**Project Documentation: Spring Boot REST API for Document Upload & QA Search**  
  
**Project Overview**

**Project Name:**Document Management and QA System  
**Technology Stack:**Spring Boot, Java, JPA (Hibernate), Apache Tika, REST API  
**Purpose:**  
This project provides REST APIs to:

* Upload multiple documents (PDF, DOCX, TXT, etc.)
* Extract and save text content using Apache Tika
* Perform question-based search across uploaded documents
* Filter documents based on author or metadata  
    
  **Features**
* **Upload Documents**  
  Endpoint to upload multiple files with author metadata.
* **Search by Question/Keyword**  
  Search documents by matching keyword in content and return snippets.
* **Filter by Author**  
  Retrieve paginated and sorted document list based on metadata (e.g., author).

**Project Structure**

src/

├── controller/ // REST Controllers

├──dto/ // Data Transfer Objects

├── entities/ // JPA Entities

├── repository/ // JPA Repositories

├── service/ // Service Interfaces & Implementations

├──DemospringrestapiApplication.java // Main Class  
  
**How to Run the Project on Another System (via ZIP or Manual)**

**Requirements**

* Java 11 or above
* Maven installed
* (Optional) PostgreSQL/MySQL if DB config is set for them; default is H2 for dev/testing  
    
  **Option 1: Deploy via ZIP**
* **Step 1: Export/Package the Project**
* Zip the complete project folder.
* Ensure pom.xml is present at the root.
* **Step 2: Copy ZIP to Target System**
* Extract the ZIP on the new system.
* Open the project using IntelliJ, Eclipse, or VS Code (with Java support).
* **Step 3: Run the Application**
* Open terminal inside the project folder and run:

./mvnwspring-boot:run

#### **Install ffmpeg (optional)**

**Windows:**

* Download from: <https://ffmpeg.org/download.html>
* Add to system PATH variable

Or build and run:  
./mvnw clean install

java -jar target/demospringrestapi-0.0.1-SNAPSHOT.jar  
  
**API Endpoints Summary**

| **Method** | **Endpoint** | **Description** |
| --- | --- | --- |
| POST | /api/documents/savedocuments | Upload multiple documents with author |
| GET | /api/documents/qa?question=... | Search content by keyword |
| GET | /api/documents/filter?author=...&page=0&size=10&sortBy=uploadDate | Filter and paginate documents |

**Testing the API**

Use Postman or CURL to test endpoints:

curl -X POST http://localhost:8080/api/documents/savedocuments \

-F "file=@sample.pdf" \

-F "author=John Doe"  
  
**Tools/Libraries Used**

* **Spring Boot** – Application framework
* **Apache Tika** – Text extraction from files
* **Spring Data JPA** – ORM layer for persistence
* **H2** – In-memory database for development