

Development of a Web-based Hiligaynon Transcription System

Detailed Technical Specifications

Software Development Requirements

The software should be developed as a web-based system and runs on any computer browser such as Google Chrome, Microsoft Edge, Firefox and others.

The Hiligaynon Transcription System shall be an implementation of speech to text (STT) Technology. The current phase of this project aims to produce a Hiligaynon Library training set for the speech to text (STT) Engine containing at least 10,000 words. The system shall adopt any of the existing speech to text engines such as those developed by Google, Microsoft, Amazon and Apple that can use the Hiligaynon Library to perform automated speech recognition. This system can process both English and Hiligaynon words as a transcription system. Primarily, the system must accept Hiligaynon spoken language, process these, and generate the word by word transcription into plain text that can be processed by a word processor in a computer device.

The web application shall include a graphical user interface with a logo and other visual branding. The user interface must provide a microphone icon in a command toggle button to start and end the real-time recording of interviews or conversations in Hiligaynon. Alternately the interface must have an import button to load the input audio data from a file that complies with the following minimum audio formats: .mp3, .mp4, .wma, & .wav.

The interface of the system must have an output text field where the generated plain text is displayed. The user may review the transcription output prior to downloading. A download button must be present to allow the user to download the output plain text.

In this phase of the project the minimum performance requirement for the transcription system is to be able to generate the plain text from a 20-minute input audio file to produce the output text within 3 minutes or less. The transcription system shall be able to process input audio of at most 3 hours. The transcription shall be a word-by-word conversion of spoken Hiligaynon word to the text of the spoken word. Future phase of this project will consider use of natural language processing to convert Hiligaynon speech to English and other language.

The accuracy performance requirement of this project requires that every input of 100 words, the accuracy of transcription from spoken word to text shall be at least 85%. Using other STT engines, this system can transcribe audio to text file using Hiligaynon or other language libraries, but the output rate must still be at least 3 minutes or less for every 20 minutes of input audio.

The activities of the development of this system shall be done using the facilities of the consultant. However, the testing of the system shall be done using the equipment of WVSU.

BUSINESS REQUIREMENTS DEFINITION

1. DETAILED BUSINESS REQUIREMENTS

Each requirement is classified as one of the following:

CLASS TYPE	DESCRIPTION
Mandatory (M)	An absolutely essential feature; project will be cancelled if not included
Recommended (R)	Individual features are not essential but together they affect the viability of the project
Desired (D)	Nice-to-have feature; one or more of these features could be omitted without affecting the project's viability

SYSTEM REQUIREMENT	DESCRIPTION	CLASS TYPE	REMARKS
SR-1	The Hiligaynon Transcription System is a web-based software that runs on most common browsers.	M	
SR-2	The Hiligaynon Library must contain at least 10,000 words as training data for the STT engine.	M	
SR-3	The system may adopt the existing speech to text engines such as those by Google, Microsoft, Amazon and Apple in making the Hiligaynon Library.	M	
SR-4	The system must accept Hiligaynon spoken language, process these, and generate word by word transcription.	M	
SR-5	The system shall complete the transcription of 20-minute spoken audio within 3 minutes or less.	M	

2. OTHER REQUIREMENTS

Interface requirements (IR)

IR No.	DESCRIPTION	CLASS TYPE	REMARKS
IR-1	The user interface must provide a microphone icon in a command toggle button to start and end the real-time recording of interviews or conversations in Hiligaynon	M	

IR-2	The interface must have an import button to load the input audio data from a file that complies with the following minimum audio formats: .mp3, .mp4, .wma, & .wav.	M	
IR-3	The interface of the system must have an output text field where the generated plain text is displayed. The user may review the transcription output prior to downloading.	M	
IR-4	The web application shall include a graphical user interface with a logo and other visual branding.	M	
IR-5	A download button must be present to allow the user to download the output plain text.	M	

Performance (PR)

PR No.	DESCRIPTION	CLASS TYPE	REMARKS
PR-1	The system shall produce a transcription for a 20 minutes audio file within 3 minutes or less.	M	
PR-2	The system shall support transcription of audio files up to 3 hours.	M	
PR-3	The transcription system shall be able to process input audio of at most 3 hours. The transcription shall be a word-by-word conversion of spoken Hiligaynon word to the text of the spoken word.	M	
PR-4	For every input of 100 words, the accuracy of transcription of to the correct text of the spoken word shall be 85% or better.	M	
PR-5	Using other STT engines, this system can transcribe audio to text file using Hiligaynon or other language libraries, but the output rate must still be at least 3 minutes or less for every 20 minutes of input audio.	M	

Security and Confidentiality (SC)

SC No.	DESCRIPTION	CLASS TYPE	REMARKS
SC-1	Include appropriate security controls to manage user access to the application	M	
SC-2	Enhance application access control regarding account creation and maintenance based on account ownership.	M	
SC-3	Enhance applications to provide audit logs for all transactional updates, especially changes made affecting the system	M	

SC-4	Provide an audit log trail screen viewer, extract and reports.	M	
SC-5	Facility to view extract and/or print all activities based on defined criteria – i.e. inclusive dates, transaction type, user ID, and user role	M	
SC-6	Provide for a separate user activity audit log – log-in attempts, password change, etc., with capability for specific record search using defined criteria	M	

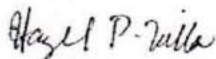
Technical Requirements (TR)

TR No.	DESCRIPTION	CLASS TYPE	REMARKS
TR-1	The web server shall be hosted in a physical server within the WVSU premises.	M	

The web-based transcription system must have a simple architecture and support deployment of webserver and application on a single mid-range server to support up to 5,000 users at various locations nationwide.

The web-based transaction system should provide a long-term strategy and capacity to scale up in terms of functionality. Future phases of the project should allow other features such as natural language processing to allow transcription and translations of Hiligaynon to other languages.

Prepared by:



HAZEL P. VILLA, PhD
Project Leaders



ELIAS C. OLIPANE, PhD

Reviewed by:



LOUIE C. CERVANTES
MIS Director

