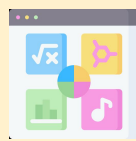
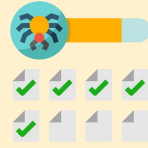


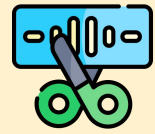
A. Audio Collection



Domain Definition



Audio Crawling

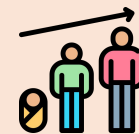


Voice Activity Detection

B. Speaker Attributes Annotation



Speaker Diarization



Age Annotation



Gender Annotation

Speaker ID: SPK_001
Age: 28
Gender: Male

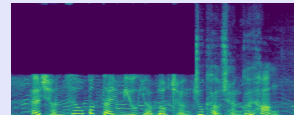
D. Automatic Speech Recognition

SenseVoice
Whisper
TeleASR



SenseVoice_Trans
Whisper_Trans
TeleASR_Trans

C. Speech Quality Annotation



Bandwidth Detection



SNR Prediction



MOS Prediction

Sampling Rate: 32000
SNR: 15.7
DNSMOS: 4.2

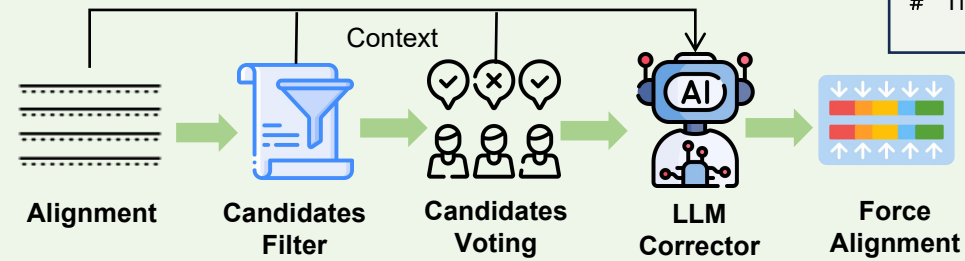
E. Text Postprocessing

Punctuation Removal
Traditional-Simplified Conversion
Text Normalization



SenseVoice_Trans_tn
Whisper_Trans_tn
TeleASR_Trans_tn

F. Recognizer Output Voting (ROVER)



Rover Text: AI 嘅 准确度
Text Conf: 92.05%
Timestamp: <1.02 1.18 嘅>