

**Project Name: Your Own Spider (Multi-threaded Web Crawler)**

**Summary:** This project aims to create a web crawler to scan webpages for links to other websites. Instead of using single threads, it takes advantage of multithreading in order to travel different paths to source as many links as possible concurrently (divide and conquer). The programming language chosen to create this project is Java, due to its multithreading capabilities and the fact that I already have a solid understanding of this language. The program uses the theme of a virtual pet to complete this task because web crawlers are also called spiders. So, the program involves some interactivity and personalization from the user.

**Input:** Using console prompts, the user will be asked to provide a name for the spider, how deep should the spider search (you are able to input a number between 1 and 3, however for testing purposes I would recommend 1 to minimize runtime), and the initial website the spider should start the scan at. The WebCrawler object will take these parameters and will begin the search. During the search, there is an animation of a spider crawling.

**Output:** Once the crawl is done, the user will be notified and given the number of links accessed. There will be another question asking if the user would like to see the list of visited links, you can respond yes or no. Finally, this is a goodbye message from your spider.

Instructions: I used Visual Studio Code to run this program, so these instructions are to be run on that IDE. Make sure you have the extension pack for Java support. In Explorer, choose the option to Open Java Project Folder (included in my zip file named "OS Project"). This should have everything you need to run the code, including the Jsoup library. If this does not work, create a Java Project Folder, which should create three folders, bin, lib, and src (Only src is important). I have included the .java files and Jsoup .jar files individually in the folder titled "Plan B" in case the direct Java Project folder does not work. Under '**Reference Libraries**', click the plus sign, and select the jar file for Jsoup. Place the .java class files under the "src" folder, and delete the premade 'App.java' file. "SpiderTester.java" is the tester file that should be run to start programs.

Project Summary  
Brooke Harris  
Operating Systems - A

Below, there is an image of how the Explorer tab should look in order to run the code as well as input I gave during my testing of the program, feel free to change any inputs to your own (Doesn't show the animation unfortunately, and list of links is not included so that it would fit on this page). Have fun!

```
----      Create Your Own Spider!      ----
What will your spider's name be? >> Arachne
How deep to do you want Arachne to search? (a number from 1 to 3) >> 1
What link do you want Arachne to start at? >> https://www.nbc.com
Successfully created webcrawler.
Starting crawl...

      ( _ )      Look at Arachne go!
...../>\oo)
Finished crawl.

There are 584 links.
Do you want to see the links visited? (Y or N) > N
Thank you, bye! -Arachne
```

