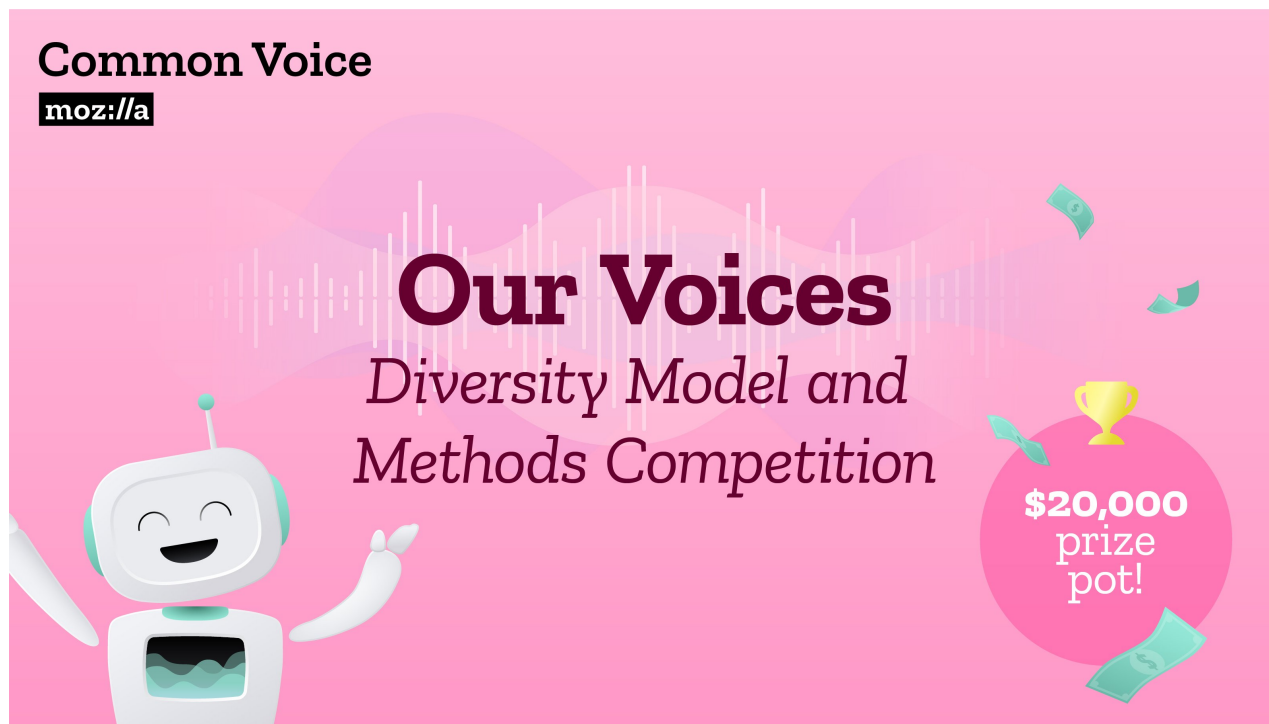


# MCV 'Our Voices' Model and Methods Competition - Taking Part

 [foundation.mozilla.org/en/blog/our-voices-taking-part](https://foundation.mozilla.org/en/blog/our-voices-taking-part)

June 29, 2022



Help us teach robots how real people speak!

---

## What is this competition?

We want to see - and incentivise! - great diversity, equity and inclusion-conscious work being done with the Common Voice dataset.

We are running a model and methods competition with three broad themes, plus an open category.

**Gender**

**Variant and Accent**

**Methodologies**

**Open call**

---

## What are we looking for?

Your entry must be a diversity, equity and inclusion-conscious Model or Method under one of the following categories.

It must primarily make use of Mozilla Common Voice data from the 11th release (September 2022).

Outside of this, we are being deliberately open-ended.

However, here are some *illustrative examples*;

**Gender:**

+ An STT model for an under-resourced language that performs equally well for women

**Variant, Dialect or Accent:**

+ Proof of concept for an under-served language variant delivered with a small 'toy' corpus

+ Accent classifiers by, and for, a community

**Methods and Measures:**

+ A benchmark bias corpus

+ Dataset audit methodology

**Open:**

+ Exciting DEI work primarily using Common Voice that doesn't fit into the categories above

**How are you making sure it's easy for all languages to participate?**

---

We are actively encouraging submissions at proof of concept stage that use a small or 'toy' corpus

Our methodology and methods category enables teams to submit outlines for tools that they do not yet have the resources to build out further

We have allowed a month of development time to accomodate those relying on CPU / slower compute

Languages will be judged within 'Bands' - high resource, medium resource and low resource - to ensure a fairer competition between languages that exist in different contexts

We are creating a flexible, holistic rubric that makes it possible for judges to look at ecosystem value-add factors beyond performance metrics like Word Error Rate

For marginalised communities who have governance concerns about releasing their model under an open source license, they are welcome to submit with an explanation to that effect, and this will be considered accordingly

## Who are the judging panel?

---

Professor Francis Tyers - Computational Linguistics Advisor, Mozilla Foundation & Academic, University of Indiana Indiana

Dr Vitaly Lavrukhin - Principal Applied Research Scientist, NVIDIA

Wiebke Hutiri- PhD Candidate at Delft University of Technology - Fairness in Voice Tech

Dr Abeba Birhane - AI Fellow Mozilla

Rebecca Ryakitimbo - Community Fellow, Kiswahili

Britone Mwasaru - Community Fellow, Kiswahili

Dr Josh Meyers- Co-Founder, Coqui

Stefania Delprete - Data Scientist and Italian MCV Community Rep

Kathy Reid - PhD Candidate at Australia National University - Bias in Speech Tech, Open Source

Gabriel Habayeb - Senior Data Engineer, Mozilla Foundation

## What is the timeline?

---

### June 30th

Competition announced publicly

### July 6th

MCV 10 released

### July-September

Teams evaluate the data, then mobilise to grow and enhance datasets as needed for their ideas. There will also be changes to ask questions of the MCV team via an Ask Me Anything.

### September 14th

MCV 11 released

### September 15th

Competition opens

## **October 12th**

Competition closes

## **Mid October (date TBC)**

Judges meet, agree and Mozilla notifies winners

## **Mid-Late October (date TBC)**

Winners announcement ceremony with demos at Speech Summit

## **What are the prizes?**

---

1st place winners for the 4 categories will each receive \$USD 2000. We have allocated a total pot of \$USD 20,000, to allow us to a) make awards within each language resource band, and b) potentially make multiple awards if there are multiple 'winning quality' entries.

## **How do I register my interest?**

---

Just [register your interest with this form](#), and you'll receive a participant pack with guidance, resources, advice and more to help you

## **What are the rules?**

---

### **COMMON VOICE - Our Voices - OFFICIAL RULES**

Sponsor. Sponsor of The Voice Challenge (the “Challenge”) is Mozilla Foundation, 2 Harrison St. #175, San Francisco, California 94105.

No Purchase Necessary; Entry Instructions. NO PURCHASE OR PAYMENT OF ANY MONEY IS NECESSARY TO ENTER. A PURCHASE WILL NOT IMPROVE THE CHANCES OF WINNING. This is a contest of skill. Odds of winning the Challenge depend on the number and quality of eligible entries received during the Challenge Period. VOID WHERE PROHIBITED AND WHERE ANY REGISTRATION, BONDING OR LOCALIZATION REQUIRED.

To enter and participate in the Challenge, you must fully complete the Challenge submission form on the Challenge page available at <https://github.com/common-voice/ourvoicesmodelcompetition> and submit the required entry submissions as provided on the Challenge page and in these Official Rules. You may enter as an individual or as part of a Team (as defined below). Teams may have no more than 5 members. By submitting your entry, you are agreeing to these Official Rules on behalf of yourself and all Team members, and by participating all Team members agree to the Official Rules as a participating entrant. A Team will collectively also be deemed an entrant. Your entry

submission information must include (i) your name and email address, (ii) your project name, (iii) the voice recognition model, and (iv) all other information listed as required in the contest description below, and, if applicable, the names of all Team members. Your initial entry must also contain a short written presentation of your idea. All requested entry information must be provided. One entry per individual or Team. An individual may only participate on one team. Any attempt by any participant to obtain more than the stated number of entries by using multiple/different email or addresses, accounts, identities, or any other methods will void that participant's entries and that participant may be disqualified.

In the event of a dispute as to any entrant, the authorized account holder of the email address associated with the entry will be deemed to be the entrant. The "authorized account holder" is the natural person assigned an email address by an Internet access provider, online service provider or other organization responsible for assigning email addresses for the domain associated with the submitted address. A potential winner may be required to show proof of being the authorized account holder. In the event of an ongoing dispute as to any entrant, Sponsor has the right to make determination as to entrant(s) in its sole discretion.

For a group of individuals to enter the contest (a "Team"), the Team must agree to and designate one person as the agent of the Team to submit the entry and to accept the prize on behalf of the Team. Each individual on a Team must agree to these Official Rules. You may not make changes to your Team once registered to participate. By submitting an entry, the submitter represents that all Team members have read and agreed to these Official Rules.

All entries must meet the following criteria:

A voice recognition model, or bias evaluation and mitigation methodology

Each entry must include a short, written description (e.g., write-up, PowerPoint slides) that presents the model and why it matters. You may also include any demonstration material you would like (e.g., a video with your own subtitles or a blog post about your process)

Entries may not contain likenesses of any individuals who are under 18 years of age, and may not contain the likeness of any individuals who have not provided their authorization. By submitting likeness of any individual, you represent that you have received permission from such person to include their likeness.

Entries may not contain material that is obscene, defamatory, libelous, threatening, pornographic, racially or ethnically offensive, or encourages conduct that would be considered a criminal offense, give rise to civil liability, or violate any law. Entries must be appropriate for viewing by the general public; appropriateness will be determined by Sponsor.

Entries must be original, exclusively created and owned by entrant, and the entrant must have all rights necessary to submit the entry.

Sponsor reserves the right to reject any entry for any reason.

**Challenge Period.** This Challenge begins at 12:01 am Pacific Time on September 15, 2022 and ends at 11:59 pm Pacific Time on October 12, 2022 (the “Challenge Period”).

Sponsor’s computer is the official time-keeping device for the Challenge. All entries must be received during the dates and times specified in the Challenge Period.

**Eligibility.** In order to be eligible, participants must be at least the age of majority in their jurisdiction of residence. Employees of Sponsor and its parent and affiliate companies as well as the immediate family (spouse, parents, siblings and children) and household members of each such employee are not eligible. Before participating in the Challenge, you should also consider whether you or your team members have any conflicting obligations that prevent participation. If an entrant is an employee of a corporation, government or an academic institution, or enrolled as a student, it is his or her sole responsibility to review, understand and abide by his or her employer’s, or academic institution’s policies regarding eligibility to participate in the Challenge. If an entrant is found to be in violation of his or her employer’s or academic institution’s policies, then he or she will be, and the Team may be, disqualified from participating in the Challenge and from being awarded or retaining any prize. Sponsor disclaims any and all liability or responsibility for disputes arising between an employee or student and his or her employer or academic institution related to the Challenge.

Contestants are not eligible to receive prizes if they are on the US Specifically Designated Nationals (SDN) list or if there are sanctions against the contestant’s country such that Mozilla is prohibited from paying them.

**Finalists and Winners.** We will select winners for (a) gender-equal performance and (b) dialect/accent performance in each of three categories of languages based on Common Voice Language Resource Bands, as designated by the Mozilla Common Voice team and the language communities themselves.

Band A consists of languages with a corpus of 750 sentences or fewer.

Band B consists of languages with a corpus of between 751 and 2,000 sentences.

Band C consists of languages with a corpus of more than 2,001 sentences.

There will also be at least one winner for a third category: methods and measures.

A winner for the Open Call will also be selected, if appropriate.

There will also be ten (10) runners up per category. Winners and runners up will be announced on or about October 20th, 2022. The overall “best” entries will be selected as winners. Winners must meet a certain minimum quality threshold. We reserve the right

to decide not to select a winner in a category if there are no entries that meet the quality threshold. Judging will be done by Sponsor or its designees, who shall have sole discretion in determining winners based on the following equally weighted criteria:

Word Error Rate (“WER”) - how many word recognition mistakes your model makes when used on a fresh dataset

WER Score when balanced by gender or Accent demographics (as per competition theme)

Utility - this is a judgment scoring by panelists evaluating how effective, original and useful your method or measure would be

Social need / ecosystem value - whether this model adds value to the universe of other models for the same language. We do not disqualify submissions that are not open source, but when considering ecosystem value-add within the wider rubric, we will consider the license under which your work is available.

Deployability rating - this is a judgment scoring by panelists evaluating how easy would this be to install in an application

Environmental impact rating (via GPU usage) - this has two components - expert panel rating plus required provision of their processing stats - aka how 'hungry' is your model? is it written to be efficient?

Prizes.

1st Place Winner Gender - Band A - \$2000

1st Place Winner Gender - Band B - \$2000

1st Place Winner Gender - Band C - \$2000

1st Place Winner Variant / Accent - Band A - \$2000

1st Place Winner Variant / Accent - Band B - \$2000

1st Place Winner Variant / Accent - Band C - \$2000

1st Place Winner Methods - \$2000

1st Place Winner Open / wild card (No bands) - \$2000

Runner up (10 per category above) will receive Common Voice merch such as a T-shirt.

Prizes will be delivered shortly after winners are announced.

The aggregate retail value of all the prize(s) is up to approximately USD20,000. Team members may determine how to divide prizes among themselves. Team Prizes will be awarded to the agent of the Team. No substitution, assignment or transfer of the prize is permitted, except by Sponsor, who has the right to substitute a prize with another of

comparable or greater value. Winners are responsible for all taxes and fees associated with the receipt and/or use of the prize, and winner may be required to provide tax information prior to receiving the prize.

Each winner agrees to self-report to applicable taxing authorities, as may be required by applicable laws.

Prize monies should be retained by individuals only in conformity with any applicable policies of his or her employers, academic institutions, or government regarding participation in and receipt of promotional consideration relating to the Challenge and receipt and retention of prize. If a government, employer's or school's policies are applicable, it is the entrant's sole and ultimate responsibility, in consultation with his or her government, employer or school, to determine how and if a prize will be retained and/or distributed and accounted for and we assume no responsibility for the decisions made by such government, employers or schools regarding this issue.

**Conditions of Participation.** By submitting an entry for or participating in this Challenge, you agree to abide by these rules and any decision Sponsor makes regarding this Challenge, which Sponsor shall make in its sole discretion. Sponsor reserves the right to disqualify and prosecute to the fullest extent permitted by law any participant or winner who, in Sponsor's reasonable suspicion, tampers with Sponsor site, the entry process, intentionally submits more than a single entry, violates these rules, engages in fraud, attempted fraud, or acts in an unsportsmanlike or disruptive manner.

By submitting an entry for or participating in this Challenge, you agree to abide by Mozilla's Community Participation Guidelines. Violation of the Community Participation Guidelines ("CPG") by any Entrant will result in disqualification for that Entrant. The CPG is available at: <https://www.mozilla.org/en-US/about/governance/policies/participation/>

Each Team is solely responsible for its own cooperation and teamwork. In no event will Sponsor officiate in any dispute between or among any Team(s) or its/their members regarding their conduct, participation, cooperation or contribution. In the event that any dispute cannot be resolved, Sponsor reserves the right in its sole discretion to make a determination as to the identity of Team members or the team agent, and it may disqualify Teams and/or Team members in its sole discretion.

**Intellectual Property.** Ownership of the pre-existing underlying intellectual property of the entrant remains the property of the entrant subject to Sponsor's rights to reprint, display, reproduce, perform, use, and exhibit the entry and materials and information submitted, for the purpose of administering and promoting the Challenge and for any business, marketing and advertising purposes for the benefit of Sponsor. By participating in the Challenge, each entrant grants to Sponsor a non-exclusive, worldwide, fully paid, royalty-free, perpetual, irrevocable, transferable, sub-licensable, license to reprint, display, reproduce, exploit, perform, use, and exhibit (including the right to make derivative works of) the entry and materials and information submitted on and in



connection with the Challenge or use or receipt of the prize for any and all purposes in any medium. Each participating entrant hereby warrants that any entry and other materials and information provided by entrant are original with entrant and do not violate or infringe upon the copyrights, trademarks, rights of privacy, publicity, moral rights or other intellectual property or other rights of any person or entity, and do not violate any rules, regulations, or laws. If the entry or information or materials provided by entrant contain any material or elements that are not owned by entrant and/or which are subject to the rights of third parties, entrant represents he or she has obtained, prior to submission of the entry and information or materials, any and all releases and consents necessary to permit use and exploitation of the entry and information and materials by Sponsor in the manner set forth in the Official Rules without additional compensation.

Each entrant warrants that the entry and materials and information provided do not contain information considered by entrant, its employees, employer, or personnel, or any other third party to be confidential, and that submission of the materials and information will not violate any agreements with third parties or policies that entrant is subject to. Entrant agrees that Sponsor has the right to verify the ownership and originality of all entries and that, upon Sponsor's request, entrant must submit a written copy of any release or permission entrant has received from a third party granting entrant the right to use such property. Entrant understands and acknowledges that in the event a submission is selected as a winning entry, and entrant's ownership, rights and the originality of the entry cannot be verified to the satisfaction of Sponsor or is in any other way ineligible, Sponsor may select an alternate winner based on the same judging criteria. Entrant acknowledges that other entrants may submit entries that are similar to yours and that they, or Sponsor, may already be considering or developing, or may subsequently consider or develop independent of the Challenge, content or ideas that are related or similar to yours. You acknowledge that this does not create in Sponsor or others any obligation or liability to you.

Disclaimer, Release and Limit of Liability. SPONSOR MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, REGARDING ANY PRIZE OR YOUR PARTICIPATION IN THE PROMOTION. BY ENTERING THE PROMOTION OR RECEIPT OF ANY PRIZE, EACH ENTRANT AGREES TO RELEASE AND HOLD HARMLESS SPONSOR AND ITS SUBSIDIARIES, AFFILIATES, SUPPLIERS, DISTRIBUTORS, ADVERTISING/PROMOTION AGENCIES, PARTNERS, AND PRIZE SUPPLIERS, AND EACH OF THEIR RESPECTIVE PARENT COMPANIES AND EACH SUCH COMPANY'S OFFICERS, DIRECTORS, EMPLOYEES AND AGENTS (COLLECTIVELY, THE "RELEASED PARTIES") FROM AND AGAINST ANY CLAIM OR CAUSE OF ACTION, INCLUDING, BUT NOT LIMITED TO, PERSONAL INJURY, DEATH, OR DAMAGE TO OR LOSS OF PROPERTY, ARISING OUT OF PARTICIPATION IN THE PROMOTION OR RECEIPT OR USE OR MISUSE OF ANY PRIZE. THE RELEASED PARTIES ARE NOT RESPONSIBLE FOR: (1) ANY INCORRECT OR INACCURATE INFORMATION, WHETHER CAUSED BY ENTRANTS, PRINTING ERRORS OR BY ANY OF THE EQUIPMENT OR PROGRAMMING ASSOCIATED WITH OR UTILIZED IN THE PROMOTION; (2) TECHNICAL FAILURES OF ANY KIND,

INCLUDING, BUT NOT LIMITED TO MALFUNCTIONS, INTERRUPTIONS, OR DISCONNECTIONS IN PHONE LINES OR NETWORK HARDWARE OR SOFTWARE; (3) UNAUTHORIZED HUMAN INTERVENTION IN ANY PART OF THE ENTRY PROCESS OR THE PROMOTION; (4) TECHNICAL OR HUMAN ERROR WHICH MAY OCCUR IN THE ADMINISTRATION OF THE PROMOTION OR THE PROCESSING OF ENTRIES; OR (5) ANY INJURY OR DAMAGE TO PERSONS OR PROPERTY WHICH MAY BE CAUSED, DIRECTLY OR INDIRECTLY, IN WHOLE OR IN PART, FROM ENTRANT'S PARTICIPATION IN THE PROMOTION OR RECEIPT OR USE OR MISUSE OF ANY PRIZE. No more than the stated number of prizes will be awarded. If someone cheats, or a virus, bug, bot, catastrophic event, or any other unforeseen or unexpected action or event affects the fairness and/or integrity of this Challenge, Sponsor reserves the right to cancel, change, or suspend this Challenge. This right is reserved whether the event is due to human or technical error. If a solution cannot be found to restore the integrity of the Challenge, we reserve the right, but are not required, to select winner(s) from among all eligible entries received before we had to cancel, change or suspend the Challenge.

Privacy and Use of Personal Information. Sponsor collects personal information from you when you enter this Challenge. Sponsor reserves the right to use any information collected in accordance with its privacy policy, which may be found at <https://www.mozilla.org/privacy/>.

GOVERNING LAW AND DISPUTES. THESE OFFICIAL RULES AND THE PROMOTION ARE GOVERNED BY, AND WILL BE CONSTRUED IN ACCORDANCE WITH, THE LAWS OF THE STATE OF CALIFORNIA, AND THE FORUM AND VENUE FOR ANY DISPUTE ARISING OUT OF OR RELATING TO THESE OFFICIAL RULES SHALL BE IN SANTA CLARA COUNTY, CALIFORNIA. IF THE CONTROVERSY OR CLAIM IS NOT OTHERWISE RESOLVED THROUGH DIRECT DISCUSSIONS OR MEDIATION, IT SHALL THEN BE RESOLVED BY FINAL AND BINDING ARBITRATION ADMINISTERED BY JUDICIAL ARBITRATION AND MEDIATION SERVICES, INC., IN ACCORDANCE WITH ITS STREAMLINED ARBITRATION RULES AND PROCEDURES OR SUBSEQUENT VERSIONS THEREOF ("JAMS RULES"). THE JAMS RULES FOR SELECTION OF AN ARBITRATOR SHALL BE FOLLOWED, EXCEPT THAT THE ARBITRATOR SHALL BE EXPERIENCED AND LICENSED TO PRACTICE LAW IN CALIFORNIA. ANY SUCH CONTROVERSY OR CLAIM WILL BE ARBITRATED ON AN INDIVIDUAL BASIS, AND WILL NOT BE CONSOLIDATED IN ANY ARBITRATION WITH ANY CLAIM OR CONTROVERSY OF ANY OTHER PARTY. ALL PROCEEDINGS BROUGHT PURSUANT TO THIS PARAGRAPH WILL BE CONDUCTED IN SANTA CLARA COUNTY, CALIFORNIA. THE REMEDY FOR ANY CLAIM SHALL BE LIMITED TO ACTUAL DAMAGES, AND IN NO EVENT SHALL ANY PARTY BE ENTITLED TO RECOVER PUNITIVE, EXEMPLARY, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING ATTORNEY'S FEES OR OTHER SUCH RELATED COSTS OF BRINGING A CLAIM, OR TO RESCIND THIS AGREEMENT OR SEEK INJUNCTIVE OR ANY OTHER

EQUITABLE RELIEF. SPONSOR'S TOTAL LIABILITY UNDER, ARISING OUT OF, OR RELATED TO THIS AGREEMENT SHALL NOT EXCEED THE TOTAL AMOUNT OF THE PRIZE OFFERED TO THE WINNER OF THE CHALLENGE.

Publicity Grant. Except where prohibited, participation in the Challenge constitutes entrant's consent to Sponsor's and Sponsor's designees' publication, broadcast, display, sharing and use of entrant's and Team's name, likeness, voice, image, persona, biographical information, entry and audio and visual content shared by entrant for any purposes in any media, worldwide, without further payment or consideration.

No Confidentiality. Submissions will be shared with others and portions of entries may be reproduced elsewhere, including in connection with Challenge-related publicity materials. You should not disclose any information in your entry that is proprietary or confidential. No confidential relationship is established between you and Sponsor in connection with your entry.

Winners List. Individuals may request the name of winners by submitting a self-addressed stamped envelope prior to August 10, 2022 to Mozilla Challenge Winners List Request, 2 Harrison St #175, San Francisco, CA 94107. Vermont residents may omit postage.

Subscribe to our newsletter

Help shape the future of the web for the public good. Join our Mozilla News email list to get action alerts & internet tips right to your inbox.