





## Automated **Transcription**



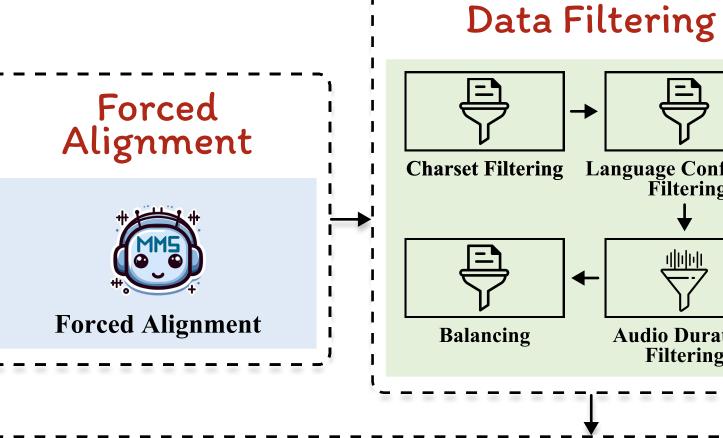
Dataset

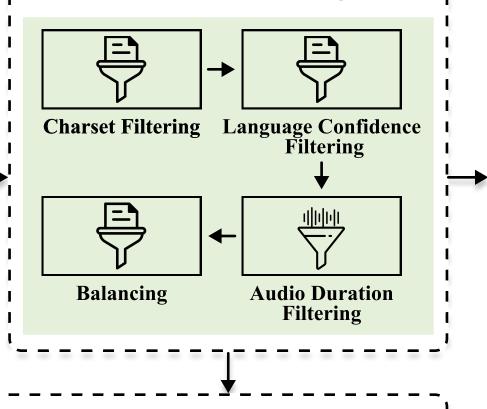
Partitioning

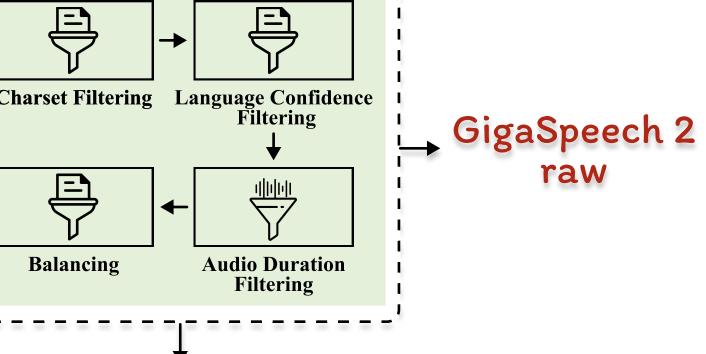
**TRAIN Splitting** 

**DEV/TEST Splitting** 





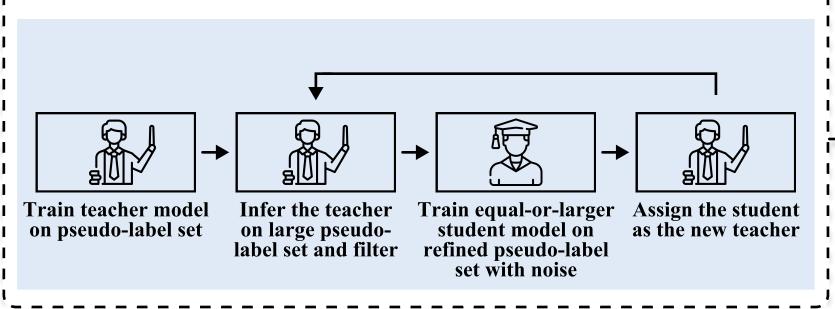












GigaSpeech 2 refined