

1

2024-03-11

☰ 태그	
👤 참석자	기흥 김

## 회의 이슈

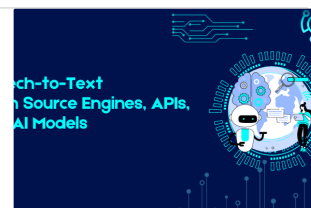
- open source
- Data 어떻게 받아올 지
- 모델은 반드시 학습할 수 있어야 한다.
- 표준화(나중에)
- 사용자가 말한 이후부터 1초 이내의 시간인지, 아니면 말하는 시간을 포함하여 1초 이내의 시간으로 동작해야 하는 지 기준이 있어야 함.
- 몇 초당 1초 이내의 시간인지 기준 필요
- 오픈소스 기반 stt모델들 비교해야 함(13개 정도. 3명에서 4개씩 분배): 3월까지.
- 앵커는 google, aws stt성능(시간, 정확도, 하드웨어 점유율)
- 왓슨보다는 높은 성능.

## 추후 계획

1. 3/11 ~ 3/17(1주차)
  - google, aws로 샘플 데이터 성능 내기
  - 측정 지표 설정하기
    - wer
2. 2주차~
  - 아래 링크에서 4개 정해서 성능 비교해 보기

### 13 Best Free Speech-to-Text Open Source Engines, APIs, and AI Models

Best 13 speech-to-text open-source engine · 1 Whisper · 2 Project DeepSpeech · 3 Kaldi · 4 SpeechBrain · 5 Coqui · 6 Julius · 7 Flashlight ASR (Formerly Wav2Letter++) · 8 PaddleSpeech (Formerly DeepSpeech2) · 9 OpenSeq2Seq · 10 <https://www.notta.ai/en/blog/speech-to-text-open-source>



- 김기흥:
- 하은지:
- 박성준: