
Coding Challenge

You are working as an AI Engineer at a company that is building a case management system. Your task is to automate the process of creating a demand letter for a client.

Task: Implement a console application using Python that takes a folder name containing medical records in PDF format as input. Based on these medical records, generate a section of the demand letter document called "Summary of Injuries."

Summary of Injuries Format: The summary should be presented in a table with the following columns:

- Date of Visit
- Diagnosis
- ICD Code
- Reference

The records should be listed in reverse chronological order, with the most recent visit displayed at the top.

Column Descriptions:

- **Date of Visit:** Extracted from the document. Can be in different format in each document
- **Diagnosis:** Extracted from the document.
- **ICD Code:** The ICD-10 code for the injury, which is a system used by physicians to classify and code all diagnoses, symptoms, and procedures for claims processing. The ICD-10 code can be found online. The information about the injury should be extracted from the document, and it is up to you to determine the ICD-10 code.
- **Reference:** A reference to the specific document and page containing the reference to diagnosis (e.g., "John Doe - MedicalVisit.pdf - p. 2").

Input and Output:

- The tool should accept two parameters: the input folder and the output folder.
- The input folder contains the medical visit documents.
- The output folder should contain a file named "summary_of_injuries.pdf" (other text formats are acceptable).

Script Naming: Name your script: generate_summary_of_injuries.py

Example of Launching the Script:

```
python generate_summary_of_injuries.py /folder/with/documents /outputfolder
```

Requirements:

- The code should be of production quality, including all the artifacts that are typically produced when writing code in a corporate environment.
- Do not use regex to extract data from the medical records – the format of the documents is different between providers. Use AI techniques to analyze the documents and extract the information.

Additional Information:

- For your convenience, we are providing sample documents that you can use for analysis.

Due to the high number of applications, we will not be able to respond with further clarifications. Only selected candidates will be contacted.