

# OpenAI Whisper - Local Live ASR service (Master Plan of Attack!)

소프트웨어 끈대 강의

노기섭 교수

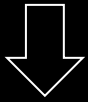
([kafa46@cju.ac.kr](mailto:kafa46@cju.ac.kr))

# Big Picture



## Project 1

Finetune ASR model (Whisper)



AI Model



Local Test

## Project 2

Local live ASR

Server-client based service

## Project 3

File service w/o progress-bar

## Project 4

File service w/ progress-bar

Realtime streaming voice

## Project 5

# 지금까지 우리가 해온 작업은?

멀고 험했던... ㅠㅠ  
데이터 전처리 작업 ㅠㅠ  
그 때 생각 나세요? ㅎ

Finetune ASR model (Whisper)

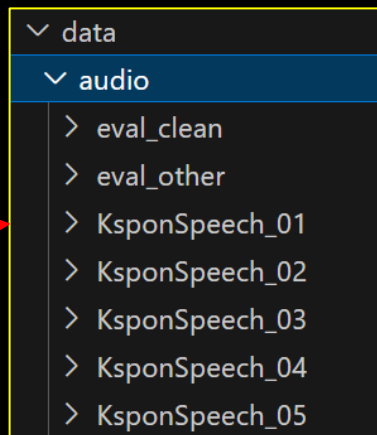


AI Model

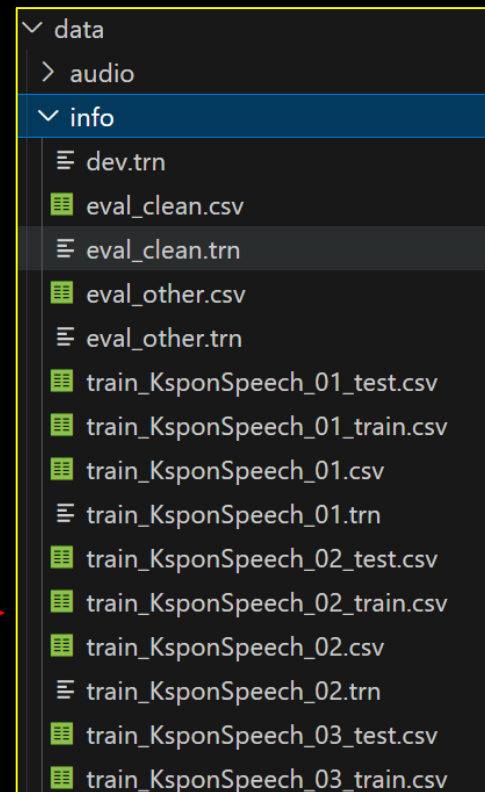


Local Test

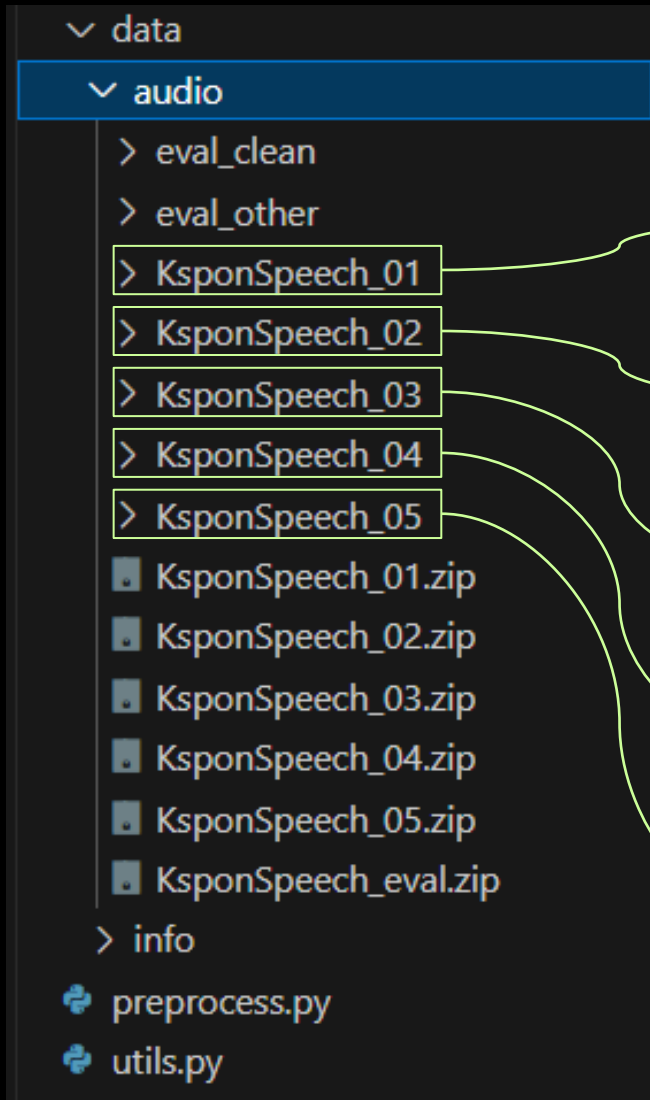
음성파일  
전처리



텍스트 파일  
(Transcript)  
전처리



# Recap: AI hub dataset → Finetuning



Pre-trained model

→ “[openai/whisper-small](#)”, “[openai/whisper-medium](#)”, etc.

finetuning #1 → Updated pre-trained model #1

finetuning #2 → Updated pre-trained model #2

finetuning #3 → Updated pre-trained model #3

finetuning #4 → Updated pre-trained model #4

finetuning #5 → Updated pre-trained model #5

Final Model for us ^^

# 그간의 Finetuning Summary

## ACIN Academy

[https://github.com/kafa46/acin\\_academy/tree/master](https://github.com/kafa46/acin_academy/tree/master)

## Whisper ASR Finetuning

[https://github.com/kafa46/acin\\_academy/tree/master/202\\_fine\\_tuning](https://github.com/kafa46/acin_academy/tree/master/202_fine_tuning)

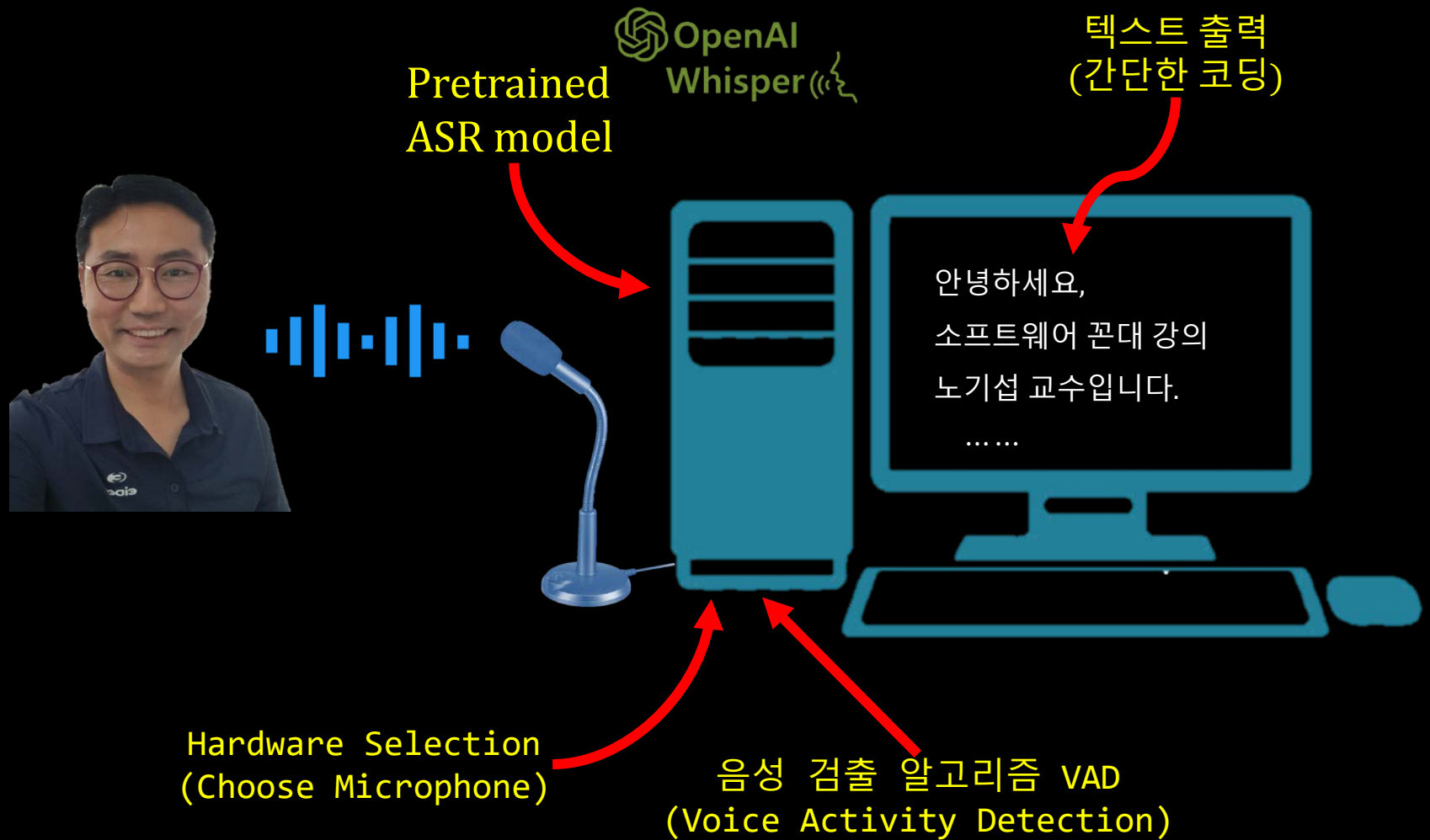
## Download for Finetuned Models

[https://github.com/kafa46/acin\\_academy/blob/master/202\\_fine\\_tuning/whisper/model\\_archive/readme\\_archive.md](https://github.com/kafa46/acin_academy/blob/master/202_fine_tuning/whisper/model_archive/readme_archive.md)

## What you need before we start

- Local Computer (Desktop)
  - In my opinion,  
server environment is not proper to this mini-project
- Microphone (USB, wireless, built-in, .....)
- Pretrained ASR model(s) → We will use Whisper model ^^

# Mini-project Configurations



# Cheating for VAD

## 음성 검출 알고리즘 VAD (Voice Activity Detection)

[https://github.com/snakers4/silero-vad/tree/master/examples/microphone and webRTC integration](https://github.com/snakers4/silero-vad/tree/master/examples/microphone_and_webRTC_integration)

py-webrtcvad: <https://github.com/wiseman/py-webrtcvad>  
(pypi: <https://pypi.org/project/webRTCvad/>)

```
Errors in "pip install webrtcvad"  
Failed to build webrtcvad  
ERROR: Could not build wheels for webrtcvad, which is  
required to install pyproject.toml-based projects
```

```
Final solution:  
pip install webrtcvad-wheels  
(pypi: https://pypi.org/project/webRTCvad-wheels/)
```



# Cheating for Hardware (microphone) Selection



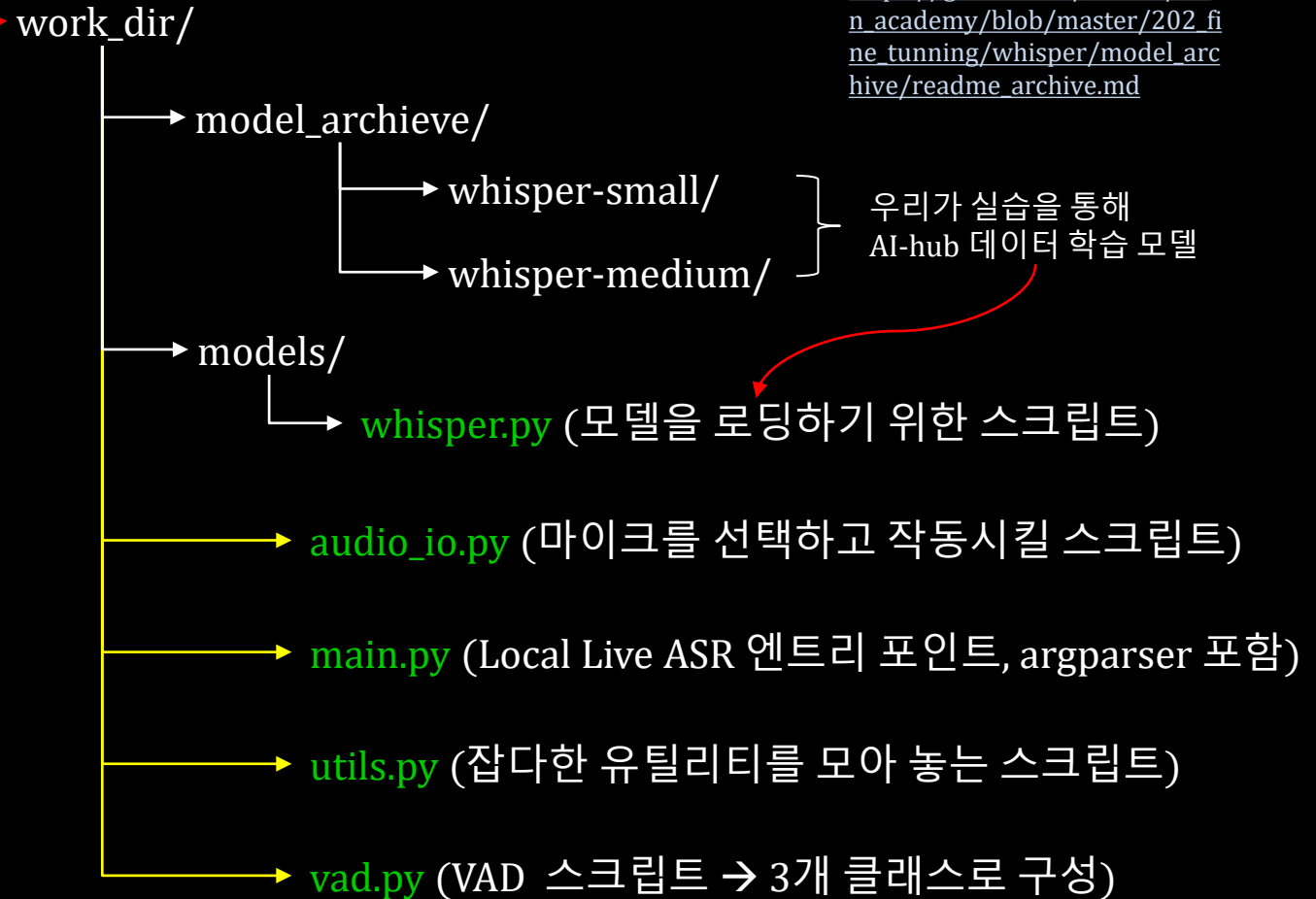
## Hardware Selection (Choose Microphone)

<https://digging-on-bytes.com/whisper%EC%99%80-python%EC%9C%BC%EB%A1%9C-%EC%8B%A4%EC%8B%9C%EA%B0%84-%EB%B2%88%EC%97%AD%EA%B8%B0-%EB%A7%8C%EB%93%A4%EA%B8%B0-part1/>

# Implementation Approach



## Workspace Structure





수고하셨습니다 ..^^..