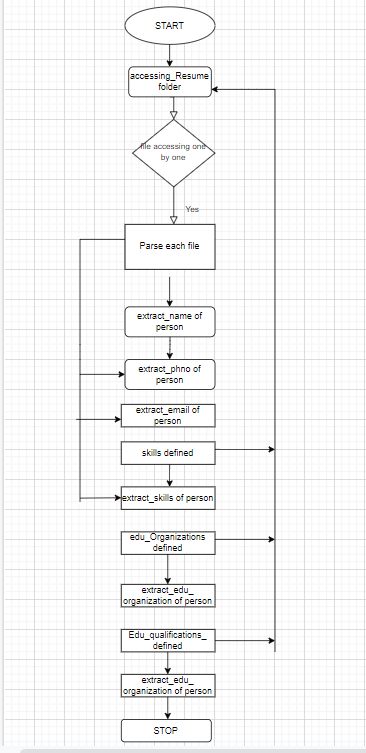
**Data Extraction from Resume**

1. Proposed Solution:

Process/Data Flow diagram



1. Stages of Implementation:

Approach:

Note: I have tried with .pdf and .doc and .docx RESUME formats separately to parse; but I have used

Tika library.

A Python port of the Apache Tika library, according to the documentation Apache tika supports text extraction from over 1500 file formats. And need not to import dependencies individually.

**Challenges faced, and resolution**: for different formats accessing we should check their formats and dependencies should import it and it’s some difficult to do.

1. Implementation Requirements  :

Hardware/Software requirements: Python 3, Spacy(NLP),tika

1. Limitations:

We should take skill set of updates data, still work on it to do further

In resumes years’ experience is key feature with relevant skills should work on it further.

1. Result/Impact:

Accuracy: As per my code I got results as per my requirements.(100%)

Benefits of the model: Yes its best to parse bunch of resumes at a time to get immediate summary of relevant person.

1. Prototype

Code given in folder