



# MyVoice: Arabic Speech Resource Collaboration Platform

Yousseif Elshahawy, Yassine El Kheir, Shammur Absar Chowdhury, Ahmed Ali

Qatar Computing Research Institute, HBKU, Doha, Qatar

info@arabicspeech.org

## Abstract

We introduce MyVoice, a crowdsourcing platform designed to collect Arabic speech to enhance dialectal speech technologies. This platform offers an opportunity to design large dialectal speech datasets; and makes them publicly available. MyVoice allows contributors to select city/country-level fine-grained dialect and record the displayed utterances. Users can switch roles between contributors and annotators. The platform incorporates a quality assurance system that filters out low-quality and spurious recordings before sending them for validation. During the validation phase, contributors can assess the quality of recordings, annotate them, and provide feedback which is then reviewed by administrators. Furthermore, the platform offers flexibility to admin roles to add new data or tasks beyond dialectal speech and word collection, which are displayed to contributors. Thus, enabling collaborative efforts in gathering diverse and large Arabic speech data.

**Index Terms:** data collection, multi-dialect Arabic, speech recognition

## 1. Introduction

The field of speech and language processing has been transformed by the accessibility of large datasets, empowering the creation of advanced models that exhibit outstanding performance. However, the data preparation process can be costly, time-consuming, and, most importantly, encounter issues due to the under-representation of the language. These challenges can also impede progress and result in a centralized advancement, limiting the accessibility and utilization of these resources.

MyVoice<sup>1</sup> aims to foster a collaborative community by building valuable resources that can further accelerate speech and language technology advancements while promoting open access to diverse and large datasets for everyone. The platform is designed to collect Arabic data to improve dialectal speech technologies and bridge the gap within the Arab world [1], with data being accessible to everyone. MyVoice enables admins to host multiple tasks while allowing contributors to select and contribute to the tasks of their choice. For each task, a collaborator can record the displayed utterances, also can validate others recordings. The platform integrates state-of-the-art voice activity detection and dialectal speech recognition system for quality assurance and provides different statistics to the contributor and administration.

## 2. MyVoice

**Platform Architecture:** The platform consists of two main components: a front-end and a back-end. The front-end is re-

<sup>1</sup><https://myvoice.arabicspeech.org/>

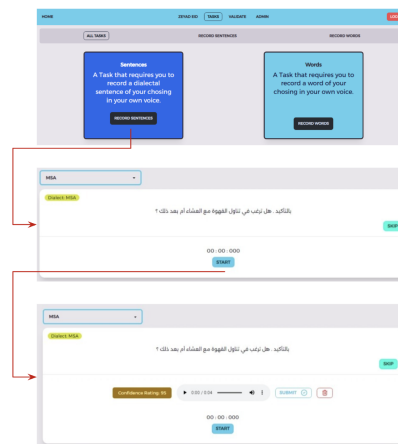


Figure 1: MyVoice Tasks Page

sponsible for providing an intuitive and user-friendly interface for recording and submitting audio segments, while the back-end handles the processing, storage, and management of the submitted audio data.

The front-end is built using the following: (i) **Nuxt3**<sup>2</sup>: which is a progressive framework for building web applications. It provides a powerful development experience with features such as automatic code splitting, server-side rendering, and static site generation; and (ii) **TailwindCSS**<sup>3</sup>: which is a styling framework that provides a set of pre-defined styles and components for building responsive and modern web interfaces.

The back-end of the MyVoice platform is built using the following technologies: (i) **FastAPI**<sup>4</sup>: FastAPI is a modern, fast (high-performance) web framework for building APIs. The Framework natively supports asynchronous programming, making it well-suited for high-traffic and data-intensive applications; (ii) **Uvicorn**<sup>5</sup>: Uvicorn is a fast ASGI server implementation that is built on top of the asyncio library. It's mainly used to deploy the FastAPI server for production; (iii) **Supabase**<sup>6</sup>: Supabase is an open-source Tool that provides a suite of back-end services, including database management, authentication, and storage. In the MyVoice platform, Supabase is used primarily for authentication and user metadata storage, allowing users to securely log in to the platform and store their

<sup>2</sup><https://nuxt.com/>

<sup>3</sup><https://tailwindcss.com/>

<sup>4</sup><https://fastapi.tiangolo.com/>

<sup>5</sup><https://www.uvicorn.org/>

<sup>6</sup><https://supabase.com/>

submission history and progress; (iv) **PM2**<sup>7</sup>: PM2 is a general process manager that is used to handle, monitor, and deploy the FastAPI server that powers the back-end of the MyVoice platform. PM2 provides features such as automatic process restarting, log management, and load balancing, which help to ensure that the server is always running smoothly and reliably. It also allows for easy deployment of updates and new features to the server, making it a key component in the development and maintenance of the MyVoice platform.

**Recording Tasks Interface:** The **Tasks** page is designed to offer contributors a range of options for recording tasks and submitting their voice data as shown in Figure 1. MyVoice provides the contributors with the flexibility to choose a specific dialect based on their experience and begin recording their voices given a displayed text.



Figure 2: *MyVoice Audio Validation Page*

**Validation Interface:** The **Validation** page is a powerful tool that allows contributors to view all of their recorded audio files in one place. This page also enables contributors to assess the quality of their recordings and make decisions on whether to submit them or redo them for better quality. By providing contributors with the ability to review their recordings and assess their quality, the **Validation** page ensures that only the highest quality recordings are submitted to the project. This additional layer of validation increases the overall quality of the voice data collected and enhances the accuracy of any research conducted using this data.

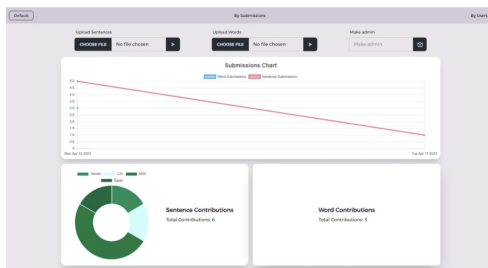


Figure 3: *MyVoice Admin Page*

**Admin Interface:** The admin page is a crucial component of the platform, it provides insights into the collected datasets and allows access to the submissions for a customizable timeline shown in Figure 3. From here, the admin can upload a new set of utterances to be recorded. The admin also has the ability to give admin privileges to any contributor.

<sup>7</sup><https://pm2.keymetrics.io/>

**Submissions Interface:** The admin can also inspect individual submissions as shown in Figure 4. Each submission is accompanied by important information, such as the task being performed and a confidence score calculated using a state-of-the-art multi-dialect Arabic speech recognition system [2] that reflects the quality of the recording. With this information at their fingertips, administrators are able to take action based on the quality of each submission. For example, if the confidence score indicates that the recording is of poor quality or outlier recordings, administrators can delete the recordings. This ensures that all data collected by MyVoice is accurate and reliable and that contributors are held to high standards of quality.

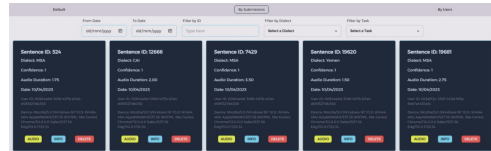


Figure 4: *MyVoice Submissions Page*

**Users Interface:** The admin can view all the current users on the page shown in Figure 5, and from there, the admin can inspect the details of a specific user. The admin can use the details for a specific user, as shown in Figure 5, to decide if the user is malicious or not, in which case they can block them from making further submissions as well as delete all their previous submissions.

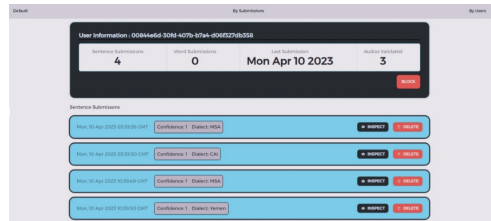


Figure 5: *MyVoice User Details Page*

### 3. Conclusion

MyVoice is a crowdsourcing platform introduced for collecting dialectal Arabic speech to enhance dialectal speech technologies. It allows contributors to record and validate large dialectal speech datasets. The platform includes an integrated quality assurance system to aid the validator and admin to assess the recordings before making them publicly available, thus ensuring the quality of the data. It also offers statistical insights into the data and promotes collaborative efforts in gathering diverse and large dialectal Arabic speech data for further advancing speech technologies.

### 4. References

- [1] A. Ali, S. Chowdhury, M. Afify, W. El-Hajj, H. Hajj, M. Abbas, A. Hussein, N. Ghneim, M. Abushariah, and A. Alqudah, "Connecting arabs: Bridging the gap in dialectal speech recognition," *Communications of the ACM*, vol. 64, no. 4, pp. 124–129, 2021.
- [2] S. A. Chowdhury, A. Hussein, A. Abdelali, and A. Ali, "Towards one model to rule all: Multilingual strategy for dialectal code-switching arabic asr," *arXiv preprint arXiv:2105.14779*, 2021.