

Indexer

@author Luke Blunden

@version Java 19

Description

This is a console-based application that creates a unique word definition index from a user specified text file or URL, using a dictionary csv file to find word definitions, and a common words txt file to identify words to ignore.

To Run

From console at .jar file directory:

```
java --enable-preview -cp ./indexer.jar ie.atu.sw.Runner
```

Then navigate through console options to set desired files to be used.

If the dictionary or stop words are not set, they will attempt to use the default files “./dictionary.csv” and “./google-1000.txt” respectively. If they cannot be found, you will be prompted to add the file path for them.

If input and output files have not been set on execution you will be prompted to add them.

If project is imported as dependency, the BookIndexer and UrlIndexer classes can be used as follows:

```
Indexer bookIndexer = new BookIndexer("[filePath]", "[dictionaryPath]",  
"[stopWordsPath]");  
bookIndexer.index();  
bookIndexer.printIndex("[outputFile.txt]", [true/false(outputOrdered)]);
```

Features

- Specify the text file or URL to be indexed.
 - Input either text or url
 - Input path of file (including extension) or URL
- Specify the dictionary and stop word files to be used or leave as default.
 - If specifying, input path to dictionary csv file or stop words file (including extension)
- Specify the text file including directory path where you want to save the index to. Indexing can be done in alphabetical or reverse alphabetical order.
 - Input name of output including .txt extension
 - Specify whether you want alphabetical or reverse alphabetical ordering, n for alphabetical and vice versa
- The index will be formatted to show word definitions indented to the right of the word, with a page reference underneath if a text file was indexed.
- The index will have a unique word count found at the top of the outputted file.