



# 230509 주간 발표 @2023년 5월 9일

🕒 생성 일시	@2023년 5월 15일 오후 5:23
👤 생성자	 아윤 노
📌 유형	팀 주간 회의
📅 이벤트 시간	@2023년 4월 23일 오전 6:00
👥 참석자	
🕒 최종 편집 일시	@2023년 5월 18일 오전 11:02
👤 최종 편집자	① 민규 황

## 완료

- ✓ ~~모델 finetuning 가능한지 ? (wav2lib, makeittalk train git code)~~
- ✓ ~~이미지 제너레이터 데이터 생성법 떠올리기 ▲ (ai 허브 데이터셋 get)~~
- ✓ ~~TTS 데이터 생성법 떠올리기 ▲ (배성재 데이터셋 get)~~

## ★앞으로 할거 (5월 2주)

- ☐ Data 생성 ( MakeItTalk / Wave2Lip → 공통 : 동영상 편집( 문장 단위), 25fps 변환 / (W2L) Text , filelist )
- ☐ TTS 모델 선정 + 데이터 준비
- ☐ 기획서 작성

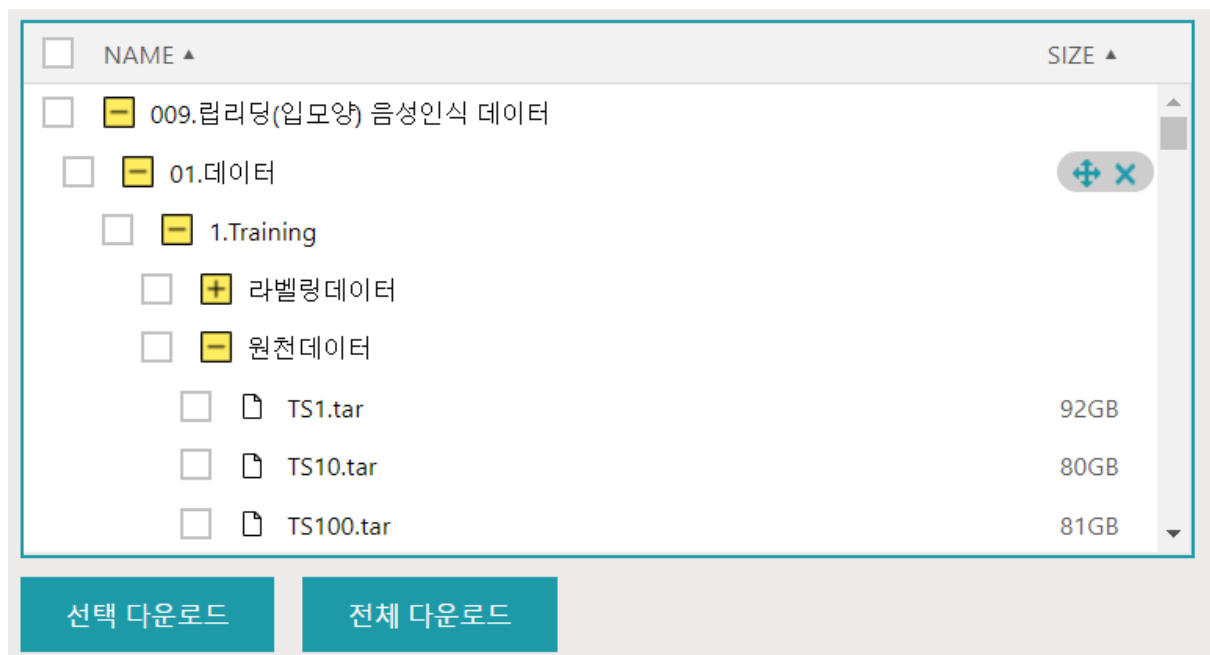
## 1. Img Generator 데이터 편집

- AI hub .json + .mp4

video frame 29 → 25 변환 할때 동영상과 음성의 싱크를 맞추어

(음성, 영상) 데이터셋을 문장단위 편집해야함 (일반적인 프레임 버리기로 frame 변경하면 안될듯)

### Wav2Lip수집



데이터 .tar 한개당 4~5명

총 무소음 57 x .tar

→ 정 좌 우 ) 돌리기 → ( 수, 목 ) 사이 끝내기 .

약 200여명 이상, 각 인원 15분, 각 인원 150여 문장,

총 60여시간의 영상데이터

## 2. 모델 finetune

- MakeltTalk : AWS 서버(ubuntu)
- Wav2lip : local 다른 환경

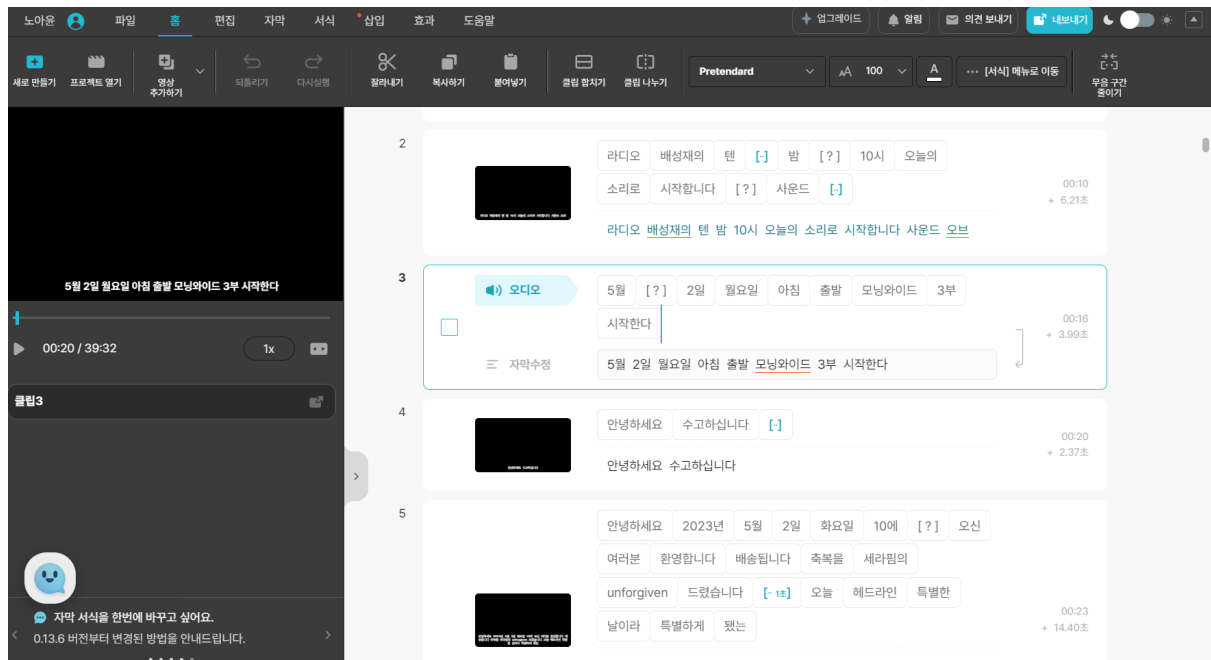
### 3. TTS 모델 관련 GIt link or 해당 모델에 있는 tool kit 정리

Ko - TTS 정리 : <https://github.com/seastar105/kr-custom-tts/tree/main/tts1>

### 4. TTS 데이터 만들기

- Vrew, naver clova note 고민

naver clova note : Text good, 음성 부분 추출 불가능, 시간 분할 정수값(정확하지 않은 시간분할)



Vrew : Text 별로지만, text에 맞춰 음성 분할 편집 가능