1. Install Poetry And set it up to path.
2. Initialize Peotry Project in powershell (“poetry init --name "doc-automation" --description "AI Document Processor Microservice" --python "^3.10" --no-interaction”)
3. Add Dependencies (“poetry add fastapi uvicorn python-multipart”)
4. Install OCR and ML tools (“poetry add google-cloud-documentai azure-ai-documentintelligence scikit-learn pymupdf pdfplumber pillow streamlit”)
5. Create Folders (“mkdir src

cd src

mkdir api, core, models, services

cd services

mkdir clarity\_check, classification, processors, ocr, mapping, validation, hitl, storage, splitting

cd ../..”)

1. Run this to create empty \_\_init\_\_.py files so Python treats folders as packages(“Get-ChildItem -Recurse -Directory | ForEach-Object {

New-Item -Path $\_.FullName -Name "\_\_init\_\_.py" -ItemType File -Force

}”)

1. Now create the FastAPI entrypoint file (“New-Item -Path .\src\main.py -ItemType "File"”)
2. Create routes.py in src/v1 (“mkdir src/api/v1 New-Item -Path src/api/v1/routes.py -ItemType "File"”)