

## ICS261: SSAD & Project

Team number	43
Project Title	Tools for Speech Recognition
Document	SSAD Project Concept Document
Creation date	23/08/2014
Created By	Adhish Singla, Shivam Khandelwal, Romil Aggarwal, Nitish Jain
Client	E. Naresh Kumar (naresh@aksharspeech.com) T. P. S. Manikanta (manikanta@aksharspeech.com) V. Hari Krishna (harikrishna@aksharspeech.com)

### **Description**

As part of Akshar Speech technology development, who need to build massive training sets with speech recordings. These speech recordings need to be pre- and post-processed, which requires the usage of tools. They have offered development of these tools to us.

### **Profile of Users**

Tool 1 : No user, i.e. Crawler to be run by us (developers).

Tool 2 : Distributed users across the world.

Tool 3 : Akshar Speech users (employees)

### **Modules**

Our project is sub-divided into 3 main categories which are described as follows:

#### **Web crawler**

1. Crawl open web data such as Wikipedia, news articles and social media.
2. Extract the text present in each article
3. Clean the text and parse it into sentences
4. Shortlist sentences based on criteria such as number of unique words per sentence
5. Create a final set of sentences and calculate statistics such as number of phonemes, unique words, etc.

#### **Voice Recorder**

A recording tool needs to be developed for Windows, Mac and Web/Mobile platforms. The tasks consists of :

1. Take in an input file with a list of sentences to be recorded
2. Display one sentence at a time. The user should be able to record the sentence, stop, play and re-record each recording. Once the user is satisfied with the recording, then he should be able to go to the next sentence.
3. The recordings should be verified using signal processing algorithms.
4. The save option should be activated when the speech signal is approved by the algorithm.

## **ITRANS**

A UI tool that prepares text according to an audio file The task consists of:

1. The tool requires four input parameters a) ITRANS script file b) Unicode script file c) Audio files
2. Renaming and normalizing the audio files.
3. Upload the modified audio files on the server in “tar” format.
4. Generate individual ITRANS and Unicode files for each audio file.