

Harvard

Yá'át'ééh 🖐️

CRESTLEX 3.0

CReating **E**ffective **ST**em
Learning **EX**periences

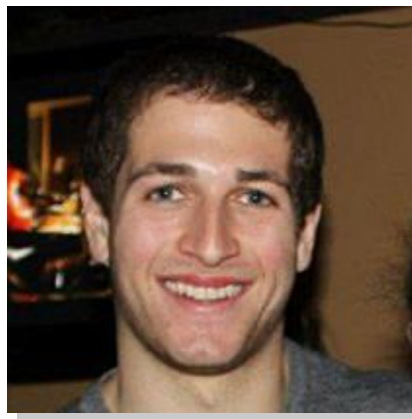
with Navajo Tech





I'm **Monsuru!**

I am an Assistant Professor
at **Navajo Technical University**



I'm Brian!


I am a Robotics PhD Candidate and
teach AI/ML at **Harvard University**.

Using Zoom

- **Ask questions**
 - Use the Zoom chat
 - Raise your hand
 - Interrupt us!
- **Cameras & Microphone**
 - Camera on if you'd like
 - **Mic muted** unless speaking




▼ Participants (1)

 Dhilan Rama... (Host, me) Unmute More ▼

- Rename
- Edit Profile Picture

The image shows a dark-themed interface for a meeting participant list. At the top, it says "Participants (1)" with a downward arrow. Below this, there is a single participant entry for "Dhilan Rama... (Host, me)". To the left of the name is a small circular profile picture. To the right of the name are two buttons: "Unmute" and "More". A dropdown menu is open from the "More" button, showing two options: "Rename" and "Edit Profile Picture".

Rename

 Enter new name here:

Cancel Rename

The image shows a "Rename" dialog box with a dark background. At the top, it has a title bar with three window control buttons (red, yellow, grey) and the title "Rename". Below the title bar, there is a blue rounded square icon containing a white video camera symbol. To the right of the icon, the text "Enter new name here:" is displayed. Below this text is a text input field with a blue border and a white cursor, containing the text "Dhilan R. (Harvard)". At the bottom of the dialog, there are two buttons: a grey "Cancel" button and a blue "Rename" button.

Our website!

tinymlx.org/CRESTLEX3/

home base for **all information!**

Workshop Agenda

Morning Session (9–11am)

The Future of AI (with Laurence)

How ML Works? (with Dhilan)

ML in the Navajo Nation (with NTU)

Responsible AI (with Susan)

Afternoon Session (12pm–2pm)

Experimenting with AI (with Dhilan)

Exploring ML (with Jenny)

Build it! Your own app (with Jenny)

What's Next?

(with guest student panel)

What's Next?

(with instructors)

Workshop Agenda

Morning Session (9–11am)

The Future of AI (with Laurence)

How ML Works? (with Dhilan)

ML in the Navajo Nation (with NTU)

Responsible AI (with Susan)

Afternoon Session (12pm–2pm)

Experimenting with AI (with Dhilan)

Exploring ML (with Jenny)

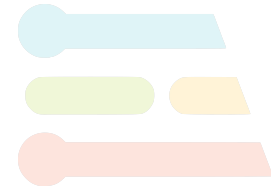
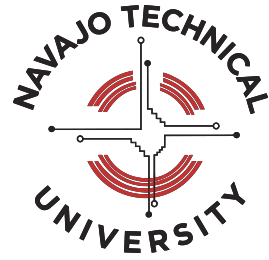
Build it! Your own app (with Jenny)

What's Next?

(with guest student panel)

What's Next?

(with instructors)





A quick recap!



AI and ML for today and tomorrow...



Laurence Moroney
@lmoroney

Harvard

Experimenting

with AI

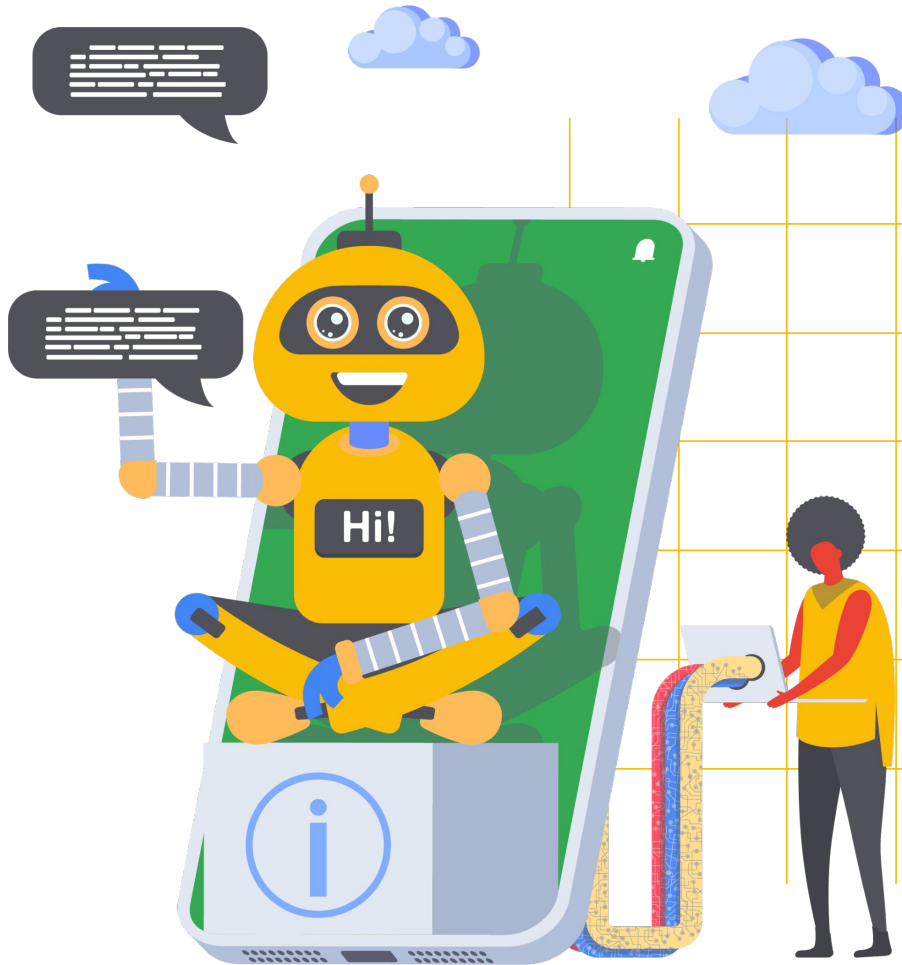
with Dhilan



Harvard

How ML works?

with Dhilan





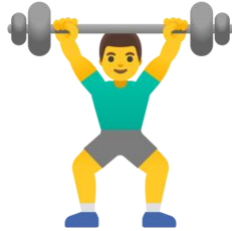
deploy

to your tiny **devices!**

Review what we've learned

Machine learning provides a computer with data, **rather than explicit instructions**. Using these data, the computer learns to **recognize patterns** and becomes able to execute tasks on its own.

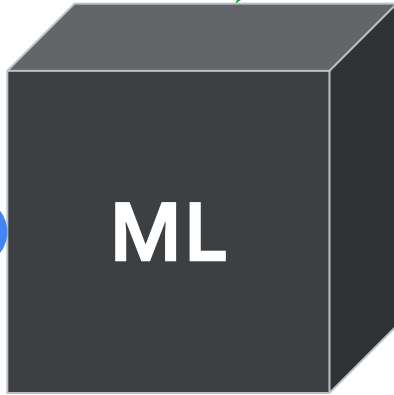
Training the machine



WE PROVIDE

LABELS

INPUTS



RULES

Good Data is Necessary for Accuracy

What problem are you trying to *solve*?

- Your data must contain useful features
- Can a human (expert) distinguish between examples of each class?
- How will you measure performance?

Good Data is Necessary for Accuracy

What problem are you trying to *solve*?

- Your data must contain useful features
- Can a human (expert) distinguish between examples of each class?
- How will you measure performance?

Both *quantity* and *quality* will influence your model's performance

- **Wide variety of training examples**
- **Correct labels (answers)**
- **Good Balance (e.g., dog, cat, random)**

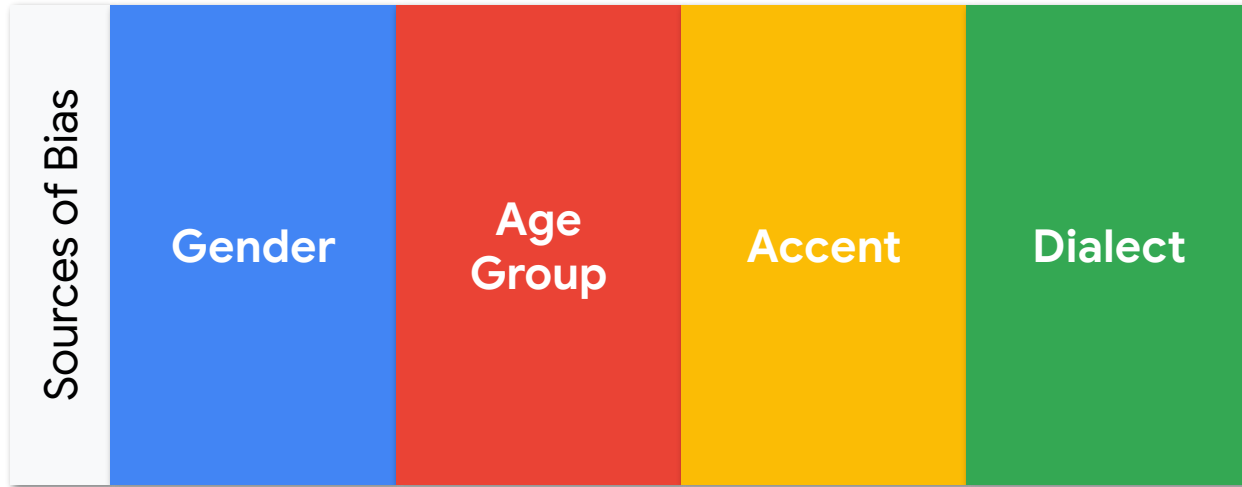
What do 50 million drawings look like?



What do 50 million drawings look like?



Potential **Bias** in Speech Recognition

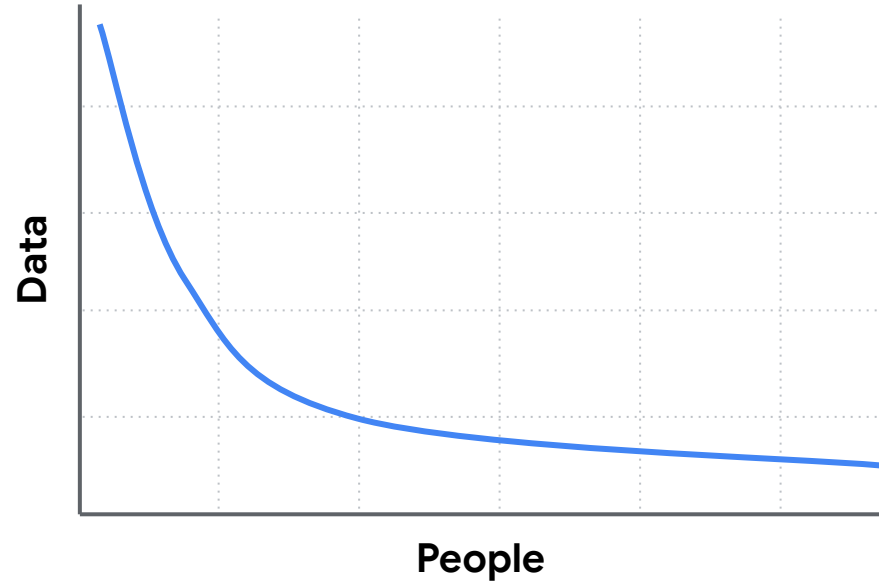


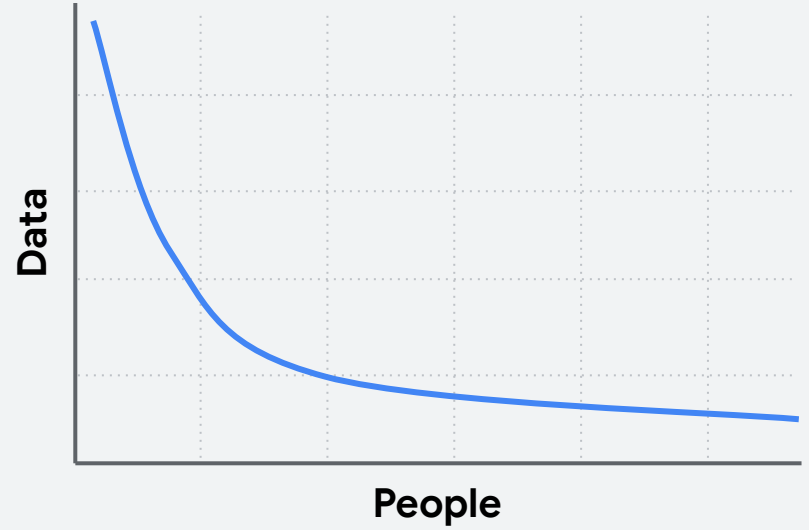
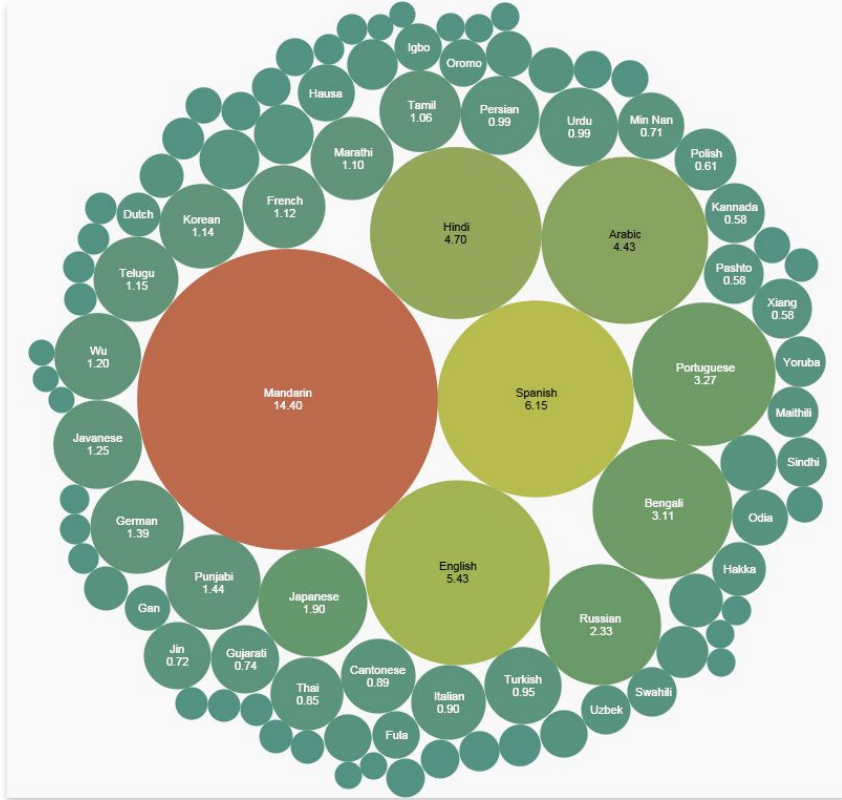
what's missing?

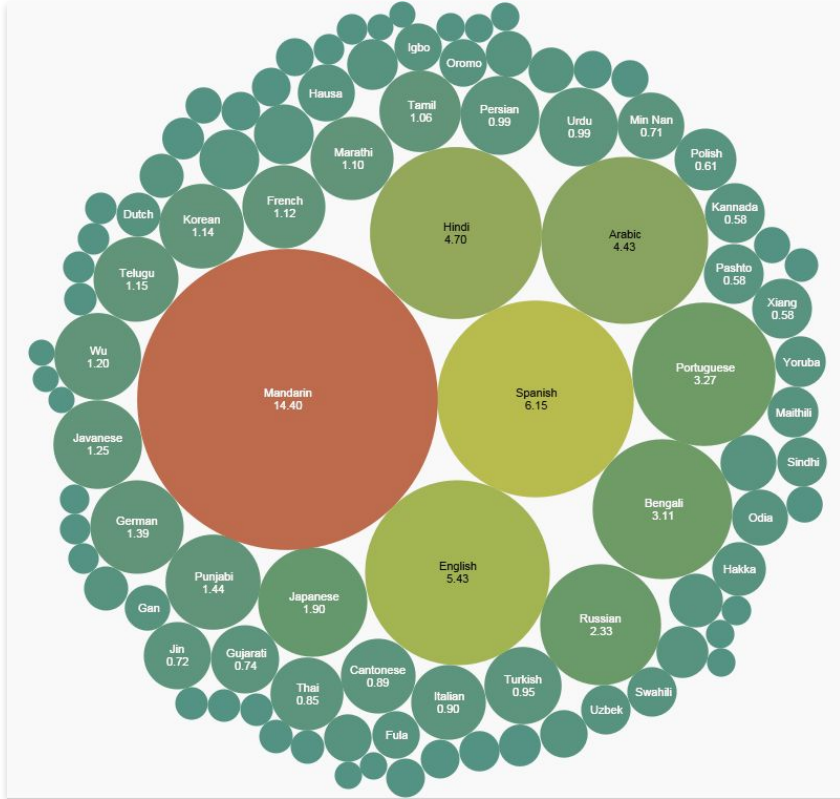


datasets are imperfect

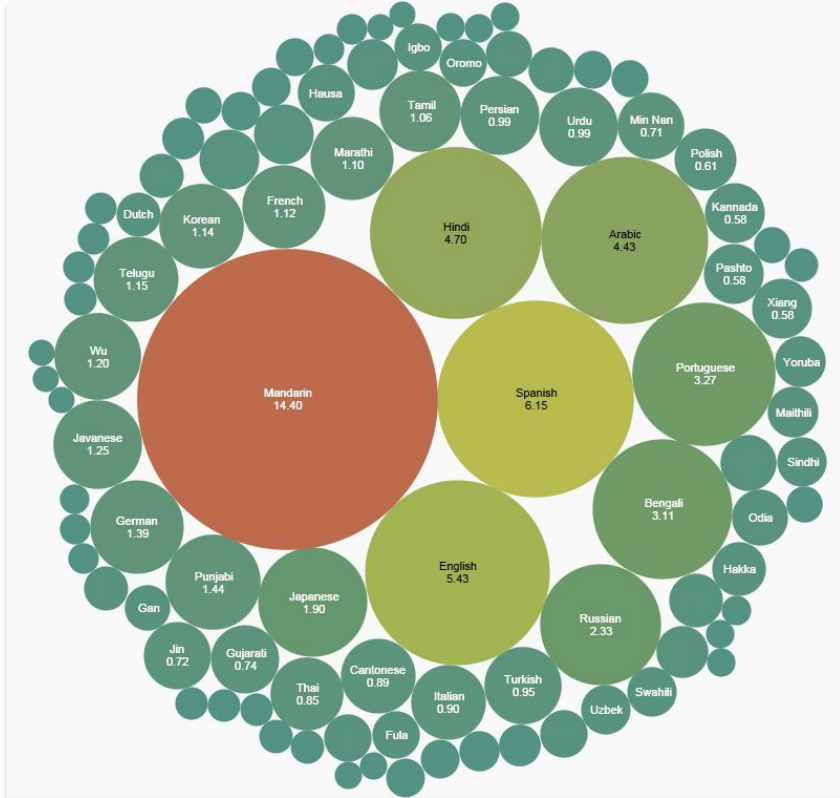
why? how can we **fix them**?







- Speech commands for the **whole planet?**



- Speech commands for the **whole planet?**
- For **more than** just voice assistants



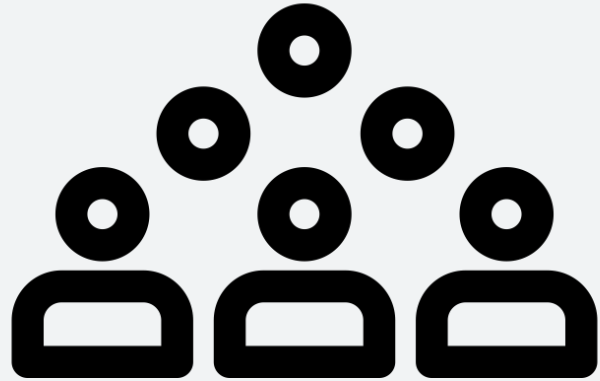
Crowdsourcing to the rescue!

Common Voice

<https://commonvoice.mozilla.org>

Common Voice

- **Crowdsourcing** platform





Speak

Listen

1/5 Clips

START RECORDING 1

2

3

4

5

Frederick and Nelson
continued the practice after
absorbing the brand.

Click  then read the sentence aloud

His wife, Edith Jane, died a few
months later and is buried
with him.



Shortcuts

Report

Skip >>

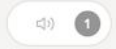
SUBMIT



Speak

Listen

1/5 clips



2

3

4

5

Blaise has since appeared in
Orville Taylor.

Click ▶ *did they accurately speak the sentence?*

The unconventional nature of
the tubular girder bridge was
not widely accepted.

👍 YES ▶ 👎 NO

Shortcuts

Report

Skip >>

Common Voice

- Crowdsourcing platform
- Over **50,000 volunteers**

The screenshot shows the Common Voice website. At the top, there is a navigation bar with the Mozilla logo, the text 'Common Voice', and links for 'CONTRIBUTE', 'DATASETS', 'LANGUAGES', and 'ABOUT'. On the right side of the navigation bar, there are icons for a microphone, a play button, and a language selector set to 'EN'. Below the navigation bar, the main content area is divided into two columns. The left column is titled 'Speak' and features a microphone icon and the text 'Donate your voice'. The right column is titled 'Listen' and features a play button icon and the text 'Help us validate voices'. Below these columns, there is a large text block that reads: 'Common Voice is Mozilla's initiative to help teach machines how real people speak.' To the right of this text is a paragraph explaining the project's goal: 'Voice is natural, voice is human. That's why we're excited about creating usable voice technology for our machines. But to create voice systems, developers need an extremely large amount of voice data. Most of the data used by large companies isn't available to the majority of people. We think that stifles innovation. So we've launched Common Voice, a project to help make voice recognition open and accessible to everyone.' Below this text is a 'READ MORE' link. At the bottom of the page, there are two data visualization charts. The left chart is a line graph titled 'Hours Recorded' and 'Hours Validated'. It shows two lines: a red line for 'Hours Recorded' and a green line for 'Hours Validated'. The red line is consistently higher than the green line, and both show an upward trend over time. The x-axis is labeled with '1y', '9mo', '6mo', '3mo', and 'Today'. The y-axis ranges from 0 to 9.6k. The red line is currently at 9.0k and the green line is at 7.1k. The right chart is a bar chart titled 'Voices Online Now'. It shows the number of voices online at different times of the day. The x-axis is labeled with times from 11PM to 08AM. The y-axis ranges from 0 to 90. The bars are blue, and the total number of voices online is 15. A dropdown menu is set to 'ALL'.

Common Voice

CONTRIBUTE DATASETS LANGUAGES ABOUT

Log In / Sign Up EN

Speak

Donate your voice

Listen

Help us validate voices

Common Voice is Mozilla's initiative to help teach machines how real people speak.

Voice is natural, voice is human. That's why we're excited about creating usable voice technology for our machines. But to create voice systems, developers need an extremely large amount of voice data.

Most of the data used by large companies isn't available to the majority of people. We think that stifles innovation. So we've launched Common Voice, a project to help make voice recognition open and accessible to everyone.

[READ MORE](#)

Hours Recorded: 9.0k
Hours Validated: 7.1k

Voices Online Now: 15

Common Voice

- Crowdsourcing platform
- Over 50,000 volunteers
- **54 different languages**

We're building

an open source, multi-language dataset of voices that anyone can use to train speech-enabled applications.

We believe that large, publicly available voice datasets will foster innovation and healthy commercial competition in machine-learning based speech technology.

Common Voice's multi-language dataset is already the largest publicly available voice dataset of its kind, but it's not the only one.

Look to this page as a reference hub for other open source voice datasets and, as Common Voice continues to grow, a home for our release updates.

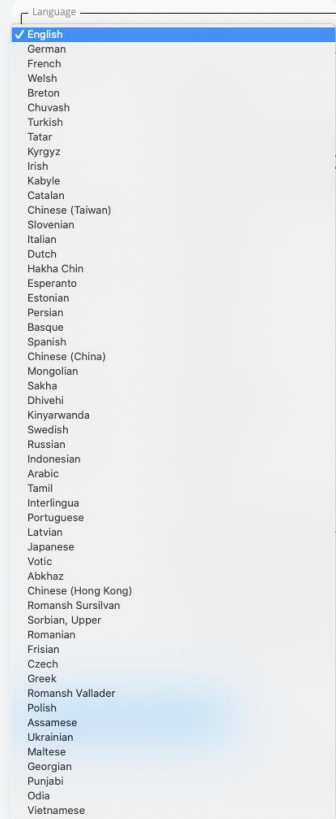
Common Voice is a multi-language dataset of voices that anyone can use to train speech-enabled applications. It is the largest publicly available voice dataset of its kind, but it's not the only one. Look to this page as a reference hub for other open source voice datasets and, as Common Voice continues to grow, a home for our release updates.

Validated Hours

5,671

Recorded Hours

7,226



What's inside the Common Voice dataset?

Common Voice


- Crowdsourcing platform
- Over 50,000 volunteers
- 54 different languages
- Goal: **speech recognition** for *all languages* on the planet

We're building
an open source, multi-language dataset of voices that anyone can use to train speech-enabled applications.

We believe that large, publicly available voice datasets will foster innovation and healthy commercial competition in machine-learning based speech technology.


Common Voice's multi-language dataset is already the largest publicly available voice dataset of its kind, but it's not the only one.

Look to this page as a reference hub for other open source voice datasets and, as Common Voice continues to grow, a home for our release updates.



Language



SIZE	50 GB
VERSION	en_1932h_2020-06-22
VALIDATED HR. TOTAL	1,469
OVERALL HR. TOTAL	1,932
LICENSE	CC-0
NUMBER OF VOICES	61,528
AUDIO FORMAT	MP3
SPLITS	Accent 23% United States English, 8% England English, ... Age 23% 19 - 29, 14% 30 - 39, ... 47% Male, 14% 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.

Validated Hours
5,671

Recorded Hours
7,226



display a menu

What's inside the Common Voice dataset?

File Structure

- **Valid**
 - At least 2 people listen to them, and the majority of those listeners say the audio matches the text



File Structure

- **Valid**
 - At least 2 people listen to them, and the majority of those listeners say the audio matches the text
- **Invalid**
 - At least 2 listeners, and the majority say the audio does not match the clip



File Structure

- **Valid**
 - At least 2 people listen to them, and the majority of those listeners say the audio matches the text
- **Invalid**
 - At least 2 listeners, and the majority say the audio does not match the clip
- **Other**
 - All other clips, i.e., fewer than 2 votes, or those that have equal valid and invalid votes, are labelled “other”



Interesting Attributes

- **Permissive license**

We're building
an open source, multi-language dataset of voices that anyone can use to train speech-enabled applications.

We believe that large, publicly available voice datasets will foster innovation and healthy commercial competition in machine-learning based speech technology.

Common Voice's multi-language dataset is already the largest publicly available voice dataset of its kind, but it's not the only one.

Look to this page as a reference hub for other open source voice datasets and, as Common Voice continues to grow, a home for our release updates.

Language
English

SIZE	50 GB
VERSION	en_1932h_2020-06-22
VALIDATED HR, TOTAL	1,469
OVERALL HR, TOTAL	1,932
LICENSE	CC-0
NUMBER OF VOICES	61,528
AUDIO FORMAT	MP3
SPLITS	Accent 23% United States English, 8% England English, ... Age 23% 19 - 29, 14% 30 - 39, ... 47% Male, 14% Female

[Enter Email to Download](#)

Validated Hours
5,671

Recorded Hours
7,226

What's inside the Common Voice dataset?

Interesting Attributes

- Permissive license
- **Many contributors**

We're building
an open source, multi-language dataset of voices that anyone can use to train speech-enabled applications.

We believe that large, publicly available voice datasets will foster innovation and healthy commercial competition in machine-learning based speech technology.

Common Voice's multi-language dataset is already the largest publicly available voice dataset of its kind, but it's not the only one.

Look to this page as a reference hub for other open source voice datasets and, as Common Voice continues to grow, a home for our release updates.

Language: English

SIZE	50 GB
VERSION	en_1932h_2020-06-22
VALIDATED HR, TOTAL	1,469
OVERALL HR, TOTAL	1,932
LICENSE	CC-0
NUMBER OF VOICES	61,528
AUDIO FORMAT	MP3
SPLITS	Accent 23% United States English, 8% England English, ... Age 23% 19 - 29, 14% 30 - 39, ... 47% Male, 14% 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.

Validated Hours
5,671

Recorded Hours
7,226

Display a menu

What's inside the Common Voice dataset?

Interesting Attributes

- Permissive license
- Many contributors
- Comes with metadata

We're building an open source, multi-language dataset of voices that anyone can use to train speech-enabled applications.

We believe that large, publicly available voice datasets will foster innovation and healthy commercial competition in machine-learning based speech technology.

Common Voice's multi-language dataset is already the largest publicly available voice dataset of its kind, but it's not the only one.

Look to this page as a reference hub for other open source voice datasets and, as Common Voice continues to grow, a home for our release updates.

Language: English

SIZE	50 GB
VERSION	en_1932h_2020-06-22
VALIDATED HR, TOTAL	1,469
OVERALL HR, TOTAL	1,932
LICENSE	CC-0
NUMBER OF VOICES	61,528
AUDIO FORMAT	MP3

SPLITS

- Accent
 - 23% United States English,
 - 8% England English, ...
- Age
 - 23% 19 - 29, 14% 30 - 39, ...
- Gender
 - 47% Male, 14% 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.

Validated Hours: 5,671

Recorded Hours: 7,226

What's inside the Common Voice dataset?

Common Voice

Download the Single Word Target Segment

This is a use case driven segment containing data to power spoken digit recognition, yes / no detection, and wakeword testing data for [Firefox Voice](#).

Validated
Hours

68



Recorded
Hours

120



Languages

18



Enter Email to Download

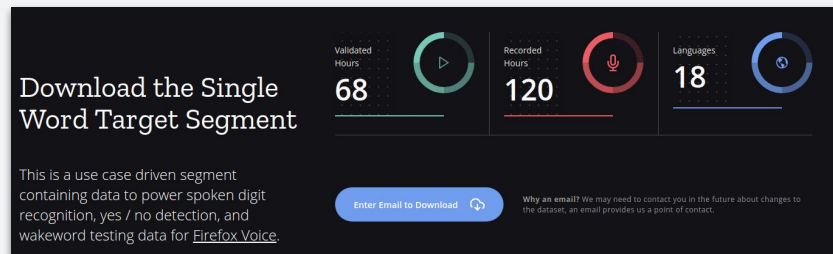


Why an email? We may need to contact you in the future about changes to the dataset, an email provides us a point of contact.

Single Word Target Segment

A *speech commands-style* dataset for **18 languages**

- “Yes” // “no”
- “hey” & “Firefox”
- **digits 0-9**



Download the Single Word Target Segment

Validated Hours: 68

Recorded Hours: 120

Languages: 18

This is a use case driven segment containing data to power spoken digit recognition, yes / no detection, and wakeword testing data for [Firefox Voice](#).

[Enter Email to Download](#)

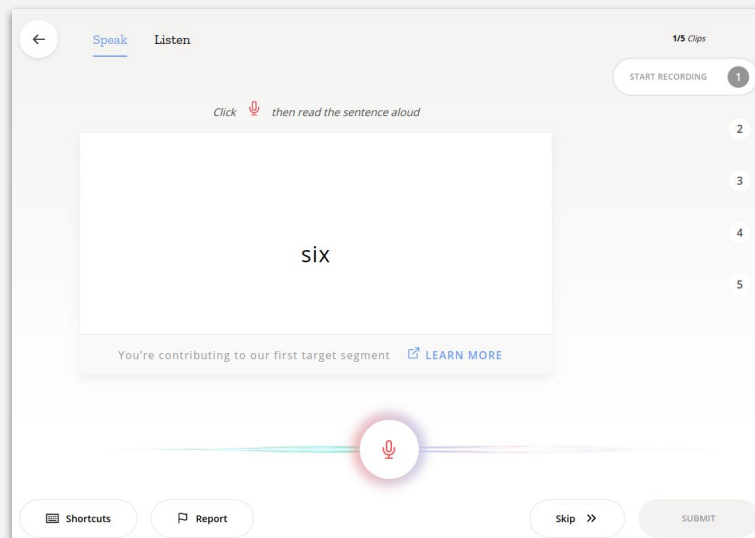
Why an email? We may need to contact you in the future about changes to the dataset, an email provides us a point of contact.

The banner features three circular progress indicators: a green play button icon for 'Validated Hours' (68), a red microphone icon for 'Recorded Hours' (120), and a blue globe icon for 'Languages' (18). Below the statistics is a blue button with a white right-pointing arrow and the text 'Enter Email to Download'. To the right of the button is a small text note explaining the email requirement.

Speech Recognition Diversity and Reach

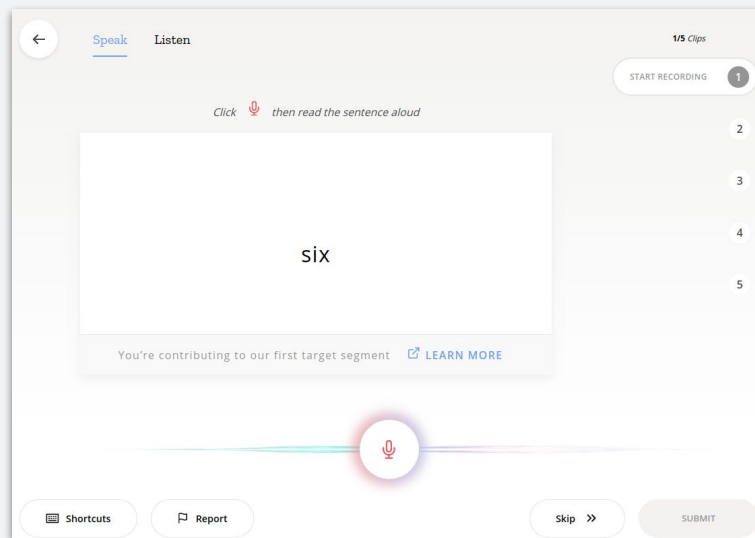
- **Common Voice**

- Permissive license
- Minority languages



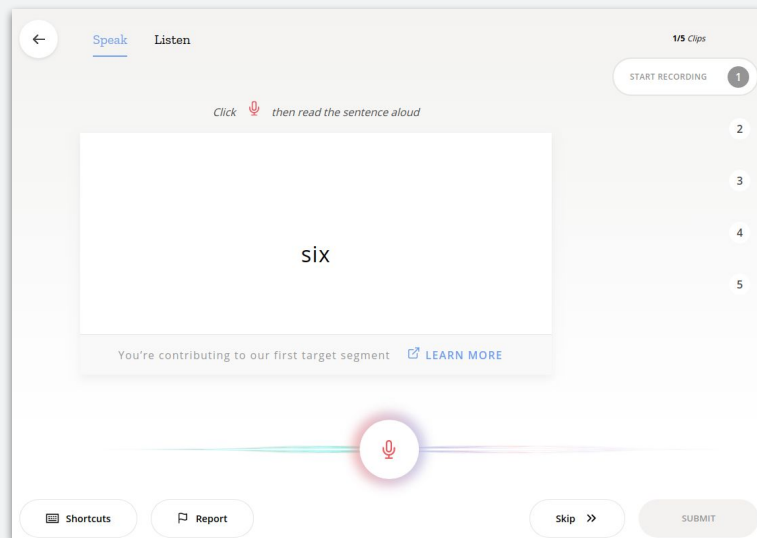
Speech Recognition Diversity and Reach

- **Common Voice**
 - Permissive license
 - Minority languages
- **Ease-of-use, wide reach**
 - Browser-based
 - Community can add new languages



Speech Recognition Diversity and Reach

- Common Voice
 - Permissive license
 - Minority languages
- Ease-of-use, wide reach
 - Browser-based
 - Community can add new languages
- You can contribute!



Common Voice

<https://commonvoice.mozilla.org>



today!

ML in the Navajo Nation

How can **you** use ML
in **your community**!

