

# Smart Virtual Agent & GenAI Realistic Avatar



ASTRI's speech recognition incorporates the researches in speech-to-text in Cantonese mixed with English (as well as professional terms, industrial jargons and slangs), speaker diarization, voiceprint verification and indexing to address the challenges of information retrieval from a large quantity of audio.



Virtual Agent

## Speech Recognition

ASTRI's NLP for analysing enquiries that are not supported by prior data training or any existing keywords database. We help to build Domain-specific database management platform for non-IT users to build and customise client-specific knowledgebase. NLP can parse Chinese sentences and understand the contextual meanings.



Supports both real time  
and offline transcribing



Supports both real time  
and offline transcribing



Cantonese / English mixed



Putonghua / English  
mixed



Accuracy 85+% in specific domain



Emotion detection



Audio indexing

# Highlights

- Large Language Model are deep learning algorithms that can recognize, summarize, translate, predict, analyse and generate content using very large datasets.
- Domain-specific Large Language Model is fine-tuned large language model with domains dataset.
- Text-to-speech support Cantonese mixed with English, Mandarin, English with voice cloning and voice conversion.
- Knowledge Graph provides a richer context and understanding of info by representing entities and their relationships, enabling both humans and machines to better interpret and interact with the data.

## Realistic Avatar Generation

- Generate realistic avatar video with matching voice
- Text or audio driven half body emotional talking head generator
- Lip, head and gesture synchronization engine to support multi-lingual translation of avatar content

