

**Project Name- Speech to Text**

**Version History**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Version #** | **Version Description** | **Date Revised** | **Author** | **Reviewer** |
| *1.0* | Speech to Text | 03 June 2020 | Bikram Sahoo & Aayushi Tiwari | Bhargav Ghatty |



**Contents**

Project Name- Speech to Text……………………………………………….1

1. Executive Summary ……………………………………………………..3
2. Project Overview …………………………………………………………3
   1. Purpose ………………………………………………………………….3
   2. Intended Audience ………………………………………………..3
3. Project Description ……………………………………………………..3
   1. Stream Recording-[Speech to text] ……………………..4
   2. Audio File Upload-[Mp3 to text] …………………………..5
4. Repository and References …………………………………………..6
   1. GitHub Repository …………………………………………………6
   2. References …………………………………………………………….6
5. Team ……………………………………………………………………………..6
6. Dev. Schedule ……………………………………………………………..6
7. Stakeholders sign off …………………………………………………….6
8. **Executive Summary**

This document describes the Development plan for Speech to text. This POC includes approaches to the Speech recognition to Text and convert MP3 to Text of the events/actions and services in Dev. environment.

1. **Project Overview**
   1. Purpose

This document outlines the scope of the Speech to text Development and provides output from Streamed Audio & MP3 recorded file.

* 1. Intended Audience

This POC is intended for use by the Dev. lead and Developers to provide guidance when preparing Dev. Scenarios, scripts and to ensure complete coverage of requirements.

Other stake holders can use this document as reference/overview of development activities.

1. **Project Description**

This POC is used to convert streamed voice into plain text and convert recorded mp3 file to text. Default language supported is English .This tool is simple and clean. Instead of typing your email, story, class or conversation, you can just speak or send audio file and this tool can convert it into text.

This POC consist of single module having two different components. They are:

* 1. Speech to text
  2. Audio to text
  3. Stream Recording-[Speech to text]:

This component is used to convert speech to text format. Here user can give input in form of speech. It requires permission from user to switch on their microphone.

**Technology used: Angular 8**

**Technology Usage:-**

* Angular 8 is used for displaying content in browser and interacting with speech recognition service.

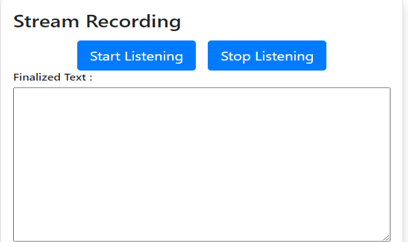
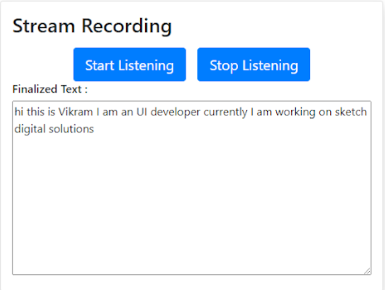
**Service Used:** Webkit Speech Recognition Service.

* Webkit Speech Recognition service is a library which will take Speech input from Users and return response as a Text.

**Example**-

1. Below Reference Image 01, When User clicks “Start Listening” button then it will activate Microphone Service and take **Speech as an input** from User.
2. When User clicks “Stop Listening ” button then it will stop Microphone Service and will give **Text as a response** in “Finalized Text”

**Image-01**

3.2 Audio File Upload-[Mp3 to text]:

This component is used to convert audio file to text. Here user can upload recorded MP3 file as an input.

**Inscope: -** Component is restricted to take only mp3 file as an input.

**Technology used:** **Angular 8, Python and node.Js**

**Technology Usage:-**

* Angular 8 is used for uploading file to backend (node.js) and also for displaying the contents in browser.
* Node.js is used for creating API services to upload files in a separate folder then send that file to python script for giving response to Angular service.
* Python is used to covert .mp3 file to text as response which comes from node service.

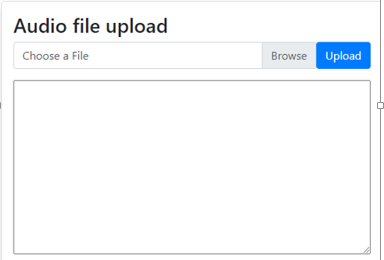
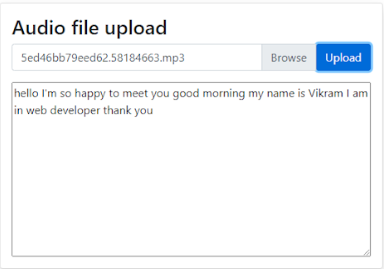
**Tools Used: FFmpeg**

* FFmpeg is an external application which needs to be imported in python file.
* This library is mandatory for converting Mp3 file to wav file.
* Should Installed in Local Machine.

**Example**-

1. Below Reference Image 02, when user click “Browse” then User can choose recorded mp3 file as **an input** from their local machine and can be able to upload file by clicking “Upload” button.
2. After uploading file User can see “Texts “in the text area as **a response.**

**Image-02**

**4. Repository and References**

**4.1 GitHub Repository:-**

* <https://github.com/biku0039/com-poc.git>

**4.2 References:**

* <https://blog.jscrambler.com/implementing-file-upload-using-node-and-angular/>
* <https://hassantariqblog.wordpress.com/2016/12/04/angular2-web-speech-api-speech-recognition-in-angular2/>
* <https://windowsloop.com/install-ffmpeg-windows-10/>

1. **Team**

|  |  |  |
| --- | --- | --- |
| Application POC | Responsibilities | Availability |
| Bikram Sahoo  Aayushi Tiwari  Guided by Uday Bhaskar | Audio File Upload  Streamed Audio | Yes  Yes |

1. **Dev. Schedule**

|  |  |  |  |
| --- | --- | --- | --- |
| **POC** | **Dev. Start date** | **Dev. Finish date** | **Point of Contact** |
| **Speech to Text** | **14/05/2020** | **TBD** | **Bhargav Ghatty** |

1. **Stakeholders sign off**

|  |  |
| --- | --- |
| Bikram Sahoo | UI Developer |
| Aayushi Tiwari | Trainee Developer |