

Software Requirements Specification (SRS)

Enterprise SOP Question Answering System (RAG-Based)

1. Introduction

1.1 Purpose

This document defines the functional and non-functional requirements for the **Enterprise SOP Question Answering System**, an AI-powered internal knowledge assistant designed to answer employee queries strictly based on company Standard Operating Procedures (SOPs).

The system leverages **Retrieval-Augmented Generation (RAG)** to ensure responses are accurate, context-aware, and free from hallucinations.

1.2 Scope

The system enables organizations to:

- Upload and manage SOP documents (Admin-only access)
- Convert SOPs into searchable vector embeddings
- Allow authenticated users to ask questions related to SOPs
- Provide answers strictly grounded in official SOP content
- Explicitly refuse out-of-scope or unsupported questions

This application is intended for **enterprise internal usage**, such as HR, Operations, Compliance, and Policy Management.

1.3 Definitions, Acronyms, and Abbreviations

Term	Description
SOP	Standard Operating Procedure
RAG	Retrieval-Augmented Generation
LLM	Large Language Model
Admin	Authorized user responsible for SOP management

Term	Description
User	Employee querying SOP information
Vector Search	Semantic similarity search using embeddings

2. Overall Description

2.1 Product Perspective

The Enterprise SOP Q&A System is a standalone internal web application built using the following technologies:

- **Backend:** Node.js, Express.js
- **Database:** MongoDB Atlas with Vector Search
- **AI Model:** Ollama (Local LLM inference)
- **Frontend:** HTML, CSS, JavaScript
- **Authentication:** Session-based authentication

The system follows a **role-based access control model** distinguishing Admin and User responsibilities.

2.2 Product Functions

Admin Functions

- Secure login
- Upload SOP PDF documents
- Trigger embedding generation
- Manage SOP lifecycle
- Logout

User Functions

- Secure login
- Submit SOP-related questions
- Receive context-aware answers
- Get explicit refusal for out-of-scope queries

- Logout
-

2.3 User Classes and Characteristics

User Type Description

Admin HR / Operations personnel managing SOPs

User Employees accessing SOP information

Admins have elevated privileges; Users have read-only access.

2.4 Operating Environment

- Operating System: Windows / macOS / Linux
 - Browser: Chrome, Edge, Firefox
 - Backend Runtime: Node.js (v18+)
 - Database: MongoDB Atlas
 - AI Runtime: Ollama Local Server
-

2.5 Design and Implementation Constraints

- Answers must be generated **only** from uploaded SOP documents
 - No cloud-based LLM APIs are used
 - All AI inference is performed locally using Ollama
 - MongoDB vector dimensions must match embedding model dimensions
 - Authentication is mandatory for protected routes
-

2.6 Assumptions and Dependencies

- Ollama server is running locally
- SOP documents are provided in PDF format
- Admin uploads SOPs before user queries
- MongoDB Atlas Vector Search index is configured correctly

3. System Features and Requirements

3.1 Authentication System

Description

The system authenticates users and enforces role-based access control.

Functional Requirements

- Admin and User login
- Session-based authentication
- Role validation for protected routes
- Logout functionality

Non-Functional Requirements

- Secure session handling
 - Prevention of unauthorized access
-

3.2 SOP Upload and Processing (Admin)

Description

Admins upload SOP PDF files that are processed and stored for querying.

Functional Requirements

- Accept PDF uploads
- Parse text content from PDFs
- Chunk SOP content into manageable segments
- Store chunks in the database

Constraints

- Only Admin users are permitted to upload SOPs
-

3.3 Embedding Generation

Description

The system generates vector embeddings for SOP chunks to enable semantic search.

Functional Requirements

- Generate embeddings using a local embedding model
 - Store embeddings in MongoDB
 - Ensure vector dimensional consistency
-

3.4 SOP Question Answering (User)

Description

Users submit natural language queries related to SOP content.

Functional Requirements

- Accept user questions
- Perform vector similarity search
- Retrieve relevant SOP context
- Generate grounded responses
- Display answers in paragraph form

Out-of-Scope Handling

If no relevant SOP context is found, the system must respond with:

“I don’t know based on the provided SOP.”

3.5 Hallucination Prevention

Description

The system prevents speculative or fabricated responses.

Functional Requirements

- Relevance threshold filtering
 - Context-only answer generation
 - Explicit refusal for unrelated queries (e.g., “Who is the CEO of Google?”)
-

4. External Interface Requirements

4.1 User Interface

Admin Interface

- SOP upload control
- Embedding generation button
- Logout button

User Interface

- Question input textarea
 - Answer display area
 - Logout button
-

4.2 Software Interfaces

- MongoDB Atlas Vector Search API
 - Ollama REST API (/api/generate)
 - Express session middleware
-

5. Non-Functional Requirements

5.1 Performance

- Initial query may have higher latency due to LLM warm-up
 - Subsequent queries respond faster
 - Designed for internal enterprise-scale usage
-

5.2 Security

- Role-based access control
 - Session-based authentication
 - Admin-only SOP management
 - No external data leakage
-

5.3 Reliability

- Graceful handling of missing SOP context
 - Clear error messaging when services are unavailable
-

5.4 Maintainability

- Modular backend architecture
 - Separation of frontend, backend, and database layers
 - Easy SOP updates without code changes
-

6. Future Enhancements

- Admin analytics dashboard
 - SOP versioning
 - User query audit logs
 - React / Next.js UI
 - Cloud deployment support
-

7. Conclusion

The Enterprise SOP Question Answering System provides a secure, reliable, and scalable internal knowledge assistant using Retrieval-Augmented Generation. It enhances employee access to official SOPs while ensuring compliance, accuracy, and hallucination prevention.

Document Status

Version: 1.0

Prepared For: Enterprise Internship Project

Status: Final
