[1] Soohwan Kim et al, "Open-source toolkit for end-to-end Korean speech recognition", Software Impacts Vol.7, 2021 [2] Gulati, Anmol, et al. "Conformer: Convolution-augmented transformer for speech recognition." arXiv preprint arXiv:2005.08100 (2020).

Topic: English & Korean Automatic Speech Recognition for Understanding Voice-Commands

[Current Task]

- Train ASR model w/ Korean & English speech dataset

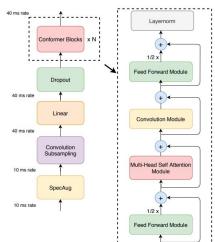
/my_ASR_package /Speech2Text_result

[This Week]

- Study ROS: Writing nodes, packages
- Related Works : recent ASR models
 - Conformer (Gulati et al. 2020), ContextNet, SqueezeFormer, etc.
- Dataset Gathering
 - LibriSpeech: 1000+ hrs, audio-books, English
 - KsponSpeech: ~1000 hrs, conversations, Korean (<u>AIHub.or.kr</u>)
- KoSpeech / OpenSpeech Library
 - Supports KsponSpeech (+ preprocessing)
 - 20+ latest SOTA models (Conformer, etc.)
 - (Problem) KsponSpeech dataset files not readable (corrupt?)
 - (Problem) Other bugs & issues -> fixing GitHub code as needed

[Next Week]

- Fix dataset issues
- Try KoSpeech model training w/ Korean dataset (KsponSpeech)
- Try other libraries: Nvidia NeMO, torchaudio, etc.



```
# audio_file_name [tab] 한글 전사 [tab] 숫자 전사

$DATA_PATH/audio_01.pcm 점심 맛있게 먹었니 320 1 3 24 56 88 3 98 124 76

...

$DATA_PATH/audio_02.pcm 너 뭐 강남 자주 가? 51 3 42 3 243 197 3 47 77 3 9 15
```