

@Documentation

@Author: Aman K Mahato

@Date & Time: 08/02/2017, 7:40 AM EST

This a documentation for the Java App which parse the PDF files of the datas and creates the Excel Files which can be fed to the backend to ingest the medical insurance plan data.

Tools Required:

Java JDK 8

Maven 3 or higher

Running this App from IntelliJ Idea:

1. Download the entire project and load it as a Maven Project in your IDE. To download from command line

```
git clone git@github.com:AmanMahato/BenefixProject.git
```

2. Put All the PDF files which needs to be parsed in the Resources Folder. It is located inside src folder.

3. Compile the Entire Project. That can be done either from Maven or Manually from Terminal.

4. Run App.java. It will create an Excel file named "BenefixResult.xlsx" in the root directory. This is a file which has all the datas parsed from the pdf files.

Dependencies:

The project depends on two Major Dependencies viz, itextpdf and apache Poi. These are some of the external libraries which I have used to parse pdf.

Maven Dependencies:

```
<dependency>
  <groupId>com.itextpdf</groupId>
  <artifactId>itextpdf</artifactId>
  <version>5.5.6</version>
</dependency>
```

```
<!-- https://mvnrepository.com/artifact/org.apache.poi/poi-ooxml -->
<dependency>
  <groupId>org.apache.poi</groupId>
  <artifactId>poi-ooxml</artifactId>
  <version>3.9</version>
</dependency>
```

All these jars are downloaded by Maven if we inject the above mentioned dependencies.

Limitations:

1. This App will only work if there is no image in the pdf copied to the Resources folder.
2. The entire app is format specific and should not be used for pdfs in format other than the one given to work on.