

# 13주차

☰ 태그	
👤 참석자	<span>기흥</span> 기흥 김Ⓟ PARKSEONGJUN <span>은지</span> 은지 하

## 회의 이슈

- Transformer study
- 앞으로의 계획

## 회의 내용(문제, 현재 진행 상황, 앞으로 계획 등)

### Retraining 가능한 OSS기반 STT 엔진 구현

**상황:** 발주사의 데이터 부제로 인해 자체 수집한 50시간 분량 내외의 dataset으로 model 을 학습시켜야 하는 상황이다. 기존의 OSS model 중 kaldi model을 base-line으로 잡으려 했으나, API 복잡성 및, data preprocessing 부분에서 시간이 부족할 것으로 판단. audio data로부터 발화자, 발화 내용 등에 대한 매핑 수작업이 필요함. 또한, 한다고 해도 기존의 STT model들이 학습시킬 때 사용하는 data 품질은 아주 좋은 환경에서 녹음한 audio이기 때문에, 팀 내에서 자체수집하여 noise cancellation한 data는 그에 비하면 저 품질의 데이터임. 여러 요건에 따라 프로젝트 방향성을 다음과 같이 수정함.

- 사용과 개발 확장성 면에서 용이한 상용 solution을 재조사하자.
- preprocessing없이 audio와 transcribe data 만으로 재학습이 가능한 model을 기반으로 구현하자.
- 해당 model로 STT base-line을 구축하고 그 다음을 생각하자.

## 결정 사항(내용, 진행 일정)

- Transformer 기반 Whisper를 벤치마킹
- Transformer paper review(13주차)
- **Whisper** paper review and Base-line code 구성(14주차)

## 특이사항

없음