

(Design) Essential Japanese words dictionary

Rationale

I want to learn Japanese.

Kanji and radicals are essential parts of the Japanese language.

The problem is kanji and radicals may have broad meanings.

Radical



さんづくり
hair-style, light rays



Moreover, some learning platforms stray from Japanese rules.

3

Triceratops

Go To

Name

Found In Kanji

Name

Primary Triceratops

Mnemonic

The three horns represent the **triceratops**, one of the most awesome dinosaurs ever.

Found In Kanji

半	当	光	消	削	鎖	尚	肖
はん	とう	こう	しょう	さく	さ	しょう	しょう
Half	Correct	Sunlight	Extinguish	Whittle Down	Chain	Furthermore	Resemblance

And some others lack features linking words by kanji or radicals.

こんにち
今日は

common word

j1pt n3

wanikani level 3

Links

1. hello; good day; good afternoon Usually written using kana alone, は is pronounced as わ

Other forms
今日は 【こんにちわ】

Notes
こんにちわ: Irregular kana usage.

Kanji in this word

4 strokes. JLPT N5. Jōyō kanji, taught in grade 2.

今 now
Kun: いま
On: コン、キン

Details ▶

Discussions

[Log in](#) to talk about this word.

4 strokes. JLPT N5. Jōyō kanji, taught in grade 1.

日 day, sun, Japan, counter for days
Kun: ひ、-び、-か
On: ニチ、ジツ

Details ▶

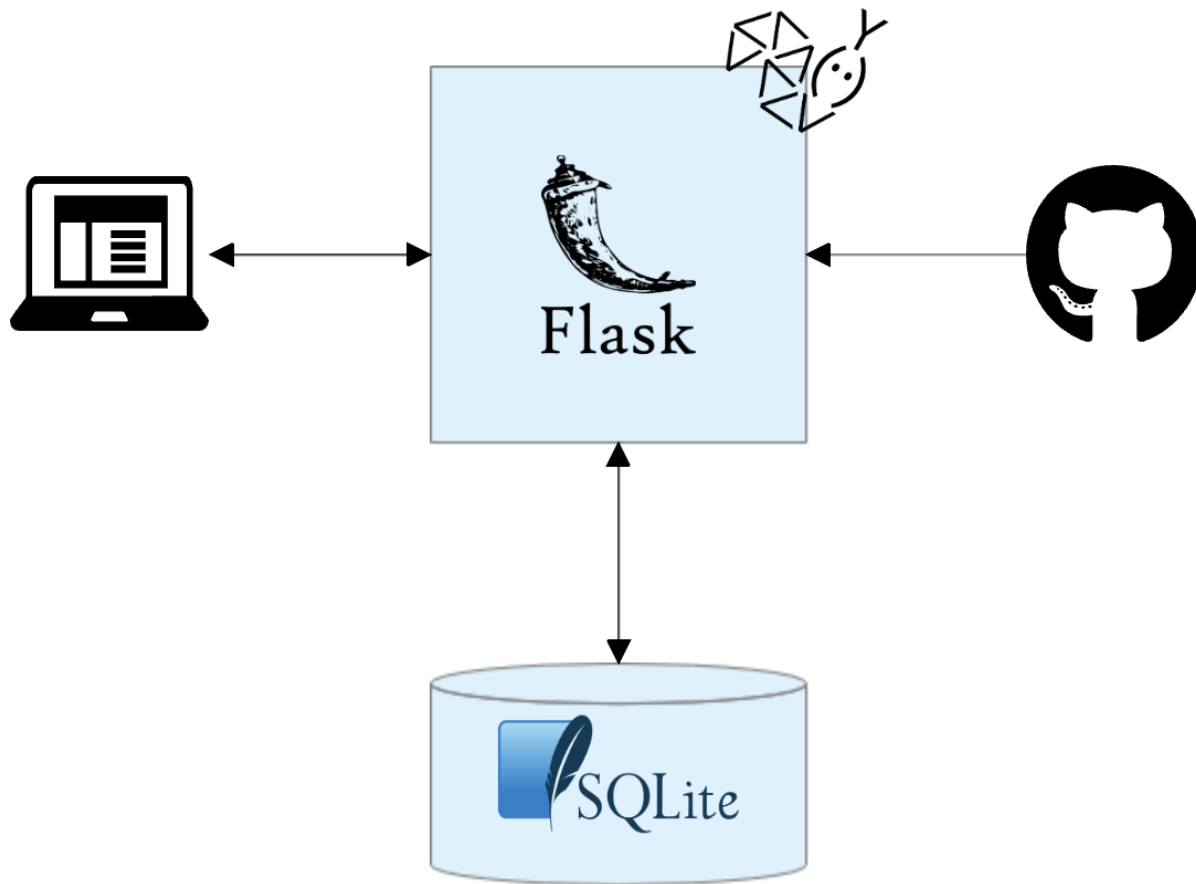
Luckily, Japanese words' definitions are more precise.

So, my solution is to create a word-indexed Japanese dictionary, where I can see essential words', along with their kanji and radicals', information. All in one page. A core feature is being able to Cmd + F in English or Japanese at any moment.

八月	はちがつ	August	八月 eight moon, month	八月 eight moon, month, period
八時間	はちじかん	8 hours	八時間 eight time interval, between	八日門 eight sun, day, time gate, door
尺八	しゃくはち	bamboo flute	尺八 shaku (unit of length, about 30 cm), measure eight	尸八 corpse, awning eight

Design

- Pull word, kanji, and radical data from the Kanji alive data repository.
<https://github.com/kanjialive/kanji-data-media/tree/master>.
- Create a Japanese dictionary as a SQLite database and display it via Flask.
- Display all 5812 words and their information in a single page.
- Deploy the Flask application on PythonAnywhere.



Implementation

- Pull radical data from the Kanji alive repository: <https://github.com/kanjialive/kanji-data-media/blob/master/language-data/japanese-radicals.csv>.
- Pull radical image data from the (forked and unzipped) Kanji alive repository: <https://github.com/leonardo-blas/kanji-alive-data-media/tree/master/radical-animations>.
- Pull word and kanji data from the Kanji alive repository: https://github.com/kanjialive/kanji-data-media/blob/master/language-data/kan_data.csv.
- Create Python dictionaries for radicals, kanji, and words.
- The radicals dictionary maps a radical to its radical and radical mnemonic image URLs.
- The kanji dictionary maps a kanji to its meanings and radical.

- The words dictionary maps a word to its hiragana spelling, English definition, and all the information in the kanji and radicals dictionaries.
- To populate the words dictionary, iterate over a word's characters to access the kanji dictionary in $O(1)$. From the kanji dictionary, use a kanji's radical to access the radical's dictionary in $O(1)$.
- Use the words dictionary to create a SQLite database. Only create it if it is not already created. Use a global flag.
 - In the words dictionary, one word may be mapped to many values, like kanji and radicals. These represent nested rows. Use `' ; '` to separate kanji within the kanji column, radicals within the radical column, etc.
 - JSON is an alternative to separators to add nested rows in the SQLite table.
- Display the SQLite table as a word-indexed Japanese dictionary using an HTML template.

What's next

- Consider displaying words only vertically, appending hiragana characters where needed.
- Consider splitting the dictionary into different pages to minimize crashes and reduce load time.