

# 7주차

☰ 태그	
👤 참석자	기흥 김(P) PARKSEONGJUN 은지 은지 하

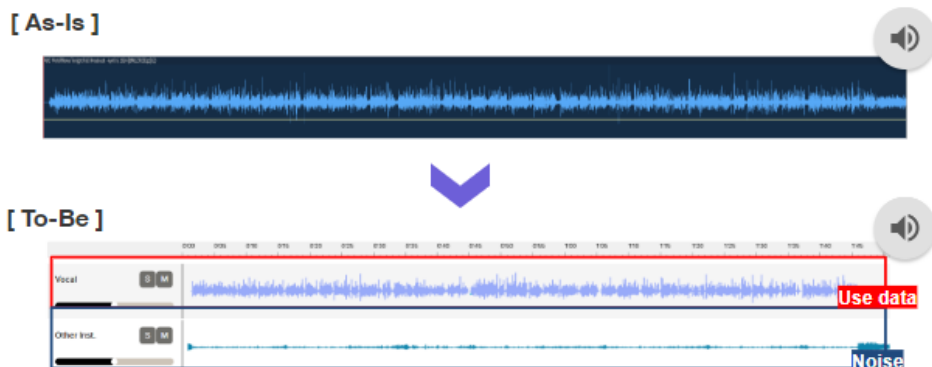
## 회의 이슈

- 상용 Solution 및 SOTA Model Evaluation Metric(WER 등) 분석 마무리
- 연구 방향성 확정 (**Implementation vs Research**)

## 회의 내용(문제, 현재 진행 상황, 앞으로 계획 등)

### 데이터 처리

- **상황:** 현재 발주사와의 계약 전
- **데이터 수집:** 도메인의 요청에 맞추어 우선 자막이 함께 있는 뉴스 영상으로부터 오디오와 전사를 추출하는 작업을 통해 dataset 구축.
- **데이터 처리(1):** 기존에 OSS tool을 활용해서 뉴스 영상으로부터 audio(wav or mp3)와 자막(txt or vtt)을 추출
- **데이터 처리(2):** audio(wav or mp3)를 Gaudio Studio 서비스를 통해 Noise Cancellation



### STT(Speech To Text)

- **Google Cloud Platform STT:** 데이터 처리가 완료된 audio를 STT
- **AWS Transcribe:** 데이터 처리가 완료된 audio를 STT
- **OpenAI Whisper:** 데이터 처리가 완료된 audio를 STT
  - **Small Model:** GPU vRam: 약 11.0GB 점유 & 10분 영상(음성)기준: 약 150초 내외 소요
  - **Large Model:** GPU vRam: 약 2.5GB 점유 & 10분 영상(음성)기준: 약 30초 내외 소요

## 결정 사항(내용, 진행 일정)

- 기존 연구 목표를 재설정
  - Implementation vs Research
- STT Engine Baseline Code 구성

## 특이사항

없음