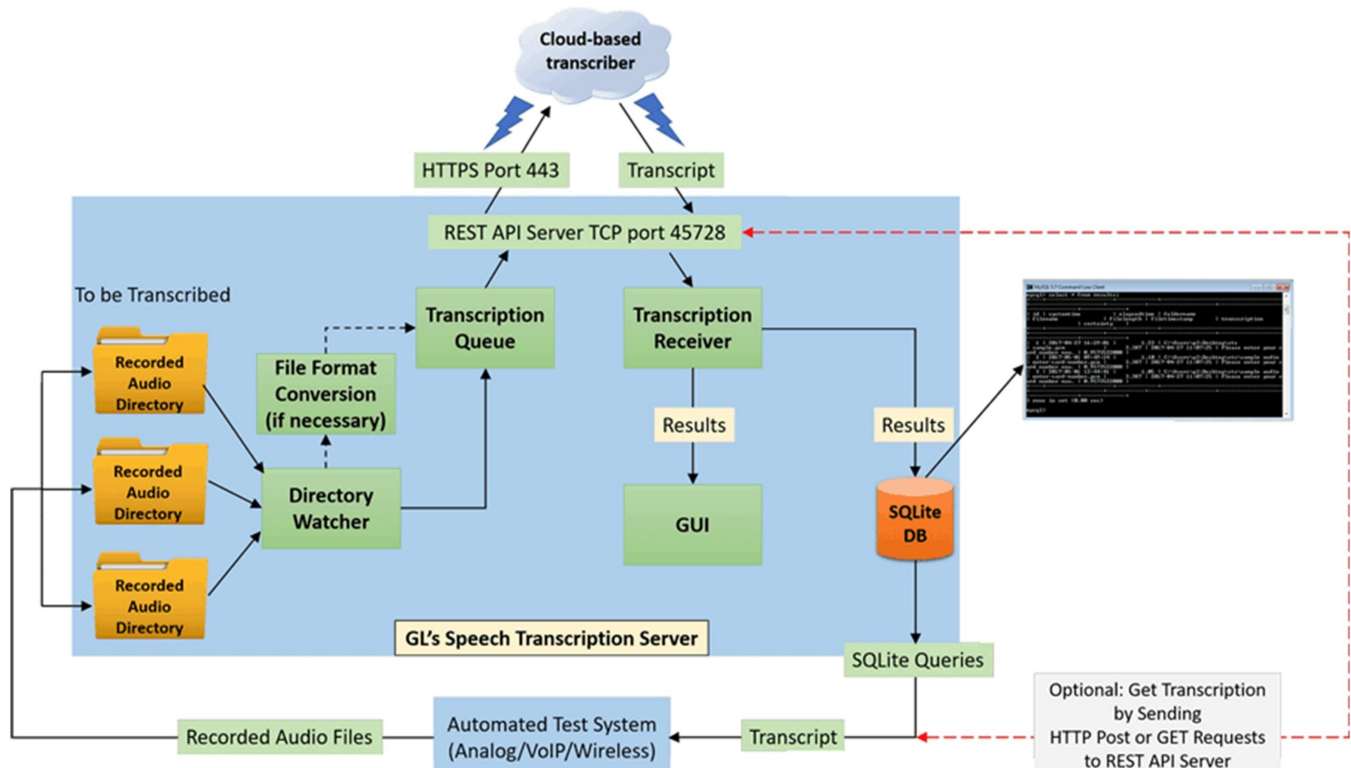# Speech Transcription Server

Overview

GL's **Speech Transcription Server** is a PC-based automated speech-to-text conversion application. Among numerous applications, the Speech Transcription Server can be used for confirming voice prompts (announcements) and aid in testing **Interactive Voice Response (IVR)** systems as well as voice transportation over any network.

Network providers use the application to record the voice prompts associated with the IVR, perform a Speech to Text conversion on the recording to confirm the prompt was proper (based on what the prompt should be), and thus confirming their IVR functioning. The application can be used to verify network quality as well as effect of different codecs on the speech quality.

This application works with GL's File Conversion Utility which supports various industry standard codecs. This application can be good companion for GL's MAPS™ and VQuad™ applications that allow transmission and recording of voice to/from audio files over many industry standards telecommunication network interfaces like FXO/FXS/4-Wire/ISDN/SS7/GSM/UMTS/VoLTE/5G.

For more details, please visit Speech Transcription Server webpage.
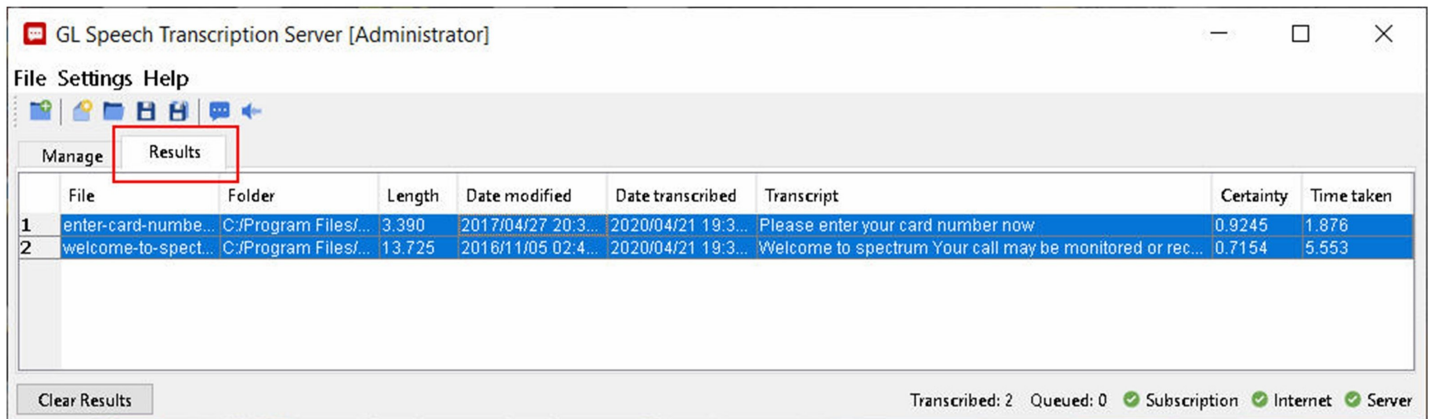
Main Features

- Ability to convert PCM, WAV, or GLW files into text format
- Transcribe up-to 20 seconds of speech files into text
- Transcribe recorded short speech files into text; there is a maximum limit of 20 seconds for any single utterance
- Single or Multiple folders containing audio files can be converted to text easily
- Cloud-based processing provides accurate translations
- Supports multi-languages such as U.S English, French, German, Italian, Japanese and more
- Easy to access transcribed text via file, API or database
- Base Software includes up to 420 hours of audio transcription per year, validity can be extended with annual support contract
- REST APIs for transcription request and transcript retrieval
- Ability to convert text to speech in PCM or WAV file format
- Base Software includes up to 420 hours of audio transcription per year, validity can be extended with annual support contract
- GL's VQuad™ application supports full automation of the Speech Transcription Server (as well as analysis of the transcribed text)

Working Principle

The **Speech Transcription Server** can convert recorded audio (*.PCM, *.WAV, or *GLW) files into text format. Single or multiple folders are monitored continuously for short audio files and once the files are detected, they are placed in the speech transcription queue for transcription. These files are sent to the cloud-based transcription service and are accurately converted to text.

Transcription Results

Results are listed in the order they are transcribed and includes source file information (such as File Name, Directory path, Length of the file, Modified Date and Time) along with the transcribed file information such as Transcribed Date and Time, Transcription output text, Certainty score, and the Time taken in seconds to transcribe the file. The certainty ranges from 0-1 where 0 indicating the lowest confidence score and 1 indicating the highest confidence score for the transcribed text.

Figure: Transcription Results

GL Communications Inc.

Text to Speech Conversion

The **Speech Transcription Server** application converts text into speech in PCM or WAV file format.
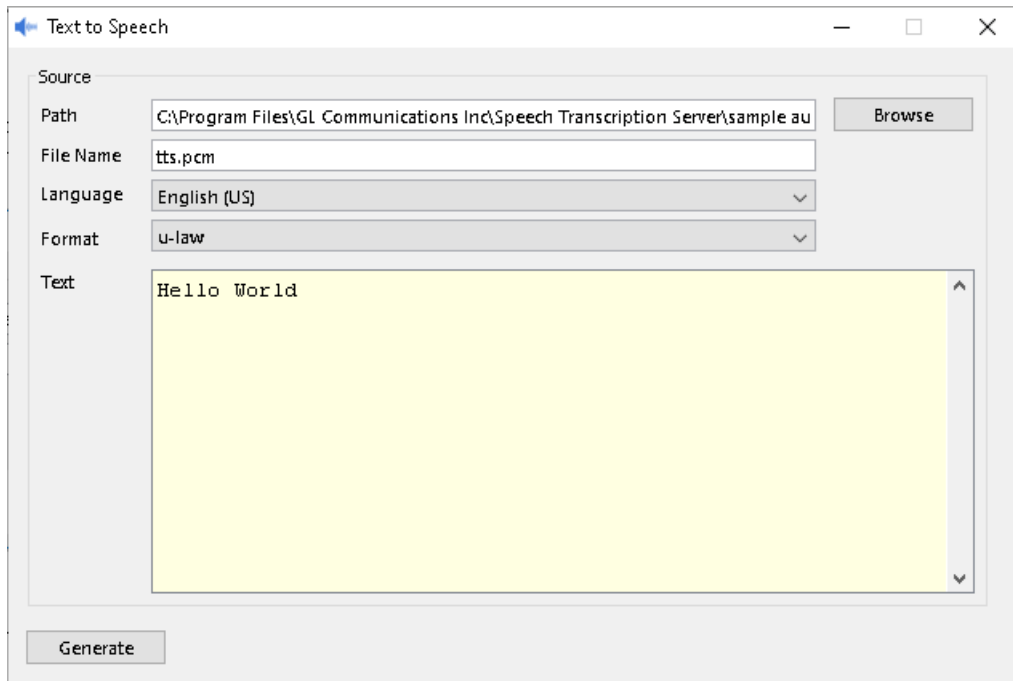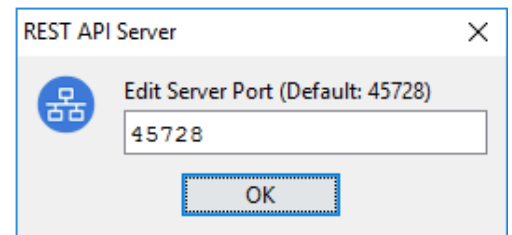

Figure: Text to Speech Conversion

REST API Server

REST API Server helps to transcribe and get the transcribed results on the Remote PC using HTTP request on the configured REST API port number.

Manual Transcription

The **Speech Transcription Server** can transcribe speech to text manually for a single PCM, WAV, or GLW files using Manual Transcription option. This option will help user to quickly transcribe required file manually.
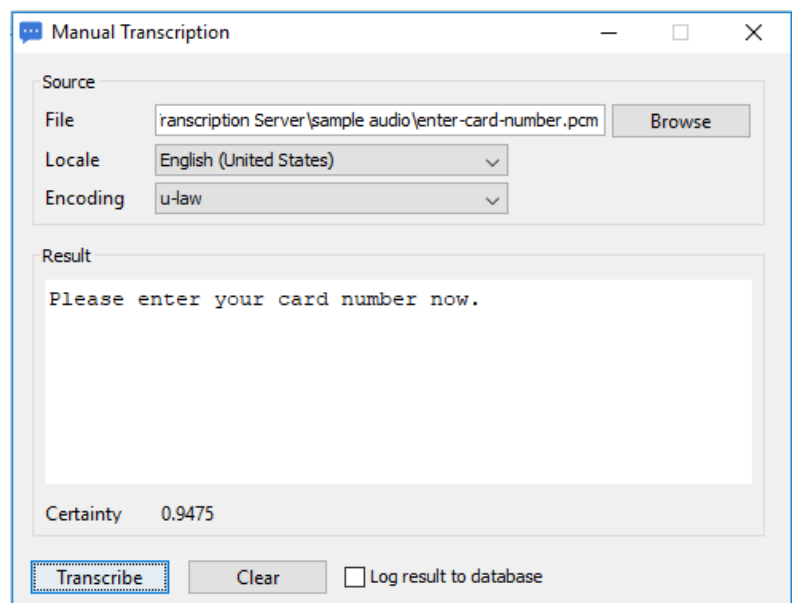

Figure: Manual Transcription

GL Communications Inc.

Buyer's Guide

Item No	Product Description
VQT009	Speech-to-Text Conversion

Item No	Related Software
VQT002	Voice Quality Testing (PESQ only)
VQT006	VQT w/ POLQA server license
VQT010	VQuad™ Software (Stand Alone)
VQT013	VQuad™ with SIP (VoIP) Call Control
VQT251	Dual UTA HD Next generation Dual UTA with FXO Wideband support
VQT252	Dual UTA HD – Bluetooth Option
PKS102	RTP Soft Core for RTP Traffic Generation
PKS120	MAPS™ SIP
PKS108	RTP Voice Quality Measurements
PKS106	RTP Video Traffic Generation

For more details, please visit Speech Transcription Server webpage.