

Local Accent Speech to Text Transcription Engine

The Ministry of Social and Family Development is a ministry of the Government of Singapore focusing on nurturing resilient individuals, strong families and a caring society in Singapore.



"The 100E collaboration has delivered a customised ASR that recognises the unique vocabulary and acronyms used in the social service sector to deliver a word error rate that is much lower than most commercial ASRs. Its deployment will allow MSF to lift the quality of the answers delivered by our voicebot on some of our hotlines and reduce the amount of after call work at our contact centre"

Yvonne Low

Deputy Director Service Excellence and Deputy Quality Service Manager

BACKGROUND

- MSF is currently adopting Natural Language Processing (NLP) technology to answer simple queries in the Baby Bonus Hotline called Ask Jamie Voice (AJV)
- Analysis shows that AJV's speech to text transcription is only about 56% accurate, and 35% of the transcriptions in a sampling of 280 call records contained errors

BUSINESS CHALLENGE

How can MSF develop a speech to text transcription engine that can better recognise local accent, account for Singlish and Code Switching, and is customisable to MSF's terminology?

OUTCOMES

Developed algorithms to train and adapt acoustic and language models for Large Vocabulary Continuous Speech Recognition (over 60 thousand words) to transcribe into Singlish and code switch between English and Mandarin

Developed a system that can transcribe live speech automatically with sentence breaks and speaker turns, approximating human listening performance

Word error rate (WER) decreased from 33.47% to 15.67% for one test set; and from 37.30% to 14.96% for another test set. The overall accuracy of the engine is 84.69%

Worked with the existing chatbot provider to integrate the engine into the chatbot, allowing it to run in parallel with the current Google Speech API solution

Automatic Speech Transcribing system could be deployed at various government agencies and companies to assist frontline officers in call center-type of work

AI SOLUTION DEPLOYED

The AI Speech Lab (launched by AISG and led by Prof Li Haizhou (NUS) and Prof Chng Eng Siong (NTU)) has an Automatic Speech Transcribing system that could interpret and process the unique vocabulary used by Singaporeans – including Singlish and dialects - in daily conversations.

- A multitask learning (MTL) framework using language identification (LID) as the auxiliary task was employed to help improve the code-switch speech recognition perform
- A word vocabulary expansion method was applied to alleviate cross-lingual data sparsity issue in language modelling
- The Asterisk framework with Nautilus was employed to demonstrate the ability of the system to intercept real phone calls and perform real-time transcription