MULTIPLE
FILE CONVERSIONS

PRESENTATION OF PROJECT WORK
By:
Ansh Sachdeva(41325502716)
Nitin Kumar(41725502716)
Kunal Wadhawan(0122502716)

Under The Guidance of :
PANKAJ SINGH YADAV
Assistant Professor
,Dept. of CSE, JEMTEC, Gr. Noida, GGSIP University

Outline.

- **Abstract**
- **Introduction**
  - Objective
  - Scope
- **Theoretical Background**
  - Existing System
  - Proposed System

- **Problem Definition**
  - Project Mission
  - Target Users

- **Feasibility Study**
  - Economic Feasibility
  - Technical Feasibility

Outline.

- Operational Feasibility
- **Problem Definition**
  - Problem Statement
  - Target Users
- **Designing and Implementation**
  - System Design
  - Data Flow Diagram
  - Use Case Diagram
- **Miscallanous**
  - Source Code and Screenshots
  - Task Log

- **Conclusion and Future Work**
- **References**

❖ **File conversion** is the conversion of computer data from one format to another.

❖ A conversion can be as simple as converting a text file from one character encoding system to another *(e.g. converting a markdown file to html or converting a colored image to grayscale)*. It can also be as complex as conversion between different image formats, which requires deeper understanding of compression and encoding algorithms**.**

❖ In this project, we have proposed and implemented a platform (*aio.io*) for converting files from one format to another **without the use of any server.**

❖ Files are converted into different formats using various conversion algorithms.

A Computer File is a blob of information stored on a memory device. These files are used for recording data discretely and are managed and organized by a File system.

Data conversion is the process of conversion of computer data from one format to another.
We often need to convert a file to a different format based on our needs
But there are many challenges and principles involved in the conversion of file from one format to another:

❖ **Simple or Complex**: A conversion can be as simple as converting a text file from one character encoding system to another(for eg converting a markdown file to html or converting a colored image to grayscale). It can also be as complex as conversion between different image formats, which requires deeper understanding of compression and encoding algorithms.

❖ **Seamless or processing involvement** :A conversion can be as seamless as upgrading a file to a newer version of a computer program(like converting a wordstar 2 file (ws2) to wordstar 3 (ws3)). Alternatively, a conversion can require processing by the use of a special conversion program.

❖ **Single or multiple stages**: A conversion could happen in a single stage or it may involve a complex process of going through intermediary stages.

❖ **Export/Import procedures:** A conversion may involve complex "exporting" and "importing" procedures.

❖ The list continues.

In this project , we researched on various concepts involved and approaches used in converting a file from one format to another, their advantages and drawbacks, feasibility,etc. We use this knowledge to build a better solution for converting files.

❖ **Internet dependency:** Websites like smallpdf.com, webresize.com, etc provide online solution to converting a file from one format to another. These solutions would not work with poor connectivity

❖ **Privacy Concern** : Giving your personal data to untrusted websites is a high risk to security as these websites can analyse and misuse your data

❖ **Memory issue** :Offline softwares like libreoffice, CoUtils ,etc provide an offline approach to this problem but they take a very large memory space

❖ **Cost** :Softwares like Acrobat and Photoshop are very expensive.

Proposed Solution.

In this project, we have proposed and implemented a platform (*aio.io*) for converting files from one format to another **without the use of any server.** Various image(png,jpeg,svg,gif), video(mp4,flv) , audio(mp3,ogg) and file formats (docx,pdf,odf,txt,html,etc) would be converted interoperability between the same file type

Features:

❖ **Completely offline**: Files are converted and stored on the user device only

❖ **Free to use**

❖ **Better Security**: No logs are created on the Website Server

Features:

❖ **Faster file conversions** : Since user's device is being used to perform conversions, we get incredible speeds without compromising on efficiency

❖ **Small Application Size**

Challenges in Data Conversion

There are many ways in which data is converted within the computer environment:

1. **Simple or Complex conversion**

2. **Seamless or processing involvement**

3. **Single or multiple stages**

4. **Export/Import procedures**

There are many ways in which data is converted within the computer environment:

5. **Single or Multiple Programs used for conversion (reduce this blob to only meaningful content** :A conversion may require multiple softwares to complete. For eg A program may be able to completely convert the data in a file of format A to B by recognizing the input them at the input stage, convert them and give the output. But there is also a chance that If the source/target format is not recognized, then a third program/software would be needed. This program would permit the conversion to an intermediate format, which can then be reformatted using the first program.

6. **Proprietary codecs involved**
7. **Resource intensive and internet intensive**

Principles of Data Conversion

Before any data conversion is carried out, the user or application programmer should keep a few basics of computing and information theory in mind. These includes:

1. Information can easily be discarded by the computer, but adding new information takes effort.
2. The computer can add information only in a rule-based fashion.
3. Up-sampling the data or converting to a more feature-rich format does not add information; it merely makes room for that addition, which usually a human must do (Automatic restoration of information that was lost through a lossy compression process would probably require important advances in artificial intelligence).
4. Data stored in an electronic format can be quickly modified and analysed.

Because of these realities of computing and information theory, data conversion is often a complex and error-prone process that requires the help of experts.

- .PNG   -> .JPG/.WEBP
- .JPG   -> .PNG/.WEBP
- .GIF   -> .PNG/.JPG/.WEBP
- .WEBP -> .JPG/.PNG
- .SVG   -> .JPG/.PNG/.WEBP

```
* 1-> Converting image to base64.
* 2-> pass base64 to function.
* 3-> create canvas with that base64 data uri.
* 4-> set with window.innerHieght and width, so that according to image size.
* 5-> read the canvas image and change image signature to respective file type
* 6-> convert canvas to image by toDataUrl() function in js.
```

Document Conversion Algorithm

.TXT -> .PDF

```
* 1-> Read txt file as plain text.
* 2-> According to length of text content, decide the number of pages.
* 3-> Then by using Open Source library jsPDF() passing arguments as text data.
* 4-> jsPDF() return a BLOB data, save it as pdf.
```

.HTML -> .PDF

```
* 1-> Read html file as plain text.
* 2-> According to length of html content, decide the number of pages.
* 3-> Then by using Open Source library jsPDF() passing arguments as text data.
* 4-> jsPDF() return a BLOB data, save it as pdf.
```

Problem Definition

Project Mission

The aim of our project is provide a platform for the file format conversion and compression without the use of server side which allows it to work offline . And keeping the size of application small making it to use anytime, anywhere.

Target Users

We are targeting every users who want to change there file formats for various purposes like online form filing, software application images need in particular format ,etc. Any user who want to particular file format without concerning about its privacy of data and want to save time and efforts.

System Design

Software and hardware used :

- Software Used: VS Code, Git Version Control Software, NPM , Google Chrome, Jest for software testing
- Languages/Frameworks Used : HTML5, BulmaCSS , React.js, JSON, NodeJS (in dev mode)
- Hardware Used : HP Pavillion laptop with 1Tb hard disk, 8gb ram and i5 7th Gen Processor running on Linux(Ubuntu)
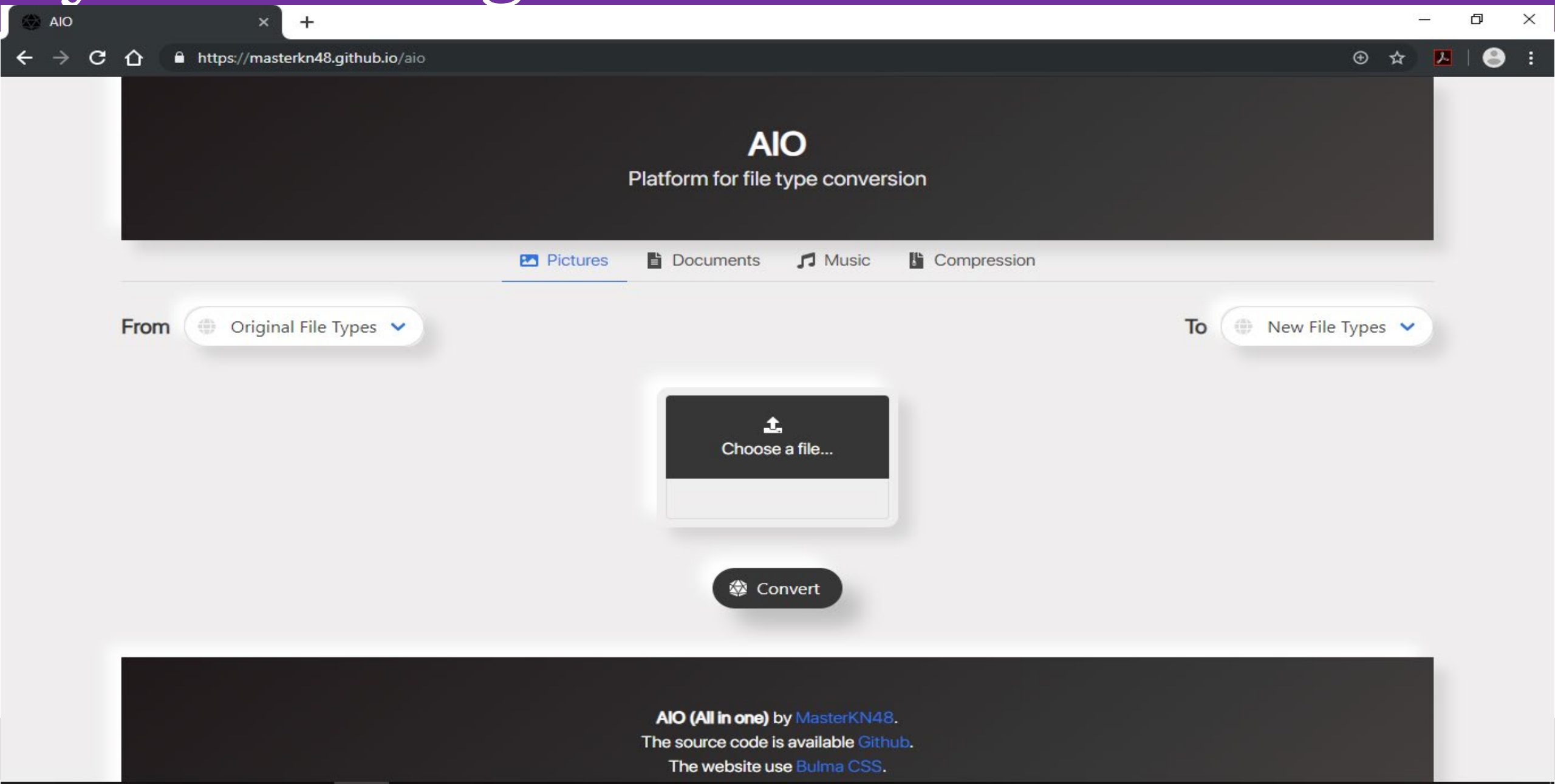
ZIP<-->TAR
RAR<-->ZIP
RAR<-->TAR

Compressed File

Image Converter

PNG <--> JPEG
WEBP <--> PNG
WEBP <--> JPEG
SVG <--> PNG
SVG <--> JPEG
SVG <--> WEBP
GIF <--> JPEG
GIF --> PNG

AIO

MP3 <--> WAV

Music Converter

Document Converter

TXT --> PDF
HTML --> PDF
DOCX <--> PDF
PPT --> PDF
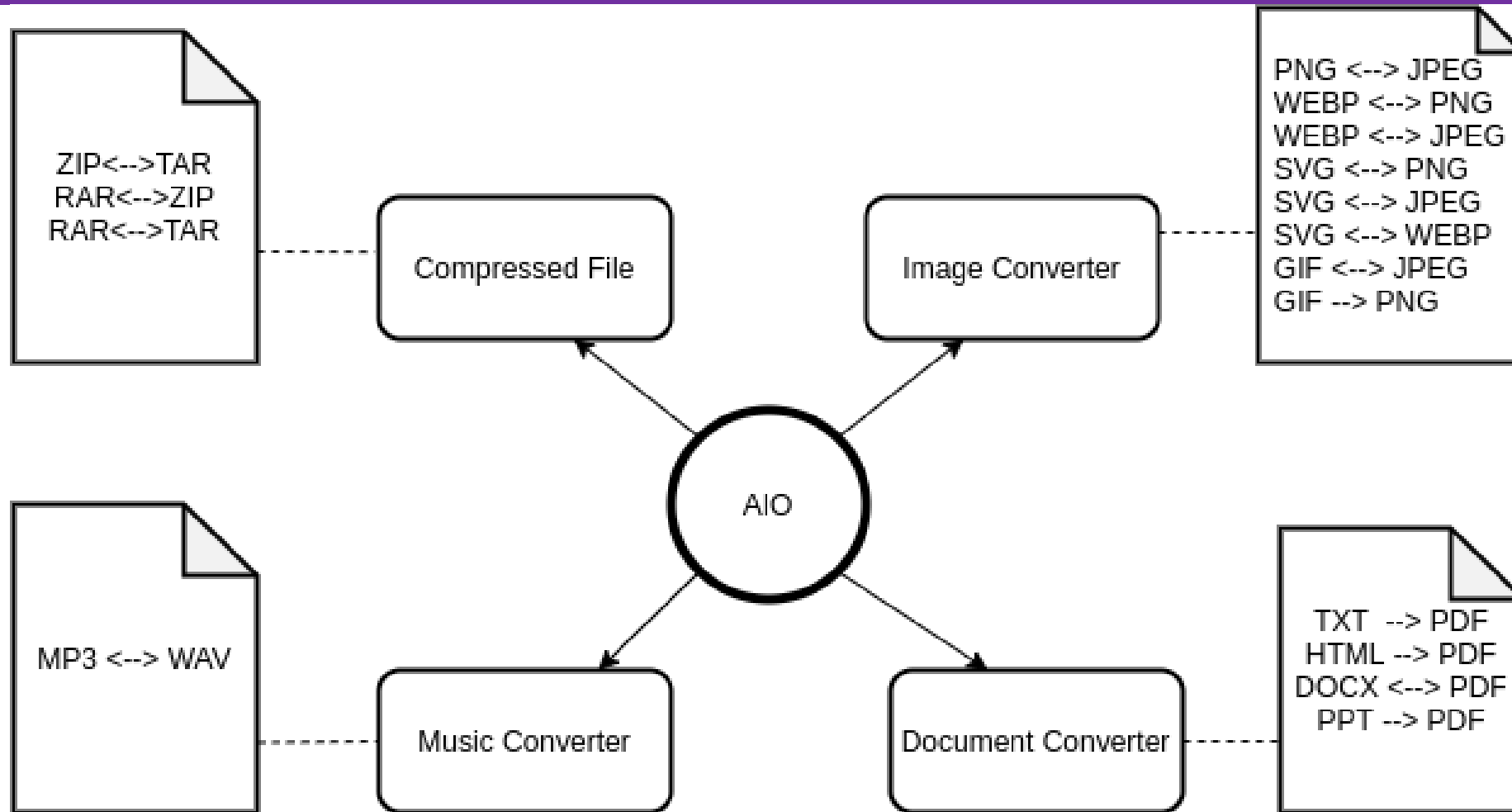
Component Diagram
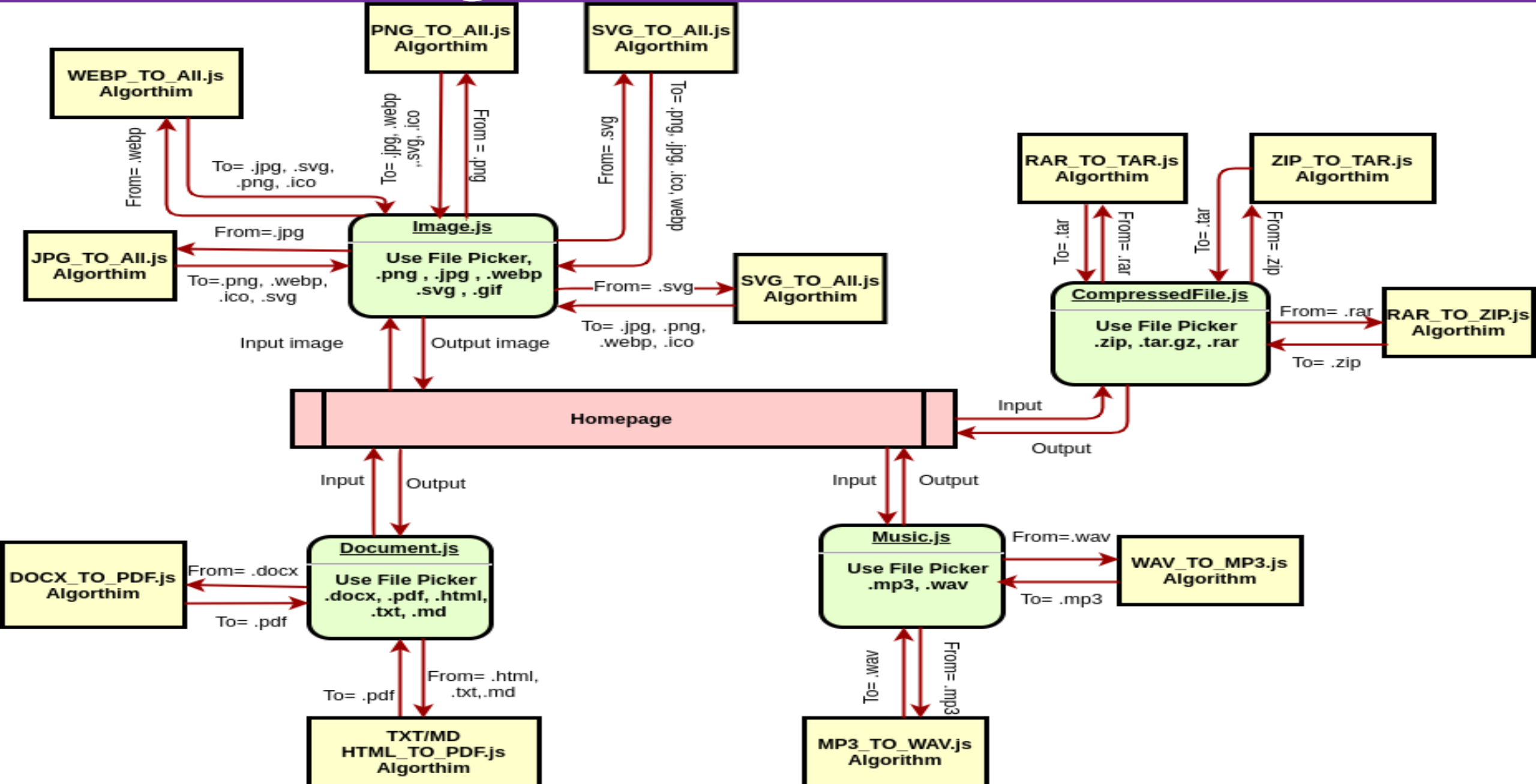
System Design: Preview

Source Code : *https://github.com/MasterKN48/aio/*
Live Demo : *https://masterkn48.github.ui/aio/*
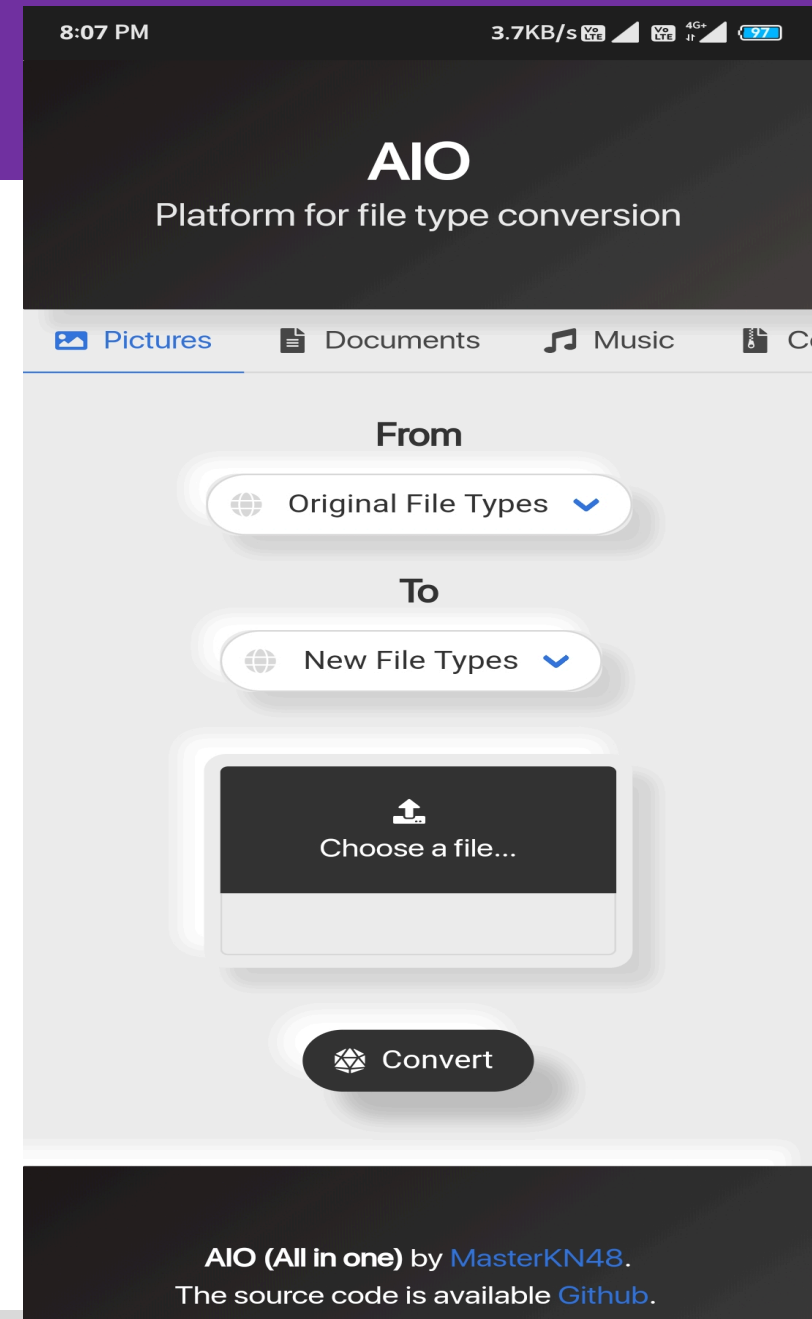
Milestones	Date Assigned	Date Completed	Task
1. Study over file conversion	5/01/2020	Till the project end	Study about various aspects of File formats and conversion.
2. Designing the application	10/01/2020	15/01/2020	Task is to create design of application like wireframe, mockups, understanding design challenges.
3. Modules Generation	16/01/2020	20/01/2020	Deciding number of modules 1-> Image Convertor, 2-> Document Convertor, 3-> Compressed File Convertor, 4-> Audio Convertor
4. Creating Image Convertor	21/01/2020	02/02/2020	Searching about images conversion algorithms, and created all image convertor

System Design: *Task Log :*

Milestones	Date Assigned	Date Completed	Task
5. Create Document Convertor	03/02/2020	10/02/2020	Creating document handling modules.
6. Document Modules	11/02/2020	15/02/2020	Study about documents file format. Creating algorithm design for txt to pdf.
7. Document Modules	16/02/2020	20/02/2020	Study about open formats, proprietary formats, worked on html to pdf and pdf to html.
8. Creating File Compression Modules and documents modules	21/02/2020	30/02/2020	Searching about compression, conversion algorithms, and creating algorithms for file compression

The Purpose and objective of our project is to provide security to user who doesn't want to share crucial data to anyone else but need a platform to convert there file without any fear of data privacy.
This project will provide a platform which can available anywhere and be use at any time.
This project will provide small size very fast file convertor.

References

❖ wikipedia.com : file formats details

❖ fileFormat.com : file formats details and its associated softwares.

❖ reactjs.org : for learning reactjs

❖ *evaluating file formats* nationalarchives.gov.uk/documents/information-management/evaluating-file-formats.pdf This guidance will help you to evaluate files.

❖ A Highly Accurate PDF-To-Text Conversion System for Academic Papers Using Natural Language Processing Approach https://www.researchgate.net/publication/327992611_A_Highly_Accurate_PDF-To-Text_Conversion_System_for_Academic_Papers_Using_Natural_Language_Processing_Approach

❖ Conversion of DICOM Image in to JPEG, BMP and PNG Image Format https://www.researchgate.net/publication/258789781_Conversion_of_DICOM_Image_in_to_JPEG_BMP_and_PNG_Image_Format