

Technical Design Document

Wikipedia page topic

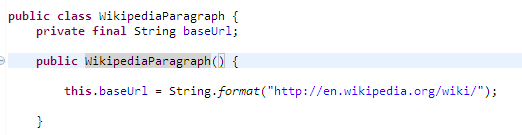
Write a command-line program that prompts the user for a topic at the terminal, and then prints out the introductory paragraph of the Wikipedia page for that topic, without HTML tags. The program should ask the user to enter the query at a prompt. If the user does not enter a topic (enters a blank one), they should be prompted again. If there is no page for that topic the program should print "Not found." and terminate. The program should also terminate after printing a successful result.

All you should need to know for this task is that English-language articles for Wikipedia are found at the following URL: <http://en.wikipedia.org/wiki/><Topic>. When there are spaces in a topic, the spaces should be replaced with underscores. You will receive bonus points if the topic can be provided as a command-line argument as well as at a prompt if no argument is provided.

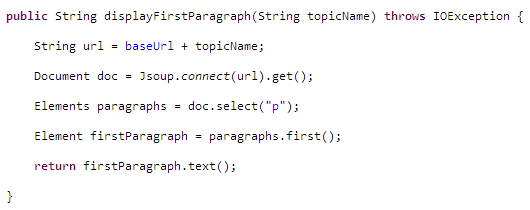
# Overview

This document describes the design of Wikipedia topic display for first paragraph by the user requirement. This program is developed with core java along third party API jsoup.

# Code Review

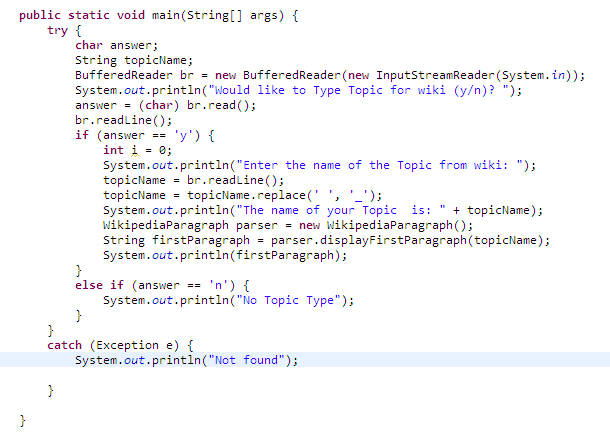


The project class name is WikipediaParagraph. I create constructor where base URL initialize at the run time.



I write one method with passing parameter named ‘Topic’ and create requirement URL by adding with base URL.

Than I used third party API named Jsoup for connect with URL and create document instance. Now from document instance I choice select method with parameter where need to pass ‘p’. Here p is the HTML tag paragraph and can get element instance. No from element instance I can get first method to display the first paragraph of the web page and return as text format.



This is the main driver method for run the program. I make buffer reader instance to read the character from key board. If you type correct topic than you can get display otherwise not found will display. I make the object of the class and invoke the dispalyFirstParagraph method.

# Conclusion

[jsoup](http://jsoup.org/) is an open source Java based HTML parser that we can use to parse HTML and extract useful information. jsoup API can be used to fetch HTML from URL or parse it from HTML string or from HTML file.