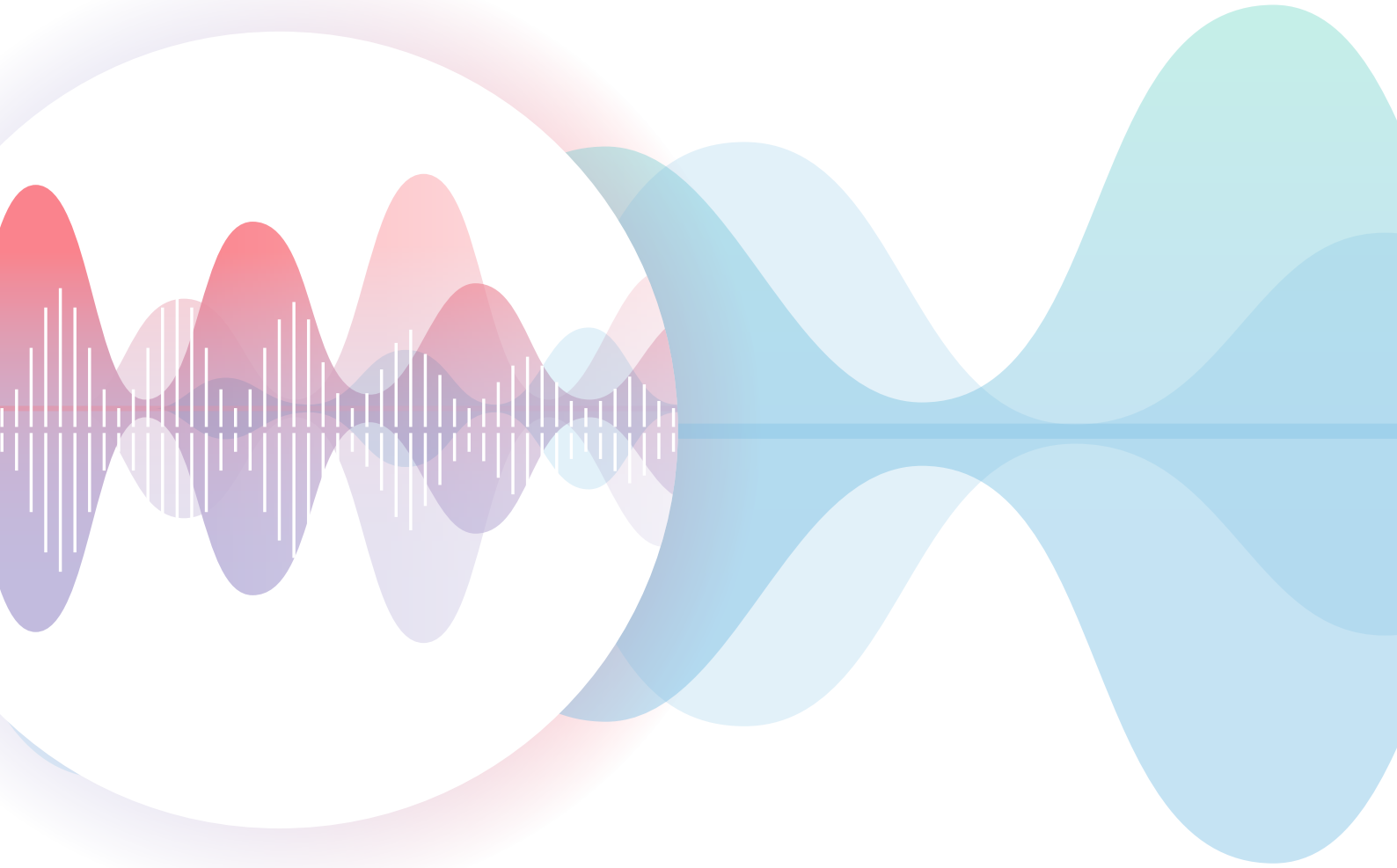
german cooperation
DEUTSCHE ZUSAMMENARBEIT

Implemented by:
giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

FAIR FORWARD
Artificial Intelligence for all.

moz://a

Creating community-driven datasets: Insights from Mozilla Common Voice activities in East Africa



Published by
Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn
FAIR Forward: Artificial Intelligence for All
Friedrich-Ebert-Allee 32 + 36,
53113 Bonn, Germany
E: info@giz.de
I: www.giz.de

Responsible
Mr. Daniel Brumund
FAIR Forward – Artificial Intelligence for All
E: daniel.brumund@giz.de

Authors
Jan Krewer, Independent consultant

Content Review
Daniel Brumund, Nadine Dammaschk, Mark Gachara, Christian Resch, GIZ
Chenai Chair, Mozilla Foundation

Design and Layout
Studio GOOD, Berlin

On behalf of
German Federal Ministry for Economic Cooperation and Development (BMZ)

GIZ is responsible for the content of this publication.

Bonn, Germany
March 2023

Contents

Executive summary

Since 2019, the Mozilla Foundation has been working with the GIZ FAIR Forward initiative to promote the creation and utilization of open voice data and technology in Kinyarwanda, Kiswahili and Luganda. These efforts include the crowdsourcing of large voice datasets together with local communities. The objective of this review is to document the strategies that have been deployed to create publicly available voice datasets in these three communities using the Mozilla Common Voice platform. It aims to provide existing and future voice communities as well as organizations who support them, with insights and recommendations, by exploring not only the necessary technical steps but also the social dynamics and structures at work.

An open voice community includes the group of people and organizations who have a shared interest in the establishment of a publicly available dataset for the future development of automated voice recognition tools in one specific language. The review shows that building a voice community first requires to have a good understanding of the socio-cultural and political context of one language. Making decisions about how language variants, dialects or accents are integrated in a dataset to build voice recognition models can be very sensitive. Similarly, defining the right vision to start mobilizing people means being able to align with the existing values, priorities and needs of potential community members.

Mobilizing contributors to build a voice dataset is difficult because the writing and validation of sentences or the recording and validation of voice clips can be a repetitive, time-consuming, and tedious task. It does not offer direct or intrinsic learning or personal development opportunities and the future benefits a publicly available voice dataset can offer to a community are indirect and uncertain. Participation in a voice community therefore is everything but a given. The time, energy and resources contributors put into the building of a dataset should always be recognized and valued, sometimes even rewarded. The Kinyarwanda, Kiswahili and Luganda communities managed to turn an activity that can be solitary into an empowering and social one by offering contributors opportunities to play, learn, connect with others, or to participate in a collective project whose objectives they share. One important difference between community-driven efforts to build a large dataset and similar private and public initiatives is that monetary or material forms of remuneration are not at the center of the value contributors perceive in their participation in the community's efforts. Contributors might have various motivations to become a member of a community and as many expectations on how they should benefit from this membership. Meeting all these expectations can be a big challenge for large communities.

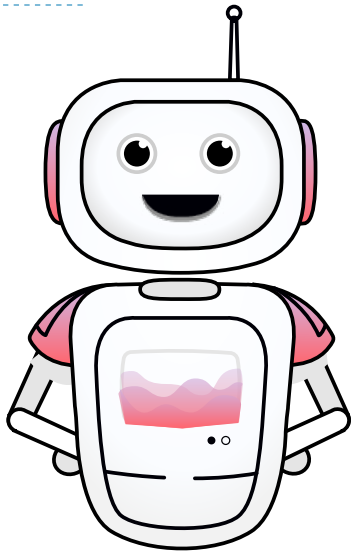
To successfully build a dataset that can be used to train effective and inclusive voice recognition models, there are two crucial factors: diversity and quantity. Contributors therefore need to represent all the speakers of the language. Diverse communities can be mobilized by building alliances with existing communities. While communication is important to raise awareness around an initiative, engagement seems to come mostly from peer mobilization through one-on-one communication or existing social structures.

Community efforts are not for free. Mobilizing a large community requires mobilizing funding, especially for the building of datasets that meet certain diversity and quality standards. Raising and managing funds requires communities to transition from spontaneous and organic distributions of roles and decision-making processes to more formalized rules of functioning. Kinyarwanda, Kiswahili, and Luganda voice communities have defined specific roles and responsibilities and separated certain tasks from others. This is crucial for covering the critical functions of open voice communities: coordination, mobilization, and contribution. Very often these distributions of roles and remuneration systems are developed along the way, allowing most active contributors to receive leading roles. However, an early open and transparent conversation about ground rules of engagement, including roles, remuneration, and visibility is essential to keep large groups engaged.

Putting a voice community together and maintaining it is time-consuming work. It requires strong and sustained commitment from local stakeholders. Partnerships with local organizations working on voice technology should be a priority of future communities, to ensure they have an even greater visibility and ownership. The Mozilla Foundation, together with its global and local partners, should further develop transparent governance frameworks that support inclusion and diverse participation, implement capacity building offers that multiply opportunities for local value creation and ensure that contributors benefit from these opportunities.

Quick guide: Too Long; Didn't Read

I want to know...



I am ...

....what are the key insights from this review for me?

Introduction

Background

Voice technology is expected to become a primary method used by many people to interact with digital devices and services. However, there are significant barriers to innovation in voice technology for most of the world's languages. This includes the lack of open and publicly available datasets necessary to build such solutions in low-resource languages, especially in the global South.

To overcome some of these barriers, the Mozilla Foundation has launched the Mozilla Common Voice (MCV) project. At its core, MCV offers a platform that supports language communities and volunteers to contribute their speech recordings to an open dataset that is published in the public domain. People wanting to build voice applications can use these datasets to build machine learning models.

Since 2019, the Mozilla Foundation has been working with the GIZ FAIR Forward initiative to promote the creation and utilization of open voice data and technology in Kinyarwanda, Kiswahili and Luganda. The initiative "FAIR Forward – Artificial Intelligence for All" is implemented by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) on behalf of the German Federal Ministry for Economic Cooperation and Development (BMZ). It strives for a more open, inclusive, and sustainable approach to artificial intelligence (AI) on the international level.

The opening of voice technology in Kinyarwanda, Kiswahili and Luganda has been achieved through efforts to build a community to create and use voice technology for each language that consists of commercial, non-commercial and government actors. These efforts include activities such as crowdsourcing voice datasets, creating speech models and developing applications with voice technology. To date, these voice communities have been mobilized to create open voice datasets including more than 2.400 hours in Kinyarwanda (one of the largest datasets on MCV), 900 hours in Kiswahili and 550 hours in Luganda. These datasets have been used to build machine learning models ready for use by local innovators, for instance a voice recognition model in Kinyarwanda.

Objective and scope of the review

The objective of this review is to **document community building strategies and efforts** that have been deployed to create open voice datasets for Kinyarwanda, Kiswahili and Luganda using the MCV platform with support from GIZ, as well as other donors such as the Bill and Melinda Gates Foundation (BMGF). It aims to provide existing and future voice communities as well as organizations who support such communities **with insights and recommendations on how to sustainably establish and maintain open voice communities.** While some recommendations might not be generalizable to all contexts, other community efforts to build open voice datasets without using the MCV platform might be able to learn from insights provided in this review.

Potential readers of this review therefore might be members of open voice communities and organizations supporting them or other individuals and organizations interested in better understanding the functioning of community-driven efforts to build publicly available digital resources, especially in the field of AI.

The review is split into four parts. The first section of this review provides a general introduction of voice communities. It offers an overview of key steps and stakeholders involved, based on the experience of MCV communities for Kinyarwanda, Kiswahili and Luganda. The second section provides guidance on mobilizing and sustaining communities that are both dynamic and diverse, looking at successful conditions, roles, and measures. The third section explores some major operation and technical challenges that were encountered by voice communities and examines solution strategies that were used to overcome these challenges. The review concludes with a list of recommendations specifically aimed at key stakeholders involved in a voice community. It also provides short lists of existing practical tools that can be reused by both established and future voice communities.

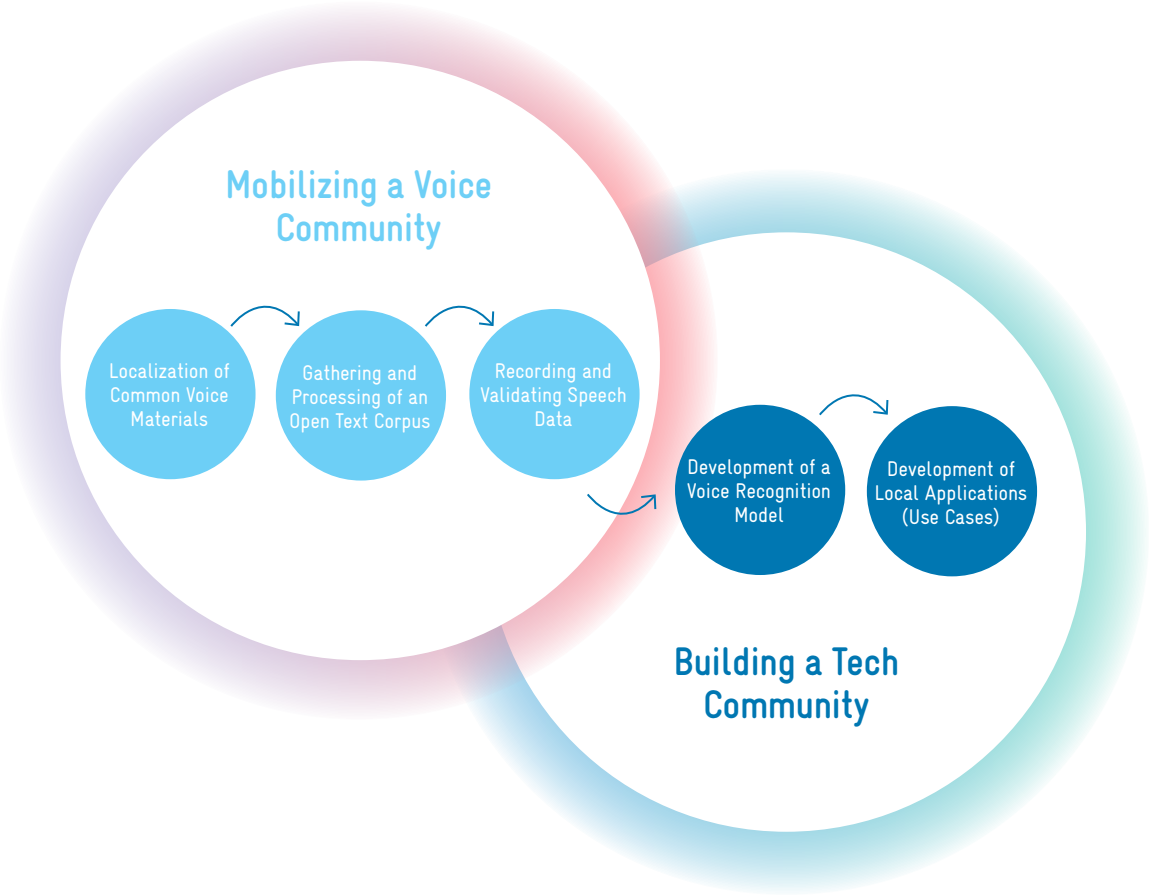
The review aims to **complement existing resources that have described the practical steps necessary for the building of open speech and voice datasets** using the MCV tools and ultimately, to develop a voice recognition model and local applications in one language. These practical steps can be summarized in the following manner:

- The localization of MCV tools and materials to be understood by contributors in their language.
- The gathering, validating, and processing of a public domain sentence corpus.
- The recording and validation of voice clips to create an open voice dataset.
- The training and publication under an open license of a machine learning model for voice recognition in the targeted language.
- The development of voice-based applications using the recognition model.

To train a near-human general speech-to-text model, 2 000 hours of verified voice clips are necessary. For this purpose, 1 800000 public domain sentences that meet certain quality standards are needed and at least 1 000 unique speakers per language. The development of an open voice dataset hence requires the mobilization of a vast community of contributors.

The review primarily focuses on lessons-learnt, insights and recommendations for **mobilizing sustainably engaged voice communities.** It focuses on the **social dynamics at work - the processes of translation, enrolment, mobilization, reappropriation, or learning**[1] at the community level, rather than focusing on the necessary technical steps. These observations shall provide detailed guidance for the social processes and structures required to allow communities to successfully follow the technical steps described in the existing documentation.

1 These processes are understood here from a sociological perspective. Translation is understood not only as the act of translating words between languages but as all the 'negotiations, intrigues, calculations, and acts of persuasion' necessary to represent something in a new form. Reappropriation is understood as the process by which a community can change or reclaim the meaning of something. Enrolment and mobilization refer to the processes of allocating roles to groups and organizing them to make them act in a certain manner. Similarly, learning is not considered simply as the act of acquiring new knowledge but as an act of dialogue and renegotiation of meaning in a given social environment.



The review does not cover the development of the voice recognition model and local applications, as these steps require the mobilization of a different community of actors, mostly involving local tech ecosystems. It will in fact look at **communities engaged in data collection and validation for speech data as well as sentence corpora.** It will **consider the tech community where it overlaps with the voice community** – e.g., for domain-specific data collection for use cases. However, this will not be the primary focus of this review.

Methodology and interviewees

This review was conducted and compiled by Jan Krewer, an independent consultant contracted by GIZ FAIR Forward, in collaboration with the Mozilla Foundation. Several existing resources were used for the review, including the MCV Community Playbook, progress reports as well as existing blog posts and publications by community members.

The consultant conducted four interview workshops with groups of stakeholders from every language community: Mozilla Foundation Fellows, community mobilizers, contributors, as well as Mozilla Foundation and GIZ staff members. These interviews were complemented by a short review of literature on community building. All people interviewed should be thanked again for their time and contributions. The observations and ideas developed in this review mainly come from them.

Chenai Chair, Mozilla Foundation

Mark Gachara, FAIR Forward – GIZ

Robert Katavi Mrima, Kiswahili contributor

Mercy King'ori, Community mobilizer for Kiswahili

Jonathan Mukiibi, Community mobilizer for Luganda

Eleanor Muyinza, Kiswahili contributor

Britone Mwasaru, Mozilla Common Voice Fellow

Audace Niyonkuru, Community coordinator for Kinyarwanda

Francis Nkurunziza, Kinyarwanda contributor

Christian Resch, FAIR Forward – GIZ

Rebecca Ryakitimbo, Mozilla Common Voice Fellow

Kathleen Siminyu, Mozilla Common Voice Fellow

What is an open voice community?

Definitions and differences with other voice communities

An open voice community includes **the group of people and organizations who have a shared interest in the establishment of a publicly available dataset** for the future development of automated voice recognition tools in one specific language. While some of the individuals contributing to the establishment of a publicly available dataset do it on a voluntary basis, others might be employed or remunerated by an organization. Voice communities can therefore be made of individuals, commercial, non-commercial and government actors.

Regarding the communities studied for this review, the Mozilla Foundation stewards the overall Common Voice project and is responsible for the platform direction and goals. Mozilla also oversees the development of tools and channels to support communities. Open voice communities differ from other private or public efforts to build large, crowdsourced datasets, as they rely on **certain principles and modes of governance.** Unlike organizations or markets, which are coordinated exclusively by either hierarchy or price mechanisms, these communities are governed using decentralized forms of communication and negotiation.

- Open voice communities agree to create **a voice dataset that is publicly available under an open data license.** This should allow anyone to use and build models with the dataset, and later develop applications, both non-commercially and commercially, in a specific language.
- Open voice communities **support local empowerment and capacity building.** They contribute to the decentralization of key technologies as they strive to make the development of locally owned, appropriate voice technology applications possible.
- Open voice communities are **self-organized communitie**s that support **participation and accountability** for the management of their activities. Since communities can be made of heteronomous networks of actors, with some actors investing resources and others contributing voluntarily, open voice communities need to develop their own governance models, ideally based on participatory and accountable coordination mechanisms, to ensure that all parties remain engaged in the long term.

Key steps and stakeholders

Four key roles have been traditionally identified by open-source software practitioners to ensure a sustainable management of open resources. These are the **critical functions** that need to be performed by any open community:

- **"Maintainers" or coordinators,** who are the organizational owners of the open resource, taking primary responsibility for its management.
- **Contributors,** who are individuals or organizations contributing to the conservation and/or the development of the open resource.
- **"Sustainers" or donors/funders and other enablers,** who provide the necessary financial and/or technical support for the resource to be built.
- **Users,** who are users of the open resource creating environmental, social, and/or economic value from it.

While it is not possible to attribute each of these functions to single groups of stakeholders participating in open voice communities, this classification can be used to provide an overview of some of the key activities and stakeholders involved in the building of a publicly available dataset. It is based on the experience of Kinyarwanda, Kiswahili, and Luganda communities and should therefore not be considered exhaustive or something to be necessarily followed by. However, critical functions that need to be performed by at least one stakeholder should be listed here. This classification is used in the final section of this review which offers recommendations for each category of stakeholders.

Critical functions	Linked activities in open voice communities	Type of stakeholders involved in open voice communities
Coordinators	Develop and execute language and diversity strategies Develop and execute language and diversity strategies Translate and localize MCV materials Provide governance framework for participation Provide legal framework for sentence and voice data collection Engage partnerships with donors and users Plan, budget, and coordinate contributor activities Provide technical support & debugging for sentence and voice data uploads Develop and execute mobilization and outreach strategies Develop and execute public relations and communication strategy Conduct quality and diversity monitoring Facilitate the development of the open voice recognition model	Mozilla Foundation and Common Voice fellows Local community mobilizers (e.g. from the tech community, academia, or civil society organizations)
Contributors	Contribute to crowdsourced sentence collection and/or donation of large text corpus Contribute to sentence verification Contribute to recording of voice clips Contribute to verification of voice clips	Student groups Local civil society and media groups, publishers, authors or experts Organizations and personalities from cultural, religious, academic sectors Local authorities
Donors and other enablers	Provide financial and in-kind support and/or technical assistance for coordinator, contributor, and user activities	Official development assistance agencies Universities (e.g. linguists and natural language processing experts) Philanthropic foundations Tech companies and tech hubs from the public and private sector
Users	Participate to the development of the open voice recognition model Develop local applications (use-cases) based on the publicly available dataset and voice recognition model	Local or international tech communities (public and/or private sector) Sector-specific communities (for application in health, agriculture, education…)

Mobilizing and sustaining an open voice community

This section provides guidance on mobilizing and sustaining open voice communities that are both dynamic and diverse, looking at successful conditions, roles, and measures that can be identified from the experience of communities for Kinyarwanda, Kiswahili and Luganda. This review complements the existing community playbook, that describes the technical steps to be followed to build an open dataset. The different processes described below should not necessarily be considered as a checklist to be followed in chronological order. The lessons are organized thematically and sometimes the activities they describe might overlap or even be repeated.

This section is mainly written for potential coordinators and mobilizers of open voice communities who are looking for advice and resources. The practical recommendations at the end of each section serve as short summaries of key learnings.

Translate your vision into the context of your local community

Building a voice community requires a good understanding of the context of one language. This understanding is necessary to properly identify the community of speakers of the language, to translate the Mozilla Common Voice vision and communication materials, but also to describe the problems that can be answered by an open voice community in terms that are adapted to local realities.

First, **it is necessary to understand if one language has variants, dialects or accents that need to be considered.** For this step, it can be necessary to work with linguists to scope the geography of one language and think about important areas that need to be covered to **build a voice recognition model that integrates a language in all its diversity.** The decisions to be made here might be politically sensitive, so it's important to be fully aware of all their implications. For Kiswahili, for instance, the voice community worked with language experts to integrate not only the different countries where Kiswahili is spoken, but also the different dialects of Kiswahili that exist within countries.

One of the first steps to launch an open voice community is **the adaptation and translation of the MCV tools and materials,** which help people to understand and engage with the community and to use the MCV platform in their respective language.

Useful measures and tools

Community Playbook – Community guidance for languages and variants
https://common-voice.github.io/community-playbook/sub_pages/Lang_Variant.html

Community Playbook – Localization of tools and materials
https://common-voice.github.io/community-playbook/sub_pages/Localization.html

Once the key elements of the MCV project are translated, **communities need to define their own purpose and vision,** which can be understood and become attractive to people and organizations in their specific context. This means formulating a problematization: what is the local problem that can be solved by this community? How can not having an open voice dataset become an issue for the potential members of one community? What are benefits of having an open voice dataset for them specifically?

Challenge

Benefits of contribution in an open voice community can be very indirect

The link between contributions to a publicly available dataset and future benefits for a community can be very indirect and uncertain. It is difficult to offer guarantees that a dataset will be used to develop certain applications while there is no voice data yet. Unlike community activities to build a school, or to develop a software for health workers, building a dataset for potential applications does not come with impact that can easily be identified, measured, or linked to the group that contributed. This challenge is particularly important to consider in a context where only a few people are already aware of the functioning and convinced of the potential of voice-recognition technologies.

One way of showing the potential benefit of voice technologies is to **use tangible examples of services that can be developed on top of an open voice dataset.** These examples should either build on services that people already know, such as popular voice assistants that might have been used in English or another language or be examples that answer some concrete problems of specific stakeholders within the community. One Mozilla Fellow, for instance, explained to contributors how voice assistants could help reduce waiting times for customer support and call centers.

Some community mobilizers also mentioned the importance of **preparing a simple explanation of how voice recognition works:** how machine learning uses data to train a voice recognition model in one language. This should help contributors to better understand the result of their contributions and the different steps needed. It should also answer questions regarding privacy risks for instance. One question that was asked frequently by potential contributors in Kenya was "Where are our voices going?".

Finally, the definition of a vision for a community should be as open and inclusive as possible, to engage many potential contributors, supporters, and allies from the start. The vision of the community could for instance be redefined according to already existing objectives of local authorities and organizations. In Rwanda, the community used vocabulary from the Smart Rwanda Master Plan by stating that they were contributing to building the infrastructure for the country's transition to a knowledge-based economy. In Uganda, the community adapted to the policy of the Buganda Kingdom to promote the Luganda language. Organizations that support the promotion of the community's language and culture, that support technology innovation and development, or that promote the inclusion of marginalized communities, for instance of people with disabilities, should be consulted to **adapt the vision to existing issues and existing efforts.**

Useful measures and tools

Create a list of concrete examples that can show the benefit of having voice-recognition technologies in the communities' language.

Develop a simple explanation for your language and context of how machine learning tools use data to develop a voice recognition model.

Engage with stakeholders from the language community to align with their vocabulary, narratives, and priorities.

Create motivations and incentives to join the community

Finding good reasons for why other people and organizations get on board of a community's project is one of the key issues of community building. This will first require keeping an open mind and remaining flexible enough to mobilize members by aligning with their own needs and priorities. It then requires having a good understanding of the interests and motivations of contributors to create the right incentives for participation and meet contributor's expectations.

It can be difficult to identify all the stakeholders who could potentially have a shared interest in the development of an open voice dataset in one language, when voice technologies are not developed yet. To this end, a **mapping of local stakeholders who are working in related areas** can be useful. Such areas include local tech ecosystems, universities, existing digital volunteering communities such as OpenStreetMap or Wikimedia local groups, NGOs working on digital inclusion, especially for people with disabilities, cultural groups or local and national government authorities working on culture and language promotion. The voice community for Kinyarwanda for instance built a partnership with the Rwanda Academy of Language and Culture (RALC) to get their support. The Luganda voice community managed to enter into a partnership with the Buganda Kingdom. Such partnerships do not only allow communities to receive data donations (e.g. for the text corpus), but also to benefit from the partners' political authority to mobilize more contributors.

Alignment with stakeholders should eventually lead to a continuous adaptation of a communities' focus and priorities. Organizations can for instance be onboarded with the prospect of a collaboration on sector-specific datasets for the development of promising applications in their field. Such organizations should be involved from the beginning to have a more goal-oriented approach. Other examples include collaboration with blogger communities for sentence writing or even the mobilization of existing public volunteering mechanisms for data collection. In Rwanda, the pre-colonial concept of Umuganda, a concept of self-help and cooperation was used for that purpose. Every last Saturday of the month, people gather in their communities, bring efforts together to build physical infrastructures such as roads, schools, and more. A Rwandan start-up, Digital Umuganda, tapped into this concept of Umuganda to crowdsource community contributions for the Kinyarwanda open voice dataset.

The role of community coordinators is to **create the space that allows for the alignment between all these stakeholders and simultaneously to balance different interests and priorities,** for instance between an international tech company funding data collection and the local tech community. These partnerships are usually established after a series of meetings. Alignment between several stakeholders can also be managed through regular meetups or even steering committees at a later stage.

Finally, community coordinators need to understand the interests of the individuals participating in the crowdsourcing of both the sentence and voice dataset. This requires **exploring and evaluating potential motivation drivers for contributors and designing the right incentives for the community.**

Challenge

Contributing to a voice dataset can be time-consuming and tedious

Writing or verifying sentences, recording voice clips, and validating them… The contribution to a voice dataset can be a repetitive, time-consuming and tedious task. It does not offer direct or intrinsic learning or personal development opportunities, unlike other volunteering activities such as taking care of a community garden or writing an article on Wikipedia. Such opportunities need to be created by the community. To do this, the community also needs to turn an activity that can be very solitary into an empowering social activity.

Participation in a voice community therefore is everything but a given. The time, energy and resources contributors put into the building of a dataset should always be recognized and valued. Contributors might have various motivations to become a member of a community and as many expectations on how they should benefit from this membership. Meeting all these expectations can be a big challenge for large communities.

One mindset many community mobilizers adopted for Kinyarwanda, Kiswahili, and Luganda is to move away from the idea that volunteering should be considered a given, and to understand that **contributors should have an interest in their participation to the communities' effort.** This means that they should always see the value in contributing. **The time, energy and resources contributors put into the building of a dataset should always be recognized and valued.** This is even more crucial as many people have understood that applications that can be developed with the crowdsourced data can be very lucrative. However, it does not mean that rewards for contributions should necessarily be monetary or material. Other motivations drive community involvement. To find the right incentives for their community, community coordinators can mobilize user-research approaches, for instance by studying personas, i.e., fictional potential contributors and their motivations. They can also conduct surveys to understand what contributors mostly expect from their participation in the community. Such evaluations could be run on a regular basis to ensure that contributors' expectations are met.

The table below shows some incentives that have been developed by open voice communities based on different motivational drivers of contributors.

Motivation drivers	Examples of incentives developed in open voice communities
Values, principles, and impact	Contribution to the MCV vision (local technology innovation and development, inclusion, promotion of local language and culture) or the vision of other related initiatives (partnership with cultural or writing festivals).
	Opportunities to participate in coordination activities and be recruited as a community champion or ambassador.
Monetary or material rewards	Small prizes or material rewards for top contributors or for contributors reaching a minimum standard / goal (during one event or for regular contributions). Rewards can include mobile data bundles (up to 5 Gb), goodies with the Mozilla brand, giveaways (earphones), snacks and drinks.
Learning opportunities	Combination of writing or recording events with keynotes or training on a matter of interest for the community. Such training could be about artificial intelligence for women and girls, but also digital literacy training in remote areas. Some communities also organized cultural events about the history of one language, to create benefits other than those offered by MCV.
Professional opportunities	Creation of spaces for networking of community members with partner organizations and personalities.
	Public recognition and advertisement of contributor's participation to the community (social media or blog posts about contributors).
	Virtual social stickers or official certifications issued by the Mozilla Foundation for top contributors.
	Opportunities to participate in coordination activities and be recruited as a community champion or ambassador.
Playfulness	Competitions (like "data collection sprints") with small rewards for top contributors.
Social connections	Public events and gatherings that leave space for social connections in a casual atmosphere, with snacks and drinks.

Today many AI solutions have been developed thanks to the labor of largely invisible and exploited workers. Millions of underpaid workers around the world perform repetitive tasks under precarious conditions to train AI models. One important difference between this approach to building large datasets and community-driven efforts is that monetary or material forms of remuneration are not at the center of the value contributors perceive in their participation in the community's efforts. Most communities have used some forms of monetary or material remuneration to create incentives for contributors. Many have distributed data bundles or goodies, or simply offered snacks and drinks to contributors, who can often be students. However, there should be **a clear internal warning system in the community to mitigate risks of exploitation,** where competing for a small prize for instance becomes the main motivation driver of large parts of contributors.

Useful measures and tools

Map stakeholders who have a potential shared interest with the community.

Organize meetings with key stakeholders to align the community's objectives with the needs and priorities of both partners and contributors.

Explore and evaluate potential motivation drivers for contributors and design the right incentives for the community.

Implement internal community warning systems to mitigate risks of exploitation.

Define roles and responsibilities among community members

Open voice communities can take a large variety of organizational forms, but to sustain themselves, they need to define roles and responsibilities among their members, separate certain tasks from others, and even create remuneration systems when necessary for the enrollment of individuals to cover critical functions. Communication channels and decision-making processes need to be flexible and decentralized enough to reward most active contributors, but transparent governance frameworks are essential to keep large groups engaged.

The three main functions that need to be covered in open voice communities, based on the experience of Kinyarwanda, Kiswahili, and Luganda communities, are **coordination, mobilization, and contribution.** All the three communities have adopted **a similar distribution of roles** according to these three functions.

- The **community coordinators** work daily on the planning, budgeting, and coordination of community activities. They are usually made of a core team of up to four or five people, who can be Mozilla Fellows and/or members of a partner organization, such as Digital Umuganda in Rwanda, which signed a Memorandum of Understanding with the Mozilla Foundation and was contracted by GIZ.

- **Community mobilizers** are the members of the community who engage directly with most of the contributors. They execute training or data collection events and enroll new contributors. They also ensure that quality standards and guidelines are followed. They are called ambassadors, "commoneers", or community champions. There can be several dozens of community mobilizers. They sometimes also cover certain contribution activities that either require certain skills or that are particularly tedious or time-consuming, such as the verification of sentences or voice clips.

- **Contributors** are the community members that participate in sentence collection and validation, as well as recording and verification of voice clips. They usually do not have coordination functions.

This distribution of roles might vary from one community to another. It allows for a **chain of outreach and mobilization** that has proven efficient in all three communities. It is particularly adapted as contributors need to get informed, trained and sometimes supported to participate efficiently. **Contributors are often encouraged to become mobilizers or to specialize on certain tasks.** Training can be provided for this purpose by coordinators or mobilizers, who will organize regular meetings among specialized teams. This applies for instance to community members who verify voice recordings, as it requires a lot of time to train voice verifiers who can ensure that recordings meet certain quality standards.

Challenge

The transition from an informal collective to an organized community can create friction among community members

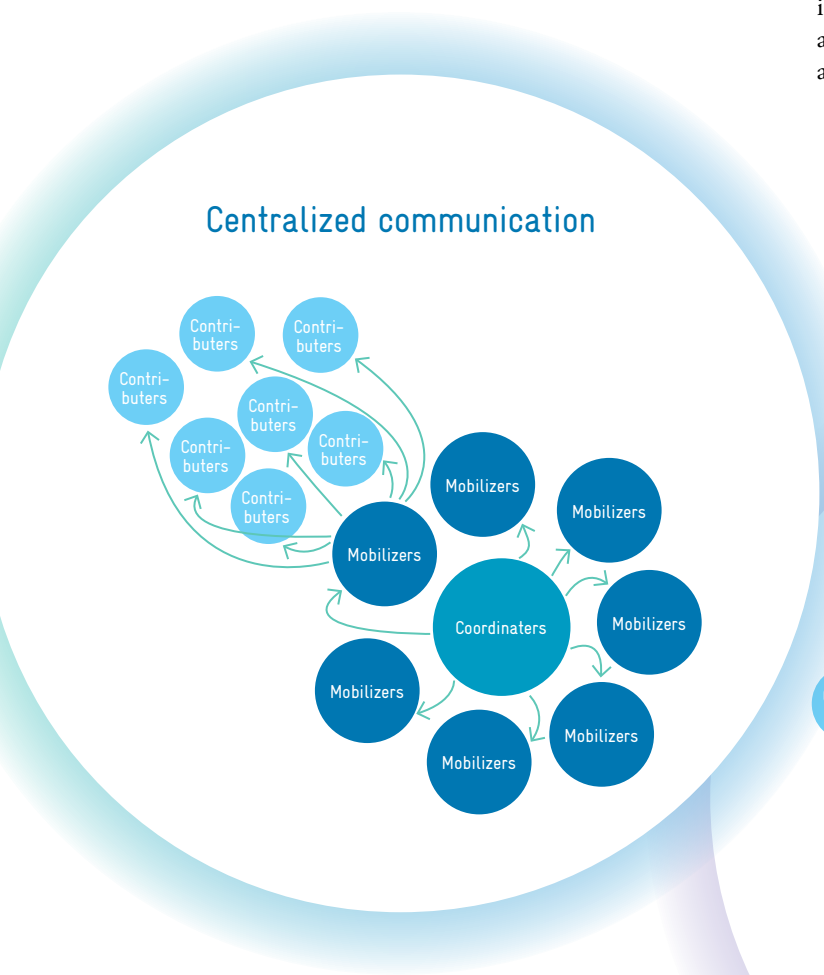
When self-organized communities grow, they usually have to transition from spontaneous and organic distributions of roles and decision-making processes to formalized rules of functioning. This formalization can create friction among members of a community who can compete for certain positions or for remuneration, when funding becomes available.

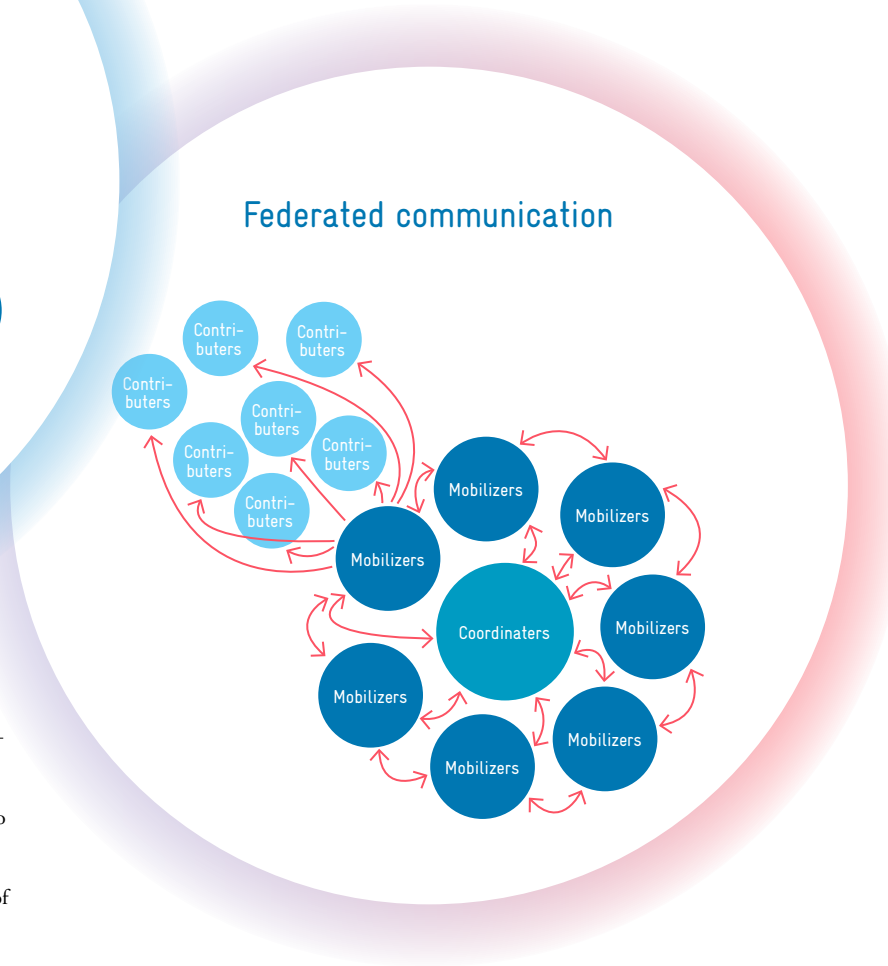
Very often these distributions of roles and remuneration systems are developed along the way: most active contributors for instance progressively gain reputation and therefore the legitimacy to influence certain decisions. This flexibility is important in a community, as it allows to reward active contributors with leading roles, as opposed to a predefined organigram. An early open and transparent conversation about ground rules of engagement, including roles, remuneration, and visibility is essential to keep large groups engaged. The Mozilla Foundation could provide governance charters to help local communities in this effort.

The structure of communication channels can vary depending on the size of the community. While a young community that is progressively formalizing its activities might have more centralized modes of communication, bigger communities become increasingly federated. In a federated type of network, community mobilizers increasingly talk to each other and coordinate themselves, instead of communicating only with one or two coordinators. **Community mobilizers are gradually trusted to innovate and develop their own solutions and mobilization strategies.** In the case of the Kiswahili community, for instance, mobilizers received an honorarium for their time and an allowance to cover activities for the management of smaller groups of mobilizers, data collection events, etc. This is reflected by groups created on messaging applications, where all the community mobilizers exchange among each other and support coordination efforts.

Decisions about budget allocation and remuneration systems, guidelines, mobilization strategies, training, or diversity strategies are mostly made by the coordinators, in consultation with some mobilizers and the Mozilla Foundation (MCV team and Mozilla Fellows). Some donors and other enablers require to be informed about these decisions.

Decisions about technical improvements of the sentence collector or the MCV platform or strategic orientations of data collection according to specific needs of use-cases are made by the Mozilla Foundation[2], in consultation with coordinators and other involved partners (donors and enablers, potential users). This **multi stakeholder approach should ensure a collective ownership and trust in the initiative.** It could be further formalized and extended to contributor and/or local coordinator representatives. One could for instance implement a steering committee that meets on a regular basis where all different community groups are represented equally.





2 As the Mozilla Foundation hosts the MCV platform, communities must accept some of its decisions, in particular the decision to publish datasets under a CC0 license. They also need to follow certain principles and guidelines. Some communities have mentioned that they have an interest in using more restrictive licensing to better reward local communities of contributors and limit the reuse by global commercial actors. The only option for communities right now is to accept these terms or to exit the global MCV community.

While there is a formalized governance model for Mozilla communities at the global level with the Mozilla Common Voice Governance Doc and the creation of a Common Voice Language Reps Council, **local language communities are self-organized and do not necessarily have frameworks that describe governance models or accountability and consultation mechanisms** at this stage. It is not clear for instance what rules govern the renewal of community coordinators and mobilizers.

Useful measures and tools

Define roles and responsibilities based on the critical functions of an open voice community: coordination, mobilization, contribution.

Progressively develop federated communication channels that allow smaller groups to execute their own strategies and solutions.

Create transparent governance frameworks that describe roles and responsibilities, decision-making, remuneration systems, accountability, and consultation mechanisms.

Build multi stakeholder steering structures for strategic decisions regarding the evolution of the open voice community.

Mozilla Common Voice Governance Doc: https://docs.google.com/document/d/1QHDhdJzBQzWRzjw88Br674OINtok09znq0vmpgVBZQo/edit#heading=h.hxndvjrxe92a

Mozilla Common Voice Community Participation Guidelines and Channels: https://github.com/common-voice/common-voice/blob/main/docs/COMMUNITIES.md

Mission Driven Mozillians – Volunteer Leadership Principles & Practices: https://discourse.mozilla.org/t/mission-driven-mozillians-volunteer-leadership-principles-practices/22386

Mobilize diverse representatives of large communities

To successfully build a dataset that can be used to train effective and inclusive voice recognition models, there are two crucial factors: diversity and quantity. Contributors need to represent all the different groups of the speakers of the language. Diverse communities can be mobilized by building alliances with existing communities and networks of potential contributors. To reach large numbers of contributors, outreach and communication strategies based on multiple channels need to be implemented, with a focus on peer engagement.

Challenge

Contributors to an open voice community might not represent the diversity of language speakers

Contributors usually participating in open voice communities are mostly urban and tech-savvy young men. As described in the Mozilla Kiswahili Project Gender Action Plan: "Women especially in developing countries have less access to technology than men, in terms of access to devices and the internet. The gaps in access and affordability are experienced more so by women in comparison to men in lower income segments, lower levels of education, limited access to resources as well as urban and rural disparities. These gaps result in less contribution to and making use of technologies for women in marginalized groups."

If women participation in the recordings of a voice dataset is low, it will automatically lead to the creation of biases that will make voice assistants less accurate for women.

Robust and inclusive voice recognition models can only be trained with data that represents all the diversity of speakers in terms of location, accent, age, income, education background, and gender diversity groups. Communities therefore need to **design and implement diversity strategies and work plans right from the start.** These strategies should detail how specific target groups can be reached out to and engaged to contribute. This could include the **adaptation of communication and publications, but also the creation of customized incentives or compensations.** Community coordinators have tried to organize data collection events that specifically focus on women, older contributors, or rural areas for instance.

One positive measure that was implemented in the Kinyarwanda, Kiswahili and Luganda communities is **the recruitment of community mobilizers who**

represent the diversity of speakers of the language. These community mobilizers will then act as ambassadors in their respective communities and will find it easier to encourage their peers to participate.

The results of a diversity work plan should be made measurable, to be able to adapt mobilization strategies. This requires implementing **accurate monitoring systems that allow a community to be informed in real-time of the diversity of the dataset** being built. One way to do this is to ensure data can be disaggregated by diversity categories. Another way is to train voice data verifiers to measure the gender diversity of collected data.

The next big challenge is to mobilize large numbers of contributors. A voice dataset should have at least 1 000 unique speakers, with more than 2 000 hours of verified voice recordings. This standard can vary as it is possible to build models with smaller numbers for certain languages or specific applications. Reaching these objectives however always requires developing an ambitious communication and outreach strategy. Community mobilizers have conducted **campaigns on multiple communication channels,** from traditional mass media like radio advertisements or reports in the news, to publications social media or instant messaging publications. Broadcasting short advertisements or creating flyers and other materials to disseminate them online can help to create visibility and attract more contributors, when these materials are customized for the identified target groups.

Useful measures and tools

Community Social Media Campaign Guide: https://github.com/common-voice/community-playbook/blob/master/assets/img/CV_Social_Media_Community_Campaign.pdf

Common Voice Graphics for Community Use: https://drive.google.com/drive/folders/1Qn5e_jKmj_Kw-tc1W7P3ZcvYCZETkEmN?usp=sharing

One key learning from all the three studied voice communities is that most contributors will join the community because someone else they know is a contributor as well. While communication is important to raise awareness around an initiative, **engagement seems to come mostly from peer mobilization through one-on-one communication or existing social structures.** Mass media advert campaigns alone indeed did not necessarily result in greater traffic and engagement (based on new accounts created).

Three strategies have been used in open voice communities to use existing social structures:

- **building alliances with other communities,**
- **recruiting community champions who are part of targeted social networks,**
- **and receiving official support from political or other institutions' authority**.

The Luganda community for instance managed to receive the public support of the Buganda Kingdom for mobilizing people, which was broadcasted in the news. Other communities have tried to reach out to the ministries of education or youth. In Rwanda, community mobilizers worked with high school professors to mobilize students while in Tanzania and the Democratic Republic of Congo they reached out to active religious groups to tap into their network of volunteers.

Community mobilizers who are part of different social structures (family, friends, student or working environments) usually directly recruit contributors at the grassroot level. They try to organize events in spaces where potential contributors already are, or to organize them during existing community gatherings. Some of the typical groups that communities have tried to target are student groups (like language or computer clubs) and professional associations or networks (members of tech hubs or mobile operator street agents).

Useful measures and tools

Develop a diversity strategy and work plan with specific target groups and appropriate monitoring mechanisms.

Develop an outreach strategy based on multiple channels that is customized for the identified target groups.

Build alliances with existing communities and networks of potential contributors.

Community Playbook - Community mobilization: https://common-voice.github.io/community-playbook/sub_pages/mobilization.html

Mozilla Kiswahili Project Gender Action Plan: https://assets.mofoprod.net/network/documents/Gender_Action_Plan.pdf

#BreaktheBias Workshop with Common Voice: https://mozillafoundation.typeform.com/to/yluCogH1?typeform-source=common-voice.github.io

Working around operational and technical challenges

Collecting and verifying sentences

The preparation of a voice dataset requires first to build a large corpus of sentences that can later be used for the recording of corresponding voice clips. Sentences need to be diverse in terms of phonemes, variants, and domains. They cannot be subject to copyright as Mozilla Common Voice datasets are released under a CC0 License[3] and are part of the public domain. They also need to follow certain criteria (for instance on length, punctuation, abbreviation and acronyms, or numbers).

Open voice communities usually either gather sentences from accessible public domain sources online, from donations of partners, or through crowdsourcing, by using the Mozilla sentence collector. The bulk upload of large files of sentences usually requires preliminary cleaning of text sources. Finally, sentences need to be verified before they are uploaded to the MCV platform and contributors can start recording voice clips.

This section describes some of the main challenges open communities have faced while collecting and verifying sentences and the positive measures they took to address them.

Useful measures and tools

Community Playbook – Text Corpus: https://common-voice.github.io/community-playbook/sub_pages/text.html

Community Playbook – CC0 Waiver: https://common-voice.github.io/community-playbook/sub_pages/cc0waiver_process.html

CHALLENGE: Finding public domain sentences

Getting legal agreements based on public domain licensing can be very difficult. Only a few organizations have an incentive to share their text corpus. Additionally, the understanding of a CC0 license can be limited, and often there is confusion with content available on the Internet.

POSITIVE MEASURE

Enroll donors for text by showing potential benefits of speech-to-text applications (for instance journalists).

Show the development benefits of the project to excite people about being part of it.

Focus on media agencies instead of publishers, who have large archives of articles that are no longer current news and therefore no longer marketable.

3 The CC0 "No Rights Reserved" License is a Creative Commons public domain equivalent license that allows anyone to freely build upon, enhance, and reuse one work for any purposes without restriction under copyright or database law. For more information, see: https://creativecommons.org/publicdomain/zero/1.0/

CHALLENGE: Find text sources in easily reusable formats Some publishers can only share text in formats that are not easily reusable, such as PDF files.	**POSITIVE MEASURE** Use web scraping and other automatic sentence extractor tools.
CHALLENGE: Diversity of sentences Sentences need to cover all types of domains and therefore come from various sources. Sentence collection also needs to target all dialects: some sentences might not work in all dialects, as there can be different writing for one specific dialect.	**POSITIVE MEASURE** Find multiple text sources. Introduce topics for writing competitions to have diversity of themes and inspire people.
CHALLENGE: Finding contributors to crowdsource sentence collection Sentence writing requires more focus than the recording of voice clips. Especially the required skills in grammar for sentence collection can create frustration and dropouts.	**POSITIVE MEASURE** Try to focus on sentence contributions based on the scrapping of public domain sources. Organize "Write-A-Thons" (writing competitions) for university students, writers, or linguists.
CHALLENGE: Verification of sentences The verification of sentences is a challenging and exhausting task. Writing sentences can sometimes be less time-consuming than editing very poor-quality sentences. Web scraping usually leads to a lower quality sentence. Sometimes meaning gets lost. The same sentences can be repeated several times. Verification of sentences cannot be fully automated.	**POSITIVE MEASURE** Develop hybrid solutions for verification: an automatic tool (script) for the deletion of repetitions, and the adaptation of raw text (break down paragraph into sentences for instance), combined with people verifying sentences, especially checking meaning. Community mobilizers can be specifically recruited to go through excel sheets to clean them / They can be recruited among core contributors or specific groups that have demonstrated their skills during writing challenges.
CHALLENGE: Sentence uploads Sentence collection is not a user-friendly tool, for bulk uploads.	**POSITIVE MEASURE** Address issues to receive support on the Mozilla Common Voice community platforms and with the Mozilla Common Voice fellows. Find guiding online on how to do bulk sentence verification and upload.
CHALLENGE: Coordination of sentence and voice collection activities Sentence uploads on the platform can sometimes take time: if not anticipated, this can lead to shortages of sentences on the MCV platform for voice recordings.	**POSITIVE MEASURE** Make estimations of the time needed for sentence writing, verification, and upload to ensure voice recording activities are not impacted.

Collecting and verifying voice recordings

Once enough sentences have been uploaded to the MCV platform, the recording of voice clips can start. Open voice communities either mobilize contributors to participate online on their own, during online group challenges or they organize data collection events. The strongest and most sustainable engagement can usually be observed during data collection events, which are organized by community mobilizers. Recordings should be made by a diversity of contributors in terms of location, accent, age, income, education background, and gender diversity groups. Once recorded, the voice clips need to be verified again by contributors to ensure that the quality guidelines for voice data are followed.

This section describes some of the main challenges open communities have faced while collecting and verifying voice data and the positive measures they took to address them.

Useful measures and tools

Community Playbook – Voice data collection: https://common-voice.github.io/community-playbook/sub_pages/voice.html

Community Drive with planning templates and tips: https://drive.google.com/drive/folders/1YsLzbzttpJ9VuyJSqwtfx9lFi4gwj6Lf?usp=sharing

Validation Parties slides: https://docs.google.com/presentation/d/1P_rvMLjiC51Y6QAqU0TM9W-EGvzuF6Qu6FwvgLGkI24/edit?usp=sharing

Common Voice Event Explainer: https://docs.google.com/presentation/d/1HatIkqvhj--4mYvEGAWHAGQ6yf3O7t6iI3LmF6lFPNc/edit?usp=sharing

CHALLENGE: Event planning and management

Finding resources to design and execute attractive data collection events in the absence of financial remuneration for contributions.

POSITIVE MEASURE

Prepare a budget for events with attractive venues, internet access, lunches, and drinks.

Organize goal-oriented data collection sprints during off-peak business hours.

CHALLENGE: Quality of recordings

During the initial recording phases in Kinyarwanda, for instance, more than half of the recordings were rejected because of their bad quality. This can be due to background noises in small rooms during data collection events, misuse of the platform and recording button, or bad microphones. Rewards targeting a high quantity of contributors can also have a negative impact on the quality of recordings. Very large data collection events can also be detrimental for the quality of data, as they are more difficult to manage for moderators, there is less room for one-on-one training and as they create more distractions.

POSITIVE MEASURE

Organize short training sessions before recordings and create awareness around bad quality recordings. Encourage people to test their audio first.

Organize frequent small events in addition to longer gatherings with large attendance.

Monitor error rates after data collection events to evaluate training. Understand the main type of mistakes that are done with trained data verifiers to better adapt collection events and train contributors accordingly.

Use an agile approach to avoid losing many recordings: start verifying while you record to adjust bad recording approaches.

CHALLENGE:	POSITIVE MEASURE
Verification of recordings Crowdsourced verification can lead to low quality of datasets: verification requires good training about pronunciation and accents. In general, contributors are also less excited about verification than recording, which can create a gap between unvalidated and validated voice clips.	Recruit specific community mobilizers focused on verification mobilization. Relying on a core team of verifiers helps avoid common mistakes and increases the speed of verification. Train verifiers on dialects to avoid having dialects flagged as mistakes. Identify potential verifiers among best contributors.
Uploads monitoring Updates of the count of unverified and verified data hours can sometimes lag behind real uploads and make monitoring of progress more difficult.	Organize your own tracking methods with contributors to regularly review progress.
Diversity monitoring Monitoring age, social background and gender during data collection events and online contribution can be difficult.	Allow self-identification as women to better monitor gender balance. Train sentence and speech validators on diversity indexes to report back to mobilizers and use this feedback to adapt mobilization strategies.
Privacy concerns Some contributors can show concerns about how their personal data (linked to their MCV accounts) and their voices can be used. Some contributors from Kenya for instance were worried that their voices could be misused once recorded and stored on the MCV platform. Some mobile service providers in the country have indeed started to develop electronic signatures based on the unique voiceprints of their users.	Ensure the privacy and security of personal data. Limit any collection of personal information. Allow contributions without logins. Be transparent and clear about how voice is used, and voice data anonymized.

Bridging the digital gap: access to broadband, devices, and digital literacy

Unequal access to the internet or to the devices needed to use the MCV platform can make contributions difficult, especially for usually marginalized communities. Different levels of literacy and digital literacy can also lead to a poor representation of these marginalized communities in the dataset. This lack of diversity could further reproduce existing discrimination as voice recognition models would be built for only a certain part of the population. Open voice communities have therefore implemented measures to bridge some of these gaps to allow more people to contribute. Such measures include collective access to Internet or devices during onsite events, offline data collection and storage, or even the association of data collection events with digital literacy training.

Useful measures and tools

Mozilla blog – Advancing platform inclusion: https://foundation.mozilla.org/nl/blog/common-voice-for-everyone-mozilla-advances-projects-with-nvidia-enabling-voice-data-donation-in-low-internet-contexts/

This section describes some of the main challenges open communities have faced regarding digital accessibility and literacy gaps and the positive measures they took to address them.

CHALLENGE: Cost of Internet connection and bandwidth

Some potential contributors can't afford the cost of Internet connection or devices necessary to participate in sentence and voice collection activities.

POSITIVE MEASURE

Provide collective access to devices and Internet connection, by organizing data collection events in venues with Wi-fi or use portable hotspots (such as university venues). Use devices of facilities or if necessary, of Mozilla fellows or community mobilizers.

CHALLENGE: Digital literacy

The use of the platform can be challenging for people with low digital literacy. Some contributors will struggle to use devices, web browsers, register to the platform and use it. All contributors won't have the technical background to completely grasp the purpose of the voice community.

Contribution can also be difficult to use for people with disabilities such as dyslexia or visual impairment, as sentences need to be read quickly.

POSITIVE MEASURE

Organize data collection events that combine data collection with digital literacy training opportunities (on use of devices or web applications). Try to explain in simple terms the objective and impact of data collection.

Adapt data collection and training approaches to the target group.

CHALLENGE: Accessibility

The MCV platform is not always accessible on all devices, systems, and browsers, who are not always updated to the latest version.

POSITIVE MEASURE

CHALLENGE: Use of data

The MCV web application is data consuming. Events in slow Internet access where mobile hotspots are used can suffer from this: some data recordings of data collection events in remote areas were lost.

POSITIVE MEASURE

Engage with the Mozilla Foundation to work on an offline or light version of the MCV platform.

Community building and social distancing

The community mobilization activities for Kinyarwanda, Kiswahili and Luganda were affected by the COVID-19 pandemic and social distancing measures established by governments. Lockdowns, curfews, or limitations of social gatherings created new challenges for community coordinators. This section provides an overview of these challenges and positive measures which the communities implemented to address them.

CHALLENGE: Restrictions for physical gathering	POSITIVE MEASURE
Many communities had to face restrictions for social gatherings and events. This led to reduced engagement by contributors.	Organize online events if access to the internet and devices is given. Create weekly challenges and competitions to attract contributors.
CHALLENGE: Limited training and support opportunities	**POSITIVE MEASURE**
Physical training and supervision on the use of the platform to ensure the quality of data is not possible.	Organize webinars and training and online sessions in groups with a supporting team.
CHALLENGE: Digital accessibility gap	**POSITIVE MEASURE**
Existing challenges around digital access are amplified by forced remote contribution.	Give out data bundles as rewards.

Recommendations

This section offers recommendations on how to sustainably establish and maintain open voice communities for the different categories of stakeholders involved. It is based on the efforts that have been deployed to create open voice datasets for Kinyarwanda, Kiswahili and Luganda using the MCV platform with support from GIZ, as well as other donors such as the Bill and Melinda Gates Foundation (BMGF). While some recommendations cannot be generalized to all contexts, other community efforts to build open voice datasets without the MCV platform might be able to learn from these insights.

For the mozilla foundation & mozilla fellows

Pursue global efforts to establish official consultation bodies that permanently strengthen the inclusion of communities in strategic decisions regarding the future of Common Voice. Increase the visibility and power of spaces for community coordinators to give feedback on their needs and challenges. Create transparency and accountability mechanisms towards the global community of contributors on orientations regarding the selection of languages that get support, partnerships, licensing options and technical developments.

Create opportunities both for organizations and individuals by giving contributor's work visibility and encouraging local entrepreneurs to participate in communities and the definition of solutions based on MCV datasets. Communication on social media or Mozilla blog posts highlighting the achievement of local community members or organizations are very important. While communities should have flexibility in finding the right motivational drivers for their community members, the Mozilla Foundation could assist them by developing standardized recognition tools such as certifications for top contributors.

Support local ownership of community coordinators by prioritizing local organizations from academia or from the public and private sector for coordination roles, especially when there are already organized groups working with voice technology in one specific language. Ensure that local organizations have the primary responsibility for the coordination of communities and delegate as many decisions and activities as possible. To ensure plurality, openness and an equal distribution of benefits within the local ecosystem, formalize these delegations of activities through partnerships that clearly state principles which community coordinators need to follow to leave space for other interested individuals and organizations to join the community.

Create focal point positions that can support communities by bringing credibility and legitimacy to their activities, notably for partnership building, giving them quick access to decision makers and resources, providing expertise and proactive support for technical issues regarding the use of tools and materials but also learnings on community mobilization strategies and challenges, for instance on the establishment of governance charters.

Continue updating tools and materials for both sentence and voice data collection to reduce barriers to entry and make participation to Common Voice even more accessible and inclusive. Technical and operational solutions to overcome the challenges of the global digital divide could for instance be better documented and made more visible in key community tools such as the Community Playbook.

For local community coordinators

Recognize and value the work and interests of contributors. Make sure that they can clearly state their expectations for the participation of your community. Develop recognition and rewarding models for contributors to make them benefit from the creation of local applications (financial or material rewards, learning or professional opportunities, applications focused on local needs, etc.). Create warning systems to prevent exploitation of contributors and ensure that value for contributors is not derived from small rewards only.

Create transparent governance frameworks that describe roles and responsibilities, decision-making, remuneration, accountability, and consultation mechanisms for contributors. Start an early conversation with contributors on how they would like the community to be organized and roles distributed. Build steering structures that meet regularly where all the different groups participating in your community are equally

represented (e.g. Mozilla Foundation, contributors, partners and donors, potential users) and where you can discuss strategic decisions regarding the evolution of your open voice community.

Develop numerous and diverse partnerships with organizations in your community to share the ownership of the dataset and voice recognition model. Try to onboard as early as possible other organizations that can see the direct value of speech recognition tools in your language and genuinely support your efforts.

For donors, partners, and other enablers

Community efforts are not for free. Mobilizing a diverse and large community requires funding. Financial support is essential for open voice communities, especially for the building of inclusive datasets. Support these efforts with grants that can cover the costs of logistics, stipends, and small rewards. This financial support should be given in priority to local organizations from academia or from the public and private sector, especially when there are already organized groups working with voice technology in one specific language. Ensure that local organizations have the primary responsibility for the management of these funds. To ensure plurality, openness, and an equal distribution of benefits within the local ecosystem, formalize these delegations of coordination activities through partnerships that clearly state principles community coordinators need to follow to leave space for other interested individuals and organizations to join the community.

Localize knowledge and expertise by building capacities for local users that allow value to be created locally from open voice resources. Connect communities with international experts but also encourage peer-learning between communities.

Build strong local communities by providing capacity building and organizational development for community coordinators and mobilizers. Align with their priorities and needs. Support their communication and public relations efforts by helping them to engage in partnerships with other institutions.

Increase efforts to promote the global commitment to fund global public goods that benefit all and contribute to the realization of sustainable development goals by advocating for increased support by other stakeholders. Use pilot projects to demonstrate the viability of community efforts and onboard more partners.

For contributors

Clearly state your expectations regarding your contribution to your voice community and make sure that you understand how your expectations can be met. Raise your voice to ensure that you receive the recognition that you deserve. Engage community mobilizers and coordinators on how they believe you can reach your goals.

Give feedback on your experience as a contributor to community mobilizers and coordinators. Ask questions on the future of your open voice communities and participate in both online and physical MCV community spaces to make your voice count.

For users of voice technology

Invest in the open voice communities that create the digital public goods that you rely on. Try to contribute directly to the dataset or to set up future reward mechanisms for the community.

Use opportunities for networking with contributors, coordinators or donors and enablers during the sentence and data collection phases. Events and social gatherings can be used to build communities that help you create user-oriented applications.

Make use of the support opportunities that can be offered by Mozilla and other enablers for the development of solutions that are relevant for sustainable development goals.