

ĐẠI HỌC ĐÀ NẴNG

TRƯỜNG ĐẠI HỌC BÁCH KHOA

KHOA CÔNG NGHỆ THÔNG TIN

ĐỒ ÁN TỐT NGHIỆP

NGÀNH: CÔNG NGHỆ THÔNG TIN

CHUYÊN NGÀNH: CÔNG NGHỆ PHẦN MỀM

ĐỀ TÀI:

DANANG INVESTBOT – CHATBOT HỖ TRỢ XÚC
TIẾN ĐẦU TƯ ĐÀ NẴNG

Người hướng dẫn: TS. TRỊNH CÔNG DUY

Sinh viên thực hiện: LÊ THỊ HỒNG NHUNG

Số thẻ sinh viên: 102210178

Lớp: 21TCLC_DT2

Đà Nẵng, 06/2025

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA CÔNG NGHỆ THÔNG TIN

ĐỒ ÁN TỐT NGHIỆP

NGÀNH: CÔNG NGHỆ THÔNG TIN
CHUYÊN NGÀNH: CÔNG NGHỆ PHẦN MỀM

ĐỀ TÀI:

**DANANG INVESTBOT – CHATBOT HỖ TRỢ XÚC
TIẾN ĐẦU TƯ ĐÀ NẴNG**

Người hướng dẫn: **TS. TRỊNH CÔNG DUY**

Sinh viên thực hiện: **LÊ THỊ HỒNG NHUNG**

Số thẻ sinh viên: **102210178**

Lớp: **21TCLC_DT2**

Đà Nẵng, 06/2025

TÓM TẮT

Tên đề tài: DaNang InvestBot – Chatbot hỗ trợ xúc tiến đầu tư Đà Nẵng.

Sinh viên thực hiện: Lê Thị Hồng Nhung

Số thẻ SV: 102210178 Lớp: 21TCLC_DT2

DaNang InvestBot là một chatbot AI thông minh, được phát triển dựa trên công nghệ RAG (Retrieval-Augmented Generation), nhằm hỗ trợ hỏi đáp về các kiến thức trong lĩnh vực xúc tiến đầu tư tại thành phố Đà Nẵng. Công cụ này cung cấp thông tin toàn diện về các chính sách đầu tư, thủ tục, cơ hội hợp tác kinh doanh và các lĩnh vực thu hút đầu tư trọng điểm, giúp nhà đầu tư đưa ra quyết định hiệu quả và sáng suốt.

Đặc biệt, DaNang InvestBot được xây dựng với hệ thống quản trị tài liệu chặt chẽ, trong đó vai trò của quản trị viên là kiểm soát và duy trì chất lượng, bảo mật của các tài liệu đầu vào. Nhân viên có thể truy vấn những tài liệu nội bộ chưa công khai, nhằm phục vụ công việc chuyên môn một cách chính xác. Trong khi đó, người dùng cuối sẽ được hỗ trợ hỏi đáp dựa trên nguồn dữ liệu cơ bản công khai hoặc thông qua việc cung cấp file tài liệu, liên kết URL để chatbot có thể phân tích và trả lời câu hỏi liên quan trực tiếp tới nội dung đó.

Việc phân quyền rõ ràng này giúp đảm bảo an toàn thông tin, đồng thời tối ưu hóa trải nghiệm người dùng theo từng nhóm đối tượng, từ đó nâng cao hiệu quả hỗ trợ và thúc đẩy sự phát triển bền vững của hoạt động xúc tiến đầu tư tại Đà Nẵng.

NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP

Họ tên sinh viên: Lê Thị Hồng Nhung Số thẻ sinh viên: 102210178
Lớp: 21TCLC_DT2 Khoa: Công Nghệ Thông Tin Ngành: Công Nghệ Phần Mềm

1. Tên đề tài đồ án:

DaNang InvestBot – Chatbot hỗ trợ xúc tiến đầu tư Đà Nẵng.

2. Đề tài thuộc diện: Có ký kết thỏa thuận sở hữu trí tuệ đối với kết quả thực hiện

3. Các số liệu và dữ liệu ban đầu:

Dữ liệu ban đầu được thu thập trên website: <https://investdanang.gov.vn/web/guest>

4. Nội dung các phần thuyết minh và tính toán:

Nội dung thuyết minh bao gồm các phần chính:

- Mở đầu: Trình bày mục đích, mục tiêu, phạm vi, phương pháp và cấu trúc đồ án.
- Chương 1: Tổng quan ứng dụng chatbot, bài toán đặt ra và ý nghĩa đề tài.
- Chương 2: Cơ sở lý thuyết về RAG, RESTful API và xác thực JWT.
- Chương 3: Phân tích yêu cầu, kiến trúc hệ thống, use case và thiết kế cơ sở dữ liệu.
- Chương 4: Mô tả chức năng hệ thống kèm sơ đồ tuần tự.
- Chương 5: Đánh giá kết quả đạt được và hiệu quả hệ thống.
- Kết luận: Tóm tắt kết quả, nêu hạn chế và hướng phát triển.

5. Các bản vẽ, đồ thị (ghi rõ các loại và kích thước bản vẽ): Không có

6. Tên người hướng dẫn: TS. Trịnh Công Duy

7. Ngày giao nhiệm vụ đồ án: 16/03/2025.

8. Ngày hoàn thành đồ án: 18/06/2025.

Đà Nẵng, ngày tháng năm 2025

Trưởng Bộ môn

Người hướng dẫn

LỜI NÓI ĐẦU VÀ CẢM ƠN

Lời đầu tiên, em xin gửi lời cảm ơn chân thành đến thầy hướng dẫn Trịnh Công Duy đã tận tình hướng dẫn và hỗ trợ em trong suốt quá trình thực hiện đề tài này. Em cũng xin cảm ơn khoa đã dìu dắt, dạy dỗ và truyền đạt kiến thức, kinh nghiệm quý báu của mình trong suốt quá trình em học tập và nghiên cứu tại trường.

Đề tài “DaNang InvestBot – Chatbot hỗ trợ xúc tiến đầu tư Đà Nẵng” là một giải pháp công nghệ tiên tiến dựa trên công nghệ RAG AI (Retrieval-Augmented Generation) nhằm tự động hoá và nâng cao hiệu quả trong việc cung cấp thông tin và hỗ trợ tư vấn trong lĩnh vực xúc tiến đầu tư Đà Nẵng. Đồng thời DaNang InvestBot phân quyền rõ ràng giữa quản trị viên, nhân viên và người dùng, bảo mật thông tin và tối ưu hoá trải nghiệm cho người dùng.

Trong quá trình thực hiện đề tài, em đã áp dụng nhiều phương pháp nghiên cứu và kỹ thuật khác nhau để xây dựng và triển khai hệ thống. Báo cáo này sẽ trình bày chi tiết về quá trình nghiên cứu, thiết kế và phát triển hệ thống.

Trong quá trình thực hiện chắc hẳn sẽ không tránh khỏi những thiếu sót nhất định. Em rất mong nhận được sự cảm thông, góp ý và tận tình chỉ bảo của quý thầy cô để đề tài của em được hoàn thiện hơn.

Một lần nữa em xin chân thành cảm ơn!

Lê Thị Hồng Nhung

LỜI CAM ĐOAN LIÊM CHÍNH HỌC THUẬT

Em xin cam đoan:

1. Nội dung trong đề án tốt nghiệp này là do chính em thực hiện, dưới sự hướng dẫn trực tiếp của thầy Trịnh Công Duy.
2. Các tham khảo dùng trong đề án đều được trích dẫn rõ ràng tên tác giả, tên công trình, thời gian, địa điểm công bố.

Nếu có những sao chép không hợp lệ, vi phạm quy chế đào tạo, em xin chịu hoàn toàn trách nhiệm.

Sinh viên thực hiện

Lê Thị Hồng Nhung

MỤC LỤC

TÓM TẮT	
NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP.....	i
LỜI NÓI ĐẦU VÀ CẢM ƠN	ii
LỜI CAM ĐOAN LIÊM CHÍNH HỌC THUẬT	iii
MỤC LỤC	iv
DANH SÁCH CÁC BẢNG, HÌNH VẼ.....	vii
DANH SÁCH CÁC CỤM TỪ VIẾT TẮT	xi
MỞ ĐẦU	1
CHƯƠNG 1. TỔNG QUAN VỀ ĐỀ TÀI	5
1.1. Tổng quan về hệ thống chatbot trong thực tiễn	5
1.2. Bài toán đặt ra cho DaNang InvestBot.....	5
1.3. Ý nghĩa thực tiễn và khoa học của đề tài	6
1.4. Tổng kết chương.....	6
CHƯƠNG 2. CƠ SỞ LÝ THUYẾT	7
2.1. Công nghệ RAG.....	7
2.1.1. Vector Data và Vector Embeddings.....	7
2.1.2. Vector Search và Vector Database.....	7
2.1.3. Giới thiệu về Qdrant.....	8
2.1.4. Công nghệ RAG (Retrieval-Augmented Generation)	8
2.1.5. Ưu và nhược điểm của RAG so với chatbot thuần.....	11
2.2. RESTful API.....	11
2.3. Xác thực và phân quyền bằng JWT	12
2.4. Tổng kết chương.....	12
CHƯƠNG 3. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG.....	13
3.1. Phân tích yêu cầu	13

3.1.1. Mô tả bài toán	13
3.1.2. Hiện trạng.....	13
3.1.3. Yêu cầu chức năng.....	13
3.1.4. Yêu cầu phi chức năng	15
3.2. Khái quát hệ thống.....	16
3.2.1. Các tác nhân chính	16
3.2.2. Kiến trúc hệ thống và các thành phần chính	16
3.2.3. Phương pháp thực hiện	17
3.3. Sơ đồ usecase	18
3.3.1. Sơ đồ usecase tổng quan.....	18
3.3.2. Quản lý tài khoản của admin.....	19
3.3.3. Quản lý tài liệu của admin	20
3.3.4. Quản lý kho kiến thức của admin	21
3.3.5. Quản lý tài liệu người dùng	22
3.3.6. Chức năng tương tác với chatbot của admin	23
3.3.7. Chức năng quản lý tài khoản của user	24
3.3.8. Chức năng tương tác chatbot với user.....	25
3.4. Công nghệ sử dụng.....	25
3.4.1. Frontend	25
3.4.2. Backend (Server).....	26
3.4.3. Database.....	26
3.5. Thiết kế cơ sở dữ liệu.....	27
3.5.1. Mô hình ERD (Entity-Relationship Diagram).....	27
3.5.2. Bảng dữ liệu	28
3.6. Tổng kết chương.....	33
CHƯƠNG 4. MÔ TẢ CHỨC NĂNG.....	34
4.1. Chức năng của Admin.....	34

4.1.1. Quản lý tài khoản.....	34
4.1.2. Quản lý tài liệu	37
4.1.3. Quản lý kho kiến thức	42
4.1.4. Quản lý tài liệu người dùng	48
4.2. Chức năng tương tác với chatbot ở Admin và User	49
4.3. Tổng kết chương.....	54
CHƯƠNG 5. TRIỂN KHAI VÀ ĐÁNH GIÁ KẾT QUẢ.....	55
5.1. Kết quả	55
5.2. Đánh giá.....	63
5.3. Tổng kết chương.....	63
KẾT LUẬN	64
TÀI LIỆU THAM KHẢO	66
PHỤ LỤC	67

DANH SÁCH CÁC BẢNG, HÌNH VẼ

Hình 2.1: Sơ đồ quy trình sử dụng RAG với LLM	9
Hình 2.2: Sơ đồ quy trình của RAG	10
Hình 2.3: Sơ đồ thành phần RESTful API	11
Hình 2.4: Cấu trúc các phần của JWT	12
Hình 3.1: Sơ đồ usecase tổng quan.....	18
Hình 3.2: Sơ đồ usecase quản lý tài khoản của admin	19
Hình 3.3: Sơ đồ usecase quản lý tài liệu của admin	20
Hình 3.4: Sơ đồ usecase quản lý kho kiến thức của admin	21
Hình 3.5: Sơ đồ usecase quản lý tài liệu người dùng	22
Hình 3.6: Sơ đồ usecase tương tác với chatbot của admin.....	23
Hình 3.7: Sơ đồ usecase quản lý tài khoản của user	24
Hình 3.8: Sơ đồ usecase tương tác chatbot của user	25
Hình 3.9: Mô hình cơ sở dữ liệu.....	27
Hình 4.1: Sơ đồ chức năng xem danh sách tài khoản	34
Hình 4.2: Sơ đồ chức năng tạo mới tài khoản	35
Hình 4.3: Sơ đồ chức năng cập nhật trạng thái tài khoản.....	36
Hình 4.4: Sơ đồ chức năng xem lịch sử hội thoại mỗi tài khoản	37
Hình 4.5: Sơ đồ chức năng xem danh sách tài liệu	37
Hình 4.6: Sơ đồ chức năng xem chi tiết tài liệu	38
Hình 4.7: Sơ đồ chức năng thêm mới tài liệu.....	39
Hình 4.8: Sơ đồ chức năng cập nhật trạng thái tài liệu	40
Hình 4.9: Sơ đồ chức năng xóa tài liệu	41
Hình 4.10: Sơ đồ chức năng quản lý kho kiến thức	42
Hình 4.11: Sơ đồ chức năng xem chi tiết kho kiến thức	43
Hình 4.12: Sơ đồ chức năng thêm mới kho kiến thức.....	44
Hình 4.13: Sơ đồ chức năng cập nhật thông tin kho kiến thức	45

Hình 4.14: Sơ đồ chức năng thêm tài liệu vào kho kiến thức	46
Hình 4.15: Sơ đồ chức năng xóa tài liệu ra khỏi kho	47
Hình 4.16: Sơ đồ chức năng cập nhật trạng thái kho kiến thức	48
Hình 4.17: Sơ đồ chức năng đồng bộ tài liệu vào hệ thống	49
Hình 4.18: Sơ đồ chức năng xem danh sách hội thoại	50
Hình 4.19: Sơ đồ chức năng xem chi tiết cuộc hội thoại	51
Hình 4.20: Sơ đồ chức năng thêm mới cuộc hội thoại	52
Hình 4.21: Sơ đồ chức năng đặt câu hỏi.....	53
Hình 4.22: Sơ đồ chức năng xóa cuộc hội thoại.....	54
Hình 5.1: Giao diện chức năng xem danh sách tài khoản	55
Hình 5.2: Giao diện chức năng tạo mới tài khoản.....	56
Hình 5.3: Giao diện chức năng cập nhật trạng thái tài khoản	56
Hình 5.4: Giao diện chức năng xem danh sách tài liệu	57
Hình 5.5: Giao diện chức năng xem chi tiết tài liệu	57
Hình 5.6: Giao diện chức năng thêm mới tài liệu	58
Hình 5.7: Giao diện chức năng xem danh sách kho kiến thức	58
Hình 5.8: Giao diện chức năng tạo mới kho kiến thức.....	59
Hình 5.9: Giao diện chức năng thêm tài liệu vào kho	59
Hình 5.10: Giao diện chức năng chỉnh sửa kho kiến thức	60
Hình 5.11: Giao diện chức năng xem danh sách tài liệu người dùng.....	60
Hình 5.12: Giao diện chức năng đồng bộ tài liệu vào hệ thống	61
Hình 5.13: Giao diện chức năng xem lịch sử hội thoại ở người dùng	61
Hình 5.14: Giao diện chức năng tìm kiếm nội dung trong hội thoại ở người dùng	62
Hình 5.15: Giao diện chức năng xóa cuộc hội thoại ở người dùng.....	62
.....	
Bảng 3.1: Các tác nhân chính của hệ thống.....	16
Bảng 3.2: Cấu trúc của bảng Account.....	28
Bảng 3.3: Cấu trúc của bảng Role.....	29

Bảng 3.5: Cấu trúc của bảng MaterialType.....	29
Bảng 3.6: Cấu trúc của bảng AccessLevel.....	29
Bảng 3.4: Cấu trúc của bảng Material.....	30
Bảng 3.7: Cấu trúc của bảng AccessLevelRole.....	31
Bảng 3.8: Cấu trúc của bảng KnowledgeStore.....	31
Bảng 3.9: Cấu trúc của bảng Conversation.....	32
Bảng 3.10: Cấu trúc của bảng QuestionAnswer.....	32
Bảng 4.1: Mô tả chức năng xem danh sách tài khoản.....	34
Bảng 4.2: Mô tả chức năng tạo mới tài khoản.....	34
Bảng 4.3: Mô tả chức năng cập nhật trạng thái tài khoản.....	35
Bảng 4.4: Mô tả chức năng xem lịch sử hội thoại.....	36
Bảng 4.5: Mô tả chức năng xem danh sách tài liệu.....	37
Bảng 4.6: Mô tả chức năng xem chi tiết tài liệu.....	38
Bảng 4.7: Mô tả chức năng thêm mới tài liệu.....	39
Bảng 4.8: Mô tả chức năng cập nhật trạng thái tài liệu.....	40
Bảng 4.9: Mô tả chức năng xóa tài liệu.....	41
Bảng 4.10: Mô tả chức năng quản lý kho kiến thức.....	42
Bảng 4.11: Mô tả chức năng xem chi tiết kho kiến thức.....	43
Bảng 4.12: Mô tả chức năng thêm mới kho kiến thức.....	44
Bảng 4.13: Mô tả chức năng chỉnh sửa kho kiến thức.....	44
Bảng 4.14: Mô tả chức năng thêm tài liệu vào kho kiến thức.....	45
Bảng 4.15: Mô tả chức năng xóa tài liệu ra khỏi kho.....	46
Bảng 4.16: Mô tả chức năng cập nhật trạng thái kho kiến thức.....	47
Bảng 4.17: Mô tả chức năng đồng bộ tài liệu vào hệ thống.....	48
Bảng 4.18: Mô tả chức năng xem danh sách cuộc hội thoại.....	49
Bảng 4.19: Mô tả chức năng xem chi tiết cuộc hội thoại.....	50
Bảng 4.20: Mô tả chức năng thêm mới cuộc hội thoại.....	51

Bảng 4.21: Mô tả chức năng đặt câu hỏi	52
Bảng 4.22: Mô tả chức năng xóa cuộc hội thoại	53

DANH SÁCH CÁC CỤM TỪ VIẾT TẮT

Từ	Viết tắt	Diễn giải
AI	Artificial Intelligence	Trí tuệ nhân tạo
API	Application Programming Interface	Giao diện lập trình ứng dụng
HTTP	Hypertext Transfer Protocol	Giao thức truyền tải siêu văn bản
HTML	Hypertext Markup Language	Ngôn ngữ đánh dấu siêu văn bản
CSS	Cascading Style Sheets	Tập tin định kiểu theo tầng
RAG	Retrieval-Augmented Generation	Kết hợp truy xuất dữ liệu và sinh câu trả lời
ERD	Entity Relationship Diagram	Sơ đồ quan hệ thực thể
JWT	JSON Web Token	Mã thông báo web JSON

MỞ ĐẦU

1. Mục đích thực hiện đề tài

Đề tài hướng đến xây dựng một chatbot thông minh có khả năng hỗ trợ hỏi đáp tự động trong lĩnh vực xúc tiến đầu tư tại thành phố Đà Nẵng. Hệ thống nhằm cung cấp thông tin tập trung, chính xác và dễ tiếp cận, qua đó nâng cao hiệu quả trong việc phổ biến chính sách, quy trình, thủ tục và cơ hội đầu tư cho các nhà đầu tư trong và ngoài nước.

2. Mục tiêu đề tài

Để đạt được mục đích nêu trên, đề tài hướng đến thực hiện các mục tiêu cụ thể sau:

- Phân tích nhu cầu và xác định phạm vi thông tin cần cung cấp trong lĩnh vực xúc tiến đầu tư tại Đà Nẵng.
- Xây dựng cơ sở tri thức và tập hợp dữ liệu đầu vào phục vụ cho quá trình hỏi đáp tự động, đảm bảo tính chính xác và cập nhật.
- Thiết kế hệ thống quản lý nội dung, cập nhật và hiệu chỉnh câu trả lời một cách linh hoạt, thuận tiện.
- Áp dụng mô hình RAG nhằm nâng cao khả năng trả lời linh hoạt, chính xác dựa trên dữ liệu thực tế.
- Triển khai cơ chế phân quyền người dùng, cho phép kiểm soát truy cập và giới hạn nội dung hỏi đáp tùy theo đối tượng sử dụng.
- Phát triển chatbot hỗ trợ song ngữ tiếng Anh và tiếng Việt, phục vụ nhu cầu đầu tư trong nước và quốc tế.
- Đề xuất các hướng cải tiến và mở rộng hệ thống nhằm nâng cao khả năng ứng dụng thực tiễn trong tương lai.

3. Đối tượng và phạm vi nghiên cứu

➤ **Đối tượng nghiên cứu** : Đề tài tập trung nghiên cứu xây dựng hệ thống chatbot nhằm hỗ trợ cung cấp thông tin trong lĩnh vực xúc tiến đầu tư tại thành phố Đà Nẵng. Hệ thống được thiết kế để phục vụ 3 nhóm người dùng chính:

- + Quản trị viên (Admin): Có quyền kiểm soát hệ thống, quản lý và cập nhật dữ liệu.

- + Nhân viên (Employee): Là cán bộ tại cơ quan xúc tiến đầu tư, có vai trò hỗ trợ, tư vấn và xử lý thông tin.
- + Người dùng cuối (User): Bao gồm nhà đầu tư, doanh nghiệp, cá nhân trong và ngoài nước quan quan tâm đến môi trường đầu tư tại Đà Nẵng.

➤ **Phạm vi nghiên cứu:**

Phạm vi nghiên cứu của đề tài bao gồm các nội dung sau:

❖ Phạm vi chức năng:

- Hỗ trợ hỏi đáp tự động về các lĩnh vực liên quan đến xúc tiến Đà Nẵng như:
 - + Lĩnh vực thu hút đầu tư, cơ hội đầu tư.
 - + Thủ tục, chính sách đầu tư: Quy định chung, thủ tục cấp giấy chứng nhận, đăng ký doanh nghiệp, văn bản pháp luật...
 - + Thông tin đầu tư và các định hướng phát triển, tầm nhìn chiến lược, chi phí đầu tư, danh mục dự án kêu gọi đầu tư.
- Hỗ trợ song ngữ: Tiếng Việt và Tiếng Anh
- Quản trị, cập nhật dữ liệu và đảm bảo hệ thống hoạt động ổn định, chính xác.

❖ Phạm vi công nghệ và dữ liệu:

- Hệ thống ứng dụng mô hình RAG để truy và sinh phản hồi tự động từ dữ liệu thực tế.
- Hệ thống được phát triển dưới dạng ứng dụng Web.
- Cơ chế hỏi đáp dựa trên 3 nguồn dữ liệu chính:
 - + Tài liệu do người dùng tải lên hệ thống.
 - + Dữ liệu được nhập trực tiếp vào hệ thống.
 - + Dữ liệu được trích xuất từ các liên kết website.

❖ Giới hạn nghiên cứu:

Đề tài không bao gồm các nội dung sau:

- Không cung cấp tư vấn pháp lý chuyên sâu: hệ thống chỉ hỗ trợ thông tin ở mức tổng quan, nhà đầu tư vẫn cần tham khảo ý kiến chuyên gia.

- Không xử lý các vấn đề liên quan đến xúc tiến đầu tư ngoài phạm vi Đà Nẵng.
- Không xử lý các giao dịch tài chính.
- Không chịu trách nhiệm pháp lý đối với tính chính xác của các tài liệu do người dùng tải lên.

4. Phương pháp nghiên cứu

Nhằm hiện thực hóa mục tiêu đề tài và xây dựng một hệ thống chatbot hoàn chỉnh, đề tài áp dụng các phương pháp nghiên cứu cụ thể sau:

- Thống kê và lập ra một danh sách các yêu cầu chức năng và phi chức năng cần thiết.
- Hợp tác với giáo viên hướng dẫn để đánh giá và nhận xét danh sách yêu cầu, từ đó rút ra được danh sách yêu cầu cuối cùng.
- Xem xét các lựa chọn công nghệ và chọn ra công nghệ phù hợp với những yêu cầu đã xác định.
- Tìm hiểu các công nghệ liên quan như các mô hình ngôn ngữ lớn, công nghệ RAG, các framework lập trình web và cơ sở dữ liệu
- Tiến hành lập trình, tích hợp các thành phần của hệ thống dựa trên kiến trúc đã lựa chọn. Triển khai chatbot trên nền tảng web, hỗ trợ phân quyền người dùng và xử lý dữ liệu đa nguồn.
- Thực hiện kiểm thử hệ thống theo các tiêu chí như tính chính xác, tốc độ phản hồi, khả năng mở rộng, và mức độ thân thiện với người dùng. Dựa vào kết quả kiểm thử để điều chỉnh và tối ưu.

5. Cấu trúc của đồ án tốt nghiệp

Đồ án bao gồm các nội dung chính sau:

Mở đầu: Trình bày mục đích, mục tiêu, phạm vi, phương pháp và cấu trúc đồ án.

Chương 1: Tổng quan ứng dụng chatbot, bài toán đặt ra và ý nghĩa đề tài.

Chương 2: Cơ sở lý thuyết về RAG, RESTful API và xác thực JWT.

Chương 3: Phân tích yêu cầu, kiến trúc hệ thống, use case và thiết kế cơ sở dữ liệu.

Chương 4: Mô tả chức năng hệ thống kèm sơ đồ tuần tự.

Chương 5: Đánh giá kết quả đạt được và hiệu quả hệ thống.

Kết luận: Tóm tắt kết quả, nêu hạn chế và hướng phát triển.

CHƯƠNG 1. TỔNG QUAN VỀ ĐỀ TÀI

1.1. Tổng quan về hệ thống chatbot trong thực tiễn

Chatbot là một hệ thống phần mềm ứng dụng trí tuệ nhân tạo, có khả năng mô phỏng cuộc đối thoại với con người thông qua văn bản hoặc giọng nói. Công nghệ này đã được triển khai rộng rãi trong các lĩnh vực như ngân hàng, thương mại điện tử, y tế, giáo dục và dịch vụ công.

Tại Đà Nẵng, chatbot chủ yếu được ứng dụng trong lĩnh vực du lịch và hành chính công, điển hình như Danang Fantasticity hỗ trợ tra cứu thông tin du lịch, hay trợ lý ảo tiếp nhận phản ánh hiện trường. Tuy nhiên, trong lĩnh vực xúc tiến đầu tư – nơi đòi hỏi cung cấp thông tin chuyên sâu, cập nhật và chính xác – vẫn chưa có hệ thống chatbot chuyên biệt.

Đề tài “DaNang InvestBot” được xây dựng nhằm lấp đầy khoảng trống này, góp phần nâng cao hiệu quả phục vụ nhà đầu tư và thúc đẩy chuyển đổi số trong công tác xúc tiến đầu tư tại địa phương.

1.2. Bài toán đặt ra cho DaNang InvestBot

Xuất phát từ nhu cầu thực tiễn, đề tài hướng đến xây dựng DaNang InvestBot – một chatbot thông minh hỗ trợ truy vấn thông tin về xúc tiến đầu tư tại Đà Nẵng. Hệ thống cần đáp ứng các yêu cầu sau:

- Truy xuất thông tin chính xác từ nhiều nguồn: tài liệu tải lên, dữ liệu nhập trực tiếp và liên kết web.
- Hỗ trợ song ngữ (tiếng Việt và tiếng Anh) cho cả nhà đầu tư trong và ngoài nước.
- Phân quyền người dùng: quản trị viên quản lý dữ liệu, nhân viên hỗ trợ nội dung, người dùng truy vấn thông tin.
- Tích hợp mô hình RAG để cải thiện độ chính xác và tính phù hợp ngữ cảnh của câu trả lời.
- Giao diện thân thiện, dễ sử dụng và được triển khai trên nền tảng Website.

1.3. Ý nghĩa thực tiễn và khoa học của đề tài

❖ Ý nghĩa thực tiễn

- Hỗ trợ nhà đầu tư bằng cách cung cấp thông tin nhanh chóng và chính xác về chính sách, quy trình cũng như cơ hội đầu tư, giúp tiết kiệm thời gian tra cứu.
- Nâng cao trải nghiệm người dùng thông qua hỗ trợ song ngữ (Anh - Việt), giúp nhà đầu tư nước ngoài dễ dàng tiếp cận thông tin.
- Tối ưu hóa hoạt động xúc tiến đầu tư bằng việc tự động trả lời các câu hỏi thường gặp, giảm áp lực và tiết kiệm nhân lực cho bộ phận hỗ trợ.

❖ Ý nghĩa khoa học:

- Ứng dụng trí tuệ nhân tạo với mô hình ngôn ngữ lớn (LLM), giúp chatbot hiểu được câu hỏi và phản hồi một cách tự nhiên, nâng cao độ chính xác của câu trả lời.
- Triển khai mô hình RAG (Retrieval-Augmented Generation) kết hợp giữa tìm kiếm và tạo sinh văn bản thông minh, giúp chatbot cập nhật thông tin linh hoạt mà không cần đào tạo lại mô hình.

1.4. Tổng kết chương

Chương này trình bày tổng quan về ứng dụng chatbot trong thực tiễn, bài toán đặt ra cho DaNang InvestBot, cũng như ý nghĩa thực tiễn và khoa học của đề tài. Đây là cơ sở định hướng cho các nội dung phân tích và triển khai hệ thống ở các chương sau.

CHƯƠNG 2. CƠ SỞ LÝ THUYẾT

2.1. Công nghệ RAG

2.1.1. Vector Data và Vector Embeddings

- **Vector Data:** Là cách biểu diễn dữ liệu dưới dạng các dãy số trong không gian nhiều chiều, nhằm giúp máy tính xử lý, phân tích và thực hiện các phép toán học dễ dàng hơn.
- **Vector Embeddings:** Là cách chuyển đổi dữ liệu như văn bản hoặc hình ảnh thành các vector số, giữ lại ý nghĩa và ngữ cảnh. Các vector gần nhau thể hiện các dữ liệu có ý nghĩa tương tự.
- Một số phương pháp embedding phổ biến bao gồm: Word2Vec, GloVe, BERT, và Sentence Transformers.
- Việc sử dụng vector embeddings cho phép thực hiện tìm kiếm tương tự (similarity search) bằng cách so sánh khoảng cách giữa các vector, từ đó nâng cao hiệu quả truy xuất thông tin liên quan trong các hệ thống hỏi đáp tự động.

2.1.2. Vector Search và Vector Database

- **Vector Search:** Là phương pháp tìm kiếm dựa trên việc so sánh các vector trong không gian nhiều chiều nhằm tìm ra các vector gần nhất (*nearest neighbors*) so với vector truy vấn. Phương pháp này được sử dụng phổ biến trong các hệ thống tìm kiếm tương tự (*similarity search*) trên dữ liệu đã được chuyển đổi thành vector embeddings, chẳng hạn như văn bản, hình ảnh hoặc âm thanh.
- **Vector Database:** Giúp lưu trữ hàng triệu đến hàng tỷ vector embeddings và truy vấn tìm kiếm tương tự với tốc độ cao và độ chính xác tốt. Một số hệ thống cơ sở dữ liệu vector phổ biến hiện nay bao gồm:
 - + Qdrant: Hỗ trợ tìm kiếm vector nhanh, dễ tích hợp với các ứng dụng AI.
 - + PineCone: Dịch vụ vector database cloud với tính năng quản lý dữ liệu và tìm kiếm nâng cao.
 - + FAISS: Thư viện mã nguồn mở giúp tìm kiếm vector hiệu quả trên máy tính cá nhân và server.

2.1.3. Giới thiệu về Qdrant

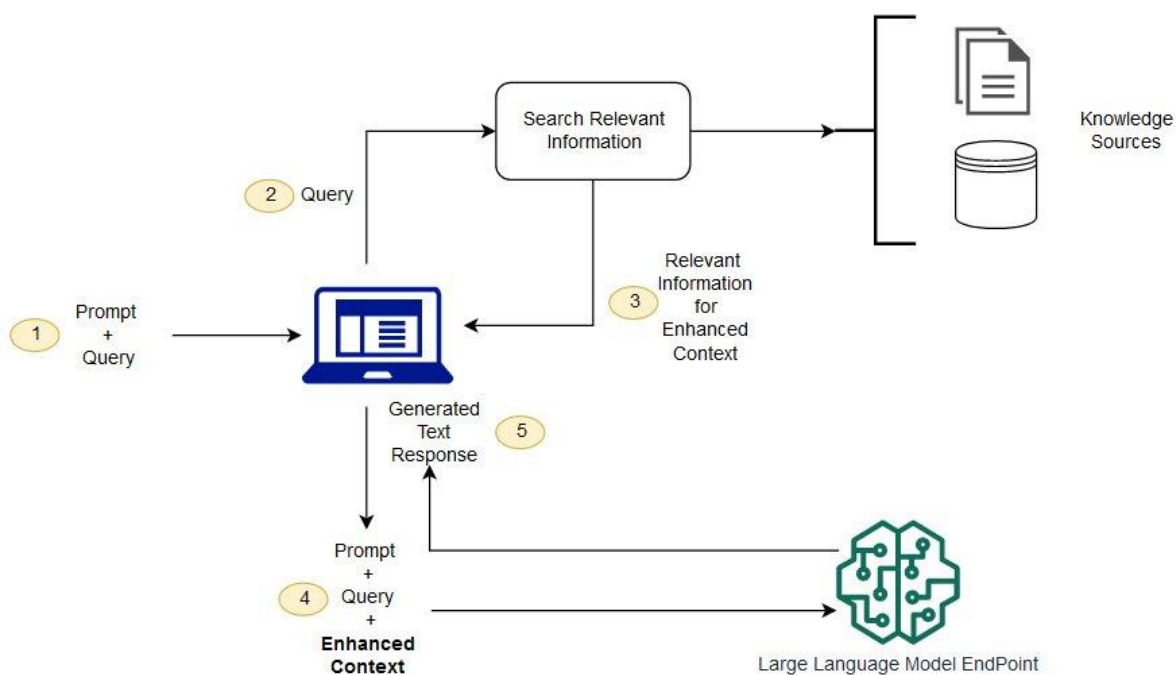
- Qdrant là một hệ quản trị cơ sở dữ liệu vector mã nguồn mở được tối ưu hóa cho việc truy vấn và lưu trữ các vector embedding. Qdrant hỗ trợ các bài toán tìm kiếm ngữ nghĩa tốc độ cao và chính xác, đồng thời dễ tích hợp với các mô hình học sâu và ứng dụng AI.
- Một số đặc điểm nổi bật của Qdrant:
 - + Hỗ trợ tìm kiếm gần đúng với độ chính xác cao.
 - + Tích hợp payload (metadata) cho mỗi vector để hỗ trợ lọc nâng cao.
 - + Giao diện API dễ sử dụng, hỗ trợ RESTful API.
 - + Tích hợp dễ dàng với các mô hình AI như OpenAI, HuggingFace...

2.1.4. Công nghệ RAG (Retrieval-Augmented Generation)

❖ Khái niệm RAG:

- Retrieval-Augmented Generation (RAG) là phương pháp kết hợp giữa khả năng tạo sinh của mô hình ngôn ngữ lớn (LLM) và khả năng truy xuất thông tin từ nguồn dữ liệu bên ngoài. Trước khi tạo ra câu trả lời, hệ thống sẽ truy vấn một cơ sở tri thức (knowledge base) để lấy thông tin có liên quan, từ đó giúp mô hình sinh phản hồi chính xác và phù hợp với ngữ cảnh hơn.
- Phương pháp này cho phép tận dụng các mô hình LLM mạnh mẽ mà không cần huấn luyện lại, giúp mở rộng khả năng ứng dụng vào các lĩnh vực chuyên biệt hoặc dữ liệu nội bộ của tổ chức. RAG được xem là một giải pháp hiệu quả về chi phí và độ chính xác, đặc biệt phù hợp cho các hệ thống hỏi đáp tự động như chatbot.

(Tham khảo: AWS – What is Retrieval-Augmented Generation)



Hình 2.1: Sơ đồ quy trình sử dụng RAG với LLM

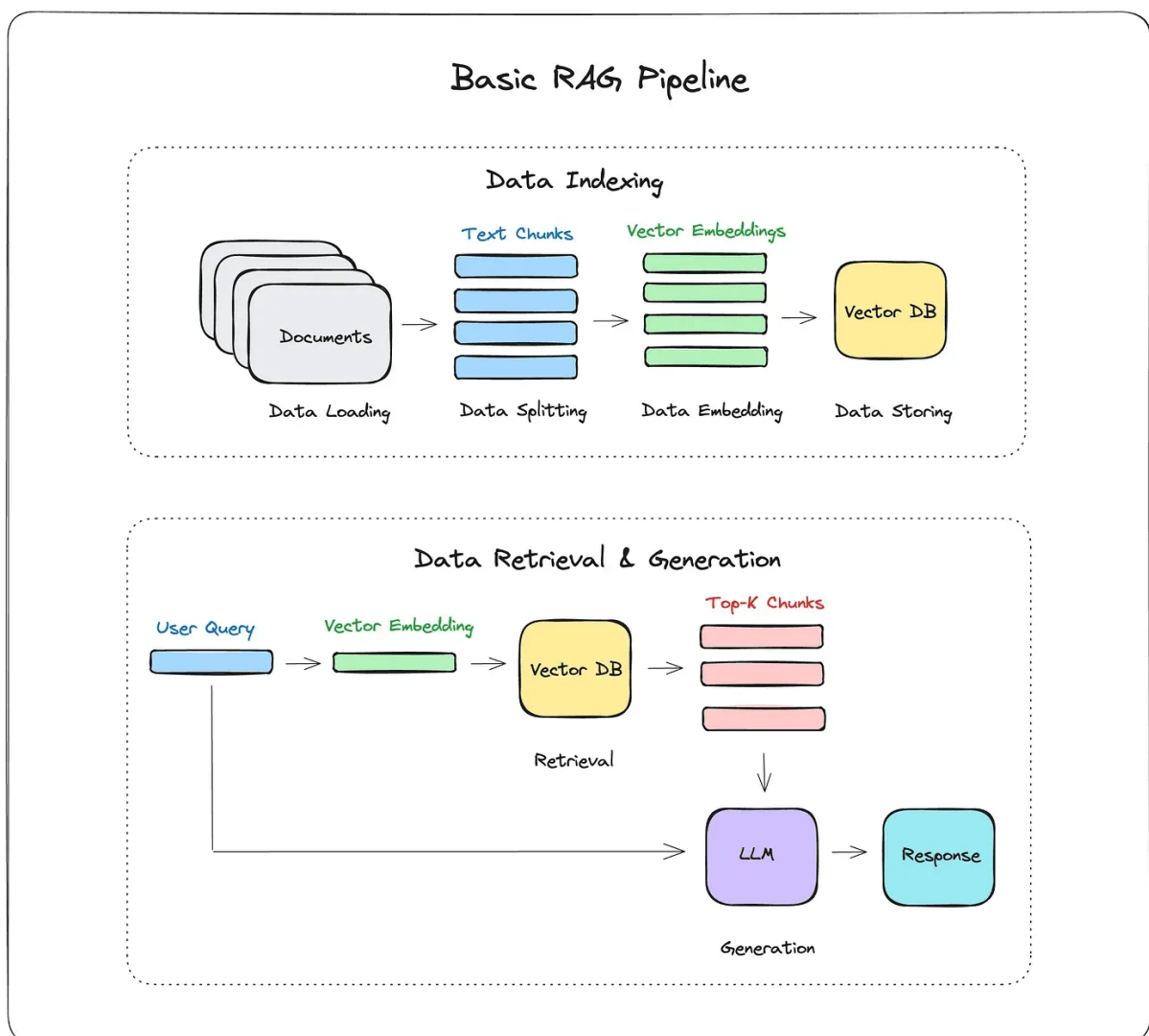
❖ **Kiến trúc cơ bản của RAG chatbot:**

- Module truy xuất thông tin (Retriever):
 - + Nhiệm vụ: Tìm kiếm và trích xuất các đoạn văn bản hoặc tài liệu có liên quan đến truy vấn người dùng.
 - + Công nghệ thường dùng: Vector database (ví dụ: Qdrant, Pinecone, FAISS) để thực hiện tìm kiếm ngữ nghĩa (semantic search).
 - + Cách hoạt động: Truy vấn của người dùng được chuyển đổi thành vector embedding và được dùng để truy vấn trong cơ sở dữ liệu để tìm các đoạn nội dung liên quan nhất.
- Module tạo sinh (Generator):
 - + Nhận truy vấn người dùng cùng với thông tin đã được truy xuất, sau đó tổng hợp và sinh ra câu trả lời đầy đủ, có ngữ cảnh.
 - + Công nghệ thường dùng: Các mô hình ngôn ngữ lớn (LLM) như GPT, DeepSeek, Gemini,...

❖ **Quy trình chính của RAG**

Quy trình RAG gồm 2 giai đoạn chính:

- Retrieval: Giai đoạn này chịu trách nhiệm tìm kiếm thông tin liên quan trong cơ sở dữ liệu (thường là vector database như Qdrant...), từ đó cung cấp ngữ cảnh cho mô hình ngôn ngữ.
- Generation: Sau khi đã có ngữ cảnh, mô hình LLM (Large Language Model) sẽ dùng các đoạn thông tin đã truy xuất để sinh ra câu trả lời hoàn chỉnh và phù hợp với truy vấn của người dùng.



Hình 2.2: Sơ đồ quy trình của RAG

2.1.5. Ưu và nhược điểm của RAG so với chatbot thuần

❖ Ưu điểm:

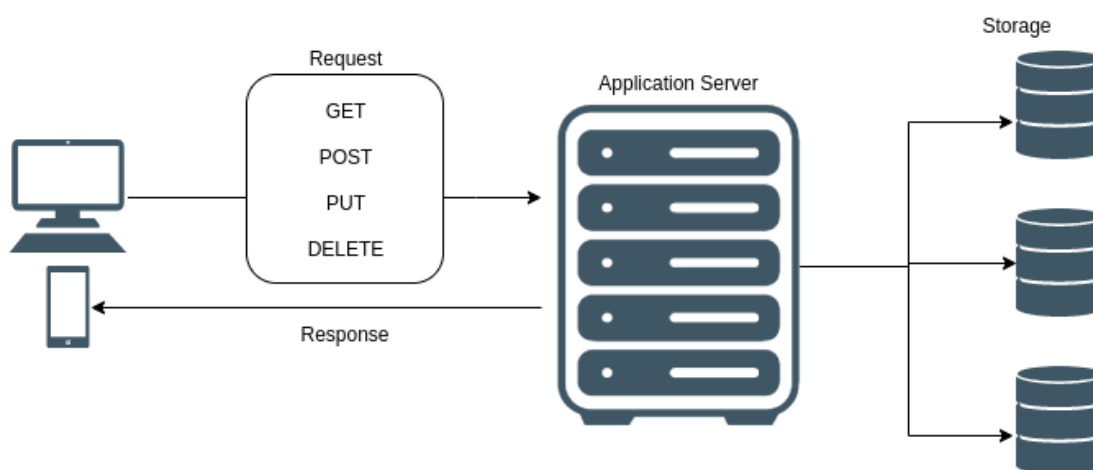
- Truy xuất thông tin chính xác, dựa trên dữ liệu thực tế, giúp tạo câu trả lời cập nhật hơn, tránh “bịa” thông tin.
- Giảm tải cho mô hình tạo sinh, không cần nhớ toàn bộ kiến thức mà dựa vào dữ liệu bên ngoài.
- Dễ dàng cập nhật kiến thức mới bằng cách thêm tài liệu mới mà không cần huấn luyện lại mô hình.

❖ Hạn chế:

- Độ phức tạp hệ thống cao hơn, cần quản lý kho dữ liệu và module truy xuất hiệu quả.
- Tốc độ phản hồi có thể chậm hơn do phải truy xuất trước khi tạo sinh.
- Chất lượng câu trả lời phụ thuộc vào chất lượng dữ liệu truy xuất.

2.2. RESTful API

RESTful API là một giao diện lập trình ứng dụng sử dụng giao thức HTTP, cho phép client và server giao tiếp với nhau bằng cách truy xuất và thao tác các tài nguyên (resource) thông qua các phương thức như GET, POST, PUT, PATCH, DELETE.

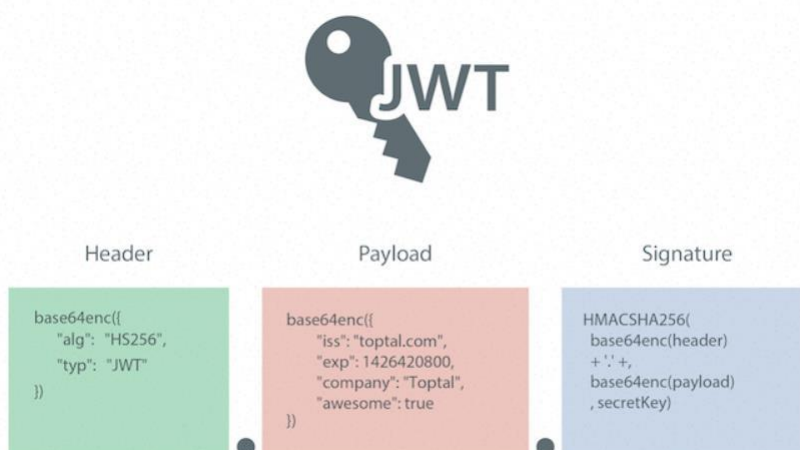


Hình 2.3: Sơ đồ thành phần RESTful API

2.3. Xác thực và phân quyền bằng JWT

JWT là cơ chế xác thực và phân quyền phổ biến, cho phép truyền thông tin người dùng giữa client và server dưới dạng token mã hóa, giúp bảo mật và không cần lưu session trên server.

Các thông tin trong chuỗi JWT được định dạng bằng JSON. Trong đó chuỗi Token phải có 3 phần là header, payload và phần signature được ngăn bằng dấu “.”



Hình 2.4: Cấu trúc các phần của JWT

2.4. Tổng kết chương

Chương này trình bày các kiến thức nền tảng về công nghệ RAG, vector database, RESTful API và cơ chế xác thực bằng JWT làm cơ sở cho việc xây dựng hệ thống RAG Chatbot.

CHƯƠNG 3. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

3.1. Phân tích yêu cầu

3.1.1. Mô tả bài toán

Hệ thống chatbot hỗ trợ xúc tiến đầu tư Đà Nẵng được xây dựng nhằm cung cấp thông tin nhanh chóng, chính xác về chính sách, dự án kêu gọi đầu tư, thủ tục pháp lý và cơ hội xúc tiến đầu tư.

Chatbot sử dụng mô hình RAG để truy xuất thông tin từ tài liệu đã lưu, sau đó kết hợp với mô hình ngôn ngữ lớn để tạo ra câu trả lời tự nhiên, đúng ngữ cảnh. Giải pháp này giúp tăng hiệu quả tư vấn, hỗ trợ nhà đầu tư và thúc đẩy chuyển đổi số trong hoạt động xúc tiến đầu tư của thành phố.

3.1.2. Hiện trạng

Hiện nay, thông tin liên quan đến xúc tiến đầu tư tại Đà Nẵng còn phân tán trên nhiều cổng thông tin, trang web và văn bản hành chính, gây khó khăn cho nhà đầu tư trong quá trình tra cứu. Một số vấn đề chính bao gồm:

- Dữ liệu không tập trung: Thông tin về dự án, chính sách, thủ tục hành chính được công bố rời rạc, không thống nhất.
- Khó truy cập và tìm kiếm: Nhà đầu tư mất nhiều thời gian để tìm đúng tài liệu hoặc quy trình cần thiết.
- Thiếu hỗ trợ tức thời: Việc đặt câu hỏi qua email hoặc hotline thường chậm trễ và không có phản hồi ngay lập tức.

3.1.3. Yêu cầu chức năng

❖ Đối với quản trị viên

- Quản lý tài khoản:
 - + Xem danh sách tài khoản có trên hệ thống.
 - + Thêm tài khoản cho nhân viên.
 - + Cập nhật trạng thái hoạt động của tài khoản.
 - + Xem lịch sử hội thoại của mỗi tài khoản.

- Quản lý tài liệu:
 - + Xem danh sách tài liệu hiện có trong hệ thống.
 - + Xem chi tiết từng tài liệu.
 - + Thêm mới tài liệu
 - + Cập nhật trạng thái hoạt động của tài liệu.
 - + Xóa tài liệu.
- Quản lý kho kiến thức:
 - + Xem danh sách và chi tiết kho kiến thức.
 - + Thêm mới kho kiến thức.
 - + Thêm hoặc xoá một hoặc nhiều tài liệu ra khỏi kho.
 - + Cập nhật thông tin kho kiến thức
 - + Cập nhật trạng thái hoạt động của các tài liệu trong kho
- Quản lý tài liệu người dùng:
 - + Xem danh sách và chi tiết các tài liệu.
 - + Đồng bộ tài liệu do người dùng cung cấp vào hệ thống.
 - + Xóa tài liệu ra khỏi hệ thống.
- Chức năng tương tác với chatbot:
 - + Xem lịch sử hội thoại.
 - + Tìm kiếm nội dung trong các đoạn hội thoại trước đó.
 - + Gửi câu hỏi, nhận câu trả lời từ chatbot (Quản trị viên có thể truy vấn tất cả tài liệu với các mức quyền: public, private, internal).
 - + Tạo mới đoạn hội thoại.
 - + Xóa đoạn hội thoại

❖ **Đối với người dùng (nhân viên và người dùng cuối)**

- Quản lý tài khoản:
 - + Đăng ký tài khoản.
 - + Đăng nhập / Đăng xuất.
 - + Xem thông tin tài khoản.
 - + Quên mật khẩu.
- Chức năng tương tác với chatbot:
 - + Xem lịch sử hội thoại.
 - + Tìm kiếm nội dung trong các đoạn hội thoại trước đó.
 - + Gửi câu hỏi, nhận câu trả lời từ chatbot (Người dùng cuối chỉ được truy vấn tài liệu với mức quyền public. Nhân viên được truy vấn tài liệu với mức quyền public và internal).
 - + Tạo mới đoạn hội thoại.
 - + Xóa đoạn hội thoại.

3.1.4. Yêu cầu phi chức năng

- Khả năng mở rộng: Hệ thống phải có khả năng mở rộng để xử lý số lượng lớn người dùng truy cập đồng thời.
- Hiệu suất: Hệ thống phải đảm bảo xử lý câu hỏi nhanh chóng và phản hồi chính xác.
- Hệ thống phải lưu trữ thông tin an toàn, tránh mất mát, sai lệch thông tin.
- Chống lại các hình thức tấn công phổ biến như SQL Injection, XSS,...
- Bảo mật: Hệ thống phải bảo mật thông tin và dữ liệu tương ứng với mỗi vai trò.
- Khả năng mở rộng: Hệ thống phải có khả năng mở rộng để xử lý số lượng lớn người dùng và dữ liệu.

3.2. Khái quát hệ thống

3.2.1. Các tác nhân chính

Bảng 3.1: Các tác nhân chính của hệ thống

Tác nhân	Mô tả
Người dùng (Nhân viên và người dùng cuối)	Thực hiện chức năng tương tác với chatbot.
Quản trị viên (Admin)	<p>Thực hiện chức năng quản lý tài khoản.</p> <p>Thực hiện chức năng quản lý tài liệu.</p> <p>Thực hiện chức năng quản lý kho kiến thức.</p> <p>Thực hiện chức năng quản lý tài liệu người dùng.</p> <p>Thực hiện chức năng tương tác với chatbot.</p>

3.2.2. Kiến trúc hệ thống và các thành phần chính

- Kiến trúc hệ thống: Kiến trúc hệ thống RAG Chatbot dựa trên sự kết hợp giữa mô hình ngôn ngữ và khả năng tìm kiếm thông tin từ cơ sở tri thức. Khi người dùng đặt câu hỏi, hệ thống sẽ truy xuất các dữ liệu liên quan và kết hợp với mô hình sinh để tạo ra câu trả lời chính xác và tự nhiên.
- Các thành phần chính:
 - + Module truy xuất thông tin (Retrieval): Tìm kiếm các đoạn văn bản liên quan từ kho dữ liệu (vector database) dựa trên câu hỏi của người dùng.
 - + Mô hình sinh câu trả lời (Generator): Sử dụng mô hình ngôn ngữ Gemini để tạo câu trả lời dựa trên thông tin đã truy xuất được. Mô hình này kết hợp với dữ liệu truy xuất, ngữ cảnh hội thoại cùng với câu hỏi để trả lời chính xác.
 - + Kho dữ liệu vector (Vector database): Lưu trữ các đoạn văn bản đã được mã hoá thành vector. Đây là nguồn dữ liệu chính mà hệ thống sử dụng để tìm kiếm thông tin phục vụ việc tạo câu trả lời.

3.2.3. Phương pháp thực hiện

❖ Bước 1: Chuẩn bị dữ liệu đầu vào

- Tài liệu người dùng cung cấp được chia nhỏ bằng CharacterTextSplitter (~1000 ký tự, overlap 200).
- Các đoạn nhỏ được biến đổi thành vector bằng mô hình all-MiniLM-L6-v2.

❖ Bước 2: Truy xuất thông tin

- Câu hỏi được chuyển thành vector embedding bằng mô hình all-MiniLM-L6-v2.
- Nếu dùng dữ liệu hệ thống đã lưu: Hệ thống truy vấn trong Qdrant (384 chiều, đo cosine) để tìm các đoạn liên quan.
- Nếu dùng tài liệu người dùng cung cấp: Hệ thống đọc trực tiếp tài liệu và đưa cùng câu hỏi vào Gemini để sinh câu trả lời, không truy vấn Qdrant. Ngược lại nếu không có tài liệu được cung cấp, hệ thống sẽ dựa vào ngữ cảnh câu hỏi để truy vấn dữ liệu đã lưu trong Qdrant.
- Kết quả là các đoạn thông tin phù hợp sẽ được dùng làm đầu vào cho mô hình sinh.

❖ Bước 3: Sinh câu trả lời

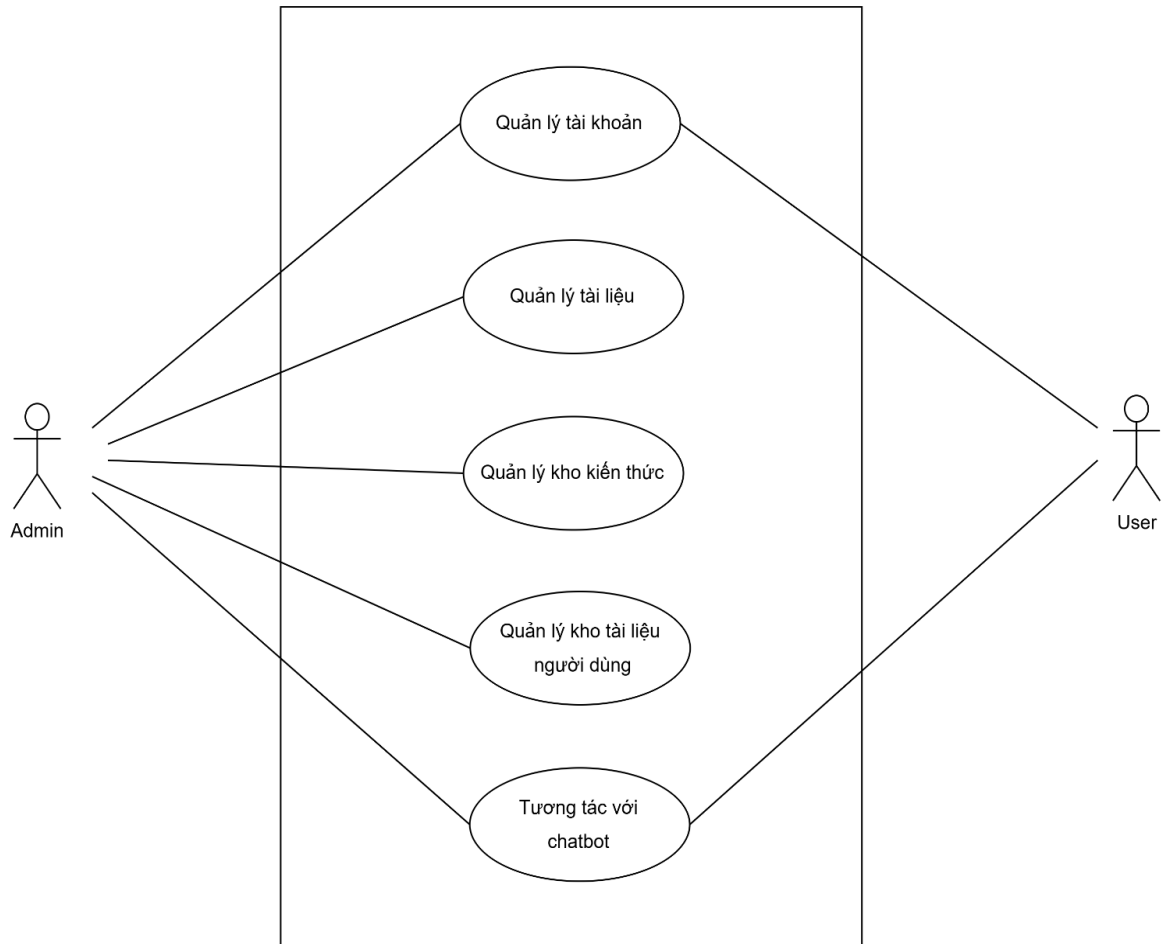
- Các đoạn được truy xuất kết hợp với lịch sử hội thoại và câu hỏi tạo thành prompt.
- Prompt được gửi vào mô hình Gemini để sinh câu trả lời tự nhiên và chính xác.

❖ Bước 4: Phản hồi

- Câu trả lời được hiển thị cho người dùng.
- Truy vấn có thể được lưu lại lịch sử cho mỗi cuộc hội thoại.

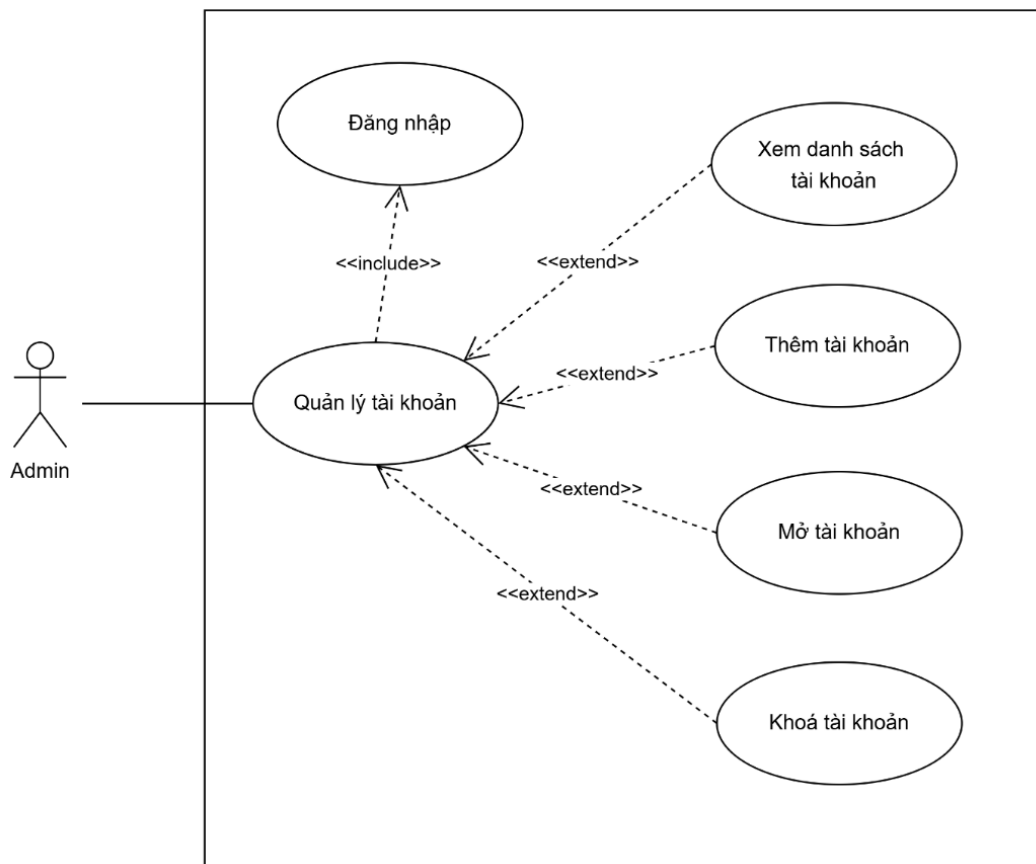
3.3. Sơ đồ usecase

3.3.1. Sơ đồ usecase tổng quan



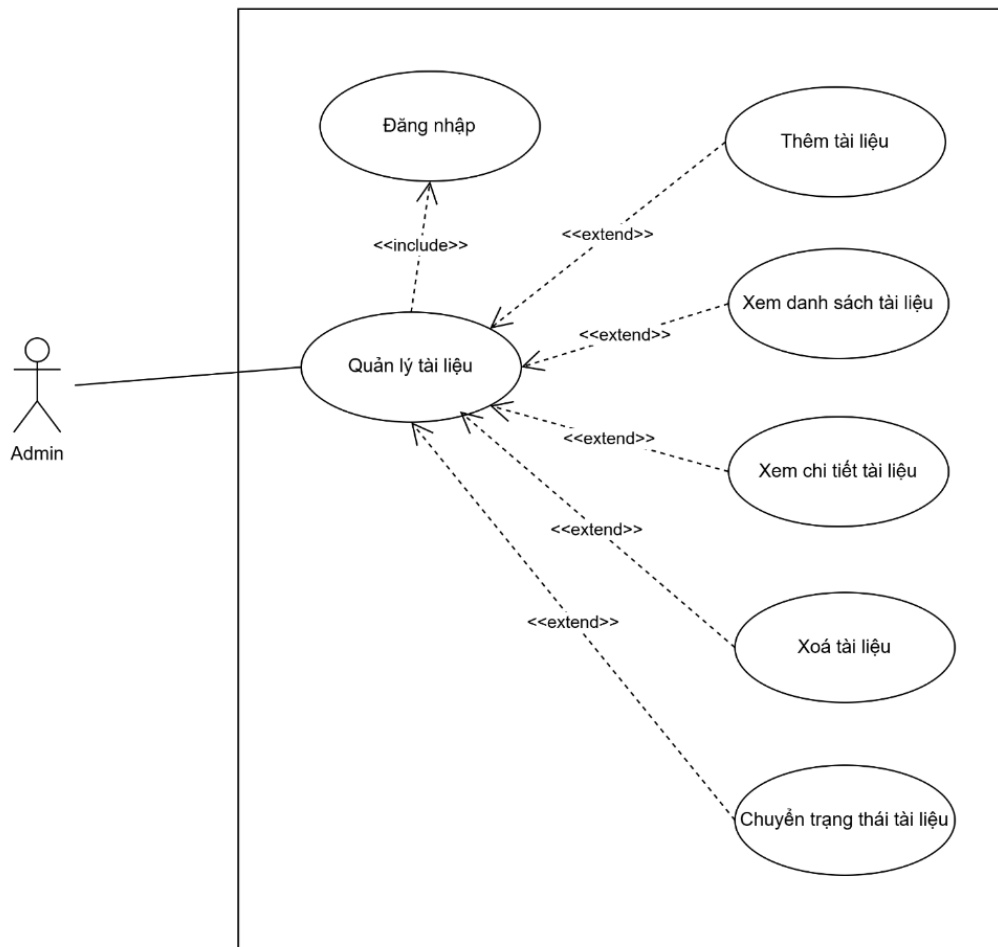
Hình 3.1: Sơ đồ usecase tổng quan

3.3.2. Quản lý tài khoản của admin



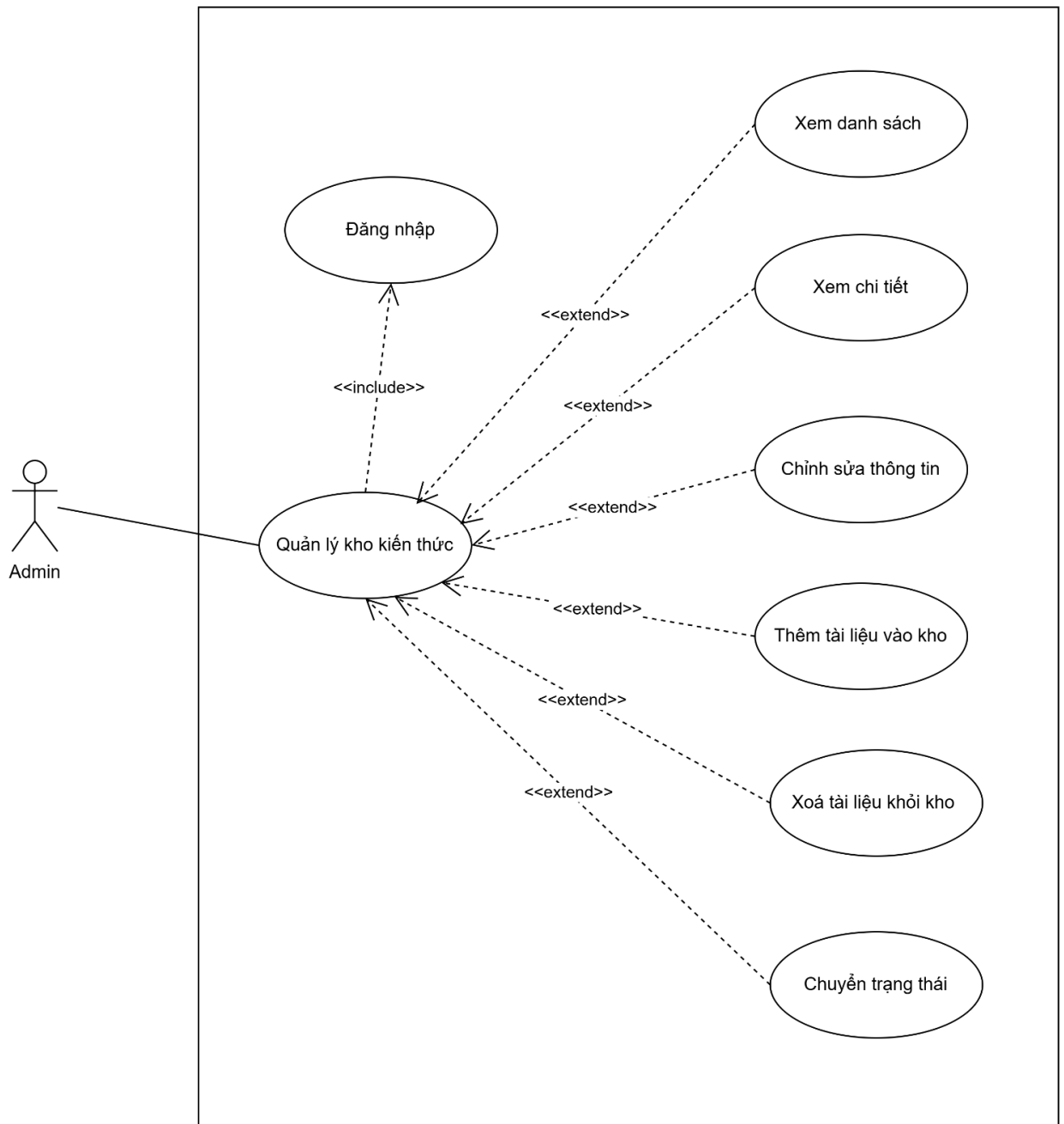
Hình 3.2: Sơ đồ usecase quản lý tài khoản của admin

3.3.3. Quản lý tài liệu của admin



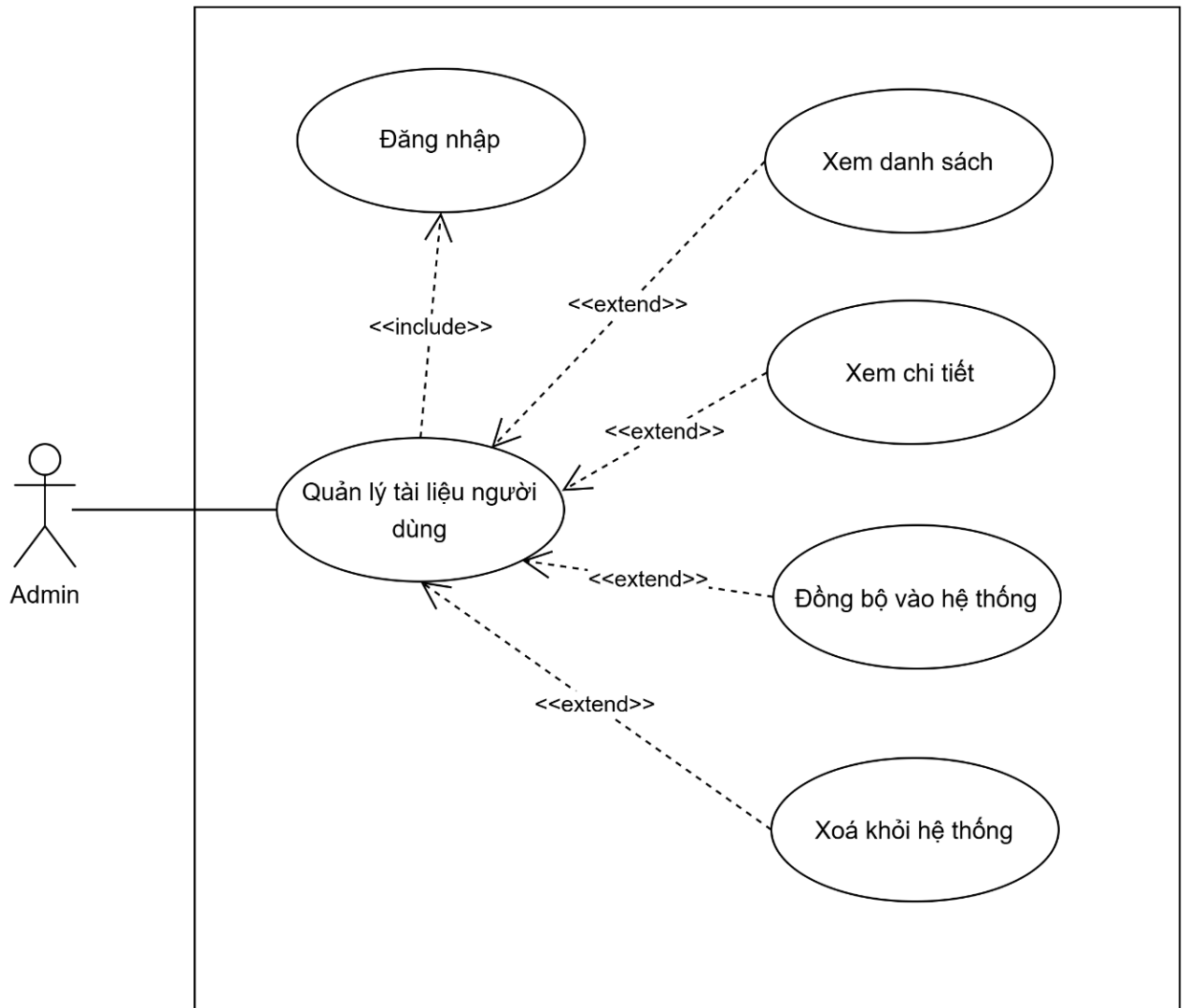
Hình 3.3: Sơ đồ usecase quản lý tài liệu của admin

3.3.4. Quản lý kho kiến thức của admin



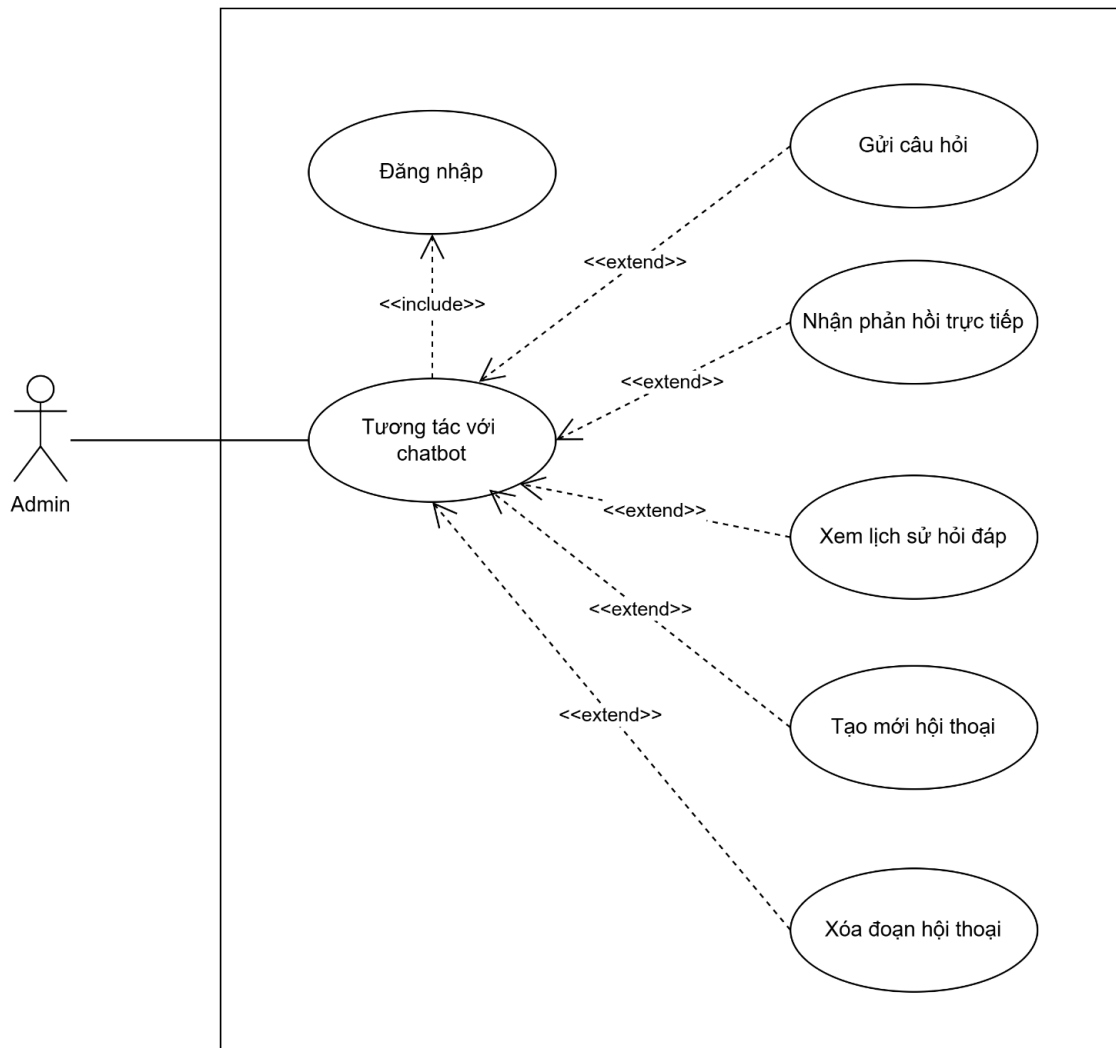
Hình 3.4: Sơ đồ usecase quản lý kho kiến thức của admin

3.3.5. Quản lý tài liệu người dùng



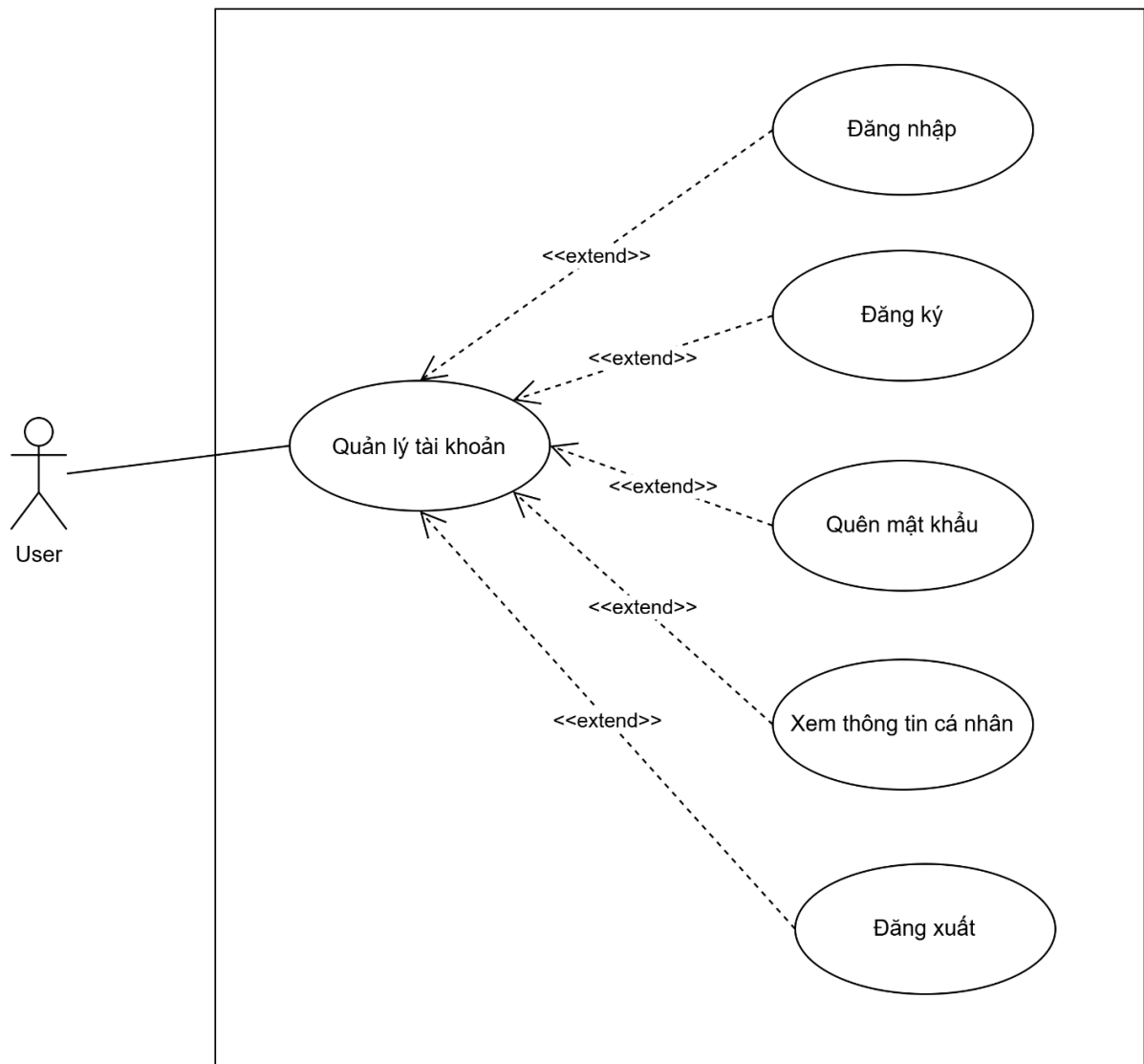
Hình 3.5: Sơ đồ usecase quản lý tài liệu người dùng

3.3.6. Chức năng tương tác với chatbot của admin



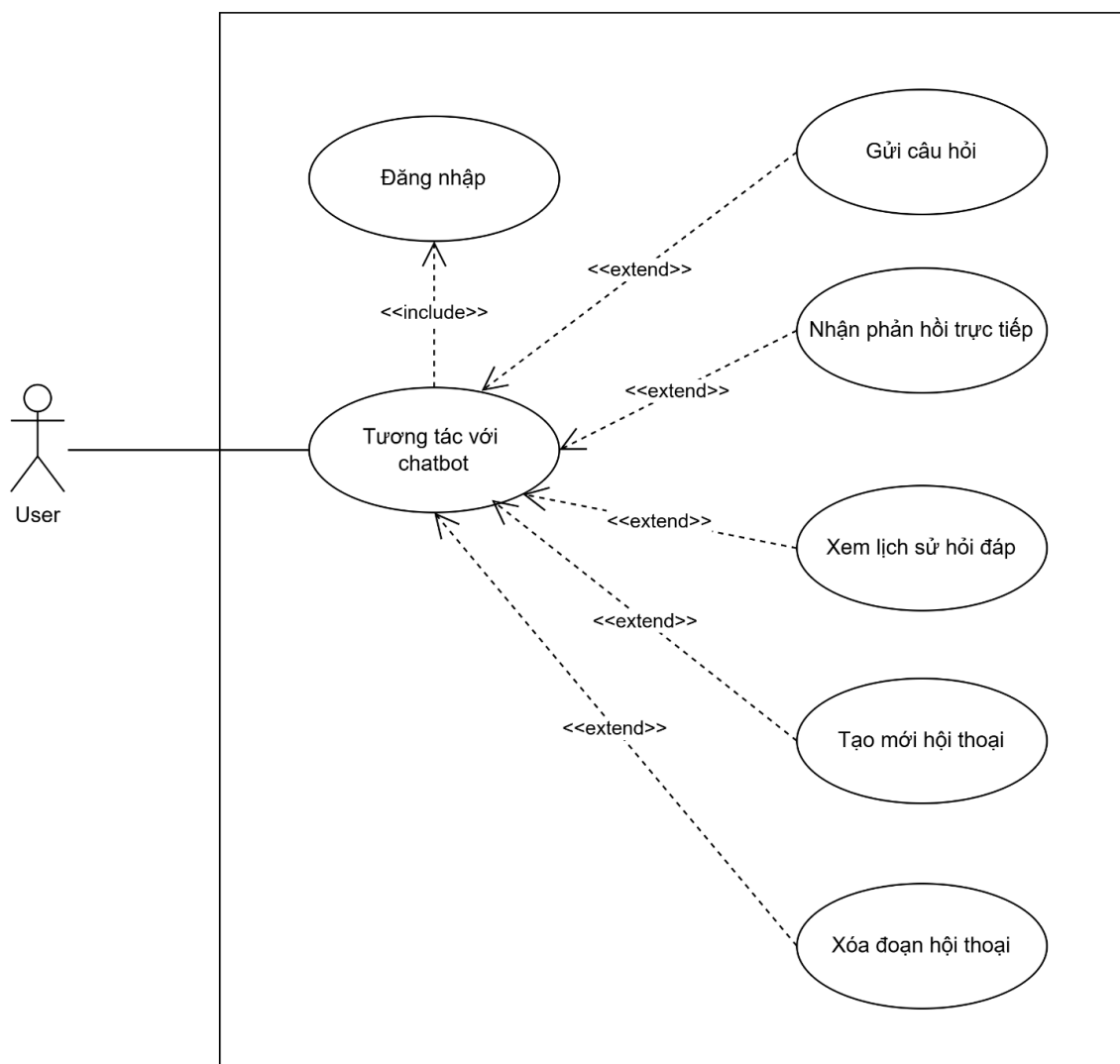
Hình 3.6: Sơ đồ usecase tương tác với chatbot của admin

3.3.7. Chức năng quản lý tài khoản của user



Hình 3.7: Sơ đồ usecase quản lý tài khoản của user

3.3.8. Chức năng tương tác chatbot với user



Hình 3.8: Sơ đồ usecase tương tác chatbot của user

3.4. Công nghệ sử dụng

3.4.1. Frontend

Hệ thống frontend được xây dựng với các công nghệ web hiện đại nhằm tối ưu hiệu suất và trải nghiệm người dùng. Các công nghệ chính gồm:

- **HTML & SCSS:** Tạo cấu trúc trang và định dạng giao diện với khả năng mở rộng và tùy biến cao.
- **JavaScript:** Tạo các tương tác động và xử lý logic phía client.
- **Vue.js:** Framework JavaScript theo hướng component, đơn giản và linh hoạt.

- Nuxt.js: Framework mở rộng từ Vue.js, hỗ trợ SSR, tối ưu hiệu suất và SEO. Tự động tạo routing dựa trên cấu trúc thư mục pages/, tích hợp sẵn quản lý trạng thái và điều hướng.
- Vite: Công cụ build thế hệ mới, hỗ trợ hot-reload nhanh và tối ưu thời gian phát triển.

3.4.2. Backend (Server)

Hệ thống backend được xây dựng với NestJS và Django, mỗi bên đảm nhận một vai trò riêng biệt:

- NestJS: Phát triển API quản lý (tài khoản, tài liệu, kho kiến thức, hội thoại) với cấu trúc module rõ ràng, sử dụng TypeORM, JWT, bcrypt, và tích hợp cấu hình môi trường qua @nestjs/config.
- Django: Xử lý pipeline RAG gồm:
 - + Nhận và phân chia tài liệu.
 - + Chuyển đổi văn bản thành vector bằng SentenceTransformers.
 - + Lưu vector vào Qdrant (vectorDB).
 - + Truy vấn vector theo embedding câu hỏi.
 - + Gửi ngữ cảnh đến LLM và trả về kết quả cho người dùng.

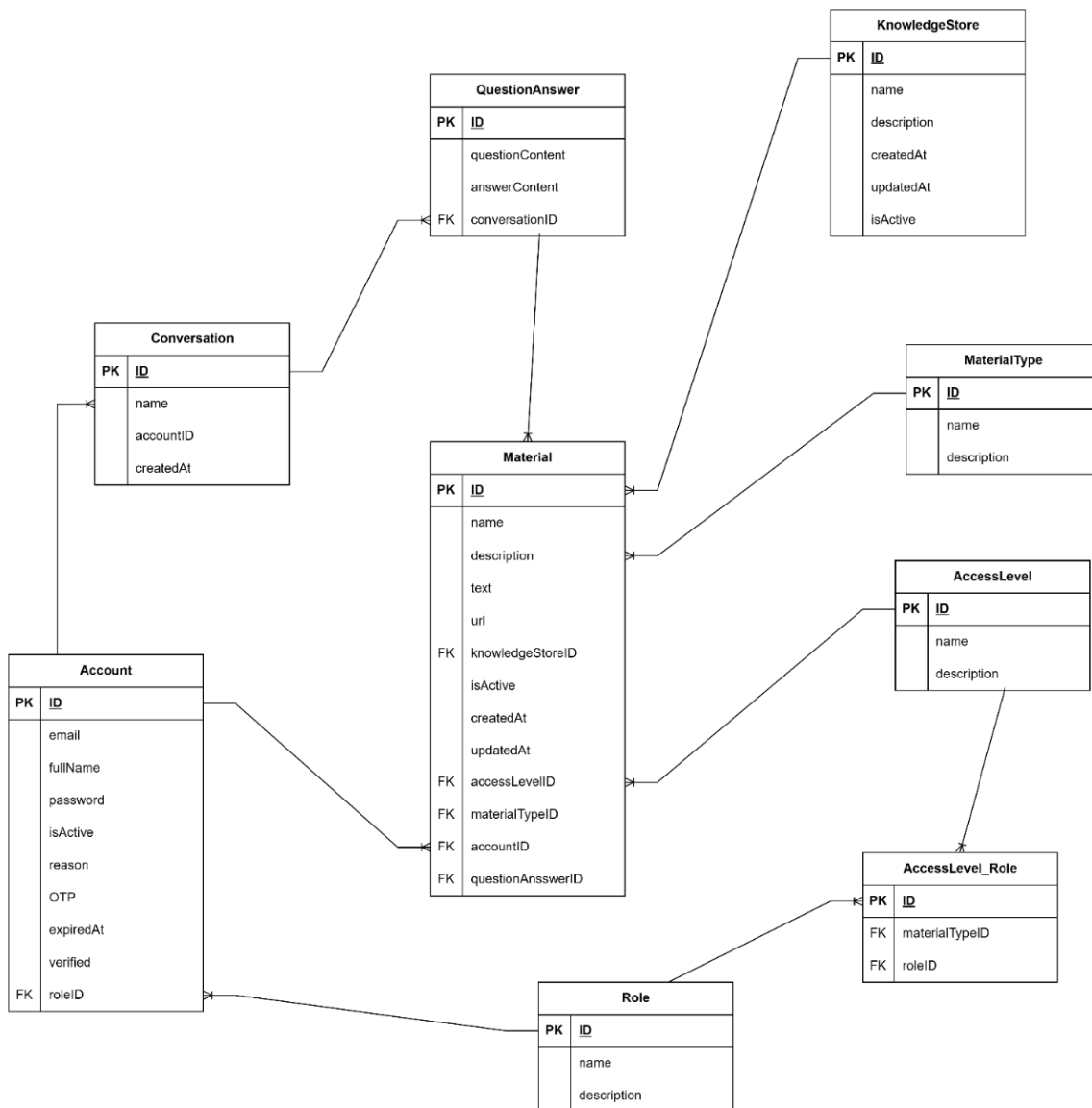
3.4.3. Database

MySQL: Hệ thống sử dụng MySQL – hệ quản trị cơ sở dữ liệu quan hệ phổ biến, có hiệu suất cao, độ tin cậy tốt và dễ triển khai, phù hợp cho các ứng dụng web.

Qdrant: Hệ thống sử dụng Qdrant – vector database mã nguồn mở, hỗ trợ tìm kiếm tương tự (similarity search) hiệu quả, thích hợp với các ứng dụng AI như RAG, tìm kiếm ngữ nghĩa và hệ thống gợi ý.

3.5. Thiết kế cơ sở dữ liệu

3.5.1. Mô hình ERD (Entity-Relationship Diagram)



Hình 3.9: Mô hình cơ sở dữ liệu

3.5.2. Bảng dữ liệu

❖ Cấu trúc bảng Account

Bảng 3.2: Cấu trúc của bảng Account

Thuộc tính	Kiểu dữ liệu	Mô tả
id	int	Khóa chính, tự động tăng, xác định duy nhất mỗi tài khoản.
email	varchar(255)	Email đăng nhập của người dùng
fullName	varchar(255)	Họ tên của người dùng
password	varchar(255)	Mật khẩu của tài khoản
isActive	tinyint	Trạng thái hoạt động của tài khoản
reason	varchar(255)	Lý do khoá tài khoản
OTP	varchar(255)	Mã OTP của tài khoản
expiredAt	datetime	Thời gian hết hạn của OTP
verified	tinyint	Trạng thái xác thực OTP
roleId	Int	Khoá ngoại liên kết với bảng Role, xác định tài khoản được đăng ký với role nào

❖ **Cấu trúc bảng Role***Bảng 3.3: Cấu trúc của bảng Role*

Thuộc tính	Kiểu dữ liệu	Mô tả
id	int	Khóa chính, tự động tăng, xác định duy nhất mỗi vai trò
name	varchar(255)	Tên của vai trò
description	varchar(255)	Mô tả vai trò

❖ **Cấu trúc của bảng MaterialType***Bảng 3.4: Cấu trúc của bảng MaterialType*

Thuộc tính	Kiểu dữ liệu	Mô tả
id	int	Khóa chính, tự động tăng, xác định duy nhất mỗi loại tài liệu
name	varchar(255)	Tên của loại tài liệu
description	varchar(255)	Mô tả loại tài liệu

❖ **Cấu trúc của bảng AccessLevel***Bảng 3.5: Cấu trúc của bảng AccessLevel*

Thuộc tính	Kiểu dữ liệu	Mô tả
id	int	Khóa chính, tự động tăng, xác định duy nhất mỗi quyền truy cập
name	varchar(255)	Tên của quyền truy cập
description	varchar(255)	Mô tả quyền truy cập

❖ Cấu trúc của bảng Material

Bảng 3.6: Cấu trúc của bảng Material

Thuộc tính	Kiểu dữ liệu	Mô tả
id	int	Khóa chính, tự động tăng, xác định duy nhất mỗi tài liệu
name	varchar(255)	Tên của tài liệu
description	varchar(255)	Mô tả cụ thể về tài liệu
url	varchar(255)	Liên kết của tài liệu
text	longtext	Nội dung văn bản của tài liệu
isActive	tinyint	Trạng thái hoạt động của tài liệu
createdAt	datetime	Thời gian tạo tài liệu
updatedAt	datetime	Thời gian cập nhật tài liệu
knowledgeStoreId	int	Khoá ngoại tham chiếu đến bảng KnowledgeStore, xác định tài liệu nằm trong knowledge store nào
materialTypeId	int	Khoá ngoại tham chiếu đến bảng Material Type, xác định tài liệu thuộc kiểu tài liệu nào (url, content, file)
accessLevelId	int	Khoá ngoại tham chiếu đến bảng Access Level, xác định tài liệu thuộc quyền truy cập nào (public, private, internal)

accountId	int	Khoá ngoại tham chiếu đến bảng Account, xác định tài liệu thuộc tài khoản nào.
questionAnswerId	int	Khóa ngoại tham chiếu đến bảng QuestionAnswer, xác định tài liệu thuộc nội dung câu hỏi, câu trả lời nào.

❖ Cấu trúc của bảng AccessLevelRole

Bảng 3.7: Cấu trúc của bảng AccessLevelRole

Thuộc tính	Kiểu dữ liệu	Mô tả
accessLevelId	int	Khoá ngoại liên kết với bảng Access Level, xác định quyền truy cập
roleId	int	Khoá ngoại liên kết với bảng Role, xác định vai trò của tài khoản

❖ Cấu trúc của bảng KnowledgeStore

Bảng 3.8: Cấu trúc của bảng KnowledgeStore

Thuộc tính	Kiểu dữ liệu	Mô tả
id	int	Khóa chính, tự động tăng, xác định duy nhất mỗi kho kiến thức
name	varchar(255)	Tên kho kiến thức
description	varchar(255)	Mô tả kho kiến thức
createdAt	datetime	Thời gian tạo kho kiến thức
updatedAt	datetime	Thời gian cập nhật kho

isActive	tinyint	Trạng thái hoạt động của kho kiến thức
----------	---------	--

❖ **Cấu trúc của bảng Conversation**

Bảng 3.9: Cấu trúc của bảng Conversation

Thuộc tính	Kiểu dữ liệu	Mô tả
id	int	Khóa chính, tự động tăng, xác định duy nhất mỗi cuộc hội thoại
name	varchar(255)	Tên của cuộc hội thoại
createdAt	datetime	Ngày tạo ra cuộc hội thoại
accountId	varchar(255)	Khóa ngoại liên kết với bảng account, xác định id của tài khoản tạo ra cuộc hội thoại

❖ **Cấu trúc của bảng QuestionAnswer**

Bảng 3.10: Cấu trúc của bảng QuestionAnswer

Thuộc tính	Kiểu dữ liệu	Mô tả
id	int	Khóa chính, tự động tăng, xác định duy nhất mỗi câu hỏi và câu trả lời
conversationId	int	Khóa ngoại liên kết với bảng conversation, xác định cuộc hội thoại mà câu hỏi và câu trả lời nó thuộc về
questionContent	longtext	Nội dung câu hỏi
answerContent	longtext	Câu trả lời của câu hỏi

3.6. Tổng kết chương

Chương này trình bày các yêu cầu hệ thống, sơ đồ use case, kiến trúc tổng thể, công nghệ sử dụng và thiết kế cơ sở dữ liệu, làm cơ sở cho quá trình triển khai hệ thống ở các chương tiếp theo.

CHƯƠNG 4. MÔ TẢ CHỨC NĂNG

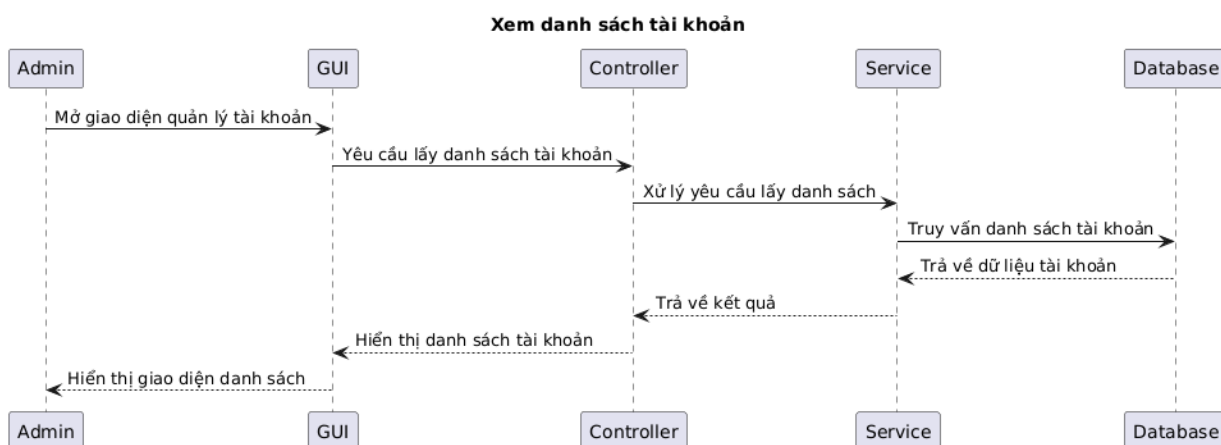
4.1. Chức năng của Admin

4.1.1. Quản lý tài khoản

❖ Xem danh sách tài khoản

Bảng 4.1: Mô tả chức năng xem danh sách tài khoản

Tên chức năng	Xem danh sách tài khoản
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài khoản
Kết quả	Hiển thị danh sách tài khoản



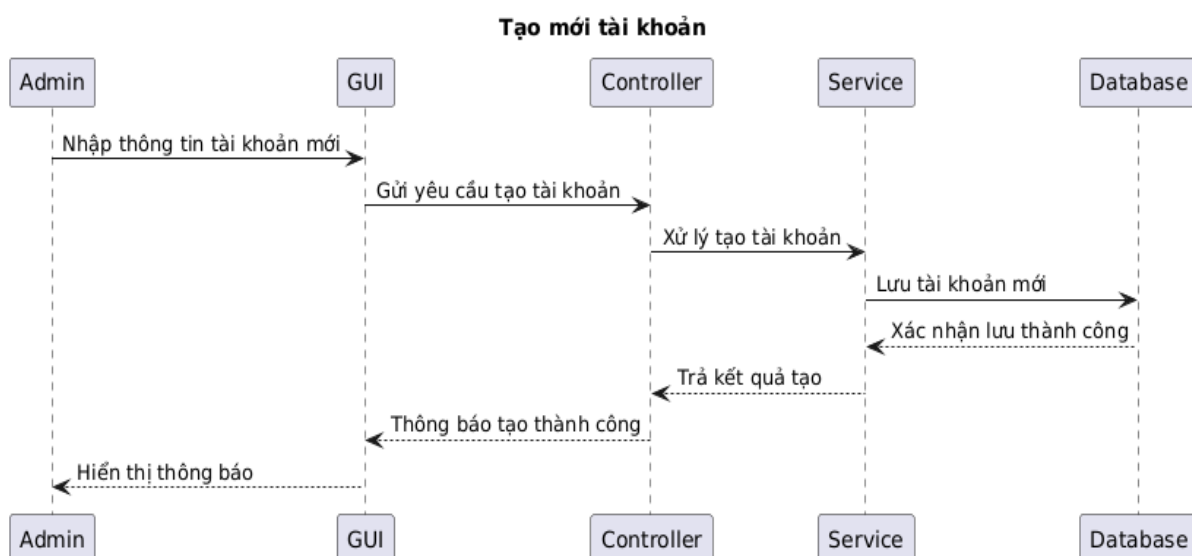
Hình 4.1: Sơ đồ chức năng xem danh sách tài khoản

❖ Tạo mới tài khoản

Bảng 4.2: Mô tả chức năng tạo mới tài khoản

Tên chức năng	Thêm mới tài khoản
Đối tượng sử dụng	Admin

Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài khoản → Nhấn vào button “Thêm mới” / “Create” → Điền thông tin và nhấn “Luu” / “Save”
Kết quả	Thông báo thành công và hiển thị danh sách tài khoản mới

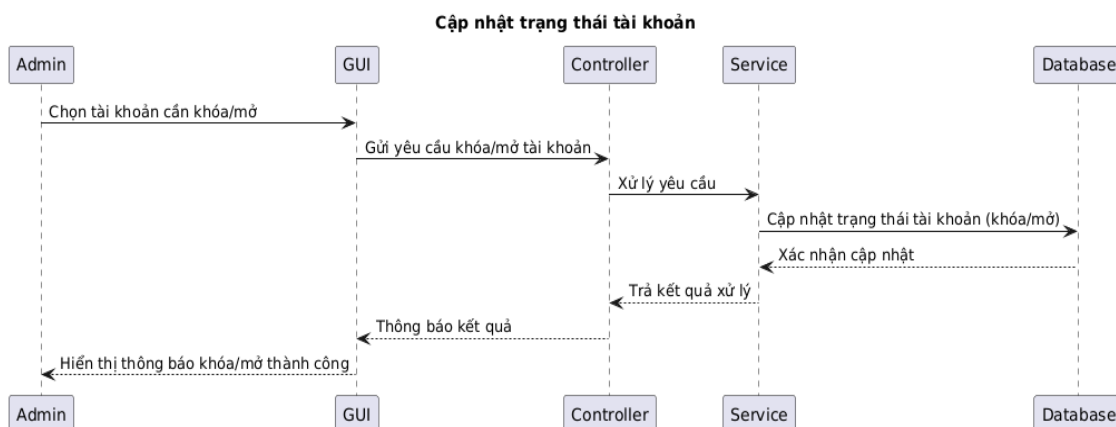


Hình 4.2: Sơ đồ chức năng tạo mới tài khoản

❖ Cập nhật trạng thái tài khoản

Bảng 4.3: Mô tả chức năng cập nhật trạng thái tài khoản

Tên chức năng	Cập nhật trạng thái tài khoản
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài khoản → Nhấn vào nút chuyển trạng thái (toggle switch) để cập nhật trạng thái
Kết quả	Cập nhật trạng thái tài khoản thành công

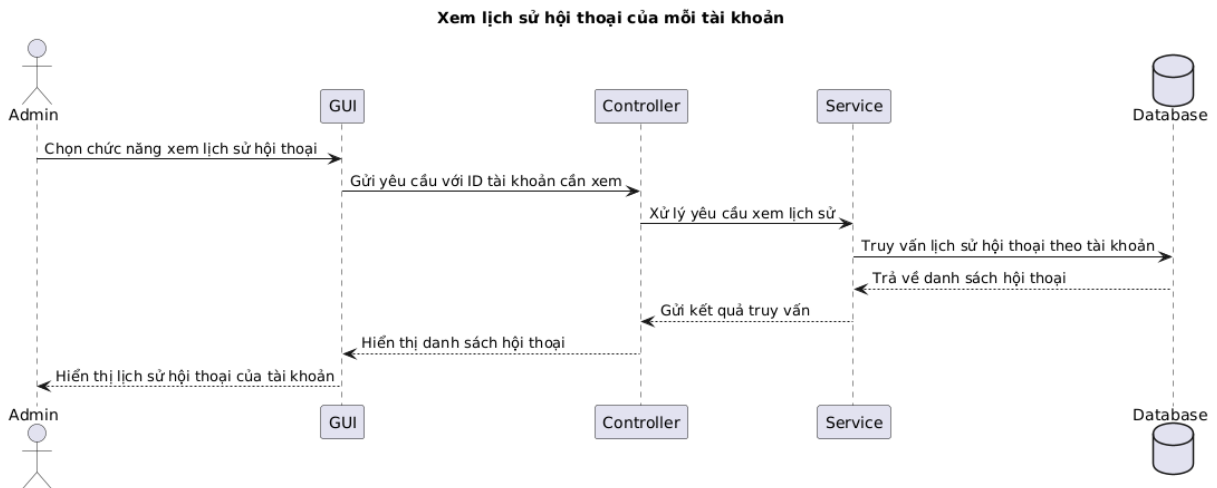


Hình 4.3: Sơ đồ chức năng cập nhật trạng thái tài khoản

❖ **Xem lịch sử hội thoại của mỗi tài khoản**

Bảng 4.4: Mô tả chức năng xem lịch sử hội thoại

Tên chức năng	Xem lịch sử hội thoại của mỗi tài khoản
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài khoản → Nhấn vào button “Xem chi tiết hội thoại” / “View detail conversations”
Kết quả	Hiển thị lịch sử các cuộc hội thoại của mỗi tài khoản



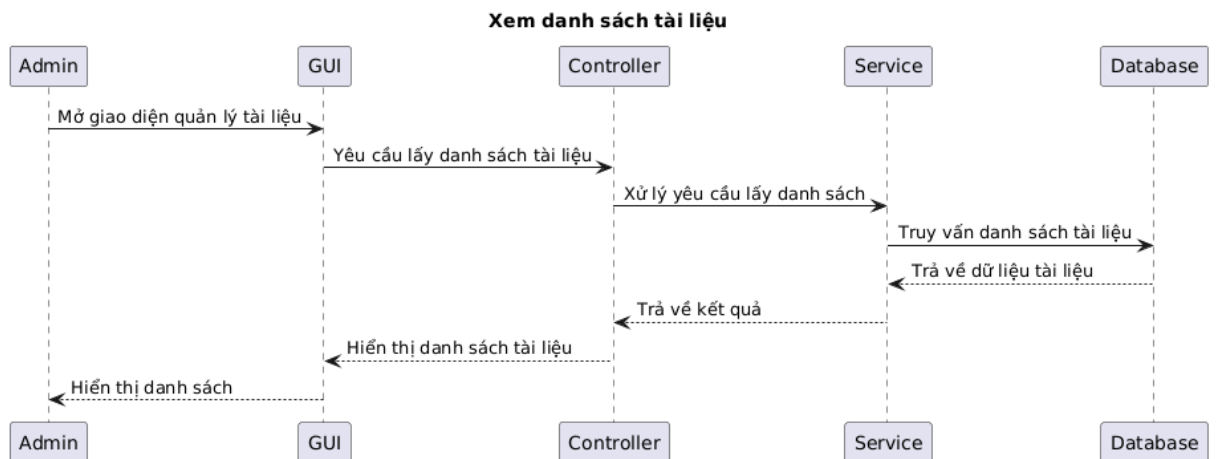
Hình 4.4: Sơ đồ chức năng xem lịch sử hội thoại mỗi tài khoản

4.1.2. Quản lý tài liệu

❖ Xem danh sách tài liệu

Bảng 4.5: Mô tả chức năng xem danh sách tài liệu

Tên chức năng	Xem danh sách tài liệu
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài liệu
Kết quả	Hiển thị danh sách tài liệu

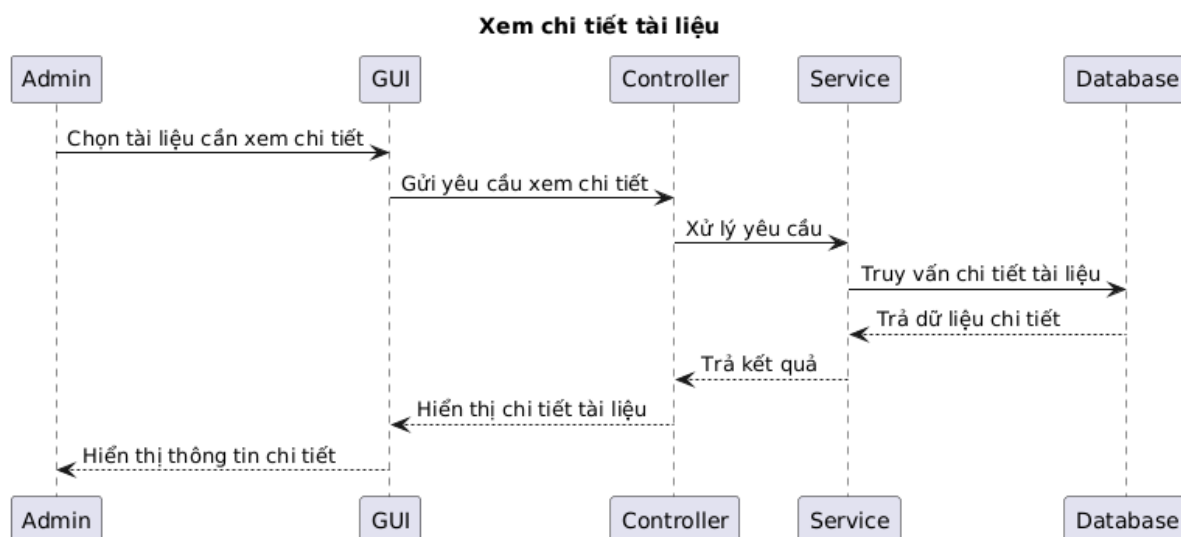


Hình 4.5: Sơ đồ chức năng xem danh sách tài liệu

❖ Xem chi tiết tài liệu

Bảng 4.6: Mô tả chức năng xem chi tiết tài liệu

Tên chức năng	Xem chi tiết tài liệu
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài liệu → Nhấn vào button “Chi tiết” / “Items”
Kết quả	Hiển thị thông tin chi tiết tài liệu

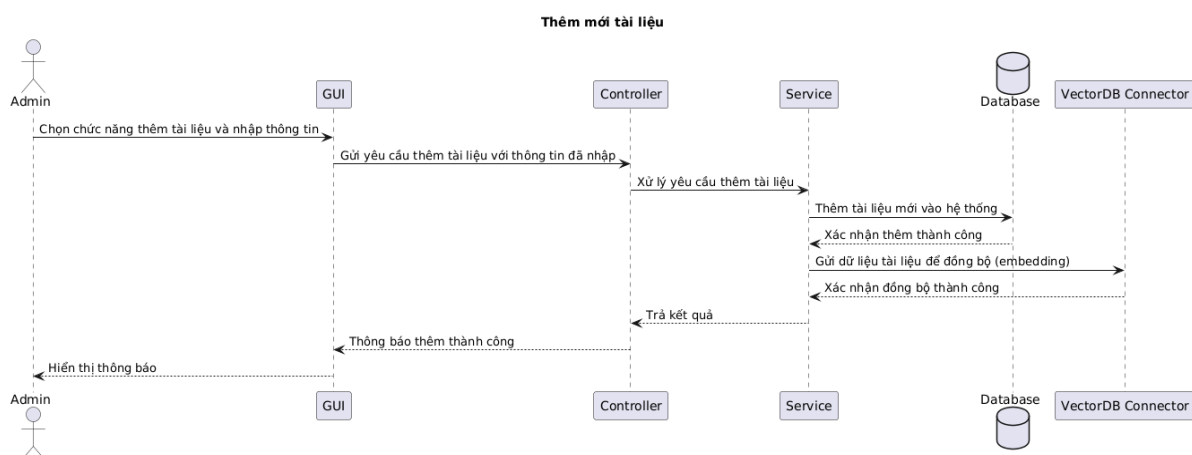


Hình 4.6: Sơ đồ chức năng xem chi tiết tài liệu

❖ Thêm mới tài liệu

Bảng 4.7: Mô tả chức năng thêm mới tài liệu

Tên chức năng	Thêm mới tài liệu
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài liệu → Nhấn vào button “Thêm mới” / “Create” → Điền thông tin và nhấn “Lưu” / “Save”
Kết quả	Hiện thị danh sách tài liệu mới.

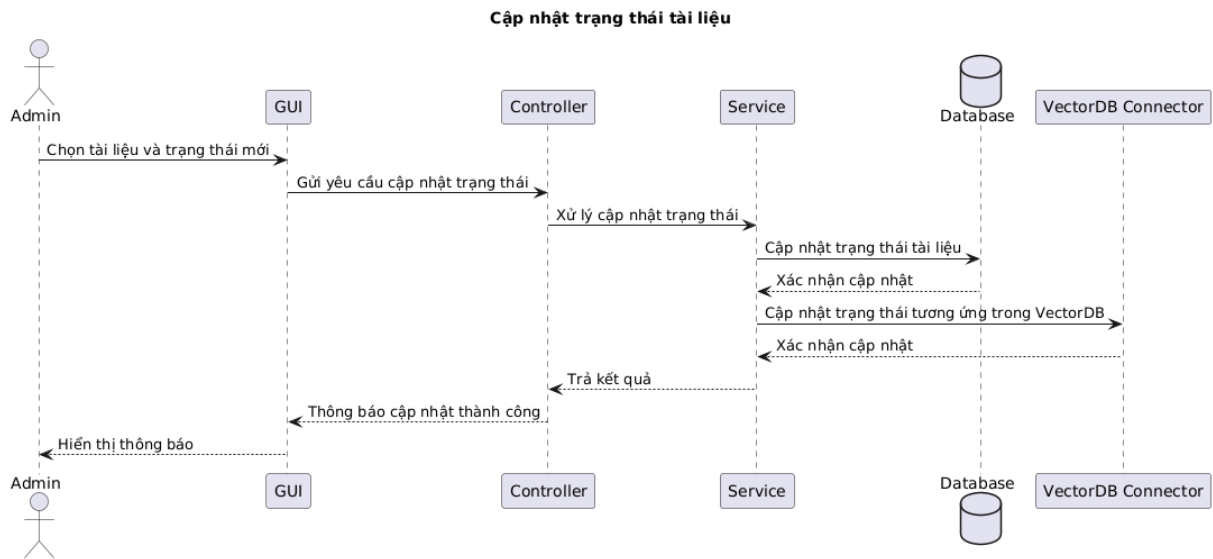


Hình 4.7: Sơ đồ chức năng thêm mới tài liệu

❖ Cập nhật trạng thái tài liệu

Bảng 4.8: Mô tả chức năng cập nhật trạng thái tài liệu

Tên chức năng	Cập nhật trạng thái tài liệu
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài liệu → Nhấn vào nút chuyển trạng thái (toggle switch) tương ứng với mỗi tài liệu để cập nhật
Kết quả	Trạng thái tài liệu bị thay đổi

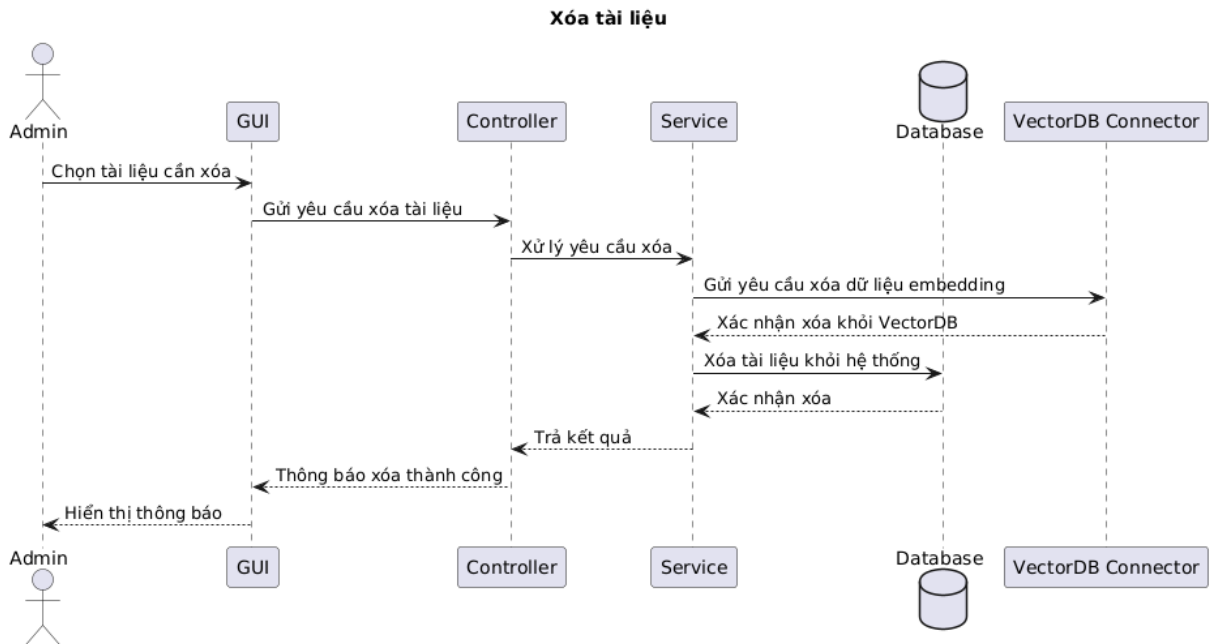


Hình 4.8: Sơ đồ chức năng cập nhật trạng thái tài liệu

❖ Xóa tài liệu

Bảng 4.9: Mô tả chức năng xóa tài liệu

Tên chức năng	Xóa tài liệu
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Truy cập trang quản lý tài liệu → Nhấn vào button “Xóa” / “Delete” tương ứng mỗi tài liệu và xác nhận xóa
Kết quả	Hiển thị danh sách tài liệu mới



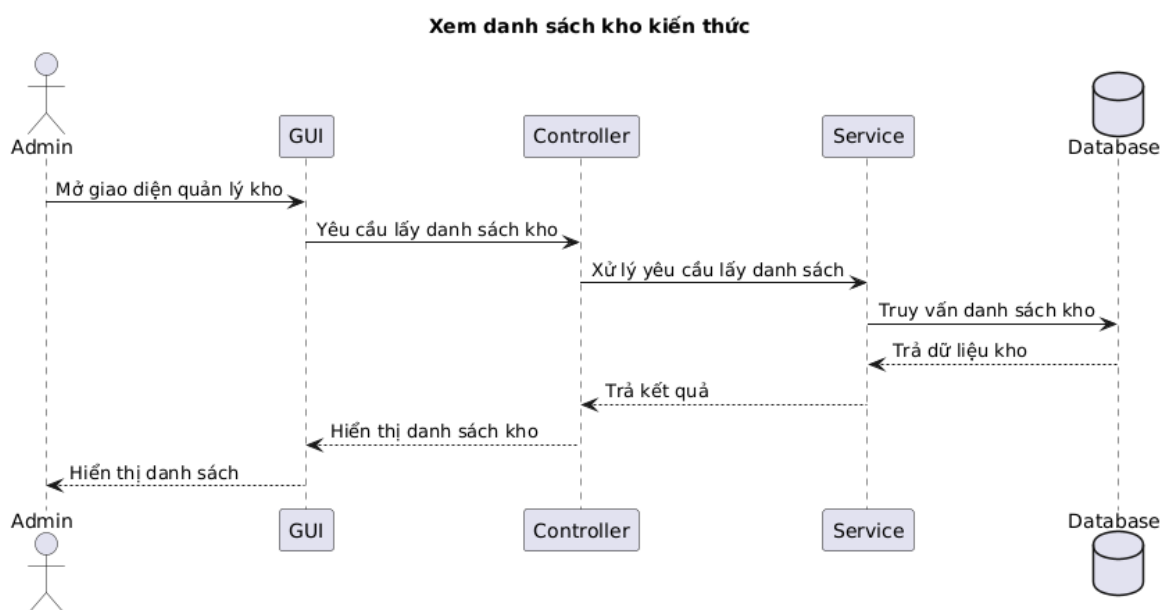
Hình 4.9: Sơ đồ chức năng xóa tài liệu

4.1.3. Quản lý kho kiến thức

❖ Xem danh sách kho kiến thức

Bảng 4.10: Mô tả chức năng quản lý kho kiến thức

Tên chức năng	Xem danh sách kho kiến thức
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang quản lý kho kiến thức
Kết quả	Hiển thị danh sách kho kiến thức

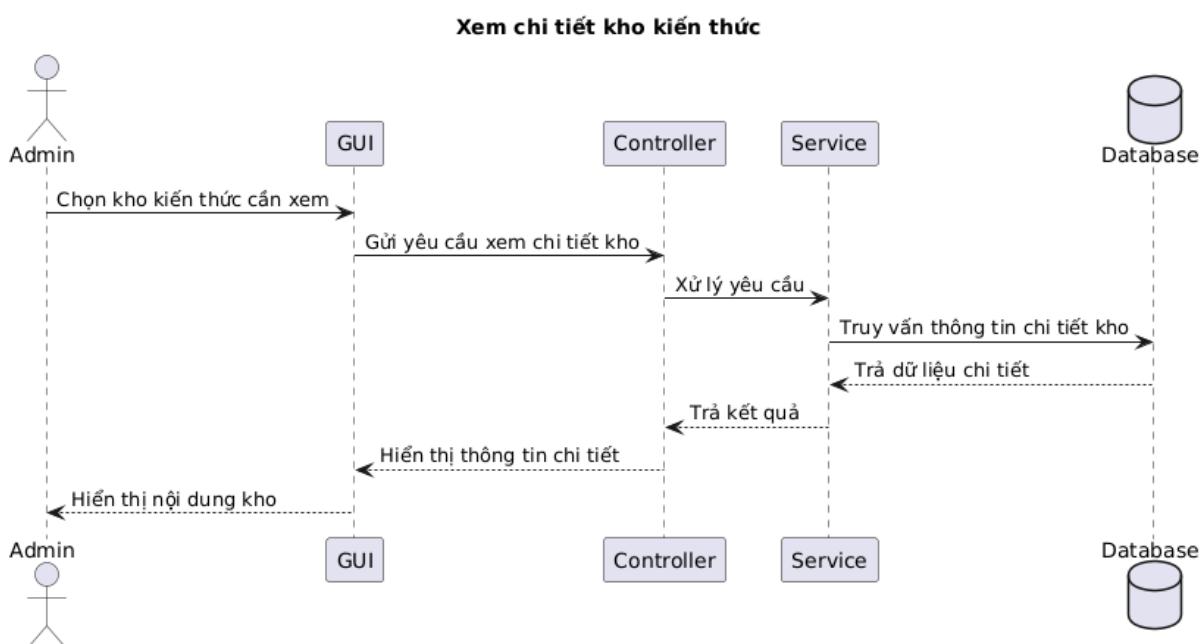


Hình 4.10: Sơ đồ chức năng quản lý kho kiến thức

❖ Xem thông tin chi tiết

Bảng 4.11: Mô tả chức năng xem chi tiết kho kiến thức

Tên chức năng	Xem thông tin chi tiết kho kiến thức
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang quản lý kho kiến thức → Nhấn vào button “Chi tiết” / “Items”
Kết quả	Hiển thị thông tin chi tiết của kho kiến thức

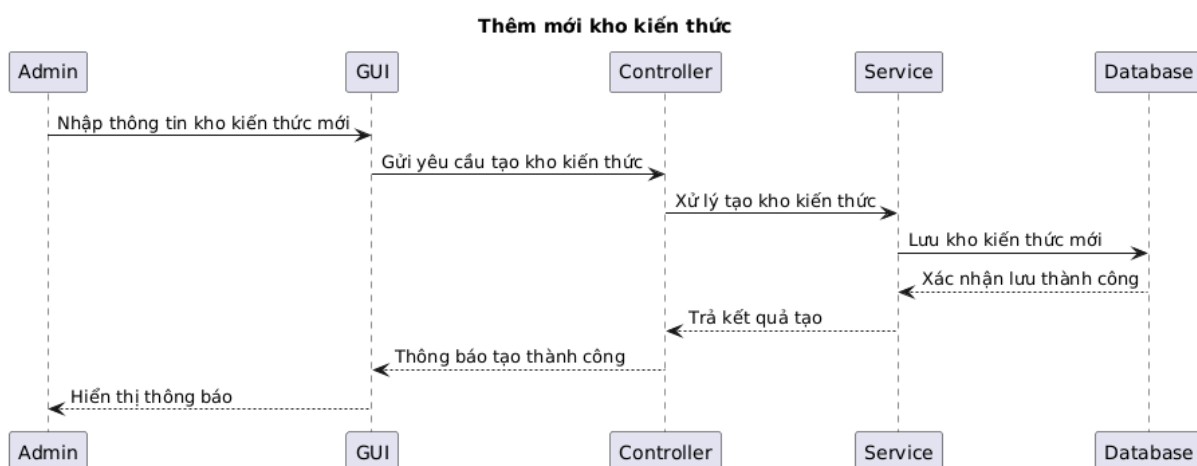


Hình 4.11: Sơ đồ chức năng xem chi tiết kho kiến thức

❖ Thêm mới kho kiến thức

Bảng 4.12: Mô tả chức năng thêm mới kho kiến thức

Tên chức năng	Thêm mới kho kiến thức
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang quản lý kho kiến thức → Nhấn vào button “Thêm mới” / “Create” → Điền thông tin và nhấn “Lưu” / “Save”
Kết quả	Hiển thị danh sách kho kiến thức mới



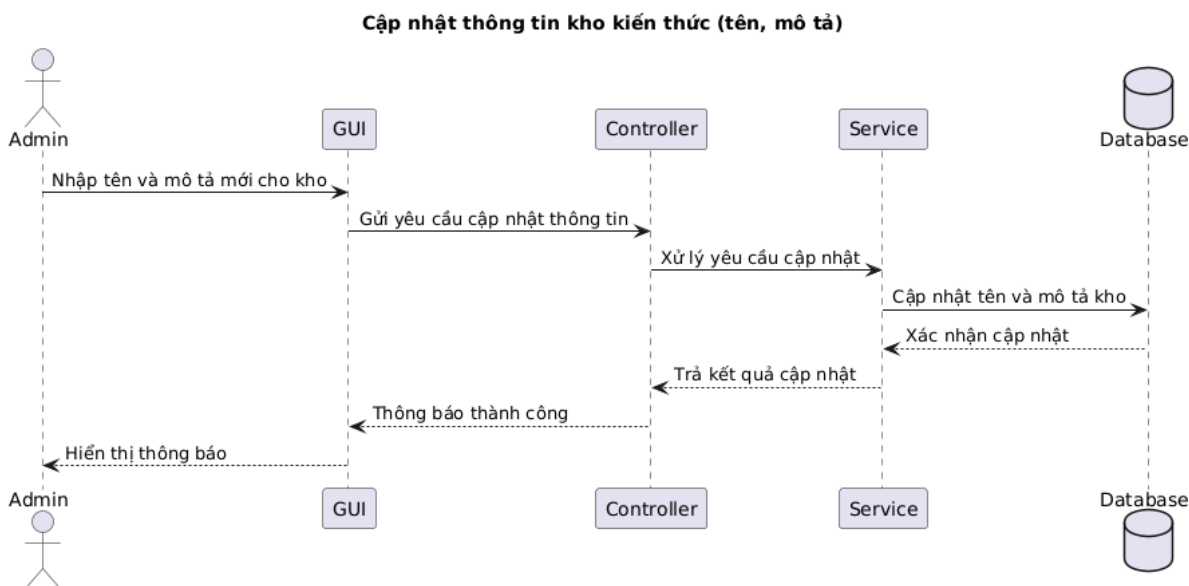
Hình 4.12: Sơ đồ chức năng thêm mới kho kiến thức

❖ Chỉnh sửa thông tin kho kiến thức

Bảng 4.13: Mô tả chức năng chỉnh sửa kho kiến thức

Tên chức năng	Chỉnh sửa thông tin kho kiến thức
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin

Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang quản lý kho kiến thức → Nhấn vào button “Chỉnh sửa” / “Edit”
Kết quả	Hiển thị thông tin mới cho kho

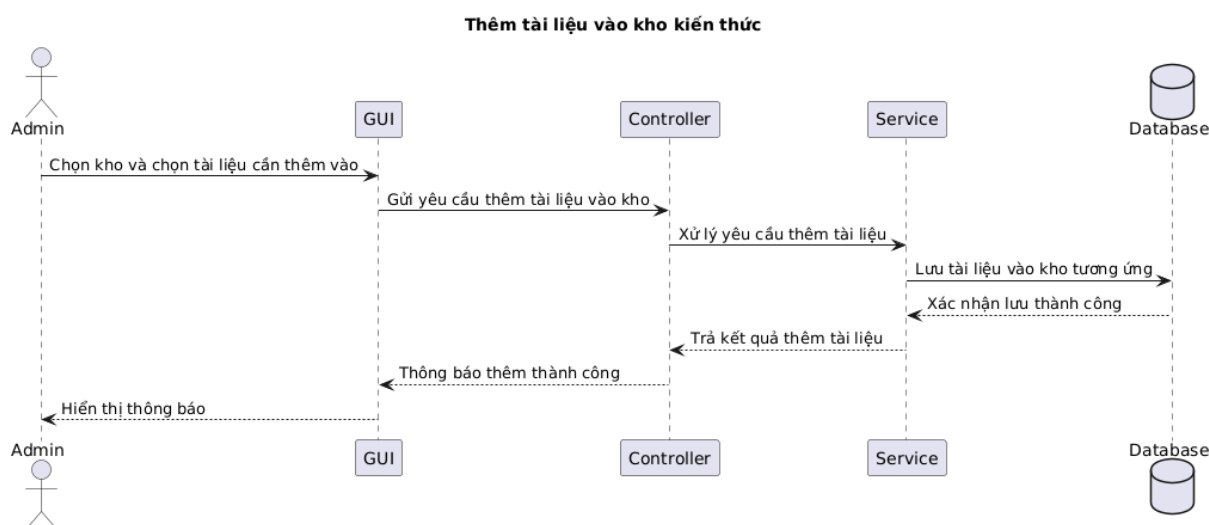


Hình 4.13: Sơ đồ chức năng cập nhật thông tin kho kiến thức

❖ Thêm tài liệu vào kho kiến thức

Bảng 4.14: Mô tả chức năng thêm tài liệu vào kho kiến thức

Tên chức năng	Thêm tài liệu vào kho kiến thức
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang quản lý kho kiến thức → Nhấn vào button “Tài liệu” / “Materials” → Checked vào những tài liệu muốn thêm vào
Kết quả	Thêm thành công tài liệu vào kho

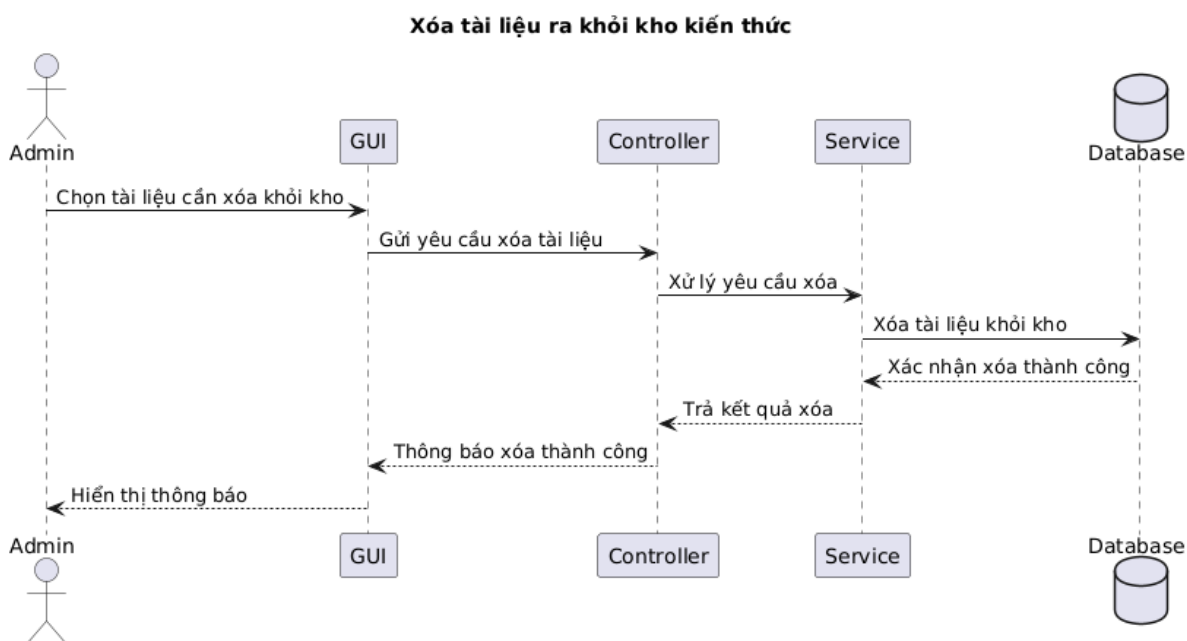


Hình 4.14: Sơ đồ chức năng thêm tài liệu vào kho kiến thức

❖ **Xóa tài liệu ra khỏi kho**

Bảng 4.15: Mô tả chức năng xóa tài liệu ra khỏi kho

Tên chức năng	Xóa tài liệu ra khỏi kho
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang quản lý kho kiến thức → Nhấn vào button “Tài liệu” / “Materials” → Uncheck với những tài liệu muốn xóa
Kết quả	Xóa tài liệu ra khỏi kho thành công

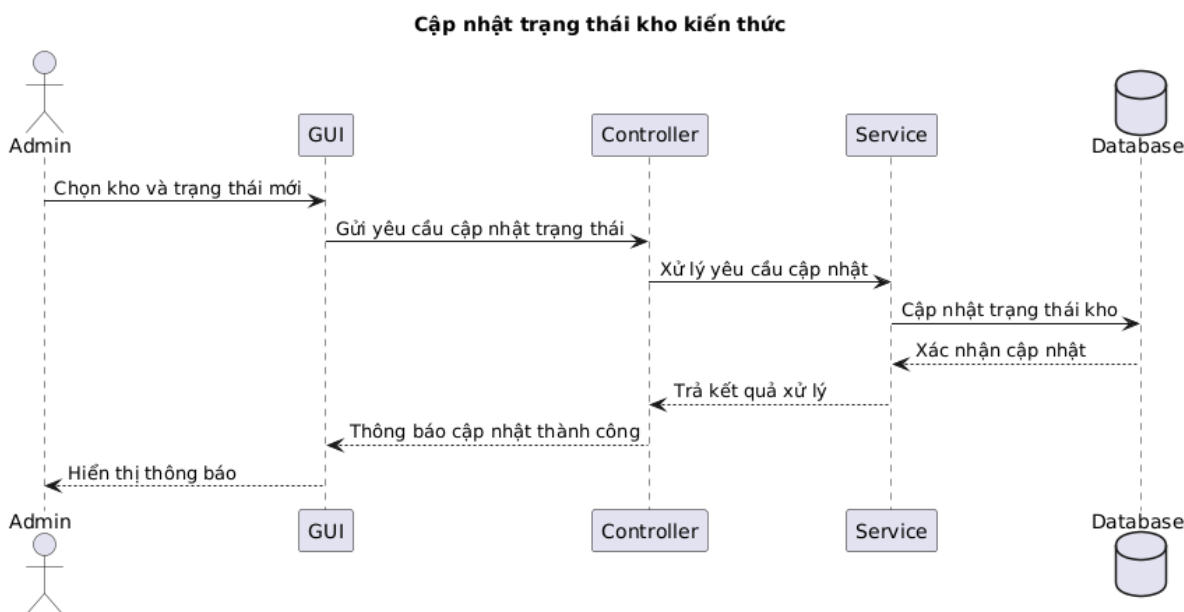


Hình 4.15: Sơ đồ chức năng xóa tài liệu ra khỏi kho

❖ **Cập nhật trạng thái của kho kiến thức**

Bảng 4.16: Mô tả chức năng cập nhật trạng thái kho kiến thức

Tên chức năng	Cập nhật trạng thái kho kiến thức
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang quản lý kho kiến thức → Nhấn vào toggle switch tương ứng mỗi kho
Kết quả	Trạng thái của kho kiến thức bị thay đổi



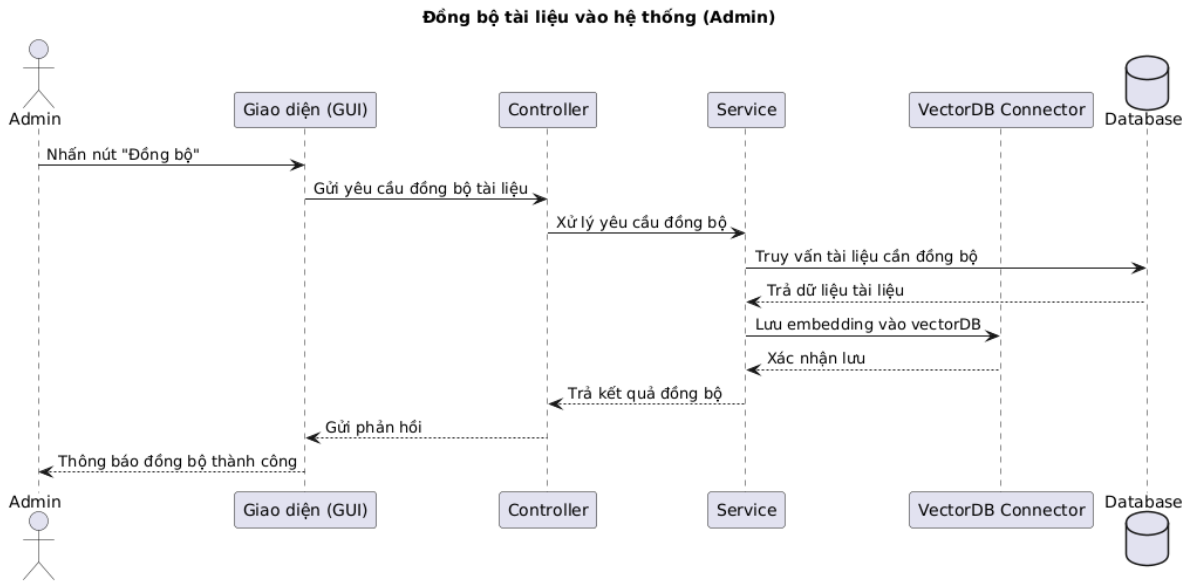
Hình 4.16: Sơ đồ chức năng cập nhật trạng thái kho kiến thức

4.1.4. Quản lý tài liệu người dùng

❖ Đồng bộ tài liệu vào hệ thống

Bảng 4.17: Mô tả chức năng đồng bộ tài liệu vào hệ thống

Tên chức năng	Đồng bộ tài liệu vào hệ thống
Đối tượng sử dụng	Admin
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang tài liệu tải lên → Nhấn vào button “Đồng bộ” / “Sync”
Kết quả	Tài liệu được lưu vào vectorDB của hệ thống



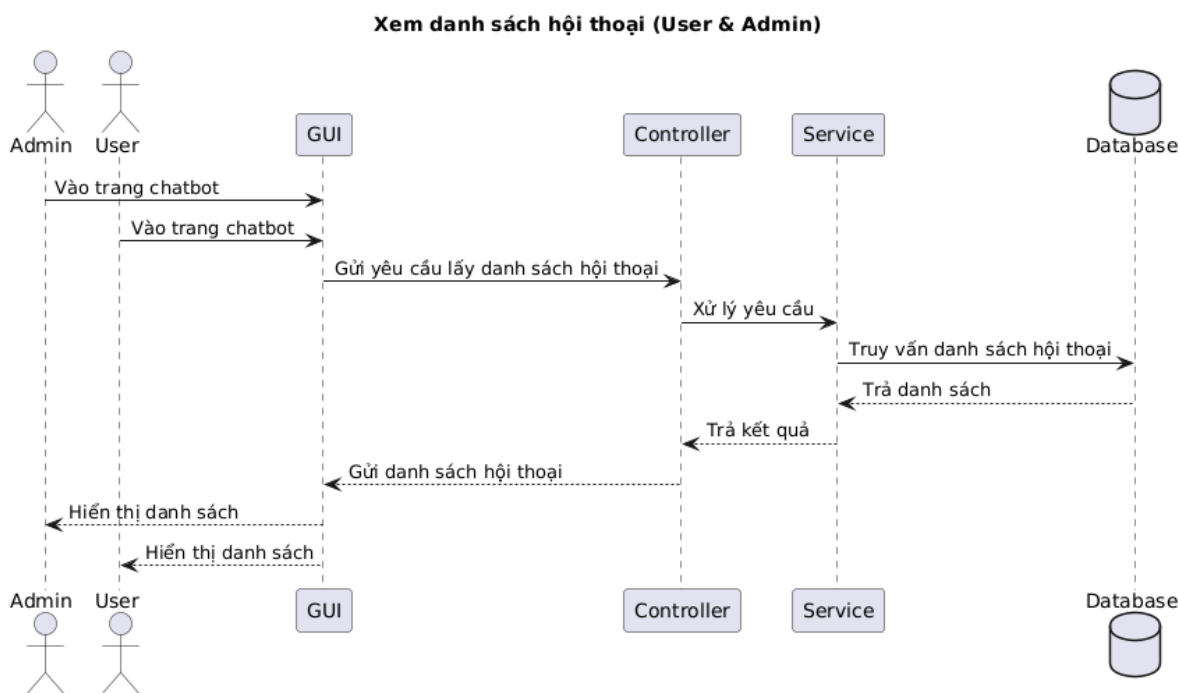
Hình 4.17: Sơ đồ chức năng đồng bộ tài liệu vào hệ thống

4.2. Chức năng tương tác với chatbot ở Admin và User

❖ Xem danh sách cuộc hội thoại

Bảng 4.18: Mô tả chức năng xem danh sách cuộc hội thoại

Tên chức năng	Xem danh sách cuộc hội thoại
Đối tượng sử dụng	Admin/User
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin hoặc User
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang chatbot
Kết quả	Hiện thị danh sách cuộc hội thoại

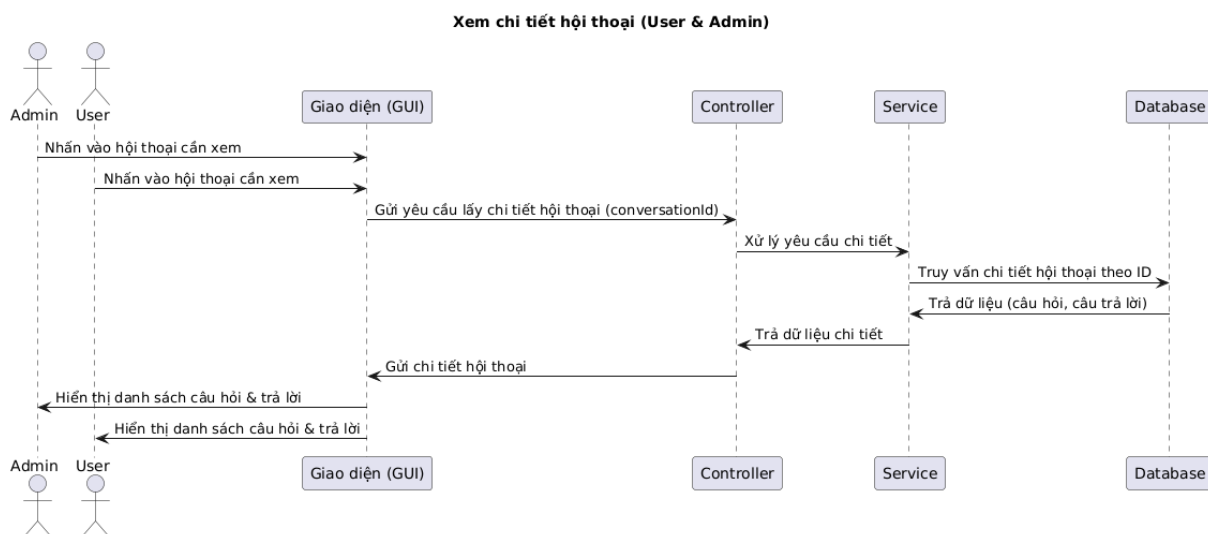


Hình 4.18: Sơ đồ chức năng xem danh sách hội thoại

❖ Xem chi tiết cuộc hội thoại

Bảng 4.19: Mô tả chức năng xem chi tiết cuộc hội thoại

Tên chức năng	Xem chi tiết cuộc hội thoại
Đối tượng sử dụng	Admin/User
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin hoặc User
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang chatbot → Chọn hội thoại cần xem.
Kết quả	Hiển thị chi tiết hội thoại bao gồm câu hỏi và câu trả lời

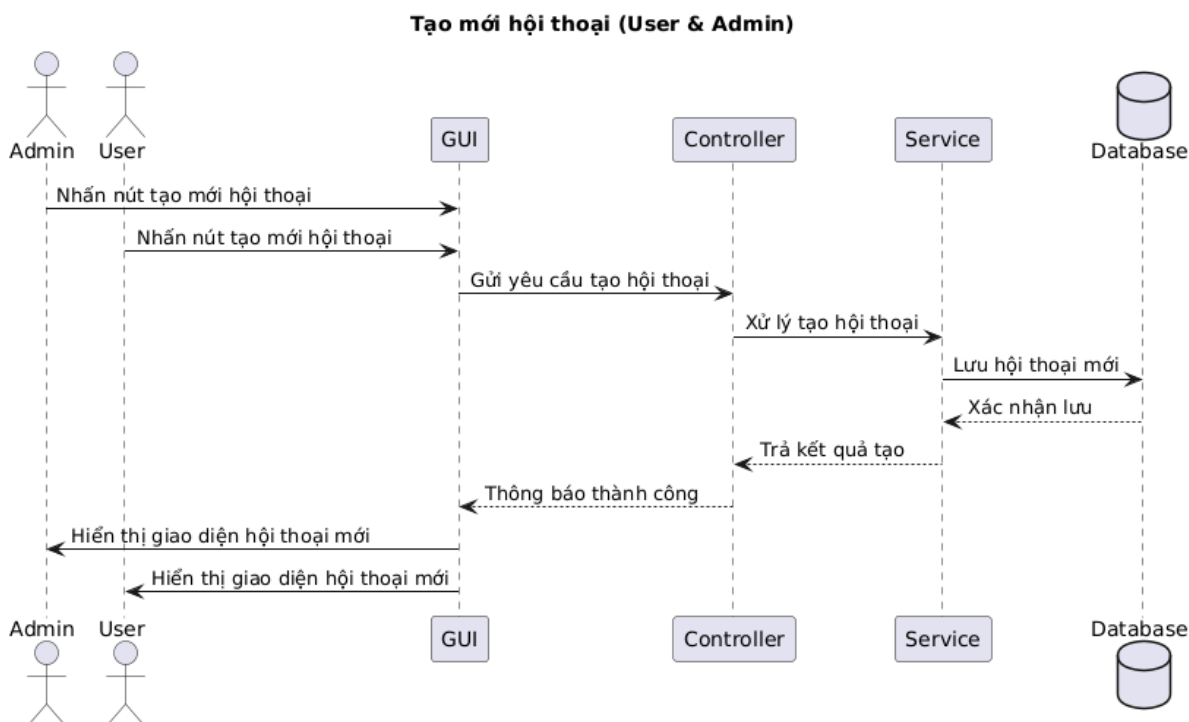


Hình 4.19: Sơ đồ chức năng xem chi tiết cuộc hội thoại

❖ Thêm mới cuộc hội thoại

Bảng 4.20: Mô tả chức năng thêm mới cuộc hội thoại

Tên chức năng	Thêm mới cuộc hội thoại
Đối tượng sử dụng	Admin/User
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin hoặc User
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang chatbot → Nhấn vào thêm mới
Kết quả	Hiển thị giao diện hội thoại mới

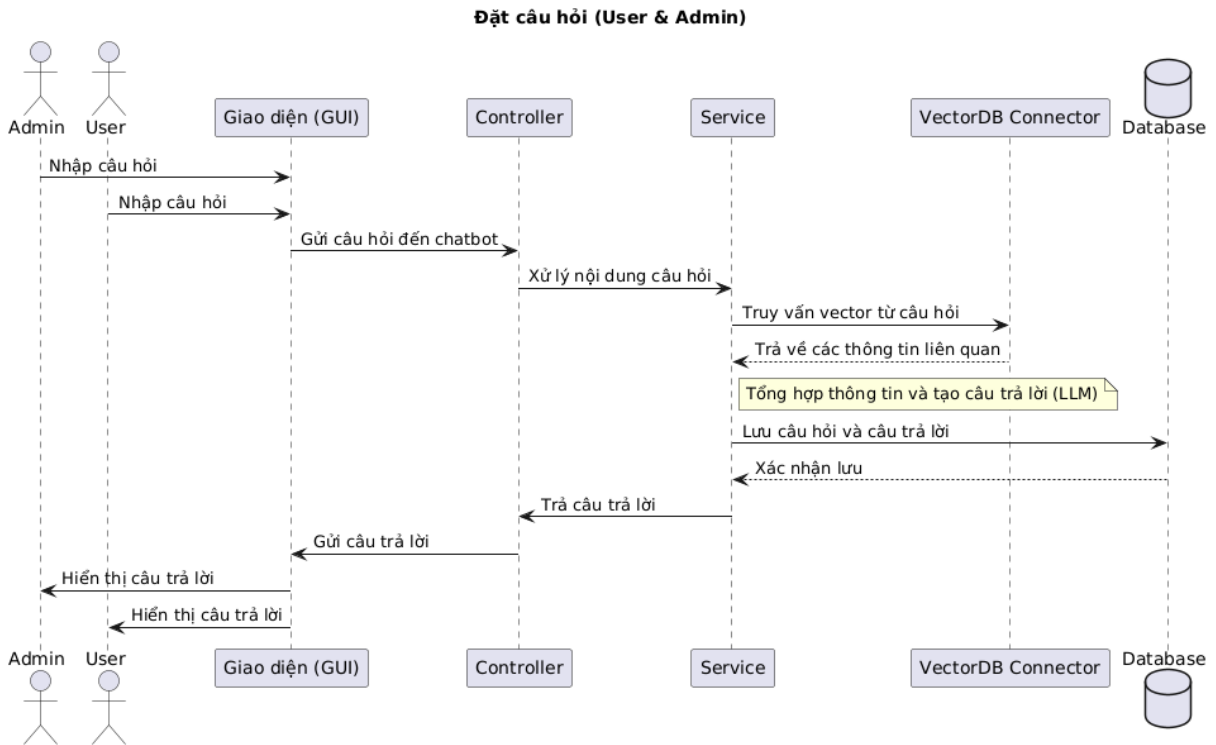


Hình 4.20: Sơ đồ chức năng thêm mới cuộc hội thoại

❖ **Đặt câu hỏi**

Bảng 4.21: Mô tả chức năng đặt câu hỏi

Tên chức năng	Đặt câu hỏi cho chatbot
Đối tượng sử dụng	Admin/User
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin hoặc User
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang chatbot → Nhập câu hỏi và gửi đến chatbot
Kết quả	Nhận và hiển thị phản hồi từ chatbot dưới dạng câu trả lời

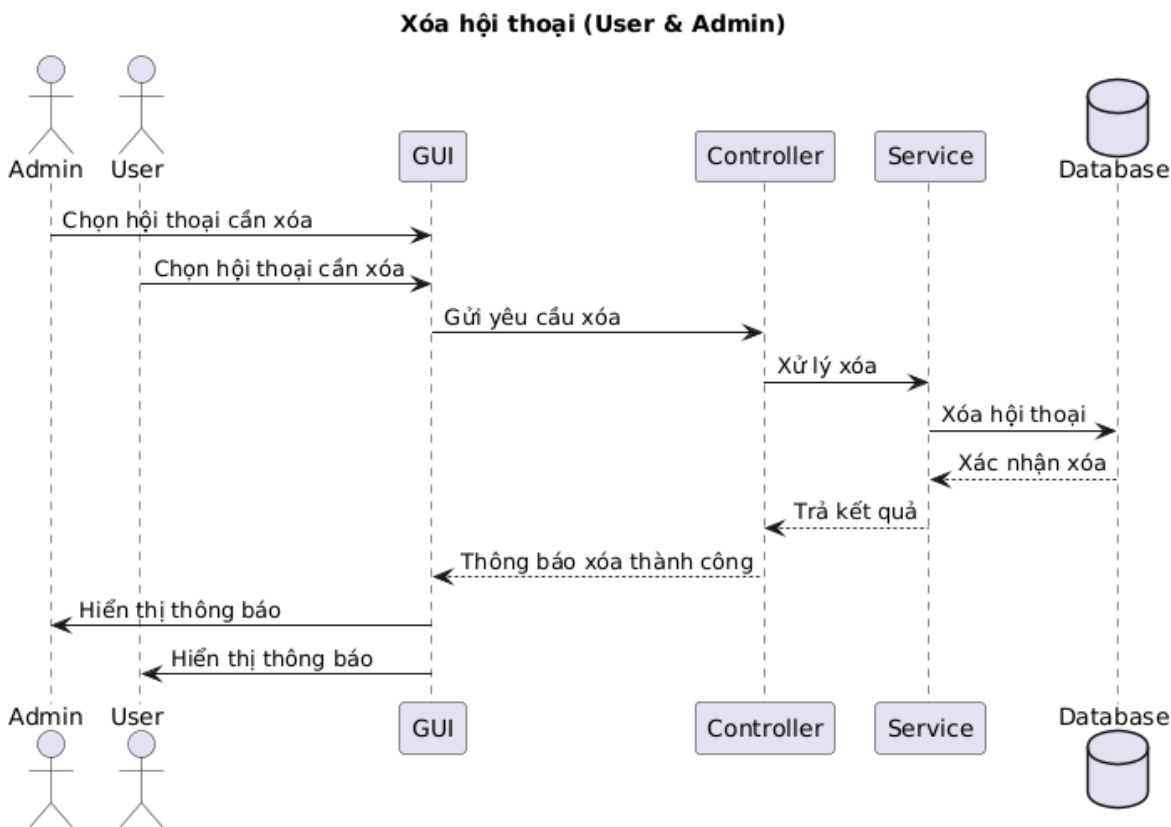


Hình 4.21: Sơ đồ chức năng đặt câu hỏi

❖ **Xóa cuộc hội thoại**

Bảng 4.22: Mô tả chức năng xóa cuộc hội thoại

Tên chức năng	Xóa cuộc hội thoại
Đối tượng sử dụng	Admin/User
Điều kiện	Tài khoản tồn tại trên hệ thống, quyền Admin hoặc User
Quy trình nghiệp vụ	Đăng nhập hệ thống → Vào trang chatbot → Chọn cuộc hội thoại cần xóa và nhấn xác nhận “Xóa” / “Delete”
Kết quả	Xóa thành công và hiển thị danh sách mới



Hình 4.22: Sơ đồ chức năng xóa cuộc hội thoại

4.3. Tổng kết chương

Chương này trình bày chi tiết các chức năng của hệ thống thông qua mô tả hành động và sơ đồ tuần tự, giúp làm rõ quy trình xử lý và tương tác giữa các thành phần.

CHƯƠNG 5. TRIỂN KHAI VÀ ĐÁNH GIÁ KẾT QUẢ

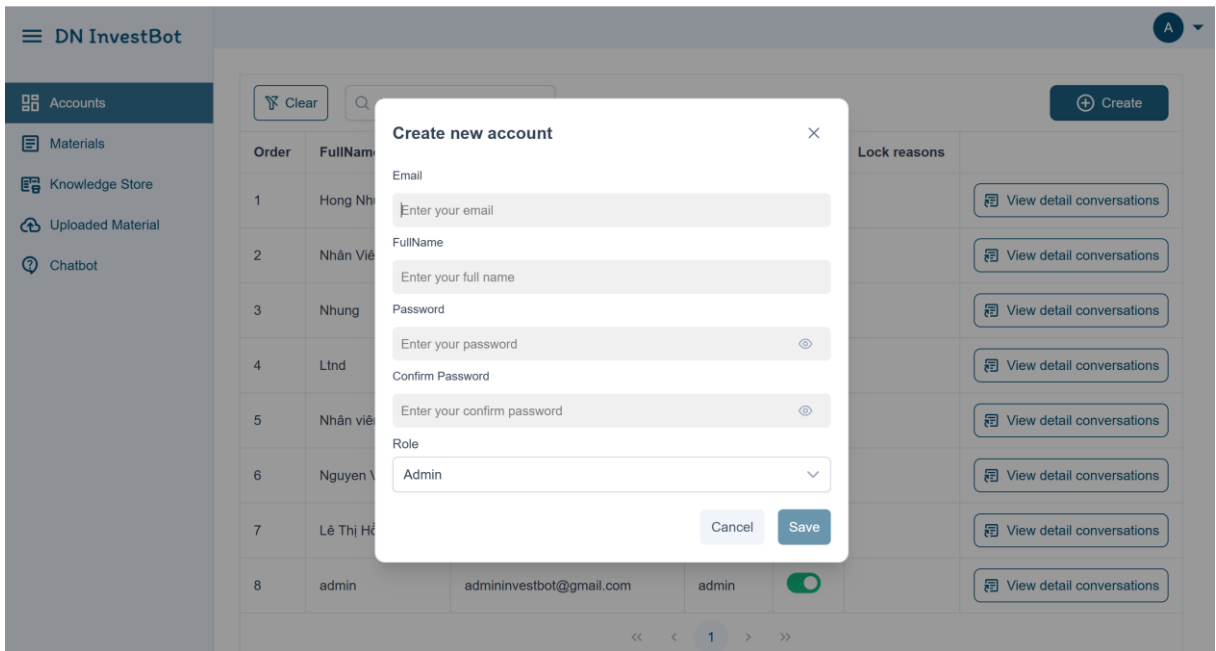
5.1. Kết quả

Dưới đây là hình ảnh minh họa giao diện trang web dành cho Admin và User. Giao diện này được thiết kế để giúp Admin quản lý các thành phần của hệ thống một cách hiệu quả, đồng thời giúp User dễ dàng tương tác và sử dụng chatbot.

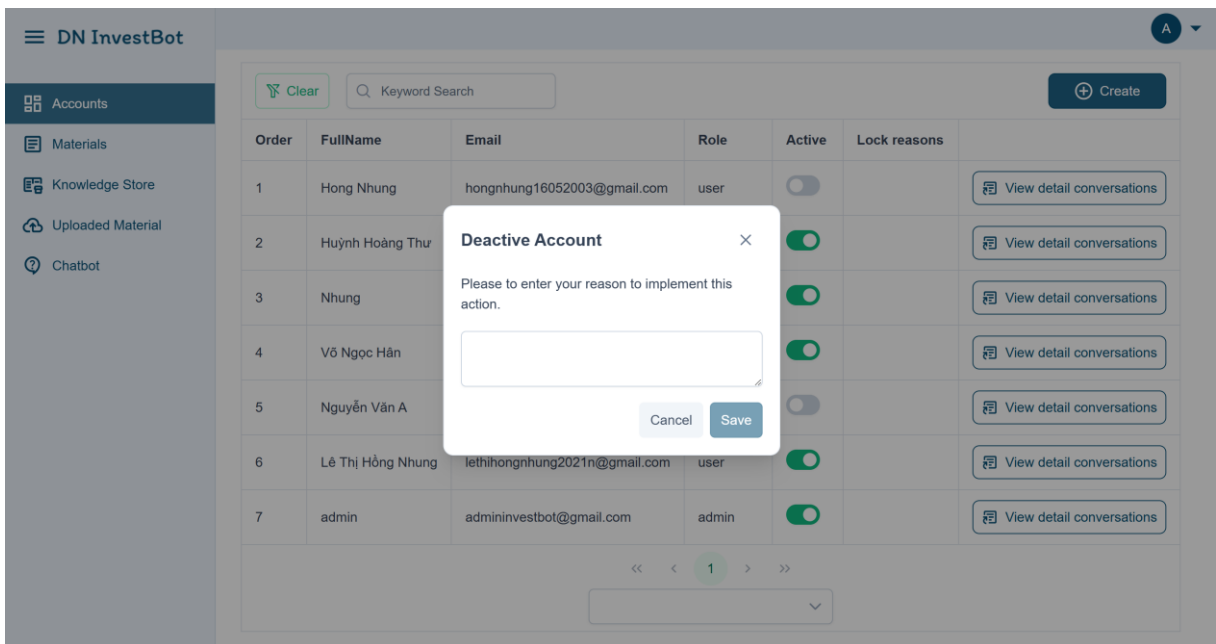
❖ Chức năng quản lý tài khoản của Admin

Order	FullName	Email	Role	Active	Lock reasons	
1	Hong Nhung	hongnhung16052003@gmail.com	user	<input checked="" type="checkbox"/>		View detail conversations
2	Huyñh Hoàng Thù	huyñhhoangthu@gmail.com	employee	<input checked="" type="checkbox"/>		View detail conversations
3	Nhung	lethihongnhung202@gmail.com	user	<input checked="" type="checkbox"/>		View detail conversations
4	Ngọc Như	ngocnhu203@gmail.com	user	<input checked="" type="checkbox"/>		View detail conversations
5	Vô Ngọc Hân	hanvo23223@gmail.com	employee	<input checked="" type="checkbox"/>		View detail conversations
6	Nguyễn Văn A	nguyenvana@gmail.com	user	<input checked="" type="checkbox"/>		View detail conversations
7	Lê Thị Hồng Nhung	lethihongnhung2021n@gmail.com	user	<input checked="" type="checkbox"/>		View detail conversations
8	admin	admininvestbot@gmail.com	admin	<input checked="" type="checkbox"/>		View detail conversations

Hình 5.1: Giao diện chức năng xem danh sách tài khoản



Hình 5.2: Giao diện chức năng tạo mới tài khoản



Hình 5.3: Giao diện chức năng cập nhật trạng thái tài khoản

❖ **Giao diện chức năng quản lý tài liệu của Admin**

Order	Name	Description	Material Type	Updated At	Access	Active	
1	collection_5 8079dc9b71 7da727f335 3c2e6fde41 4	Đà Nẵng đóng vai trò là hạt nhân tăng trưởng cho Vùng Kinh tế trọng...	uri	12:00:12 08-06-2025	public	<input checked="" type="checkbox"/>	Item Delete
2	collection_a 9fd21e4816 2b692ced61 48ec0de187 a	Đà Nẵng có tiềm năng du lịch phong phú nhờ bờ biển đẹp, khí hậu...	uri	12:00:12 08-06-2025	public	<input checked="" type="checkbox"/>	Item Delete
3	collection_6 c2091297f2 25a0bffe0a9 2a7d396267	Trang web cung cấp thông tin về đầu tư vào Đà Nẵng, bao gồm: *...	uri	12:00:12 08-06-2025	public	<input checked="" type="checkbox"/>	Item Delete
4	collection_6 7635829aae 15bebcd8b2 70a6c901d1 3	Văn bản cung cấp thông tin về thủ tục cấp Giấy chứng nhận đăng ký...	uri	12:00:12 08-06-2025	public	<input checked="" type="checkbox"/>	Item Delete

Hình 5.4: Giao diện chức năng xem danh sách tài liệu

Detail Material

Name: Viet-Healthcare 2024.pdf Access: public

Description: Viet-Healthcare 2024.pdf Knowledge store: None

Active: Active Updated At: 00:45:16 14-05-2025

View material:

ĐÀ NẴNG - TRUNG TÂM Y TẾ HÀNG ĐẦU MIỀN TRUNG VIỆT NAM

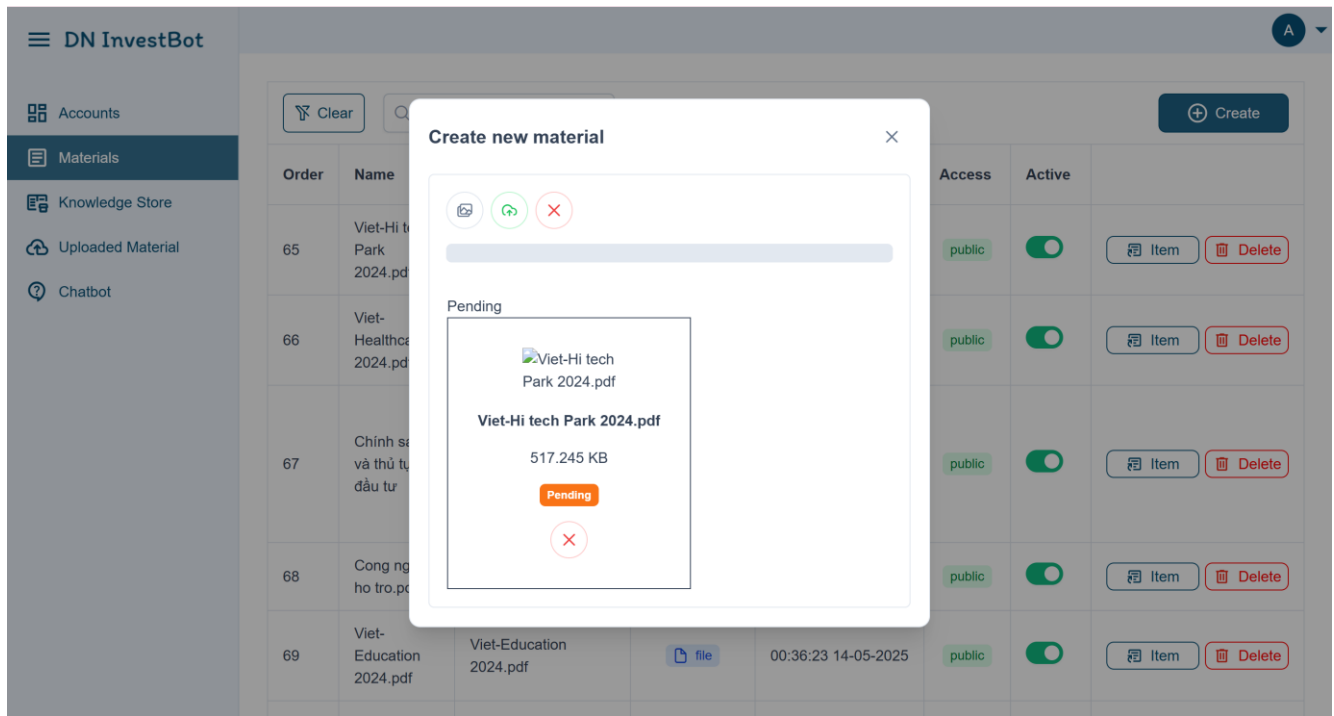
> CƠ SỞ Y TẾ

- 11 Bệnh viện tuyến thành phố
- 04 Bệnh viện trực thuộc Bộ và Trung ương
- 07 Trung tâm y tế quận, huyện
- 07 Bệnh viện tư nhân
- >1.500 Cơ sở y tế tư nhân
- 74 Số giường bệnh trên 10.000 đơn
- 100% Trăm y tế phường xã có bác sĩ chính quy, tăng cường
- 97,71% Tỷ lệ người tham gia bảo hiểm y tế

> LỰC LƯỢNG LAO ĐỘNG NGÀNH Y TẾ

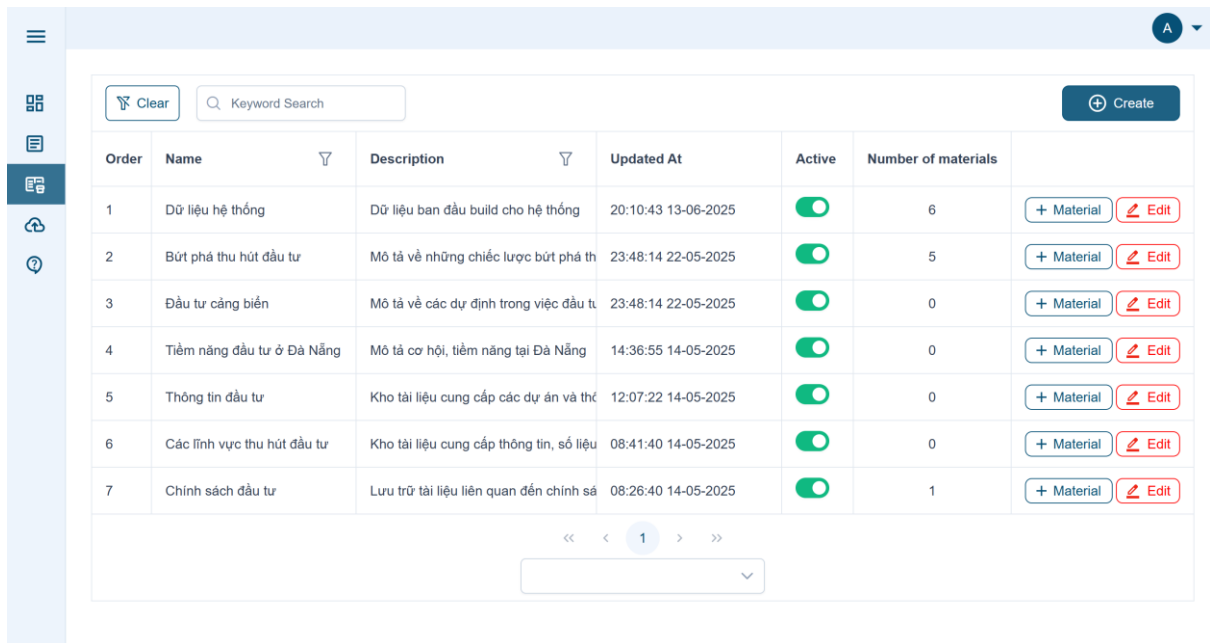
11.833 Nhân viên y tế Theo ngành nghề: 2.615 191

Hình 5.5: Giao diện chức năng xem chi tiết tài liệu

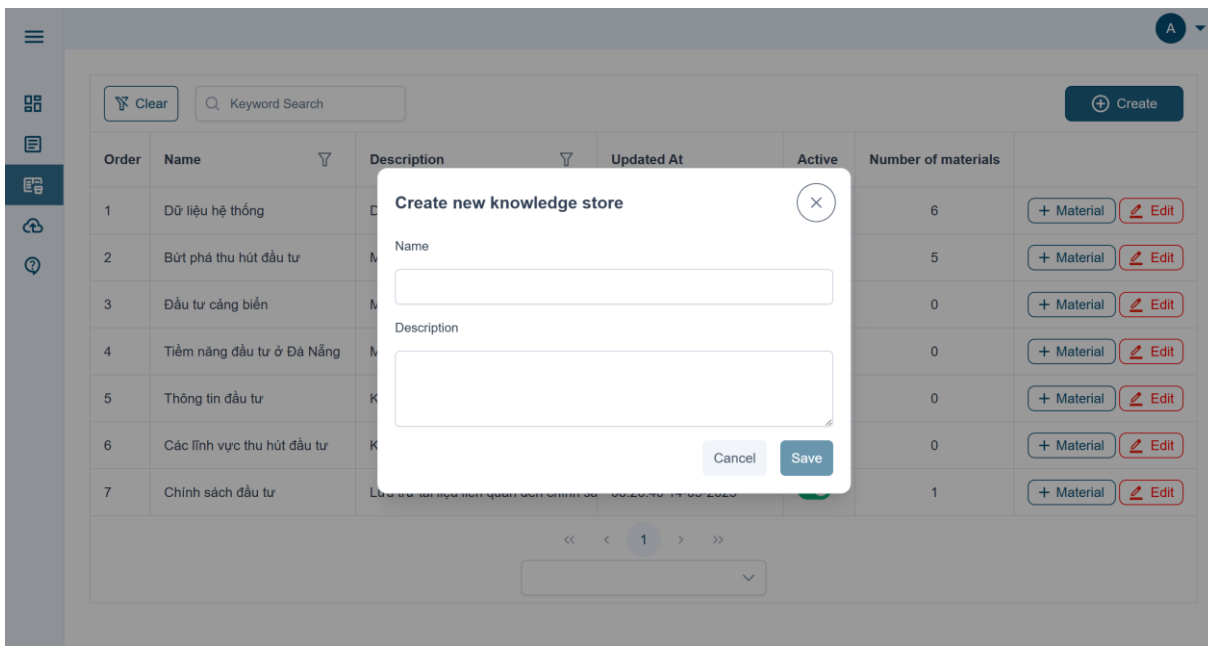


Hình 5.6: Giao diện chức năng thêm mới tài liệu

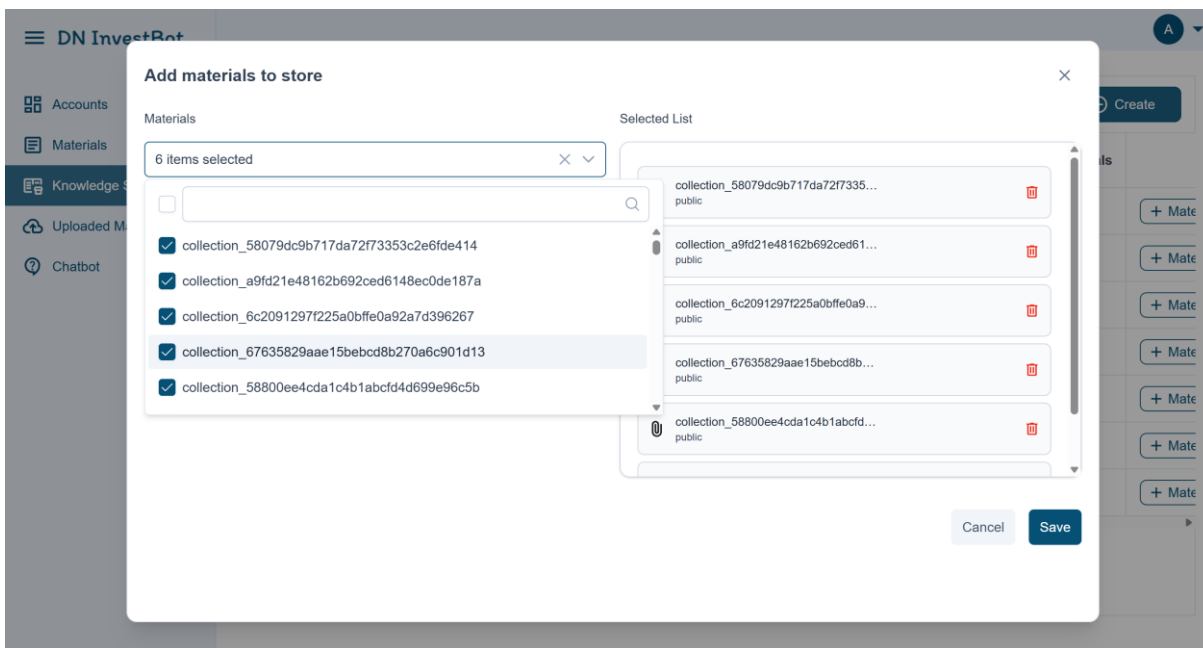
❖ **Giao diện chức năng quản lý kho kiến thức của Admin**



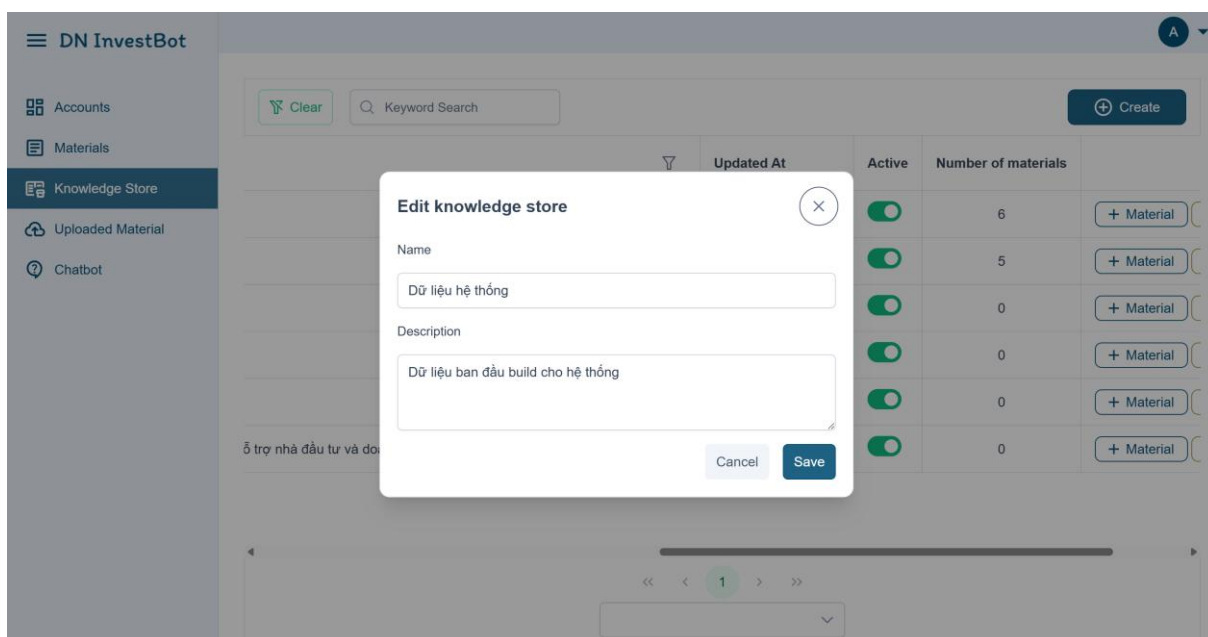
Hình 5.7: Giao diện chức năng xem danh sách kho kiến thức



Hình 5.8: Giao diện chức năng tạo mới kho kiến thức

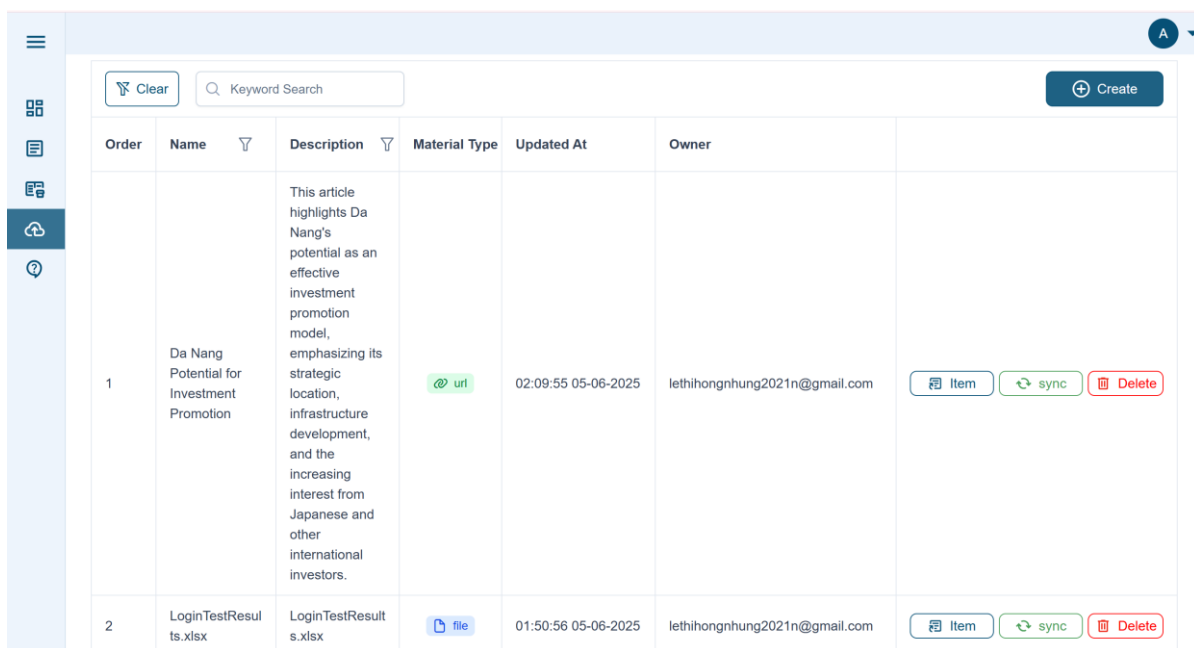


Hình 5.9: Giao diện chức năng thêm tài liệu vào kho

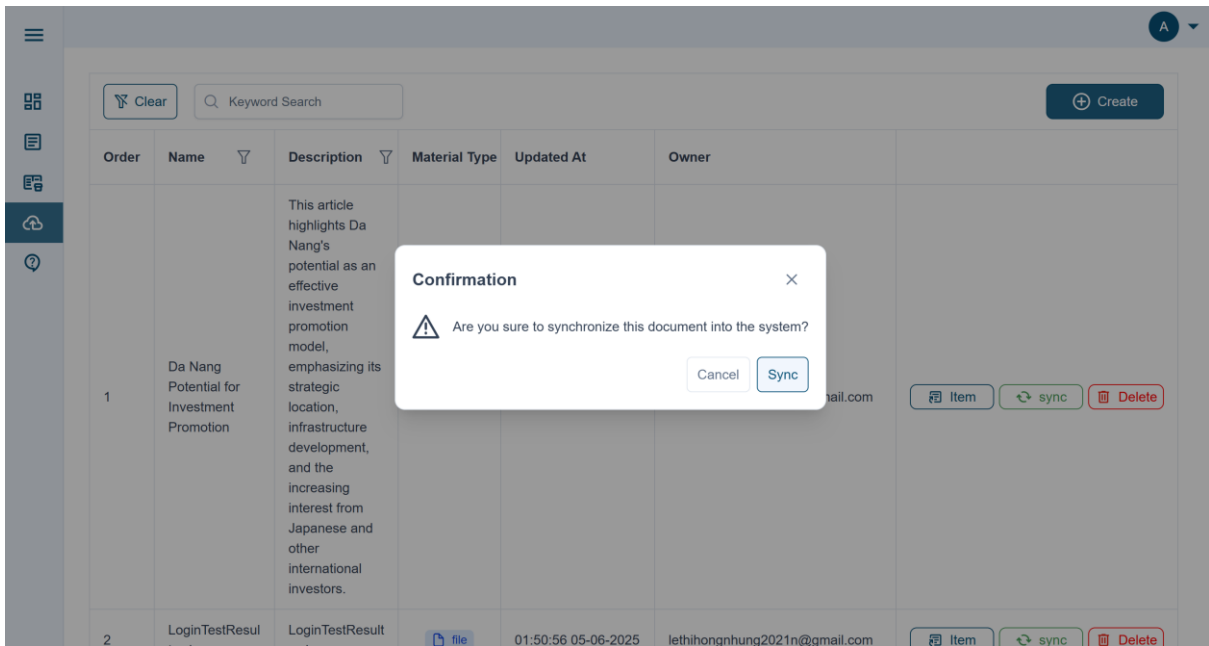


Hình 5.10: Giao diện chức năng chỉnh sửa kho kiến thức

❖ **Giao diện chức năng quản lý tài liệu người dùng của Admin**

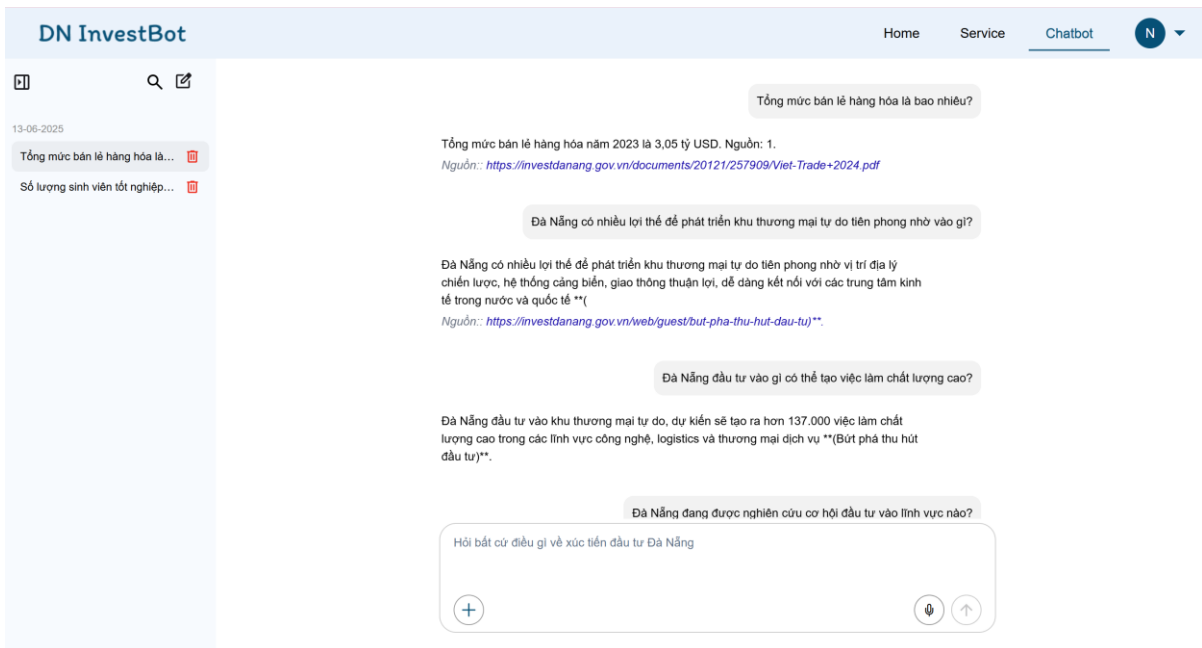


Hình 5.11: Giao diện chức năng xem danh sách tài liệu người dùng

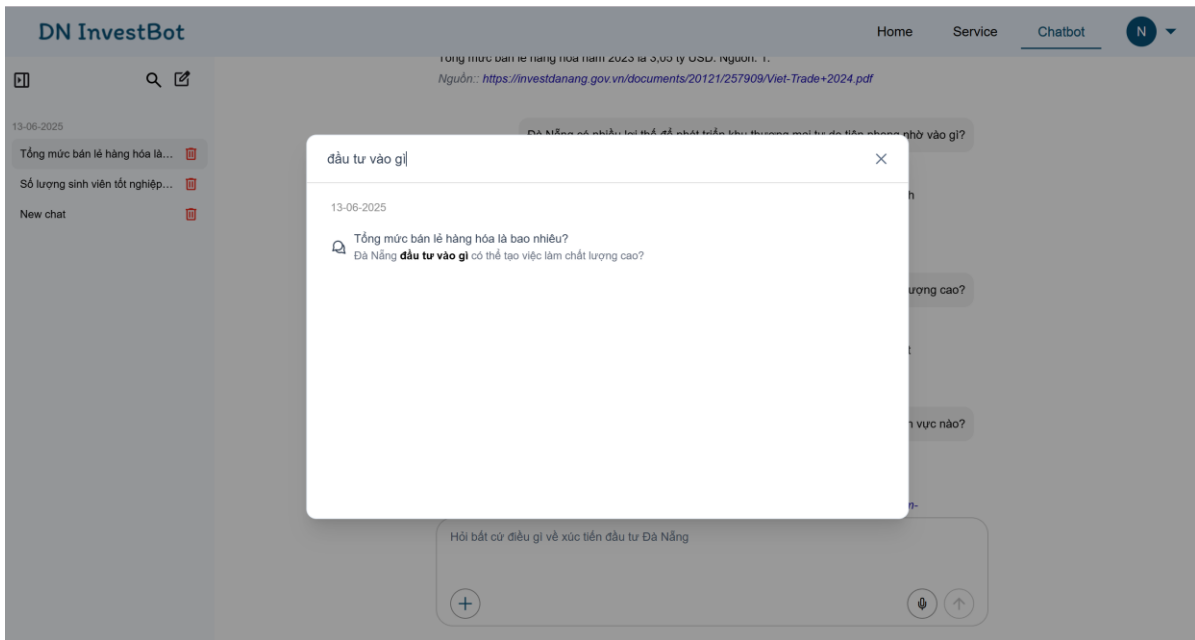


Hình 5.12: Giao diện chức năng đồng bộ tài liệu vào hệ thống

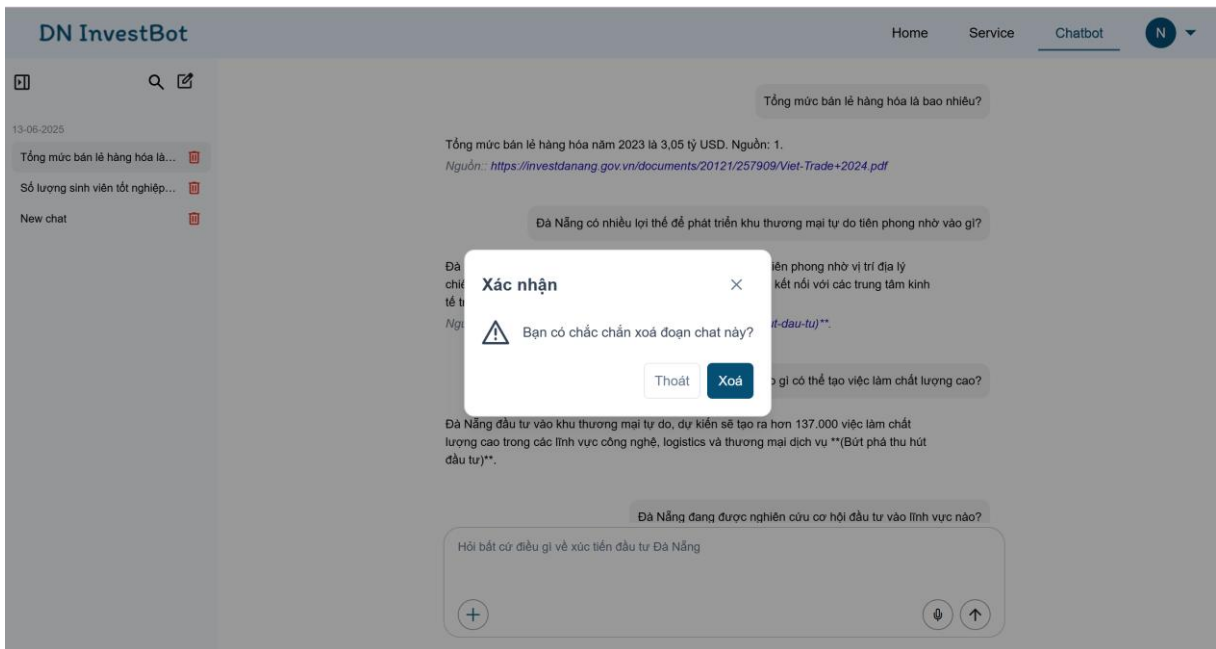
❖ Giao diện chức năng tương tác với chatbot của User



Hình 5.13: Giao diện chức năng xem lịch sử hội thoại ở người dùng



Hình 5.14: Giao diện chức năng tìm kiếm nội dung trong hội thoại ở người dùng



Hình 5.15: Giao diện chức năng xóa cuộc hội thoại ở người dùng

5.2. Đánh giá

Ưu điểm:

- Về giao diện người dùng: Giao diện (UI) được thiết kế thân thiện, bố cục rõ ràng, dễ sử dụng ngay cả với người dùng không chuyên. Trải nghiệm người dùng được đảm bảo ở mức cơ bản.
- Về ứng dụng RAG trong hệ thống: Việc tích hợp mô hình RAG giúp chatbot trả lời chính xác hơn, bám sát ngữ cảnh và nội dung truy vấn. Đặc biệt phù hợp trong các lĩnh vực đòi hỏi thông tin chuyên sâu như xúc tiến đầu tư.
- Về hiệu suất hệ thống: Cơ chế truy vấn vector hoạt động hiệu quả khi lượng truy vấn ở mức trung bình, thời gian phản hồi nhanh và ổn định.

Hạn chế:

- Về dữ liệu: Lượng dữ liệu đầu vào còn hạn chế, chưa phong phú, ảnh hưởng đến khả năng phản hồi trong một số trường hợp.
 - Về xử lý dữ liệu: Hệ thống hiện mới chỉ xử lý được dữ liệu từ các trang đơn lẻ (URL do người dùng cung cấp) và các tệp như .xlsx, .png, .docx, .pdf, chưa hỗ trợ crawl sâu nhiều tầng, xử lý dữ liệu phức tạp hoặc tài liệu với dung lượng lớn.
 - Về khả năng sinh ngôn ngữ: Mô hình đôi khi phản hồi chậm hoặc thiếu chính xác. Hạn chế về số lượng token lưu trữ ngữ cảnh cũng ảnh hưởng đến chất lượng hội thoại dài.
- Tổng kết: Mặc dù hệ thống vẫn còn một số điểm cần cải thiện, nhưng nhìn chung đã đáp ứng được mục tiêu đề ra trong việc hỗ trợ người dùng tra cứu thông tin hiệu quả trong lĩnh vực xúc tiến đầu tư.

5.3. Tổng kết chương

Chương này tổng kết kết quả đạt được và đánh giá hệ thống.

KẾT LUẬN

1. Kết quả đạt được

- ❖ Trong quá trình học tập và thực hiện đồ án tốt nghiệp, em đã có cơ hội tiếp cận thực tiễn, rèn luyện tư duy và ứng dụng nhiều công nghệ hiện đại. Qua quá trình triển khai đề tài, em rút ra một số kết luận sau:
 - Về mặt lý thuyết:
 - + Hiểu rõ quy trình triển khai một hệ thống phần mềm hoàn chỉnh, từ khâu thiết kế, xây dựng đến kiểm thử và đánh giá.
 - + Nắm bắt và vận dụng được các công nghệ mới vào thực tiễn, đặc biệt là mô hình RAG, ngôn ngữ lập trình, framework và công cụ hỗ trợ phát triển hệ thống.
 - + Nâng cao kiến thức thực tiễn trong việc kết hợp AI vào hệ thống hỗ trợ người dùng.
 - Về kỹ năng cá nhân:
 - + Rèn luyện tư duy giải quyết vấn đề, khả năng tự học và nghiên cứu độc lập.
 - + Nâng cao kỹ năng lập kế hoạch, phân chia và quản lý công việc theo từng giai đoạn cụ thể.
 - + Cải thiện khả năng tổ chức công việc, trình bày, báo cáo khi làm việc độc lập
 - Về mặt thực tiễn ứng dụng, đồ án đã đạt được:
 - + Hệ thống có khả năng tương tác hiệu quả với các chức năng cơ bản, giao diện thân thiện, dễ sử dụng với người dùng.
 - + Việc tích hợp mô hình RAG giúp tăng độ chính xác và tính liên quan của câu trả lời, đặc biệt hữu ích với kho dữ liệu lớn và đa dạng
 - + Hệ thống hoạt động ổn định trong môi trường có lượng truy vấn vừa phải, tuy nhiên, cần tối ưu thêm để mở rộng quy mô trong tương lai.
 - + Việc kết hợp nhiều công nghệ hiện đại cho thấy tiềm năng ứng dụng mạnh mẽ của chatbot trong việc hỗ trợ nhà đầu tư tiếp cận thông tin nhanh chóng, tin cậy và hiệu quả.

- ❖ Đề tài không chỉ mang lại nhiều kiến thức, kỹ năng quan trọng cho bản thân mà còn góp phần vào việc ứng dụng công nghệ AI trong thực tiễn, đặc biệt là trong lĩnh vực hỗ trợ xúc tiến đầu tư tại Đà Nẵng.

2. Hướng phát triển

- ❖ Mở rộng cơ sở dữ liệu ngữ cảnh: Tích hợp thêm thêm các nguồn dữ liệu đáng tin cậy như văn bản pháp luật, báo cáo thị trường, quy hoạch, chính sách thu hút đầu tư,...
- ❖ Nâng cao khả năng xử lý dữ liệu: Nâng cao xử lý dữ liệu: Phát triển công cụ crawl sâu nhiều tầng, kết hợp thêm các công cụ xử lý tài liệu lớn như LangChain Document Loaders. Chia nhỏ tài liệu theo nội dung logic thay vì chỉ theo số kí tự để tối ưu ngữ cảnh
- ❖ Tối ưu hiệu suất tư vấn RAG: Cải thiện tốc độ và độ chính xác của truy xuất vector bằng cách tối ưu thuật toán tìm kiếm hoặc thử nghiệm các công cụ khác ngoài Qdrant như FAISS, Weaviate,...
- ❖ Tùy biến theo đối tượng người dùng:
 - Phát triển chatbot phân loại người dùng (nhà đầu tư nước ngoài, trong nước, doanh nghiệp nhỏ, startup,...) để đưa ra câu trả lời phù hợp hơn.
 - Tích hợp đa ngôn ngữ (Tiếng Việt – Anh – Trung – Nhật...) để phục vụ cho nhiều nhà đầu tư quốc tế.
- ❖ Đánh giá chất lượng và cải tiến liên tục:
 - Xây dựng cơ chế thu thập phản hồi người dùng để cải tiến nội dung và độ chính xác của chatbot.
 - Sử dụng các chỉ số như độ chính xác, thời gian phản hồi, tỷ lệ hài lòng để đo lường hiệu quả.

TÀI LIỆU THAM KHẢO

[1] NestJS, *A Progressive Node.js Framework for Building Efficient, Reliable and Scalable Server-side Applications*. [Trực tuyến]. Link: <https://docs.nestjs.com>

[2] Nuxt, *Getting Started - Introduction*. [Trực tuyến]. Link: <https://nuxt.com/docs/getting-started/introduction>

[3] Django Software Foundation, *Django 5.2 Documentation*. [Trực tuyến]. Link: <https://docs.djangoproject.com/en/5.2/>

[4] Qdrant, *Qdrant Documentation*. [Trực tuyến]. Link: <https://qdrant.tech/documentation>

[5] Amazon Web Services, *What is Retrieval-Augmented Generation?*. [Trực tuyến]. Link: <https://aws.amazon.com/vi/what-is/retrieval-augmented-generation>

[6] lakeFS, *What is RAG pipeline?*. [Trực tuyến]. Link: <https://lakefs.io/blog/what-is-rag-pipeline>

PHỤ LỤC