

DialogueLive

ASTRI is developing speech recognition and synthesis technologies tailored to Hong Kong's unique linguistic and contextual characteristics. When integrated with advanced digital human generation technologies, these capabilities enable platforms for real-time conversational interaction, intelligent customer service, and automated digital human video production. The solution addresses the growing demand for natural, engaging human-computer interaction whilst preserving local language nuances and cultural context essential for effective communication in Hong Kong's multilingual environment.



Industry Challenges

- Traditional chatbots often deliver generic, text-based responses, resulting in limited user engagement and poor experience
- Lack of human-like presence reduces trust, emotional connection, and conversational naturalness
- Real-time enquiry systems require fast response; however, achieving low latency (<5s) with multimodal processing (ASR, LLM, TTS, avatar rendering) is technically challenging

Major Benefits

- Realistic avatars with emotion, gesture, and voice interaction significantly improve engagement compared to traditional chatbots
- Human-like conversational presence makes interactions more natural and intuitive
- Multimodal interaction combining text, voice and visual avatars enables richer real-time Q&A experiences
- Native support for Cantonese, English, and Putonghua, with extensibility to other languages
- Real-time conversational workflows for customer service, job interviews, and enquiry systems

Application Areas

Major departments of HKSAR Government, public organisations, SMEs, system integrators

Collaborator(s)

Real estate industry



Virtual agent integrates researches in speech-to-text, text-to-speech, text-to-talking avatar video generation, emotional expression and gesture generation, which can be applied in real-time voice customer service

