

Common Voice is a Mozilla initiative to address biases in speech recognition and make building voice technologies more accessible and inclusive. We're building an open-source, publicly available dataset of transcribed voice clips. We want to gather up to 10,000 hours of validated speech clips for as many different languages as possible, representing a diversity of accents, ages, and gender. That means we need voices like yours!

By publishing our dataset publicly, anyone can use it for research and to build and train voice-enabled applications. Our goal is to create a dataset representative of the real diversity of speech around the world to help meet the growing demand for innovative voice-based solutions. With the help of our contributors, we can empower local communities to build inclusive voice technologies in their own languages.

Want to help?

If you or your organization are interested in open voice interfaces, we'd love to talk (and listen).

REACH OUT

commonvoice@mozilla.com

CONTRIBUTE YOUR VOICE

voice.mozilla.org

Common Voice

moz://a

