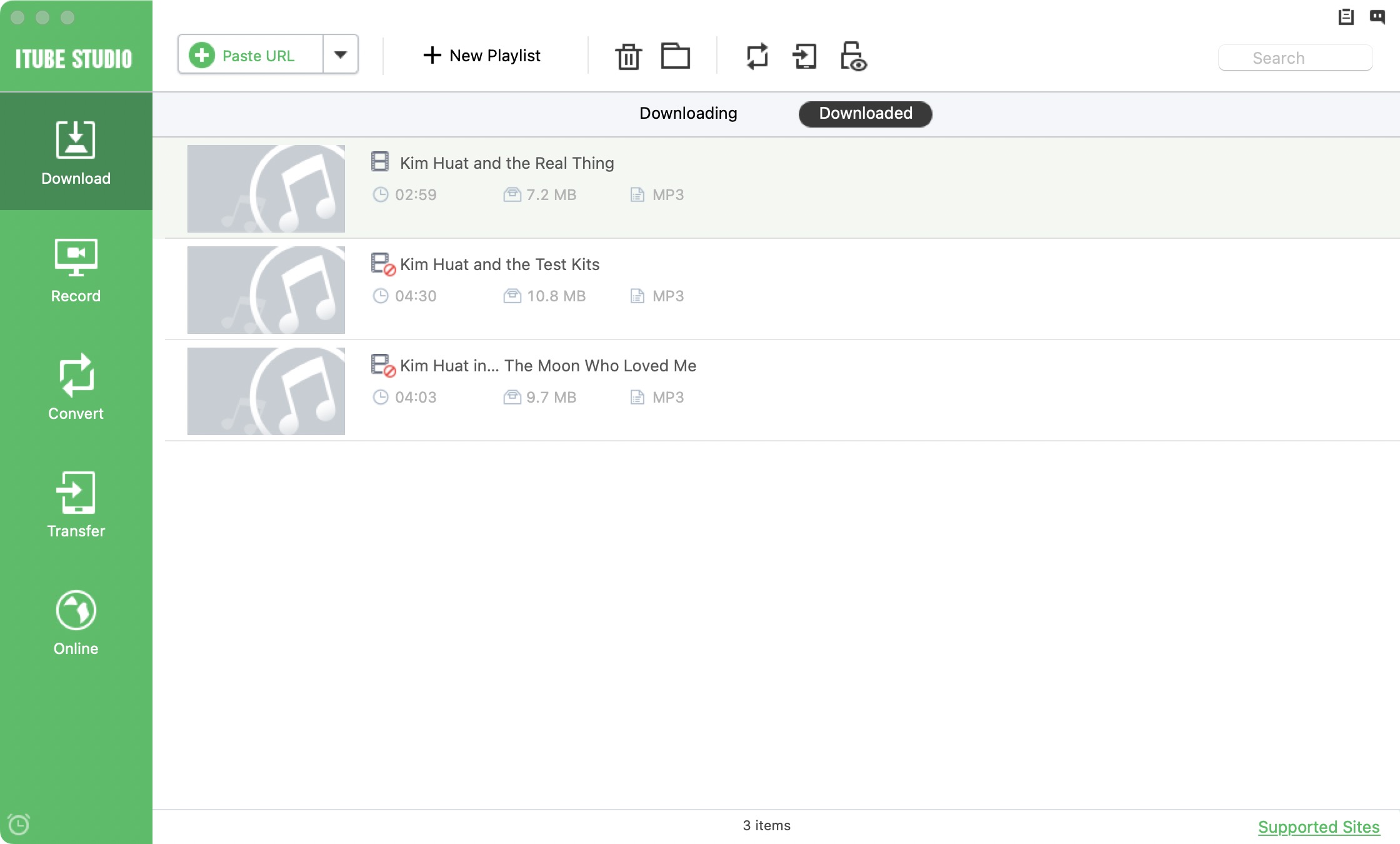
Fine-tuning data set making:

1. Download and use “iTube studio”(https://itube.aimersoft.com/) to download video to audio format. (If you operate more than two audios, you need to pay. Have tried some free python code but encountered some issue.)



## 2. use the “fine\_tune\_dataset\_making.ipynb” to cut the audio into slices by silence.

## 3. Prepare a csv file, each line contains the name of the audio slices and its corresponding text.

Example: for the audio get from:https://www.youtube.com/watch?v=a2ZOTD3R7JI

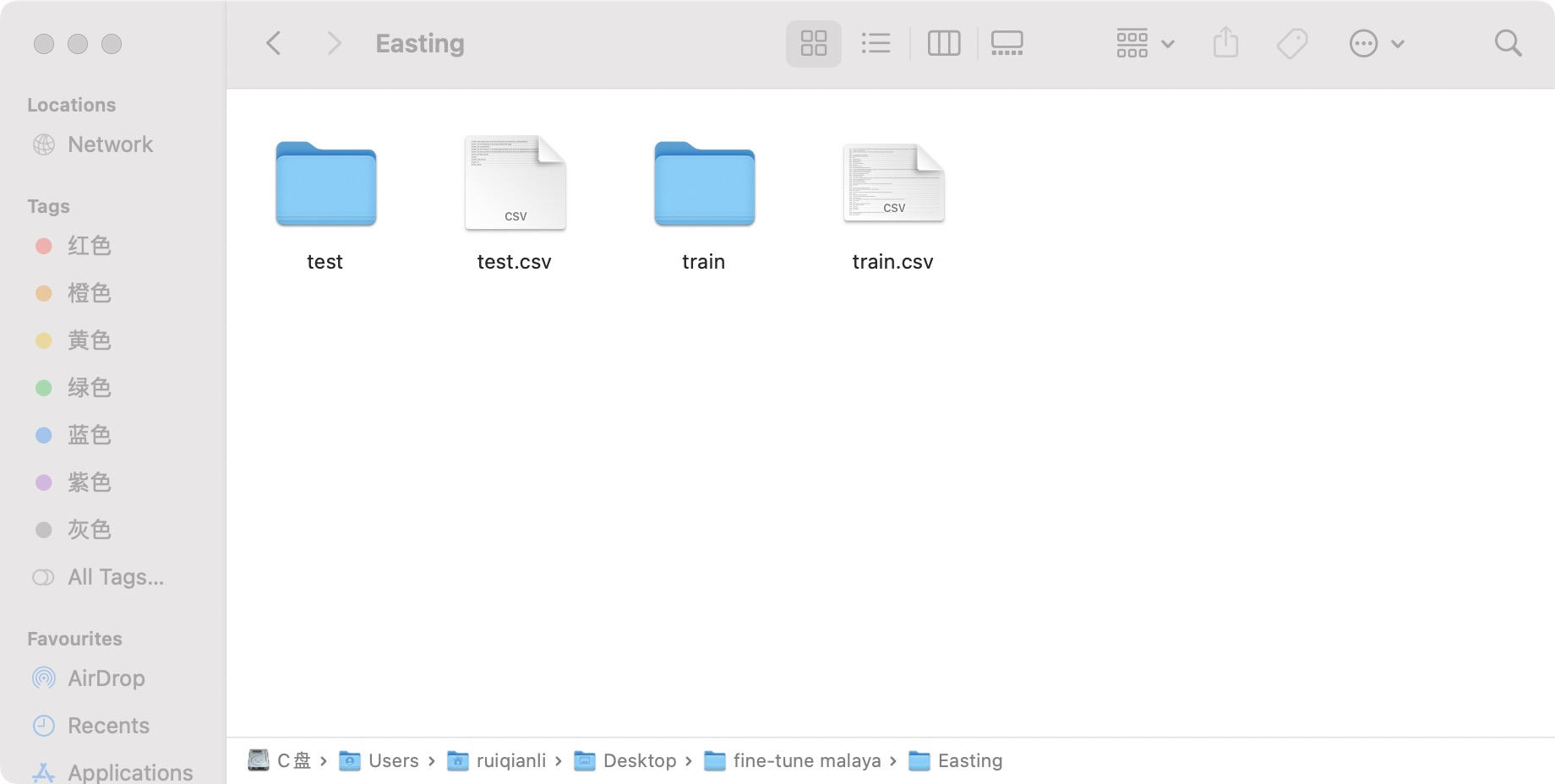
## “Kim Huat and the Easing”

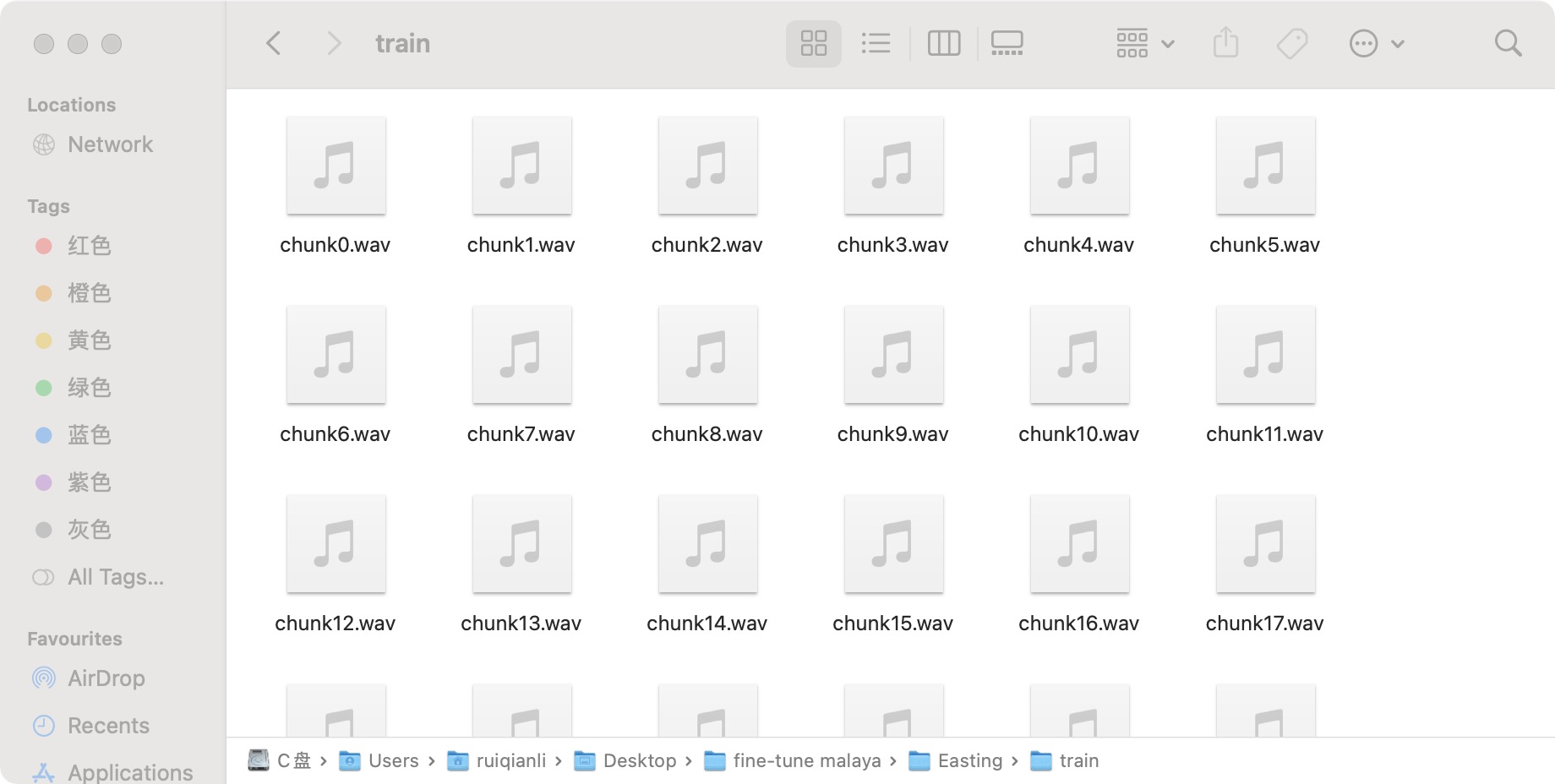
## It have been cut into 0-55 wav file slices, can write the corresponding csv file like this:

## 

## In the process of arranging the subtitles into csv according to the corresponding audio slices, we found that there are some proper nouns in them, and Singaporean and Malay cannot be heard and spelled correctly without subtitles.

1. split the training set and testing set, put all training audios into one folder called “train” and one corresponding csv file. Put all testing audios into one folder called “test” and one corresponding csv file.





5. Zip the folder, later we will use uob\_dataloading.py script to get the data