# PDF Data Inserter and Retriever Project Documentation

## 1. Introduction

This Java-based utility is designed to read data from PDF files, insert the extracted data into a MySQL database, and subsequently retrieve and store this data in an ArrayList. It's suitable for automating data entry tasks from PDF documents into a structured database format.

## 2. System Requirements

• Java JDK 8 or higher  
• MySQL Database Server  
• Apache Maven (for dependency management)

## 3. Dependencies

The project relies on the following primary dependencies:  
• Apache PDFBox 2.0.24: A library for processing PDF documents.  
• MySQL Connector/J 8.0.23: A JDBC driver for MySQL.

## 4. Setup and Installation

To set up the project, follow these steps:  
1. Database Setup: Create a MySQL database and a table for storing PDF data. Ensure the table schema matches the data format.  
2. Application Configuration: Configure the database connection settings in the utility. Set the file path to the PDF file to be processed.  
3. Build the Project: Use Maven to build the project by running 'mvn clean install'.

## 5. Usage

Run the main class, 'PdfDataInserter', to start the utility. It will process the specified PDF file, insert the data into the MySQL database, and then retrieve and store the data in an ArrayList.

## 6. Code Structure

The main class is 'PdfDataInserter'. It includes methods for PDF processing, database operations, and data retrieval.

## 7. Error Handling

Basic error handling is implemented for file I/O and database operations. For a production environment, it's recommended to enhance the error handling and include detailed logging.

## 8. Security Considerations

It's important to secure database credentials and avoid hardcoding sensitive information. Use environment variables or a secure configuration file for managing these credentials.

## 9. Future Enhancements

Future improvements can include a user interface for file selection, batch processing capabilities, and advanced error handling and logging.