2nd June, 2020

Common Voice Overview

Common Voice is a crowdsourcing platform. We're building an open source, multi-language dataset of voices anyone can use to train speech-enabled applications. This past week Common Voice achieved a new milestone with 5k validated voice hours contributed. As of today we're collecting voice donations across 54 languages.

2020 has been a busy and successful year for the project. Today I'll highlight three of the biggest updates as we round out H1...

Moving to Kubernetes

In the first quarter Common Voice engineers, Jenny Zhang and Riley Shaw, undertook substantial improvements to the infrastructure of Common Voice. With the help of Alberto del Barrio in the IT-SRE group, our engineers successfully migrated the Common Voice deployment infrastructure to Kubernetes in mid-April. This resulted in immediate site stability and performance improvements.

Enabling recording for Common Voice via mobile Safari

At the end of Q1 we were able to overcome a long-standing platform limitation for contribution via iOS devices. With the MediaRecorder standard WebAPI still not supported in various Safari browsers, Riley Shaw implemented a lightweight polyfill for MediaRecorder that has enabled recording on Safari. This has also allowed us to decommission the iOS Common Voice app. We are now redirecting visitors, contributors, and developers interested in the Common Voice iOS app to our mobile website enabling the team to focus our development efforts on the web platform.

Collecting a target dataset segment for Single Words

On May 20th Common Voice released the capability to collect voice data for a specific purpose or use-case. We're putting this ability to the test with data collection for a single word target segment. This will eventually enable 1) *spoken digit* recognition, 2) *yes* and *no* detection and 3) data for *Hey Firefox* wake word testing.

This targeted data collection will immediately benefit two collaborations: the first being with Mozilla Fellow, Josh Meyer, who is working to benchmark the accuracy of Mozilla's open source voice recognition engine, Deep Speech, in multiple languages for a similar task.

Our colleagues in Emerging Technologies are also testing and training wake word options for Firefox Voice. They reached out, curious if Common Voice communities could help generate voice data for *Hey Firefox* in multiple languages.

Our first week of collection proves that the answer to that is yes. In the first 7 days of collection we saw over **5700 speakers** contributing over **81K clips** in 13 languages. The overall **traffic to the site saw ~10x growth** compared to a typical week and the rate of **overall contribution increased by ~3x**. As of today, our 13th day of collection, we're at >8400 unique speakers contributing >136K clips.

Thanks to our partners in Marcomms who launched a new tab snippet to help drive that traffic our way and shout out to our Open Innovation Community Development team, who launched a community campaign to engage the Rebel Alliance in this effort.

Follow along with us in slack at #common-voice or #voices and go spend 3 mins contributing your voice at voice.mozilla.org.