

Task - Invoice OCR Extraction Mini App

Objective

Build a simple web application that allows users to upload invoice files (PDF or Image), automatically extract the invoice data using an open-source OCR engine, and display the extracted fields in an editable form before final submission.

Scope of the Task

You are required to develop a minimal end-to-end solution that includes:

1.1. Backend:

- Basic API or service that handles file upload (PDF/Image).
- Integrates with an **open-source OCR engine** (e.g., Tesseract OCR or similar).
- Extracts key invoice data from the file:
 - Invoice Number
 - Invoice Date
 - Customer Name
 - Total Amount
 - VAT
 - List of Items (description, quantity, unit price, line total)

1.2. Frontend:

- Upload screen to select invoice files (image/PDF).
- A dynamic form that is auto-filled with the extracted data after OCR is processed.
- The user should be able to **edit** the values before submitting.
- A **Submit** button to save the invoice.

1.3. Database Schema:

Use a basic schema with at least two tables:

- **Invoice:** General invoice data
- **InvoiceDetails:** Line items per invoice

Technical Requirements:

- **OCR Engine:** Must use open source (e.g., Tesseract, EasyOCR, or similar). No paid APIs are allowed.
- **Backend Language/Framework:** You must use .NET 9.
- **Frontend:** Angular 20.
- **Database:** SQLite / MySQL / PostgreSQL.
- **PDF/Image Handling:** System should support PDF and image formats (JPG/PNG).

Deliverables:

1. Source code uploaded to a **GitHub public or private repository**
2. Basic instructions to run the app locally (README)
3. Screenshot or short video showing how it works
4. SQL script to generate the DB schema

Evaluation Criteria:

- Correctness and completeness of the OCR pipeline
- Code structure and readability
- Simplicity of the UI
- Proper extraction and mapping of fields
- Functionality of the editable form and saving process
- Clean Git commit history and documentation

Timeframe:

Expected duration: 5 to 7 days max



نخطو معًا للمستقبل
experience next

Thank You

ejada Systems Company Limited شركة إجادة للنظم المحدودة

www.ejada.com | info@ejada.com