# GLOSSARY

## Glossary creation using NLP

*"Glossary is an important part of any book; it makes a book complete."*

## INTRODUCTION:

A glossary, also known as vocabulary, is a list of alphabetically well-arranged words. A glossary is also known as a **lexicon, wordbook.** Glossary plays an important role as a tool to learn new words.
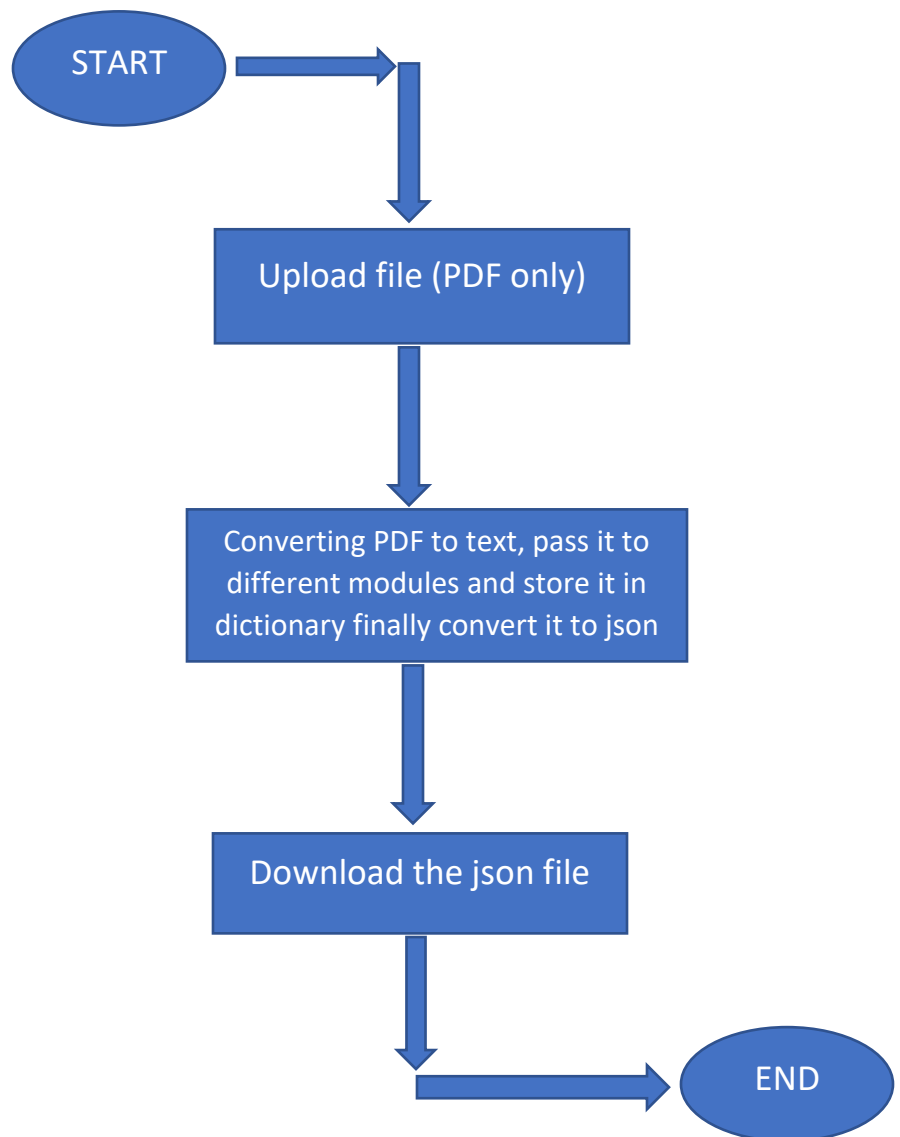
This project discusses the result of the work done in development of **GLOSSARY CREATION**. In this project we demonstrate a prototype solution in which we built a dedicated platform for students, or high-tech readers. The aim is to enable them to reach their requirements using a convert feature.

## Importance of our project:

In this current digital scenario, any delay in information is not acceptable. Moreover, digitalization demands less paperwork. Thus, we feel the requirement of this project which is any time accessible platform by anyone. This digitalization at this small level will be very useful for the students and high-tech readers.

- Find a glossary of any Book in Json format.
- We can find synonyms of words which is not easy to understand.
- Accuracy of our project is approximately 85%.

# BLOCK DIAGRAM:

```
        START  ──────────┐
                         │
                         ▼
              ┌────────────────────────┐
              │  Upload file (PDF only) │
              └────────────────────────┘
                         │
                         ▼
              ┌──────────────────────────────────┐
              │ Converting PDF to text, pass it to│
              │ different modules and store it in │
              │ dictionary finally convert it to  │
              │ json                              │
              └──────────────────────────────────┘
                         │
                         ▼
              ┌────────────────────────┐
              │ Download the json file  │
              └────────────────────────┘
                         │
                         ▼
                      ─────────►  END
```

1. **Start: -** *In this project first we run the program.*
2. **Validation: -** *Here we are taking login info from user if there is not account than it should be create account than user can use.*
3. **Upload File: -** *Here we will upload a file i.e. Book in pdf format which glossary we want to create.*
4. **Process: -** *Converting PDF to text, pass it to different modules and store it in dictionary finally convert it to json.*
5. **Download: -** *Here we can download a file which has converted into glossary in json format.*

# PYTHON (Version 3.7.3)

## LIBRARIES USED AND THEIR VERSION:

- **NumPy –** (numpy 1.17.3)
- **PdfMiner –** (pdfminer.six 20191020)
- **nltk –** (nltk 3.4.5)
- **Re –** (regex 2019.11.1)
- **Json –** (jsonlib 1.6.1)
- **Flask –** (1.1.1)