

Dataset Description

Each row of a csv file represents a single audio clip, and contains the following information:

- filename - relative path of the audio file
- text - supposed transcription of the audio
- up_votes - number of people who said audio matches the text
- down_votes - number of people who said audio does not match text
- age - age of the speaker, if the speaker reported it
 - teens: '< 19'
 - twenties: '19 - 29'
 - thirties: '30 - 39'
 - fourties: '40 - 49'
 - fifties: '50 - 59'
 - sixties: '60 - 69'
 - seventies: '70 - 79'
 - eighties: '80 - 89'
 - nineties: '> 89'
- gender - gender of the speaker, if the speaker reported it
 - male
 - female
 - other
- accent - accent of the speaker, if the speaker reported it
 - us: 'United States English'
 - australia: 'Australian English'
 - england: 'England English'
 - canada: 'Canadian English'
 - philippines: 'Filipino'
 - hongkong: 'Hong Kong English'
 - indian: 'India and South Asia (India, Pakistan, Sri Lanka)'
 - ireland: 'Irish English'
 - malaysia: 'Malaysian English'
 - newzealand: 'New Zealand English'

- scotland: 'Scottish English'
- singapore: 'Singaporean English'
- southatlantic: 'South Atlantic (Falkland Islands, Saint Helena)'
- african: 'Southern African (South Africa, Zimbabwe, Namibia)'
- wales: 'Welsh English'
- bermuda: 'West Indies and Bermuda (Bahamas, Bermuda, Jamaica, Trinidad)'

The audio clips for each subset are stored as mp3 files in folders with the same naming conventions as it's corresponding csv file. So, for instance, all audio data from the valid train set will be kept in the folder "cv-valid-train" alongside the "cv-valid-train.csv" metadata file.