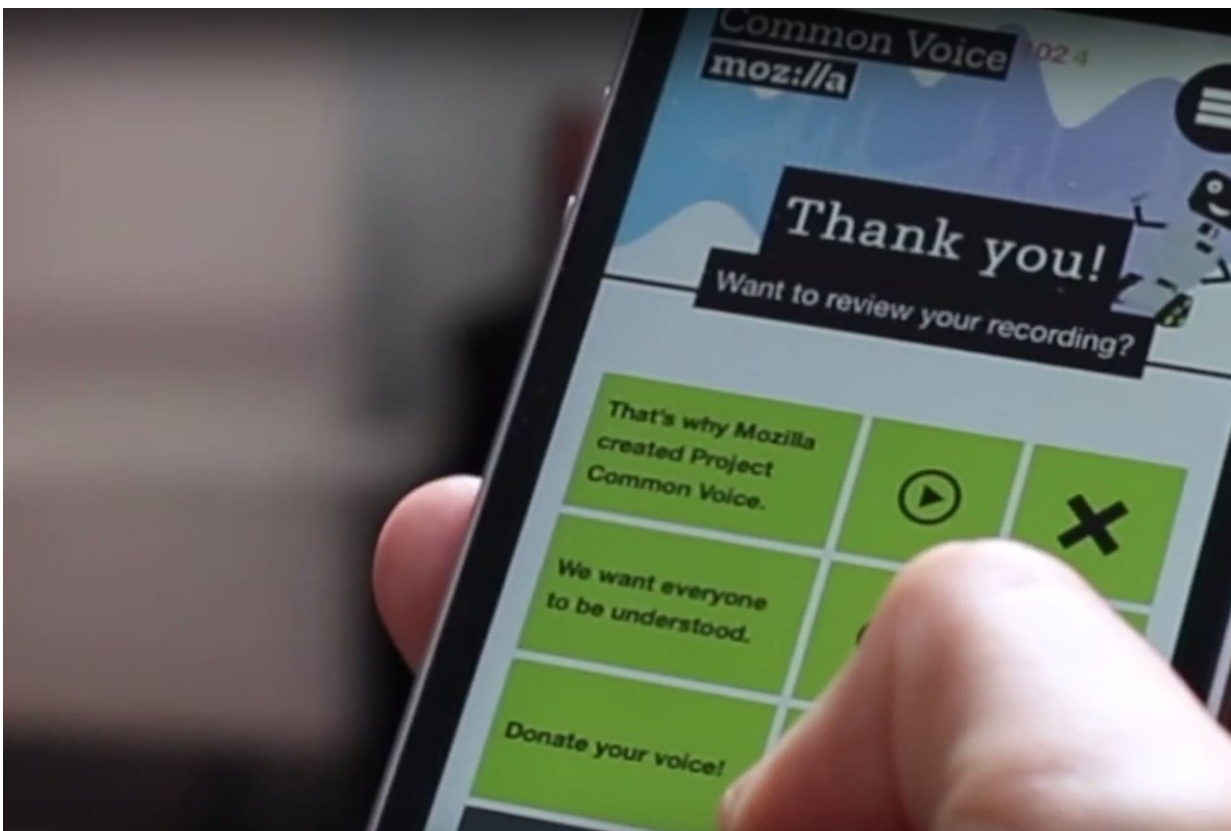


Mozilla releases transcription model and huge voice dataset

November 30 2017, by Bob Yirka



Credit: Mozilla

(Tech Xplore)—Mozilla (maker of the Firefox browser) [has announced](#) the release of an open source speech recognition model along with a large voice dataset. The release marks the advent of open source speech

recognition development. Sean White, chief executive of Mozilla, suggests in the announcement that it will "result in more internet-connected products that can listen and respond to us than ever before."

Up until now, virtually every commercially available [speech recognition](#) product has come from a major company, such as Microsoft or Google. This, White notes, is because such applications require a huge investment and an equally huge voice [dataset](#) to learn how to recognize and interpret human speech. Mozilla, he adds, promotes efforts to make technology more available to developers and users alike. To that end, the company set a goal of developing a [speech recognition](#) model that could be made publicly available for free, which it calls Project DeepSpeech. Along with that goal, the company created Project Common Voice, a website where people can volunteer to record their voices and to transcribe recordings made by others. White claims the dataset now holds voice data for over 20,000 people with 400,000 samples that can be downloaded, making it the second-largest publicly available dataset in the world.

Project DeepSpeech is based on work done by Baidu's Deep Speech [project](#) and uses Google's TensorFlow machine learning tool, which is open source. The newly released model allows developers to create applications with voice recognition abilities without having to pay royalties, and the Project Common Voice dataset allows it to be trained using a huge free voice dataset. The end result could be an onslaught of new applications, some likely in the form of apps available for smartphone users. White claims that the transcription engine has an error rate of just 6.5 percent, which is very close to what humans can do, which means new apps should be better at recognizing what users have to say than earlier products.

White also notes that currently, the model and [voice](#) dataset only work for English, but promises that multiple languages will soon be supported

as well, some as early as next year. He also encourages people to visit the Common Voice website to add to the dataset, making it better for everyone.

© 2017 Tech Xplore

Citation: Mozilla releases transcription model and huge voice dataset (2017, November 30)
retrieved 19 April 2024 from

<https://techxplore.com/news/2017-11-mozilla-transcription-huge-voice-dataset.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.