

CSCI 4030/ CSCI 6030/ DASC 6030: IR

Programming Assignment: 02 - Response

1 Assignment Goal

Develop expertise in constructing positional indexes and processing phrase and proximity queries.

2 Instructions for Compiling and Running the Program

Please provide instructions for compiling and running your program.

Please run the program with a space followed by the path of the folder to the corpus.
Exp: "pythoncode.py corpusFolder"

Note: depending on what os you are running you might need to change a segment of the code to accept the type of slashes needed for the file path.

Simply open the python code in a text editor and search for "switchsos". The comment should tell you what to do for what os you are using.

The program should then run; the program will ask for a phrase/proximity query, please do not add a query phrase to the command line.

3 Testing the Solution

Briefly describe how you tested your program. List your test cases and the expected results. Did your program pass all the test cases?

Your answer goes here.

Here is a way to show your execution results.

```
Enter a query, or type '~help' for info > document
'austen-persuasion.txt' has 1 match(es)
'melville-moby-dick.txt' has 1 match(es)
'shakespeare-hamlet.txt' has 1 match(es)
'chesterton-brown.txt' has 5 match(es)
'chesterton-ball.txt' has 4 match(es)
```

```
Enter a query, or type '~help' for info > armadillo
```

No results are found for the query.

Enter a query, or type '~help' for info > document /5 ink
Document 'chesterton-brown.txt':
--14879

Enter a query, or type '~help' for info > hand reached
Document 'milton-paradise.txt':
--35139 --68531

Document 'melville-moby-dick.txt':
--62956

Enter a query, or type '~help' for info > screamed /20 angrily
No results were found for the query.

Enter a query, or type '~help' for info > youngest /4 daughter
Document 'austen-sense.txt':
--52841

Document 'austen-persuasion.txt':
--45774 --45787

Document 'austen-emma.txt':
--53 --52143 --59262

Enter a query, or type '~help' for info > youngest /1000 daughter
Document 'austen-sense.txt':
--52841 -110937 -112712

Document 'austen-persuasion.txt':
--1481 --15411 --45774 --45787

Document 'bible-kjv.txt':
--33758 --33808 --33920 -172778 -225345 -279285 -348036 -348157

Document 'whitman-leaves.txt':
--69509

Document 'edgeworth-parents.txt':

--533 -127024

Document 'blake-poems.txt':

--5443

Document 'austen-emma.txt':

--53 --31623 --52143 --59262