameese-phobert-base để học biểu diễn câu tiếng



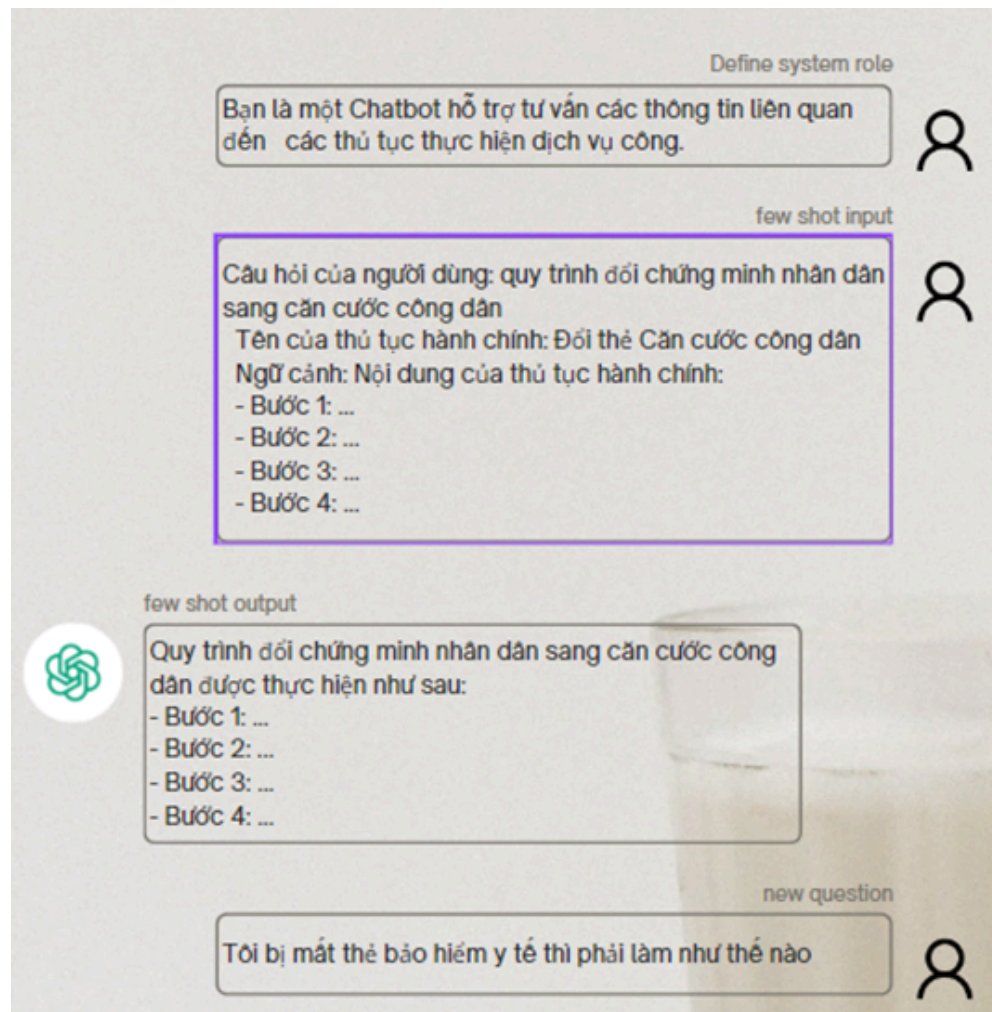
Việt, cung cấp khả năng hiểu biểu diễn ngôn ngữ tự nhiên trong tập dữ liệu của chúng tôi.

- Giai đoạn rerank: nhóm sử dụng mô hình huấn luyện trước 'unicamp-dl/mMiniLM-L6-v2-mmarco-v2' để cân nhắc lại thứ tự của các văn bản được chọn từ giai đoạn Retrieval, dựa trên việc hiểu sâu và đánh giá lại sự tương đồng giữa câu truy vấn và văn bản.

### \* Generator

Để tối ưu hóa khả năng tương tác của hệ thống, chúng tôi giới thiệu một bộ phần generator độc đáo và mạnh mẽ, sử dụng sự kết hợp hài hòa giữa Langchain và mô hình ngôn ngữ lớn GPT-3.5.

- Langchain: Xác định vai trò và tạo ra câu trả lời hiệu quả, ổn định. Langchain như một phần quan trọng của hệ thống, chịu trách nhiệm xác định rõ vai trò của mô hình ngôn ngữ trong cuộc trò chuyện. Bằng cách này, chúng ta có khả năng tối ưu hóa trải nghiệm người dùng thông qua sự hiểu biết rõ ràng về mục đích và nhiệm vụ của hệ thống. Hơn nữa, Langchain không chỉ đóng vai trò trong việc xác định vai trò, mà còn kết hợp thông tin từ “few shot learning” để tạo ra các câu trả lời mạnh mẽ và ổn định.



*Hình 5. Dùng Langchain để xác định vai trò và truyền vào few shot*

- Mô hình ngôn ngữ lớn GPT-3.5: độ chính xác cao và khả năng triển khai linh hoạt.

+ Nhóm sử dụng GPT-3.5, một mô hình ngôn ngữ lớn và mạnh mẽ, để cung cấp khả năng tạo văn bản và phản hồi ngôn ngữ vô song. Sử dụng API của GPT-3.5, chúng ta đạt được một độ tin cậy cao và khả năng triển khai linh hoạt, mang lại trải nghiệm tương tác độc đáo và thú vị cho người dùng.

+ Vì việc sử dụng GPT-3.5 API sẽ tính phí theo lượng sử dụng nên trước khi quyết định chọn mô hình ngôn ngữ lớn là GPT-3.5 API nhóm cũng đã cân nhắc về các mô hình ngôn ngữ mã nguồn mở và đưa ra bảng so sánh ở Bảng 1. Nhưng xét tới ứng dụng mà nhóm đang xây dựng là dịch vụ công, các thông tin được công bố công khai và quy mô đầu tư ban đầu còn hạn chế trong đồ án nên nhóm đã quyết định sử dụng API của OpenAI.

<b>Sự lựa chọn large language model</b>		
	<b>LLM mã nguồn mở</b>	<b>OpenAI API</b>
<b>Ưu điểm</b>	<ul style="list-style-type: none"> <li>- Độ bảo mật cao</li> <li>- Không tốn chi phí khi sử dụng</li> </ul>	<ul style="list-style-type: none"> <li>- Yêu cầu ít tài nguyên để vận hành</li> <li>- Khả năng mở rộng tốt</li> <li>- Hiệu quả cao, ổn định</li> </ul>
<b>Nhược điểm</b>	<ul style="list-style-type: none"> <li>- Khả năng mở rộng thấp</li> <li>Chi phí đầu tư ban đầu lớn</li> <li>Hiệu quả không cao</li> <li>- Cần nguyên tài nguyên để vận hành</li> </ul>	<ul style="list-style-type: none"> <li>- Phải trả phí theo lược sử dụng</li> <li>- Phụ thuộc vào bên thứ 3</li> <li>- Tính bảo mật không cao</li> </ul>

*Bảng 1: Bảng so sánh mô hình ngôn ngữ lớn*

### **3. Xây dựng các tính năng cho ứng dụng**

#### **3.1. Bộ phân loại nhị phân**

Để tối thiểu chi phí trong quá trình sử dụng cũng như tập trung tài nguyên để hỗ trợ các vấn đề liên quan đến dịch vụ công, nhóm xây dựng một bộ phân loại nhị phân để xác định một câu hỏi của người dùng có thuộc lĩnh vực dịch vụ công hay không.

Nếu câu hỏi thuộc lĩnh vực dịch vụ công thì sẽ được chuyển đến bước xử lý tiếp theo, ngược lại thì trả về không hỗ trợ.

- Xây dựng bộ dữ liệu
- + Các câu hỏi, thủ tục hành chính trong Corpus: gán nhãn là 1
- + Thu thập các câu hỏi ở các lĩnh vực khác: gán nhãn là 0

	<b>Nhãn 1</b>	<b>Nhãn 0</b>
<b>Tập huấn luyện</b>	25848	21733
<b>Tập kiểm tra</b>	6472	5424
<b>Toàn bộ</b>	32320	27157

*Bảng 2: Phân phối dữ liệu huấn luyện mô hình phân loại*

- Huấn luyện mô hình
- + Sử dụng phương pháp TF-IDF để biểu diễn văn bản
- + Huấn luyện các mô hình phân loại như Suppor Vector Machine, Naive Bayes, Logistic regression. Và nhóm sử dụng mô hình cuối cùng cho kết quả tốt nhất là Suppor Vector Machine.

	<b>Accuracy</b>	<b>F1 - weighted</b>
<b>SVM</b>	99.44	99.44
<b>Naive Bayes</b>	97.91	97.9
<b>Logistic regression</b>	98.87	98.87

*Bảng 3: Kết quả huấn luyện các mô hình phân loại*

### 3.2. Bộ thêm dấu

- Để tăng thêm trải nghiệm và hỗ trợ người dùng, nhóm đã xây dựng một bộ thêm dấu để có thể hiểu được những câu hỏi không dấu.
- Nhóm huấn luyện bộ thêm dấu là một mô hình có dạng sequence2sequence, tính chỉnh từ mô hình huấn luyện trước model vinai/bartpho-syllable-base.
- Dữ liệu huấn luyện có dạng (x\_input, y\_output) với:
  - + y\_output: câu hỏi, thủ tục hành chính trong Corpus
  - + x\_input: câu hỏi, thủ tục hành chính trong Corpus đã được bỏ dấu

🗨️ Text2Text Generation

thu tục lam can cuoc cong dan nhu the nao

Bảng 6: Minh họa bộ thêm dấu

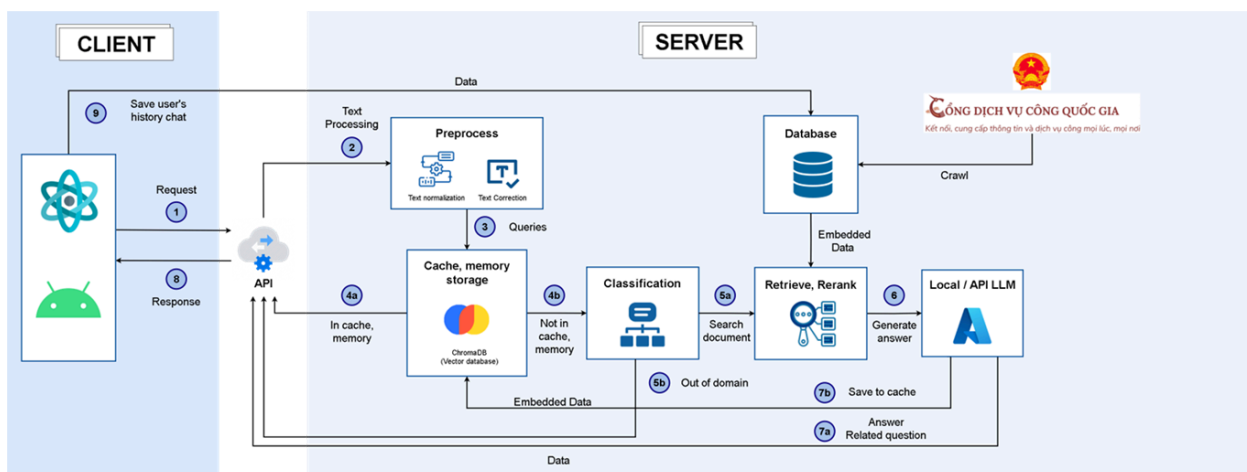
### 3.3. Lưu bộ nhớ đệm

Nhóm sử dụng Chroma DB để lưu lại các câu hỏi đã được trả lời trước đó, giúp tiết kiệm chi phí, giảm thời gian phản hồi.

Việc sử dụng Chroma DB để lưu bộ nhớ đệm trong hệ thống truy vấn văn bản có một số lợi ích quan trọng sau:

- Tăng tốc độ truy vấn: Chroma DB là một cơ sở dữ liệu cột, được thiết kế để tối ưu hóa các truy vấn quét lượng lớn dữ liệu. Khi sử dụng Chroma DB để lưu bộ nhớ đệm, các truy vấn truy dữ liệu trong bộ nhớ đệm sẽ được xử lý nhanh hơn.
- Hỗ trợ các phương pháp nhúng (embedding) văn bản hiện đại cùng với nhiều phương pháp tính độ tương tự giữa các văn bản
- Có khả năng kết hợp linh hoạt với Langchain - một thư viện hỗ trợ mạnh mẽ trong lĩnh vực Xử lý ngôn ngữ tự nhiên

### 4. Cấu trúc hệ thống



Hình 7: Cấu trúc hệ thống của ứng dụng

Hệ thống Chatbot của chúng tôi được thiết kế như Hình 7, với chi tiết quy trình như sau:

\* Giai đoạn 1 - Tiền xử lý: Trước khi đưa câu hỏi người dùng vào hệ thống để tìm trong dữ liệu cũng như tạo câu phản hồi thì chúng tôi cho câu hỏi qua các bước tiền xử lý như sau:

- Chuẩn hoá: Chúng tôi chỉnh sửa các câu hỏi về một dạng thống nhất cũng như dạng giúp chúng tôi đảm bảo được hiệu suất cao khi dùng để tìm kiếm trong database ở giai đoạn sau. Các bước trong chuẩn hoá có thể kể đến như:

- + Sửa định dạng từ

+ Sửa dấu từ

+ Bỏ các khoảng trắng trùng lặp

- Thêm dấu cho câu. Để khiến câu hỏi thêm tin cậy cũng như chính xác cho các giai đoạn sau, câu hỏi tra qua một bộ thêm dấu để hoàn tất việc thêm dấu cho câu

\* Giai đoạn 2 - Gợi nhớ: Ở giai đoạn này các câu hỏi của người dùng sẽ được kiểm tra có ở trong bộ nhớ đệm hay không từ đó quyết định trả về câu trả lời có sẵn hoặc tiếp tục tìm kiếm thông tin. Điều này là rất cần thiết bởi nhờ đó chúng ta có thể giảm thiểu thời gian, chi phí khi vận hành.

\* Giai đoạn 3 - Phân loại: Nếu câu hỏi người dùng không ở trong bộ nhớ của server ta sẽ tiến hành phân loại câu để kiểm tra câu hỏi có thuộc lĩnh vực dịch vụ công hay không, nếu thuộc thì tiếp tục bước sau, không thì trả về kết quả luôn cho người dùng

\* Giai đoạn 4 - Tìm kiếm: Sau khi nhận được các câu truy vấn từ giai đoạn 3, ta bắt đầu tìm kiếm trên toàn database để đưa ra kiến thức, dữ liệu cho giai đoạn 5.

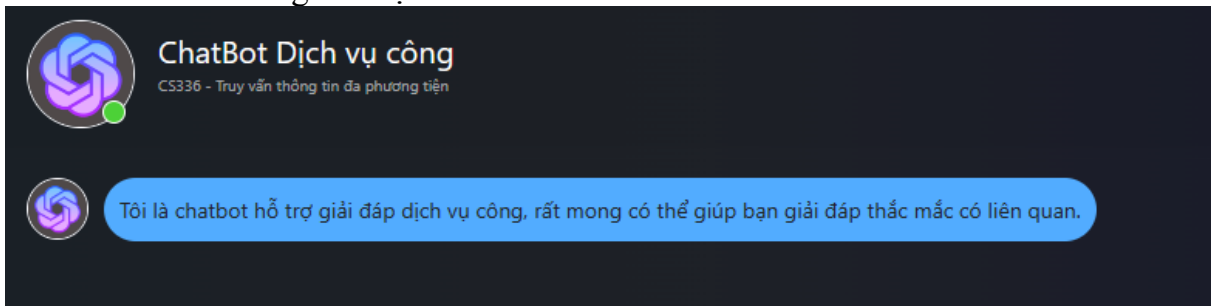
\* Giai đoạn 5 - Kiến tạo: Bằng cách tận dụng sức mạnh của mô hình ngôn ngữ lớn - GPT 3-5, ta đưa vào câu hỏi gốc của người dùng và các kiến thức có được từ giai đoạn 4 để có thể tạo ra một câu trả lời tự nhiên, gần gũi, có tính logic với người đọc.

## CHƯƠNG 3. ỨNG DỤNG DEMO

Nhóm sử dụng ngôn ngữ lập trình python, Flask API, Javascript, HTML, CSS để xây dựng một ứng dụng demo.

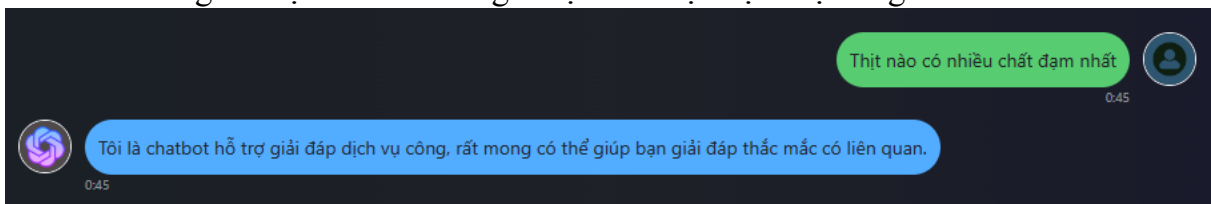
Link github: <https://github.com/thuynghuyen2003/CS336.O11-IR>

- Lời chào và giới thiệu của Chatbot



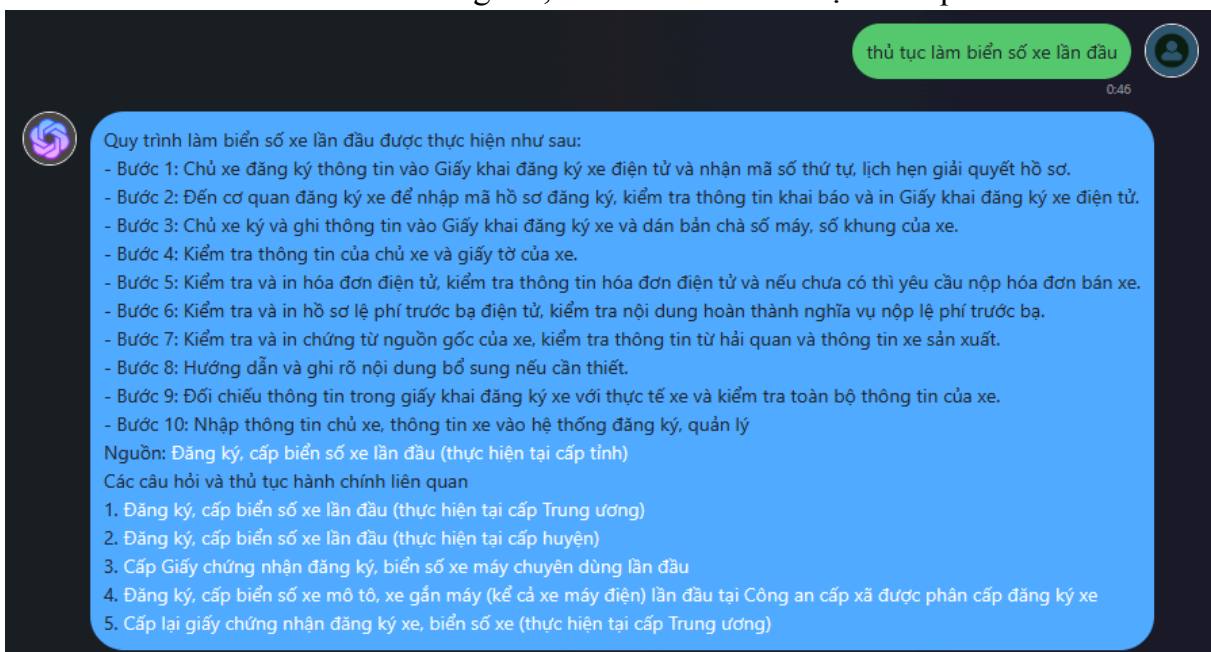
Hình 8: Minh họa ứng dụng: Mở đầu cuộc trò chuyện của hệ thống

- Không hỗ trợ câu hỏi không thuộc lĩnh vực dịch vụ công



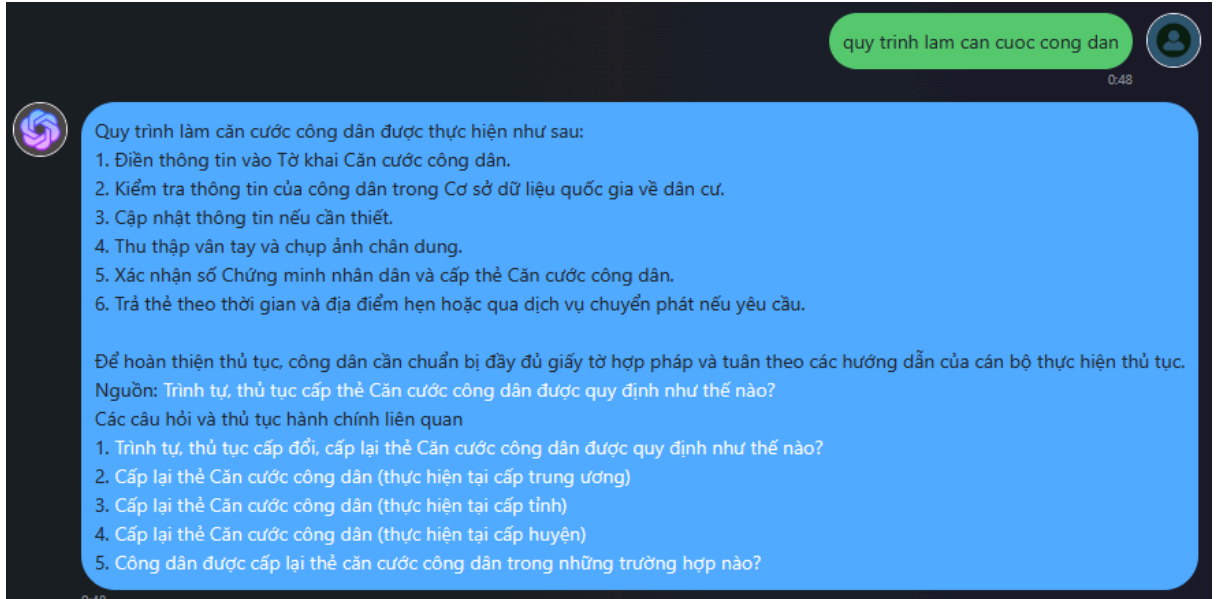
Hình 9: Minh họa ứng dụng: Không hỗ trợ câu hỏi không thuộc lĩnh vực dịch vụ công

- Câu trả lời có trích xuất nguồn, có đề xuất các thủ tục liên quan



Hình 10: Minh họa ứng dụng: Câu trả lời có trích xuất nguồn, có đề xuất các thủ tục liên quan

- Có thể hiểu câu hỏi không dấu



Quy trình làm căn cước công dân được thực hiện như sau:

1. Điền thông tin vào Tờ khai Căn cước công dân.
2. Kiểm tra thông tin của công dân trong Cơ sở dữ liệu quốc gia về dân cư.
3. Cập nhật thông tin nếu cần thiết.
4. Thu thập vân tay và chụp ảnh chân dung.
5. Xác nhận số Chứng minh nhân dân và cấp thẻ Căn cước công dân.
6. Trả thẻ theo thời gian và địa điểm hẹn hoặc qua dịch vụ chuyển phát nếu yêu cầu.

Để hoàn thiện thủ tục, công dân cần chuẩn bị đầy đủ giấy tờ hợp pháp và tuân theo các hướng dẫn của cán bộ thực hiện thủ tục.

Nguồn: Trình tự, thủ tục cấp thẻ Căn cước công dân được quy định như thế nào?

Các câu hỏi và thủ tục hành chính liên quan

1. Trình tự, thủ tục cấp đổi, cấp lại thẻ Căn cước công dân được quy định như thế nào?
2. Cấp lại thẻ Căn cước công dân (thực hiện tại cấp trung ương)
3. Cấp lại thẻ Căn cước công dân (thực hiện tại cấp tỉnh)
4. Cấp lại thẻ Căn cước công dân (thực hiện tại cấp huyện)
5. Công dân được cấp lại thẻ căn cước công dân trong những trường hợp nào?

*Hình 11: Minh họa ứng dụng: Câu trả lời có trích xuất nguồn, có đề xuất các thủ tục liên quan*

## CHƯƠNG 4. TỔNG KẾT

### 1. Kết quả đạt được

Xây dựng được một hệ thống hỏi đáp để hỗ trợ tìm kiếm thông tin dịch vụ công mang lại nhiều ưu điểm:

- Chatbot có thể hoạt động liên tục 24/7 mà không cần nghỉ ngơi, giúp người dùng có thể tìm kiếm thông tin và giải quyết câu hỏi bất kể thời gian nào
- Chatbot cung cấp câu trả lời đáng tin cậy, tạo điều kiện cho người sử dụng kiểm chứng thông tin thông qua nguồn và tìm hiểu về các vấn đề liên quan một cách triệt để thông qua việc gợi ý các thủ tục có liên quan
- Chatbot có khả năng trả lời nhanh chóng và tức thì các câu hỏi cơ bản, giúp người dùng tiết kiệm thời gian so với việc tìm hiểu thông tin trên các trang web hoặc gọi điện thoại.
- Chatbot có khả năng hiểu ngữ nghĩa tốt bao gồm cả khi câu không có dấu, có khả năng phân biệt một câu có trong lĩnh vực dịch vụ công hay không để đưa ra phương án xử lý, kết quả phù hợp.

### 2. Ưu điểm, hạn chế và hướng mở rộng

#### 2.1. Ưu điểm

- Đề tài này có thể áp dụng được trong thực tế
- Chatbot có nhiều chức năng để hỗ trợ người tăng trải nghiệm người dùng
- Ứng dụng được nhiều kỹ thuật mới trong lĩnh vực Xử lý ngôn ngữ tự nhiên

#### 2.2. Điểm hạn chế

- Hiệu quả của câu trả lời phụ thuộc chủ yếu vào thông tin ngữ cảnh nên cần xây dựng tập dữ liệu đủ lớn, chi tiết và xác thực để chatbot có thông tin để trả lời các câu hỏi của người dùng.
- Hạn chế về mặt dữ liệu nên bộ phân loại nhị phân câu hỏi thuộc lĩnh vực Dịch vụ công và khả năng truy vấn và trả lời câu hỏi chưa hoàn hảo.
- Mặc dù chatbot có thể cung cấp trợ giúp nhanh chóng, nhưng có thể gây ra sự phiền hà khi phản hồi không chính xác hoặc không thể giải quyết các vấn đề phức tạp.
- Ứng dụng chỉ mới được triển khai ở local.

#### 2.3. Hướng mở rộng

Hoàn thiện và triển khai ứng dụng trên internet để mọi người có thể sử dụng



Xây dựng thêm dữ liệu để tăng tính tổng quan, độ chính xác của bộ phân loại và truy vấn

Xây dựng các hệ thống cập nhật thủ tục hành chính tự động.

## TÀI LIỆU THAM KHẢO

- [1] M. R. Glass, G. Rossiello, Md. F. M. Chowdhury, A. R. Naik, P. Cai, and A. Gliozzo, “Re2G: Retrieve, Rerank, Generate,” ArXiv, vol. abs/2207.06300, 2022, [Online]. Available: <https://api.semanticscholar.org/CorpusID:250391085>
- [2] N. L. Tran, D. M. Le, and D. Q. Nguyen, “BARTpho: Pre-trained Sequence-to-Sequence Models for Vietnamese,” ArXiv, vol. abs/2109.09701, 2021.
- [3] G. Izacard and E. Grave, “Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering,” 2020. [Online]. Available: <https://api.semanticscholar.org/CorpusID:220302360>
- [4] P. Lewis et al., “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks,” ArXiv, vol. abs/2005.11401, 2020, [Online]. Available: <https://api.semanticscholar.org/CorpusID:218869575>
- [5] X. Liu et al., “GPT Understands, Too,” ArXiv, vol. abs/2103.10385, 2021, [Online]. Available: <https://api.semanticscholar.org/CorpusID:232269696>