OpenAI Whisper (Plan of Attack!)
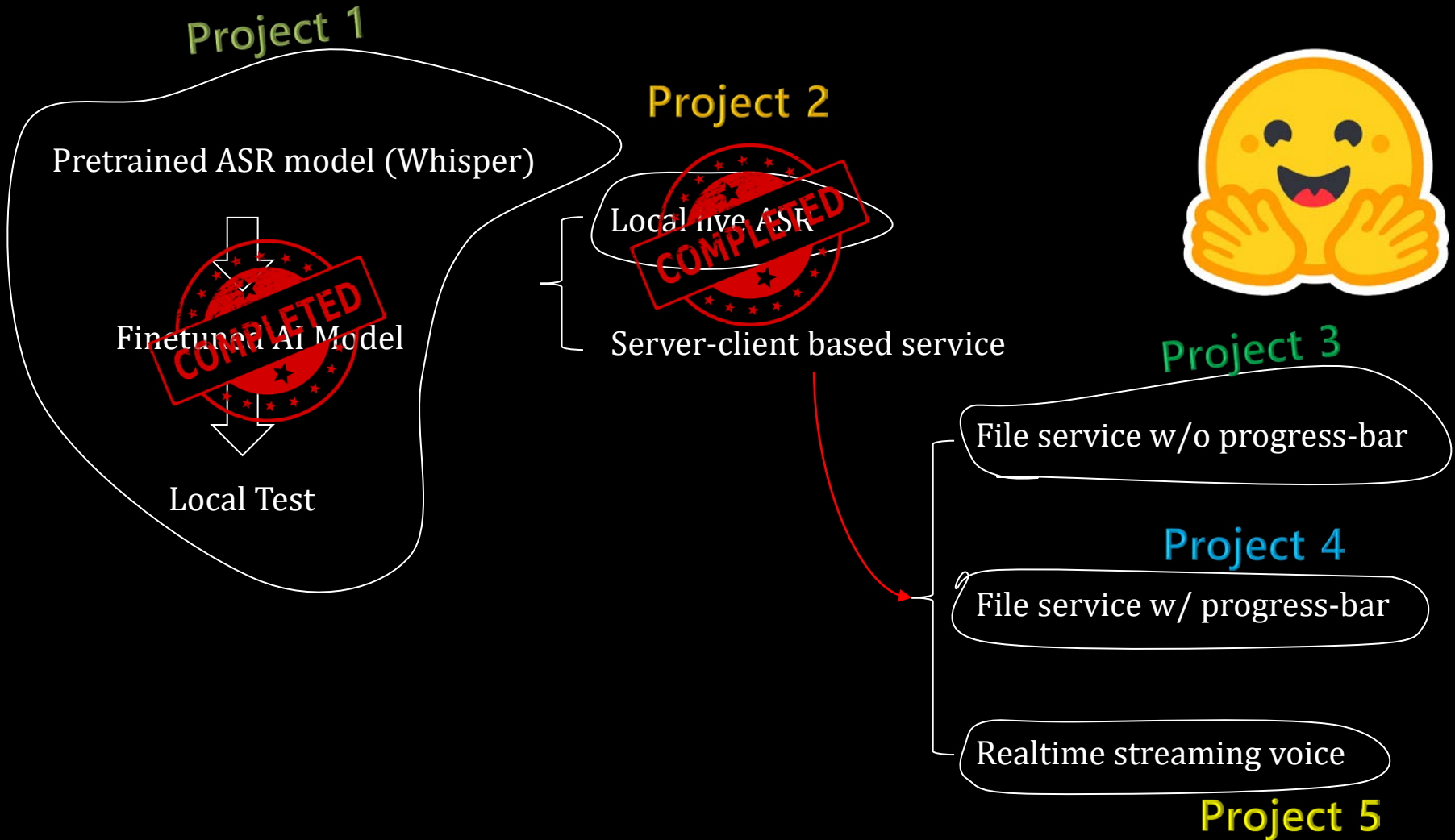
Client-server-based File ASR service

소프트웨어 꼰대 강의

노기섭 교수

(kafa46@cju.ac.kr)
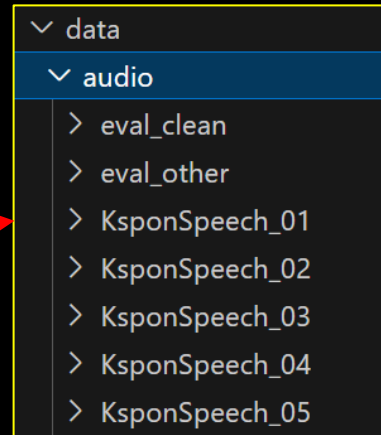
Big Picture

Project 1

Pretrained ASR model (Whisper)

Finetuned AI Model

~~COMPLETED~~

Local Test

Project 2

Local live ASR

~~COMPLETED~~

Server-client based service

Project 3

File service w/o progress-bar

Project 4

File service w/ progress-bar

Realtime streaming voice

Project 5

Recap: Finetuning from Pretrained Models

A few good memories...

멀고 험했던... ㅠㅠ

데이터 전처리 작업 ㅠㅠㅠ

그 때 생각 나세요? ㅎ

음성파일 전처리

텍스트 파일 (Transcript) 전처리
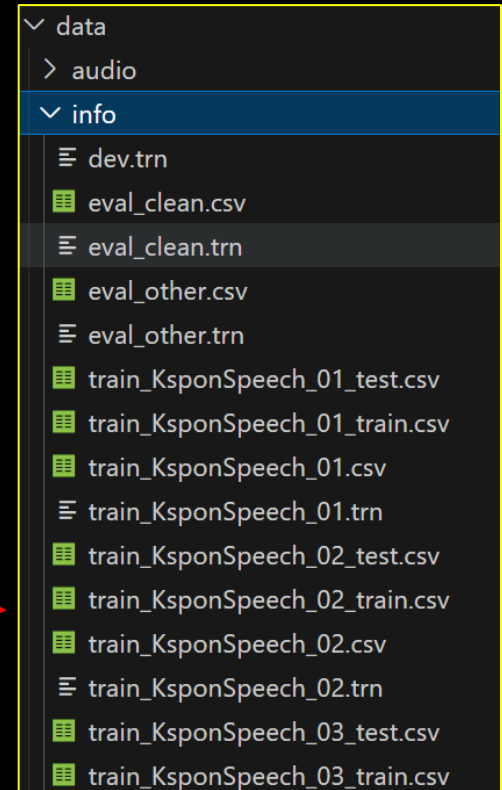
Pretrained ASR model (Whisper)

⬇

AI Model

⬇

Local Test

```
∨ data
  ∨ audio
    > eval_clean
    > eval_other
    > KsponSpeech_01
    > KsponSpeech_02
    > KsponSpeech_03
    > KsponSpeech_04
    > KsponSpeech_05
```

```
∨ data
  > audio
  ∨ info
    ≡ dev.trn
    ▦ eval_clean.csv
    ≡ eval_clean.trn
    ▦ eval_other.csv
    ≡ eval_other.trn
    ▦ train_KsponSpeech_01_test.csv
    ▦ train_KsponSpeech_01_train.csv
    ▦ train_KsponSpeech_01.csv
    ≡ train_KsponSpeech_01.trn
    ▦ train_KsponSpeech_02_test.csv
    ▦ train_KsponSpeech_02_train.csv
    ▦ train_KsponSpeech_02.csv
    ≡ train_KsponSpeech_02.trn
    ▦ train_KsponSpeech_03_test.csv
    ▦ train_KsponSpeech_03_train.csv
```

그간의 Finetuning Summary

ACIN Academy

https://github.com/kafa46/acin_academy/tree/master

Whisper ASR Finetuning

https://github.com/kafa46/acin_academy/tree/master/
202_fine_tunning

Download for Finetuned Models

https://github.com/kafa46/acin_academy/blob/master/202
_fine_tunning/whisper/model_archive/readme_archive.md

Mini-project #2: Configurations

DONE!

OpenAI Whisper

Pretrained
ASR model

텍스트 출력
(간단한 코딩)

FINISHED

안녕하세요,
소프트웨어 꼰대 강의
노기섭 교수입니다.
……
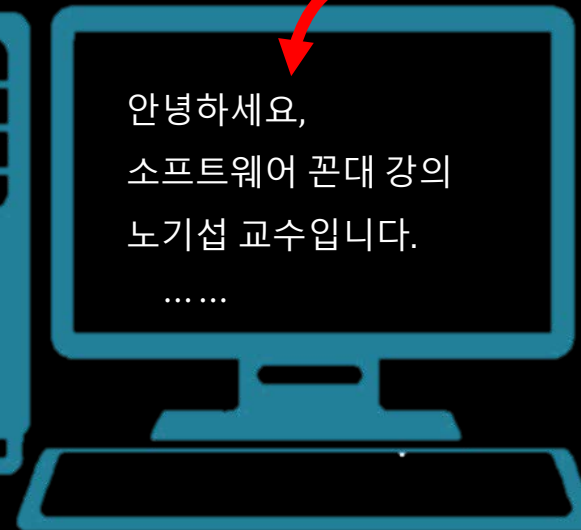
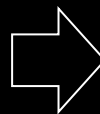Hardware Selection
(Choose Microphone)

음성 검출 알고리즘 VAD
(Voice Activity Detection)

Project #2
강의 자료

https://github.com/kafa46/a
cin_academy/tree/master/201
_deployment#local_live_asr

소프트웨어 꼰대 강의 | 노기섭 교수@Cheongju Univ.

Mini-project #3: Scenario

Request ASR service
(send .mp3)

OpenAI
Whisper

Pretrained
ASR model

1. Attach file(s)
2. Click send btn
3. Check results

GPU GPU

SERVER

Response (send text)

[Notice]

1. No WS/WAS
2. Could be ugly ㅠㅠ.
3. But it will get beautiful more and more
   in project #4 & #5

소프트웨어 꼰대 강의 | 노기섭 교수@Cheongju Univ.
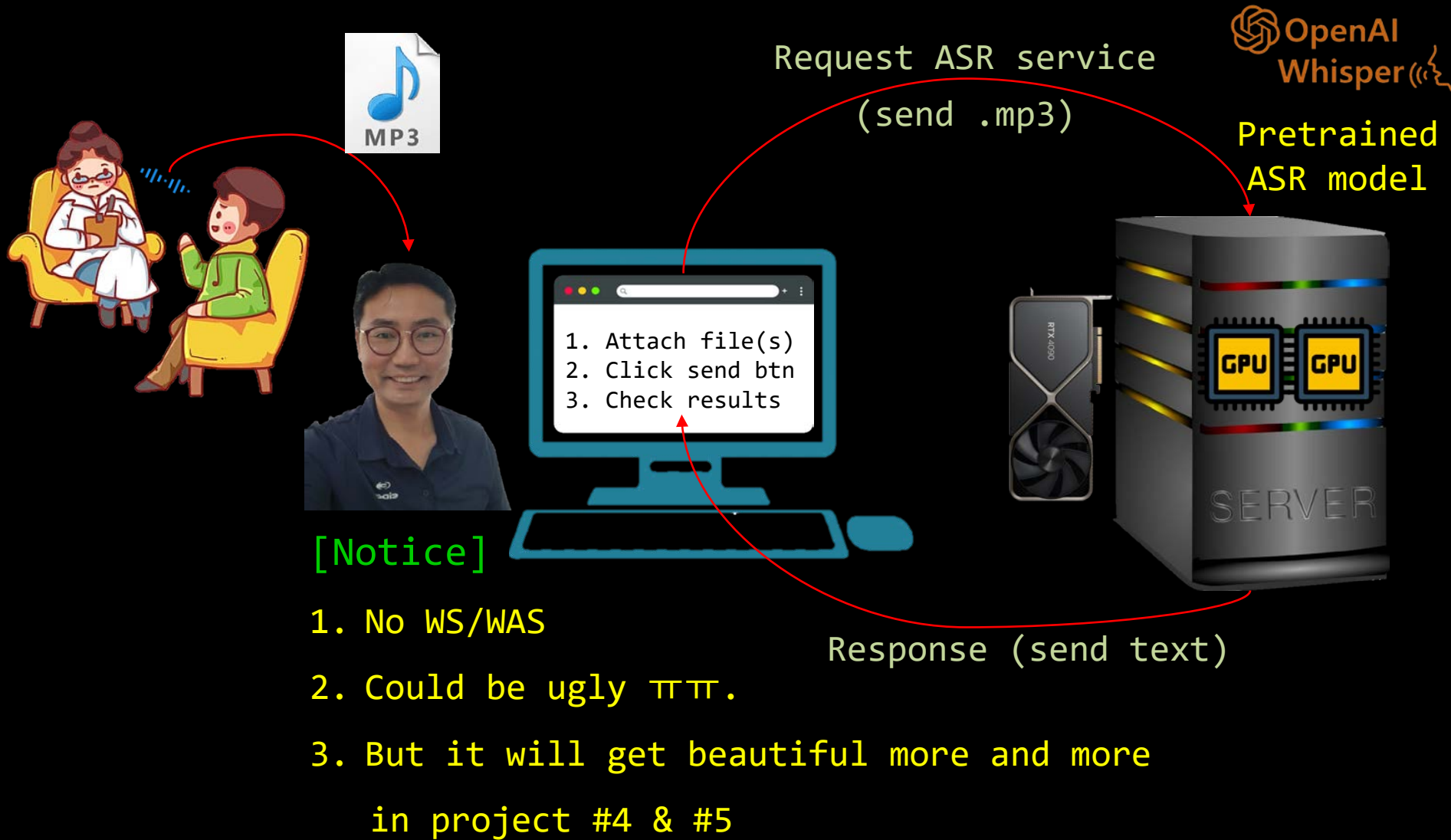
What you need before we start

- Local Computer (desktop, laptop, etc.)

    - I don't care the computing power.

- Linux server (lab, cloud, company, …)

    - Recommend Ubuntu OS

    - However, it's up to you.

- Virtual environment for dev.

    - Same as Project # 1 & #2

- Dependencies

    - flask, werkzeug, flask_wtf, wtforms,

    - Maybe more (just-in-case)

소프트웨어 꿈대 강의 | 노기섭 교수@Cheongju Univ.

Implementation Approach

Workspace
Structure

work_dir/

`FYI, 04_file_asr_no_socket in my case`

ai_asr/

models/

whisper-small/

whisper-medium/

우리가 실습을 통해
AI-hub 데이터 학습 모델

inference.py (whisper 모델을 이용하여 추론 수행)

server/

static/

css/

files/

imgs/

js/

templates/

[다운로드 링크]
https://github.com/kafa46/acin_academy/blob/master/202_fine_tunning/whisper/model_archive/readme_archive.md

config.py (플라스크 서버가 사용할 설정)

run_server.py (플라스크 서버 실행을 위한 파이썬 스크립트)

secret.py (서버에서 사용할 비밀번호 등 민감 정보를 보관)

export_env.sh (플라스크 환경 설정 값을 export 해주는 shell script)

수고하셨습니다 ..^^..