

COMP4801

Web based tool for Chinese character evolution

Project plan

Siu Lik Hang, Leo(3035125201)
Wong Tsz Hing, Chris (3035068908)

Project Background

Currently, there are some websites and online dictionaries that contain information about chinese characters. Please refer to appendix 1.

However, most of these tools have only one single focus. For example, some only focus on the pronunciation of the character, some only focus on meanings of a single character and some focus on the stroke order of a character. Users may need to browse through several websites to gather all information they are looking for. Also, these online resources may not be updated frequently. As a result, these resources may not accurately reflect the evolution of Chinese.

The target of this project is to create a web tool that contains as much information as possible about chinese characters. Moreover, those information have to be updated frequently to show the latest status of the characters, word and phrases. To achieve this, we allow users to contribute their own knowledge to the website.

Project Objective

The goal of this project is to create a website that can

1. provide a convenient tools for users to search for information they need.
2. users can also contribute to update the website so that we can keep the latest information.
3. Implement the website using a flexible database structure to allow user to enhance our database. User may input new type of data related to chinese character. For example, originally, there are only meanings and pronunciation about a chinese character. User may add a new column 'big5 code' if they wish to to store the big5 code in the system.
4. Open our source code to the public for the others to improve our website and create other similar database system for public interest.

Project Methodology

This project will act as a open source web based tool for any visitors of this website. The major feature list will be as follow:

1. Character directory

User can check different information of a Chinese character, such as pronunciation, radical, component, meaning and stroke. Also, user can check different forms of the same character (such as the archaic form and the simplified form) .

2. Word directory

User can check the meaning of a word or a phrase and the usage. Also, example sentences is given to demonstrate the usage of the word / phrase. For some words / phrases, we may also include the story behind to explain how they are created.

3. Search function

User can search for characters or words by giving different conditions, such as pronunciation, radical, component, meaning, number of stroke and other user input criterias.

4. Games

Base on the information in the database, we provide some games for users to test their understanding of Chinese characters. For example, there may be a game that required users to write a character using the correct order of strokes and a game that required users to answer the meanings of a word. Moreover, game content will be updated automatically when any information inside character directory or word directory has been modified.

5. User contribute function

We provide a platform for users to update the information on the website. The system can handle input of csv files, videos, images, sounds and hyperlinks. Moreover, the system can handle batch input.

We plan to implement the system using a flexible database structure. User can add new criteria of a character. After user added a new criteria, the dictionary layout and the search function can display the new changes correctly.

6. Revision history

Since we allow user contribution, we should record the audit trail to check the changes that have been made. Web administrator can roll back the changes if necessary.

Programming Language

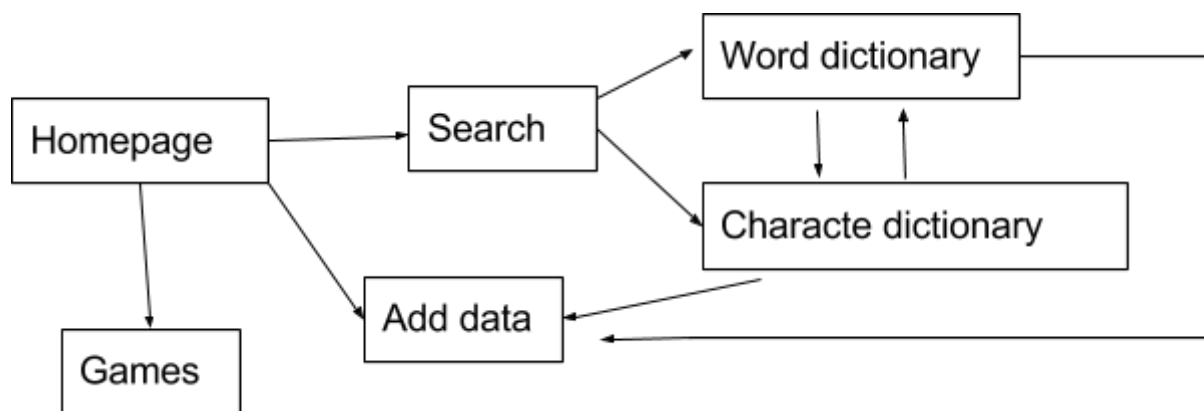
To build the website, we plan to use html, php and javascript to create the user interface. At the back end, we will use Mysql database to store the data.

The html and javascript will be used to define the basic layout of a page, just like a template. When user need to read the data of a specific character, the php script will be called to retrieve data from the database and place the data in the correct position .

However, since the database structure of our system has not been finalised, we may change to use nosql instead of Mysql, depending on which language is more efficient to implement our design.

Website structure

Basic Flow of the website



When user access our website, the first page he will access is the homepage. The homepage contain basic information of our system and links to different functions. He can access to the game page, the add data page or the search page.

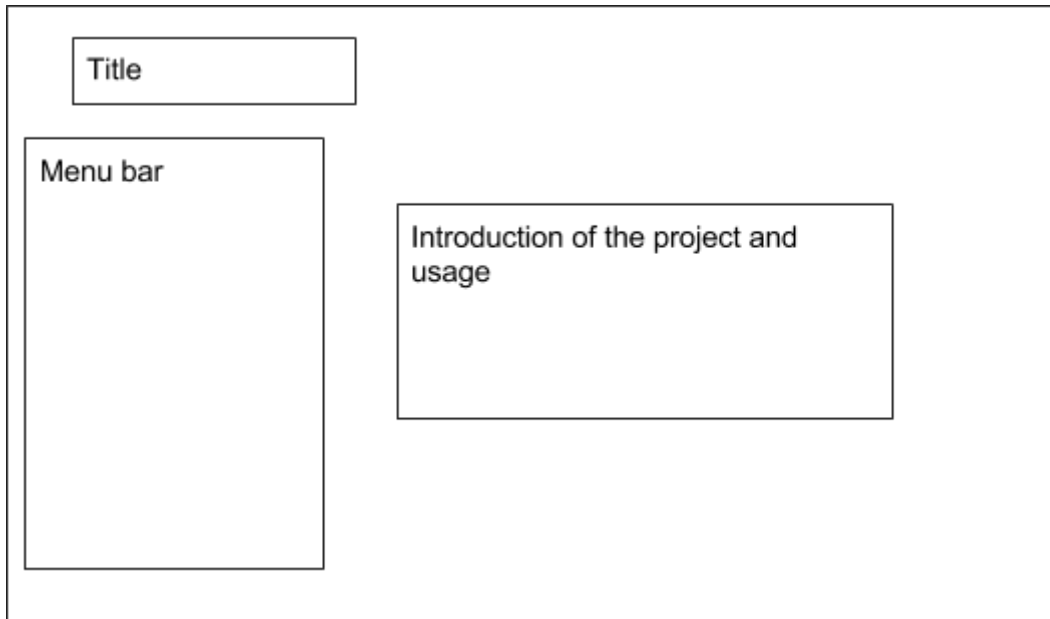
When user need to search for a character or a word, he can input the specific conditions on the search page. The system will guide the user to the correct page of the character or word. At the character dictionary page, user can access to the word dictionary page that is related to that character and vice versa.

User can enter the add data page either from home page or from the dictionary pages. When user access add data function from home page , he can choose to create a new record or edit an existing record. However, when user access the add data function from the dictionary pages, this indicates that user wants to edit

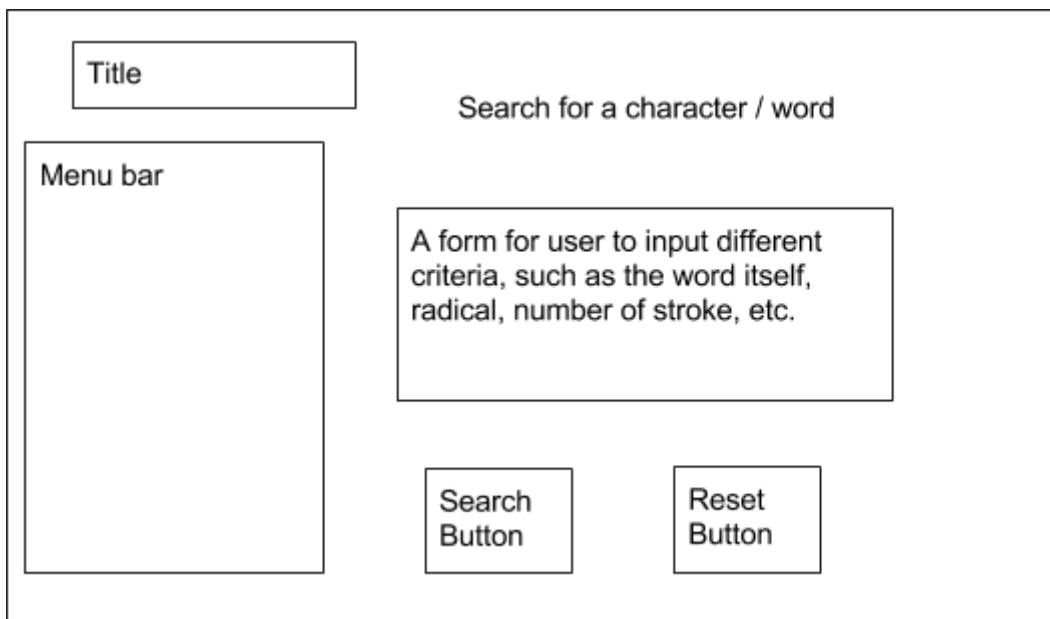
that record.

Sample UI design

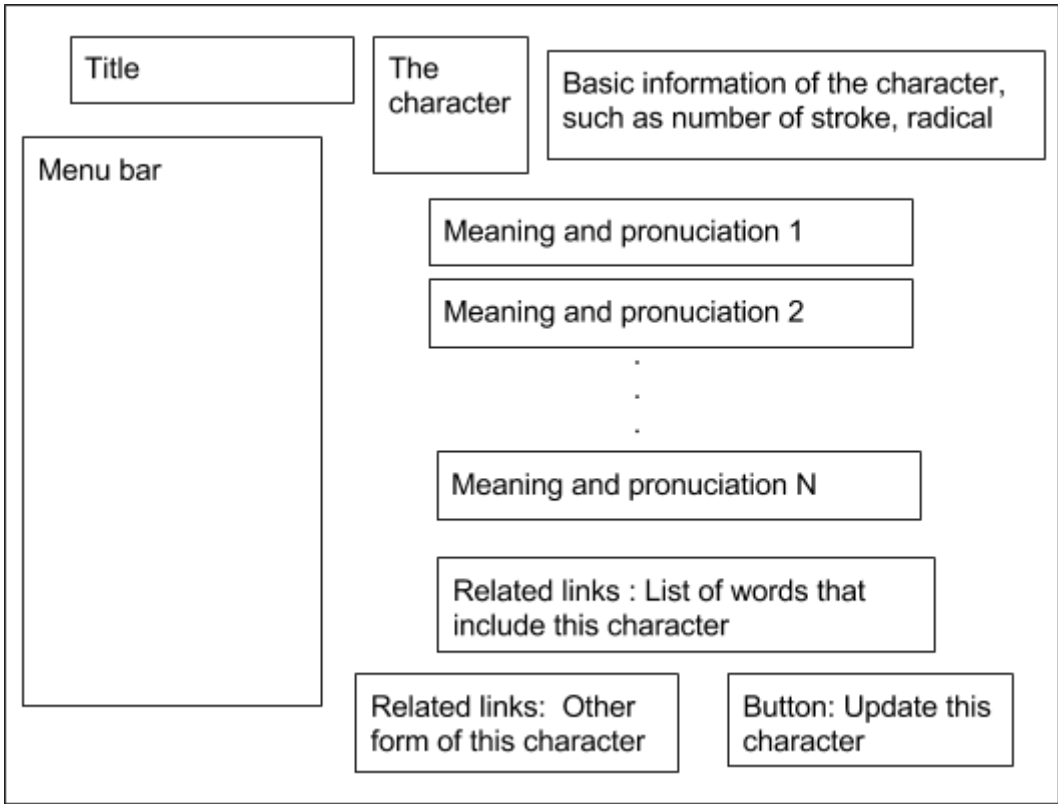
Homepage



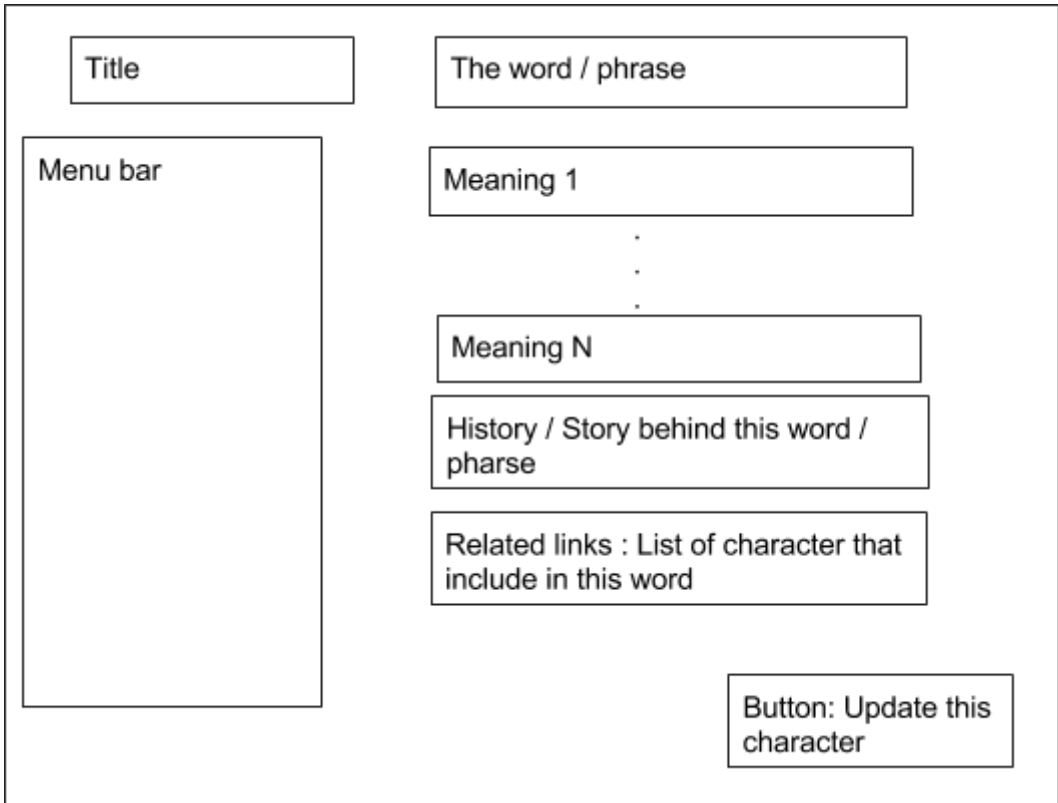
Search



Information of a character



Information of a word



User input

The diagram illustrates a user input form with the following components:

- Title**: A text input field.
- The word / phrase**: A text input field.
- Add**: A button to submit the word/phrase.
- Menu bar**: A large rectangular area on the left side.
- Meaning 1**: A text input field with an associated **Edit remove** button.
- Meaning N**: A text input field with an associated **Edit remove** button.
- History / Story behind this word / phrase**: A text input field with an associated **Edit remove** button.
- Button Submit**: A button at the bottom right of the form.

Major challenges

The major challenge of this project is how to store the information of the Chinese character in the database properly as the relationship between pronunciation, meanings and character are complicated.

Even for the same character, it may have different pronunciation and meaning according to the situation. For example, the character ‘樂’, is pronounced as *lok*⁶ when it means ‘happy / happiness’. However, when it means ‘music’, its pronunciation is *ngok*⁶.

For some situation, there are some words that have the same meanings and the same pronunciation, but they have different appearance. In fact, they can be replaced by each other. For example, ‘線’ and ‘綫’. Both of them mean ‘line’. They are identical. Therefore, using either one is correct. In our database, we need to record this relationship.

Moreover, some information of a character is difficult to express in database. For example, the correct order of stroke. We will need to design a solution of how to express the correct order of stroke of a character and store the data into

the database.

The other challenge is how to handle user input. As we plan to implement a flexible database system, we allow user to input multimedia data and add new columns. These input have to stored in the database correctly and display in the website. Also, if user added a new column, that column should also be abled to be searched.

Therefore, we need to design our interface and database structure well to handle these data so that users can update the content and express the relationships of chinese character in an easy way.

Project Schedule and Milestones

Timeline	Task
September	<ol style="list-style-type: none">1. Detailed project plan2. Project webpage
October to December	<ol style="list-style-type: none">1. Design and set up database2. Integrate with existing materials3. Website templete4. Game
December to January	<ol style="list-style-type: none">1. Search function2. Interim report
Febuary to March	<ol style="list-style-type: none">1. Improve user interface2. Testing
April	<ol style="list-style-type: none">1. Final report2. Final presation
May	<ol style="list-style-type: none">1. Project exhibition

Division of work

Detailed project plan	Leo Chris
Project webpage	Leo
Design and set up databse	Leo Chris
Intergate with exiting materials	Leo Chris
Website templete	Chris
Game	Leo
Search function	Leo Chris
Interim report	Leo Chris
Improve user interface	Leo Chris
Testing	Leo Chris
Final report	Leo Chris

System Resources

Webserver

A web server is needed for hosting the website. Currently, we decide to use the web server offered by the faculty.

Database

A database is needed to storing the data. Currently, we plan to use the MySQL server offered by the faculty. If we decide change to use nosql, we will use mongodb instead.

Appendix 1 Online resources about Chinese word and characters

Provide meanings of words and pronunciation in mandarin :

<http://www.zhongwen.com/>

Provide games for user to learn the special features of chinese character:

<http://www.dragonwise.hku.hk/dragon2/>

Provide articles to introduce special features of chinese character:

<http://chineseeffect.com/lessons/characters/>

Provide information of chinese character which primary students have to learn:

http://www.edbchinese.hk/lexlist_ch/

Provide pronunciation of Chinese Characters:

<http://humanum.arts.cuhk.edu.hk/Lexis/Canton/>

Provide information about stroke-order :

http://stroke-order.learningweb.moe.edu.tw/mobiles/word_query_form.tiles