**Longhealth Approaches**

**PDF to Text**

**Method 1:**

Package: PyPDF2

Status: Working

Process: Converting pdf into text

**Method 2:**

Package: pdfminer

Status: Working

Process: Converting pdf into a text using TextConverter method.

**Method 3:**

Package: pdfx

Status: Working

Process: Converting pdf into text using PDFx method

**PDF to XML**

**Method 1:**

Package: pdfminer

Status: Working not accurate

Process: Converting pdf to xml using XMLConverter

Reason: conversion is taking place character wise.

**Method 2:**

Package: Requests

Status: working

Process: Converting pdf to xml using URL “https://api.pdf.co/v1/pdf/convert/to/xml”

**PDF to Json**

**Method 1:**

Package: PyPDF2

Status: Not Working

Process: Converting pdf to json is not accurate and not relevant.

**Method 2:**

Package: Requests

Status: Working

Process: Converting pdf to json using URL “https://api.pdf.co/v1/pdf/convert/to/json2”