

Design and Implementation of a Small Program for Collecting Ethnic Voice

Siyi Zheng

Southwest Minzu University, Chengdu 610000, China

Abstract: The Ethnic Language Collection Mini Program is an application based on mobile internet technology, designed to help users collect and record voice and text information of ethnic languages. The use of ethnic language collection mini programs can effectively promote the inheritance and protection of ethnic languages, helping users better understand and learn about ethnic culture. Language researchers can share their voice data with other users or research institutions to promote language research and protection.

Keywords: Audio Acquisition; Corpus; WeChat Mini Program.

1. Introduction

At present, Tibetan speech researchers have established many databases, but there are few open-source databases. On the one hand, the Tibetan corpus can promote the progress of building Tibetan database technology, on the other hand, it can save production costs and reduce the difficulty of Tibetan speech recognition. It can also enable researchers to focus on improving algorithm performance and exploring the application of Tibetan speech recognition. Speech technology is an important application technology in current fields such as artificial intelligence and intelligent speech interaction. However, in the Tibetan language field, due to the lack of relevant speech data and technical support, the application and development of speech technology are greatly limited. Establishing a Tibetan speech database can provide important basic data support for the application and development of Tibetan speech technology.

The application of WeChat mini programs is extremely extensive, and as a lightweight application, WeChat mini programs are the main form of artificial intelligence applications nowadays. At the same time, the user base of WeChat has laid the foundation for the development of WeChat mini programs. Applets are developed in two ways: WeChat native development and third-party framework development. Developers can use the development tools provided by WeChat for development and debugging. Applets are an important part of WeChat [1].

Therefore, designing a mini program based on WeChat platform to provide Tibetan language acquisition mini programs for Tibetan language researchers, establishing a Tibetan language speech database can provide important basic data support for the application and development of Tibetan language speech technology, and promote the further development and application of Tibetan language speech technology [2].

2. Methodology

2.1. WeChat Developer Platform

The WeChat developer platform is a platform specifically designed to provide services for WeChat developers. Developers can create their own WeChat official account, applets, enterprise WeChat applications, etc. on the platform,

and carry out relevant development, debugging and publishing. [3] The platform provides rich development tools and resources, such as API interfaces, SDKs, development documents, sample code, etc., making it convenient for developers to quickly develop and integrate WeChat related functions. In addition, the platform also provides management tools such as data statistics, user management, and payment management, making it easy for developers to manage their own applications. Through the WeChat developer platform, developers can easily develop various WeChat related applications, providing users with a better service experience [4].

2.2. WXML and WXSS

WXML is a set of tag languages designed in a framework, which can be used to construct the basic structure of a page by combining event systems, basic components, and other components. WXML supports functions such as data binding, list rendering, and conditional rendering, and can achieve dynamic rendering by interacting with JavaScript code through specific syntax structures [5]. WXSS (WeiXinStyleSheet) is a style language for WeChat mini programs, similar to CSS, used to set the style and layout of mini program pages. WXSS supports functions such as style inheritance, style coverage, and size units, and can also interact with JavaScript code to achieve dynamic styles through specific syntax structures. Unlike traditional CSS, the style rule names of WXSS cannot use uppercase letters, and do not support all CSS properties and selectors. They need to be used according to official documentation regulations [6].

2.3. Back-end Design

Java, for example, is the most popular server-side programming language currently used in fields with strong business logic, such as mall systems, ERP, finance, and other large-scale enterprise projects that require complex calculations and data processing. However, the overall development and learning costs are relatively high. The backend logic layer is developed using the SpringBoot and MyBatis Plus frameworks. Spring is a mainstream open-source Java framework, providing a container with Inversion of control and aspect programming capabilities. SpringBoot is an open-source framework based on the Spring framework, specifically designed to quickly create and run standalone

Spring applications. It allows developers to easily create a standalone, production level Spring application without the need to manually configure a large number of Spring configurations. MyBatis is an open-source persistence layer framework that maps Java objects and data in relational databases through XML or annotations to achieve data persistence operations. MyBatis Plus is an enhancement tool for MyBatis, which is an extension and enhancement of MyBatis and provides a series of practical functions, making the use of MyBatis more convenient, fast, and efficient

2.4. MySQL Database

MySQL is an open-source relational Database management system. It is one of the most popular open-source databases at present. Because of its stability, reliability, security and efficiency, MySQL is widely used in Web application development, data storage and other fields. MySQL data is stored in a table, which consists of columns and rows, with each column having a data type and each row representing a data record. MySQL's SQL language is a standard SQL language that supports various basic SQL operations such as querying, inserting, updating, and deleting. MySQL is widely used, especially in web applications such as blogs, forums, e-commerce platforms, etc. MySQL has become the preferred database for web developers[7].

2.5. Tencent Cloud Server

Tencent Cloud Server provides a variety of instance types with different configurations, including basic, standard, memory, computing, etc. Users can choose the appropriate instance type according to their needs. In addition, Tencent Cloud Server also provides various operating systems, data disks, security reinforcement and other services to ensure user data security and business stability.

3. System Analysis

3.1. System Function Analysis

Through research and analysis, the system mainly provides users with main functions such as corpus selection, corpus recording, corpus uploading, and corpus playback. After a user logs in to the system, the system will display the corpus they need to record based on their permissions. The main functions of the system are shown in Fig 1.

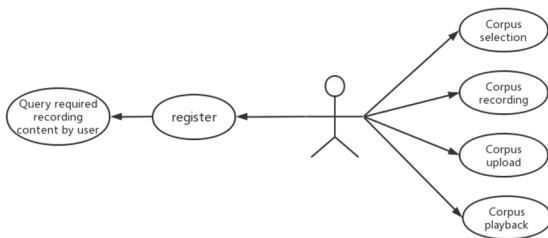


Fig 1. System Use case diagram

3.2. System Functional Module Division

Based on the above system functional analysis, the system is mainly divided into four functional modules: user login, query based on user permissions, recording audio, and uploading to the server. The system functional module diagram is shown in Fig 2.

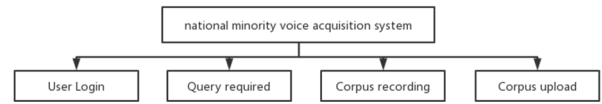


Fig 2. System Function Module Diagram

3.2.1. User Login

Before logging in, users need to enter the correct username and password. The system will verify the input information to ensure the legitimacy of the user's identity. After successful login, the system will fetch the account name and view the corresponding corpus information granted to the user with user permissions. If the user has not logged in, they cannot view the corresponding corpus.

3.2.2. User Recorded Corpus Query

After logging in, the system will display the user's corresponding corpus, and when the user logs in, they can see their own recording records. The recorded corpus will show that the recording is complete and there is no need to repeat the recording. After logging into the mini program again, the user finds the last recorded node and starts a new corpus recording.

3.2.3. Recorded Voice

Users can enter the recording interface by clicking on the corpus they need to record. After familiarizing themselves with the expected content, they can click the recording button to start recording. When the recording is completed, click Stop Recording, and the system will cache the file in the mini program. You can click the play button to listen to the recording effect. If the effect is not ideal, record again. If you are satisfied with the audio, click the upload button to upload it to Tencent Cloud Server.

3.2.4. Audio Upload

After the user clicks the upload button, the mini program will upload the audio file to Tencent Cloud Server. The file name rule takes the recorder+recording corpus number+recording time, ensuring the uniqueness of the audio name. Tibetan speech researchers log in to the Tencent cloud server backend and collect the transmitted audio corpus files to create an audio corpus.

4. System Design and Implementation

4.1. User Login Module Design

The user did not log in using the official OpenID provided by WeChat, but logged in using an account password. The account password entered by the front-end will be passed into the backend code login () on Tencent Cloud Server for verification, and the verification result will be returned to the front-end. This makes it more convenient to manage permissions and allocate corpus to users. If the account password is entered incorrectly, a password error message will be displayed and the page will be refreshed, prompting the user to log in again

4.2. User Recorded Corpus Query

After a user successfully logs in, the system will determine their permissions and access levels in the system by querying their role information, providing them with relevant corpus information to ensure that each user records different content. After receiving the request, the backend will use mybatis plus's list () method to find the user's corpus. After the user

completes recording the current corpus, the system sets the corpus to the completed state and automatically jumps to the next corpus to ensure system recording efficiency.

4.3. Recorded Voice

The user recording voice module is mainly divided into two functions:

Audio recording: When the user clicks the recording button, the mini program component used by the mini program is RecorderManager. The RecorderManager provides basic

recording operations, such as starting recording, pausing recording, continuing recording, stopping recording, and other functions. At the same time, the RecorderManager also supports setting recording parameters, as shown in the following figure. Such as recording format, sampling rate, audio quality, etc. After the recording is completed, the temporary path of the recording file can be obtained through the RecorderManager. When the upload button is clicked, the file will be saved to Tencent Cloud Server.

The following are the parameters used in system.

Table 1. Parameter Table

Property	Type	Value	IsNull	Details
Duration	number	6000	False	The duration of the recording
SampleRate	number	1600	False	Sampling rate
NumberOChannels	number	1	False	Default 2
encodeBitRate	number	96000	False	Default 48000
format	string	Wav	False	Audio format, valid values aac/mp3
frameSize	number	50	False	Specify frame size
audioSource	string	Auto	False	Specify the audio input source for recording

Audio playback: After the user has recorded the audio, they can first click the playback button to check the recording quality and determine whether a new recording is needed. After clicking the button, the mini program will call the innerAudioContext. Play () method to play the audio in the cache. When the user does not record an audio file, clicking the button will prompt: "Please record first".

4.4. User Upload

When the user clicks the upload button, the front-end calls the wx.uploadFile() method to send a URL request to the backend, passing in the username and audio file to the backend. The backend interface adopts a restful design style and is deployed on Tencent Cloud Server, with the interface parameter of Multipartfile to receive audio file data. The backend saves the file to the server file through the file.transferTo (newFile) method.

5. Database Design

5.1. User Table Database Design

Table 2. User Table

fieldName	Type	Size	IsNull	IsKey	Details
id	Int	11	false	true	id
username	varchar	255	true	false	username
password	varchar	255	true	false	password

The user information table includes three fields. The field ID is the user number, which records the number of users in the database; The field username is the username used to distinguish user information; The field password is the user password, as shown in Table 2.

5.2. Corpus Management Table

The corpus management table has six fields. The field ID is the corpus ID, which records the number of corpora in the database; The field Typeid is the type number of the corpus in the database, used for classifying different types of corpus; The field Typename is the corpus type name; The field title is

the title of the corpus, displayed at the beginning of the corpus, prompting the user to enter the title of the corpus; The field content is the corpus content, through which users record audio files; The field URL is the image address, and the title image of each corpus is different, As shown in Table 3.

Table 3. Corpus Management Table

fieldName	Type	Size	IsNull	IsKey	Details
id	int	11	false	true	Id
typeid	int	11	true	false	Type id
typename	varchar	255	true	false	Corpus name
title	varchar	255	true	false	Title
content	varchar	255	true	false	Corpus content
url	varchar	255	true	false	Photo url

5.3. User Corpus Recording Table

The user corpus recording table has four fields. The field ID is the user record number, which records every piece of data recorded by the user; The field userid is the user ID, and the mini program assigns different recording content based on different user IDs; The field 'sentimentid' is the corresponding corpus ID, which corresponds to the corpus ID recorded by the user; The field 'sound' is used to determine whether the user recorded this record. If it was not recorded, it is 0, otherwise it is 1, as shown in Table 4.

Table 4. User corpus recording Table

fieldName	Type	Size	IsNull	IsKey	Details
id	int	11	false	true	Id
userid	int	11	true	false	User id
sentimentid	int	11	true	false	Corpus id
sound	int	11	true	false	Recording or not

6. Conclusion

Applying WeChat mini programs to minority voice collection has two major advantages compared to the

development of apps and websites: firstly, it is easy and fast to develop, and can easily achieve powerful functions. Secondly, it can inherit the huge user base of nearly 1.2 billion on WeChat and reduce marketing costs. Users in ethnic minority areas do not need to install an app and can start voice recording by opening WeChat. The minority voice acquisition applet can help record and protect the voice and culture of these languages, prevent them from being forgotten and disappearing, and promote the inheritance of Minority language. The minority voice acquisition applet can make it easier for young people to learn and understand their own language and culture, and provide more ways and resources for the inheritance of Minority language. Minority language researchers can label the audio later, including text transcription, phonetic transcription, word segmentation, etc. Audio corpora and annotation information can be shared with other users or research institutions to promote language research and protection. The minority voice acquisition mini program can greatly facilitate and accelerate the process of language research, while also helping to protect and inherit the language and culture of ethnic minorities.

Acknowledgments

This work was financially supported by Southwest Minzu University Graduate Innovative Research Project (Project No. ZY20222555) fund.

References

- [1] Han Bo. Analysis of the Development Status and Prospects of WeChat Mini Programs [J]. *Research on Digital Media*, 2020, 37 (8):5-9.
- [2] Pengcui D, Ben D. Research on the Construction of Multimodal Corpus of Tibetan Teaching Taking Gannan Tibetan Middle School as an Example [C]. *Proceedings of the First International Conference on Education: Current Issues and Digital Technologies (ICECIDT 2021)*. Atlantis Press, 2021: 409-413. DOI:10.26914/c.cnkihy.2021.031631.
- [3] Wang Xiaoxing, Huang Jianchang. Analysis of Application Development Based on WeChat Mini Programs [J]. *Information Technology and Informatization*, 2021(03):23-25.
- [4] Liu Yanyan, Li Danyu, Ruan Haishan, Hu Yun, Shen Nanping. Development and usability test of a symptom management WeChat Mini Program for parents of children with cancer [J]. *Asia-Pacific Journal of Oncology Nursing*, 2022, 9(12).
- [5] Li Guangmin, Chen Wenjing, Yan Xiaowei, Wang Li. Modeling and Analysis of Group User Portrait through WeChat Mini Program [J]. *Wireless Communications and Mobile Computing*, 2022, 2022.
- [6] Ding Qifeng. Application of WeChat Mini Program in Xindong Gate Project [J]. *IOP Conference Series: Earth and Environmental Science*, 2021, 760(1).
- [7] Dang Tran Khanh, Huy Ta Manh, Dang Ly Hoang, Le Hoang Nguyen. An Elastic Data Conversion Framework: A Case Study for MySQL and MongoDB [J]. *SN Computer Science*, 2021, 2(4).