

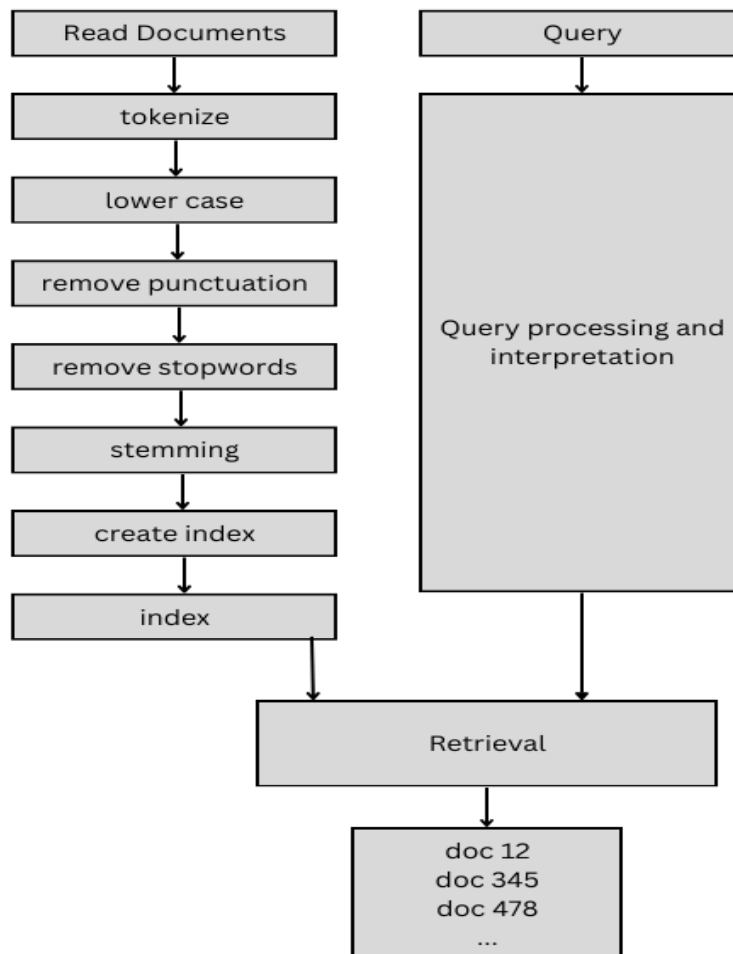
CSC790: Information Retrieval and Web Search
Spring 2024, Assignment #2
Date Assigned: Monday, February 12, 2024
Due Date: Wednesday, February 21, 2024 at 11:59 pm
(on Brightspace) (130 points)

Objectives:

- Build a simple search system.

Tasks

1. Modify the code for homework 1 in order to perform the processing steps as shown in the figure below (left side).



2. Write new code to load the inverted index from question 1 (homework 1) and answer user queries as follows

- Load the user queries from a file.
- Each query will have a maximum four words
- For each query provide the answer using all possible combination. if the query is: q= A B C, the answer should be as follows:

```
===== User Query 1: ..... =====
===== Results for: A and B and C =====
file 12
file 59
...
=====
===== Results for: A and B or C =====
file 12
file 59
...
=====
===== Results for: A or B and C =====
file 12
file 59
...
=====
===== Results for: A or B or C =====
file 12
file 59
...
=====
===== User Query 2: ..... =====
...
```

Important

- Use one function to display course and student information as follows:
===== CSC790-IR Homework 02 =====
First Name: your first name
Last Name : your last name
=====
- Document your code: write comments to explain the role each function and block of code; what are the input parameters and the output/return parameters.
- You must use NLTK.

Submission

1. Write your own code. Use as many functions as you can.
2. Make sure you writing you name and assignment number on all files you submit.
3. Your python code and the instructions on how to run it. You should submit two .py files, one for the modified version of homework 1 and one for homework 2.

4. Enclose all your files in a folder named **HW02_yourlastname.zip**.
5. Submit the zip file using Brightspace.