adambom / **dictionary**

⊙ Watch   12   ★ Star   185   ⑂ Fork   69

<> Code   ⓘ Issues **1**   Pull requests **1**   Pulse   Graphs

A JSON representation of Webster's Unabridged Dictionary

⊙ **12** commits   **1** branch   **0** releases   **2** contributors

Branch: **master** ▾   New pull request   New file   Find file   HTTPS ▾   https://github.com/adambom/di   ⬇   Download ZIP

adambom Merge pull request #6 from kazazes/master …   Latest commit a2b58a2 on Jul 15, 2014

.gitignore	Initial commit	3 years ago
LICENSE.md	MIT License/Project Gutenberg License	2 years ago
README.md	MIT License/Project Gutenberg License	2 years ago
_.jl	contains deprecated, use in	2 years ago
dictionary.json	Fixes #2.	2 years ago
dictionary.txt	JSON files validate now fixes #1 fixes #2	2 years ago
graph.json	JSON files validate now fixes #1 fixes #2	2 years ago
main.jl	newline fix	2 years ago

▤ README.md

English Language Dictionary

This repository houses the contents of Webster's Unabridged English Dictionary.

The dictionary can be found in plain text form here

You'll also find some julia files that were used to parse the text and organize it into the nice json you see here.

Contents

- dictionary.json: This is the raw data scraped from the dictionary. Unsurprisingly, it's in the format of a dictionary, i.e. `{ "Word": "Definition" }`
- graph.json: This is a graph representation of the dictionary. Each word is paired with a list of the words that define it
- dictionary.txt: This is the plain text file (I converted it from ISO-8859-1 to UTF-8)
- main.jl: The julia script that parses the data
- _.jl: My in-progress implementation of underscore in julia

How to run

If you want to run the code yourself, I would recommend downloading Julia Studio, which is and IDE for Julia. It comes with the binaries pre-installed. You can also get the binaries or build from source from julialang.org.

License

The works in this repository are licensed under the MIT License, with the exception of the contents of dictionary.txt, which are licensed under the terms of the Project Gutenberg License:

From Project Gutenberg:

> This eBook is for the use of anyone anywhere at no cost and with almost no restrictions whatsoever. You may copy it, give it away or re-use it under the terms of the Project Gutenberg License included with this eBook or online at www.gutenberg.net