

# Introduction to Mozilla Voice Tools

Josh Meyer | Mozilla Fellow

@\_josh\_meyer\_



# Roadmap

- DeepSpeech
  - Usage
  - Training
- Common Voice
  - Download
  - Contributing
- TTS
  - Usage
  - Training

# DeepSpeech

- DeepSpeech
  - Usage
  - Training

# DeepSpeech

- DeepSpeech
  - Usage
  - Training

`deepspeech.readthedocs.io`

# DeepSpeech

- Usage

CLI client

C API

Python package

NodeJS package

.NET client

# DeepSpeech

- Usage

## CLI client

```
deepspeech \  
  --model deepspeech-0.7.4-models.pbmm \  
  --scorer deepspeech-0.7.4-models.scorer \  
  --audio audio_input.wav
```

# DeepSpeech

- Training

# DeepSpeech

- Training

Python scripts w/ Tensorflow

```
python3 DeepSpeech.py \  
  --train_files train.csv \  
  --dev_files dev.csv \  
  --test_files test.csv
```



# DeepSpeech

- Training

## Requirements

- GPUs
- Transcribed Audio ( 5-15 seconds longs )

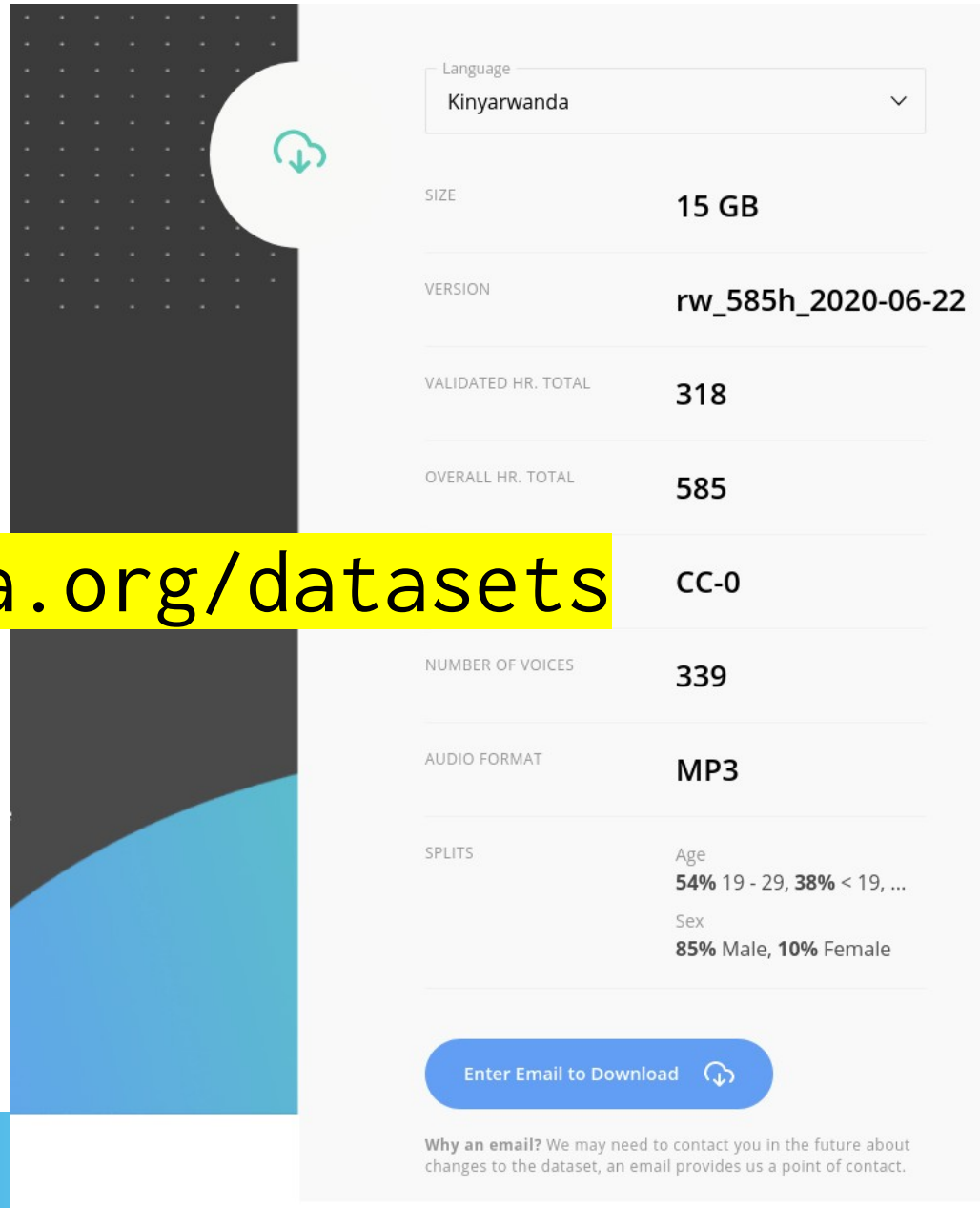
# Common Voice

- Common Voice
  - Download
  - Contributing

# Common Voice

- Download

[voice.mozilla.org/datasets](https://voice.mozilla.org/datasets)



The image shows a screenshot of the Common Voice dataset download page. On the left, there is a dark blue vertical bar with a white cloud icon and a green download arrow. The main content area is white and contains a form for downloading the dataset. The form includes a language dropdown menu set to 'Kinyarwanda', a size of '15 GB', a version of 'rw\_585h\_2020-06-22', a validated hours total of '318', an overall hours total of '585', a CC-0 license, a number of voices of '339', an audio format of 'MP3', and splits for age (54% 19 - 29, 38% < 19, ...) and sex (85% Male, 10% Female). At the bottom, there is a blue button labeled 'Enter Email to Download' with a green download arrow icon. Below the button, there is a small text block explaining why an email is required.

Language	Kinyarwanda
SIZE	15 GB
VERSION	rw_585h_2020-06-22
VALIDATED HR. TOTAL	318
OVERALL HR. TOTAL	585
	CC-0
NUMBER OF VOICES	339
AUDIO FORMAT	MP3
SPLITS	Age 54% 19 - 29, 38% < 19, ... Sex 85% Male, 10% Female

Enter Email to Download

**Why an email?** We may need to contact you in the future about changes to the dataset, an email provides us a point of contact.

# Common Voice

- Contribution



## Mozilla Voice Community Playbook V1.1

- *Last update: July 23th 2020*
- [Edit this document and propose a change](#)
- *Content coordination: Rubén Martín [rmartin@mozilla.com](mailto:rmartin@mozilla.com)*
- [License](#)

Mozilla Voice communities empower the collection of machine-learning based voice technologies -- including software, tools, and data -- that Mozilla stands behind.

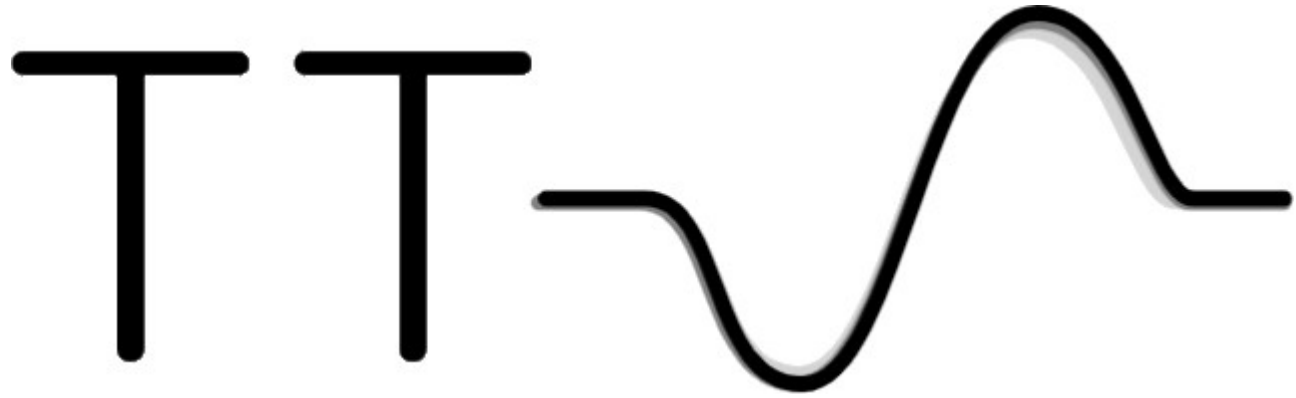
### Goals for this playbook

- Provide people interested in contributing to Mozilla Voice goals and mission with clear guidelines and expectations on how to set up and run a self-sustaining Mozilla Voice community.
- Unify existing community knowledge previously documented in different places.
- Be the central place to understand the whole voice community journey.
- Communicate what brings value to the project and how communities can support it.

[github.com/Common-Voice/community-playbook](https://github.com/Common-Voice/community-playbook)

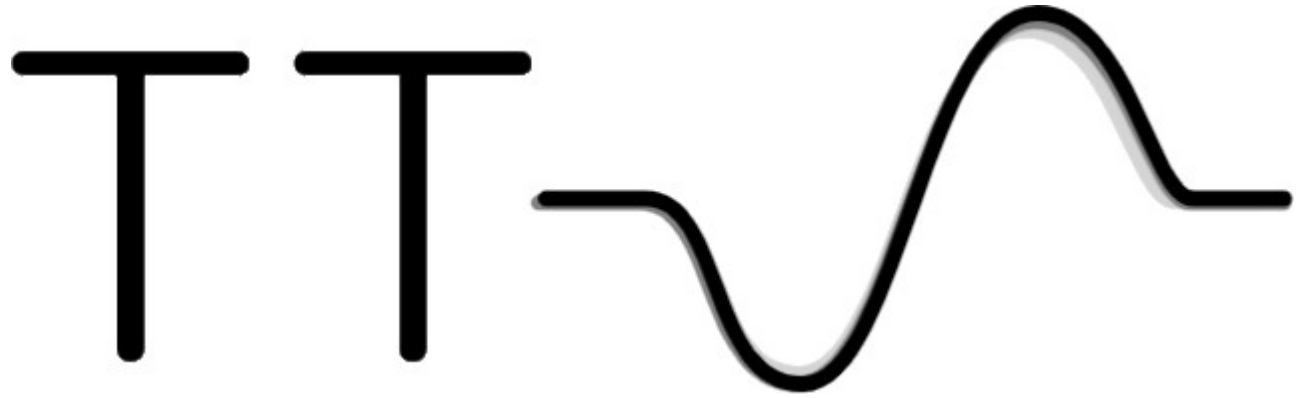
# TTS

- TTS
  - Using
  - Training



# TTS

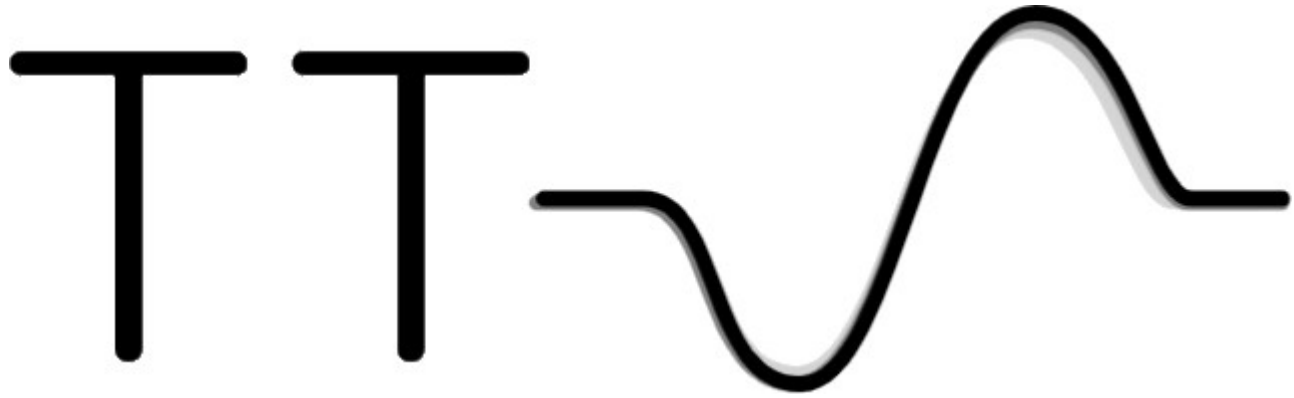
- Usage



- Python + Torch or TensorFlow

# TTS

- Training



- 10+ hours of voice acting
- more GPUs

# Introduction to Mozilla Voice Tools

Josh Meyer | Mozilla Fellow

@\_josh\_meyer\_

