

Exercise Sheet 1

Submit until Tuesday, October 29 at 4:00pm

Exercise 1 (5 points)

Please register at our course system Daphne (using your RZ Account + password for authentication). Make sure that your data is correct, in particular, that you are reachable under the specified e-mail address.

Exercise 2 (10 points)

Extend the *InvertedIndex* class from the lecture by two methods *intersect* and *processQuery*.

The method *intersect* should compute the intersection of two given inverted lists, as explained in the lecture.

The method *processQuery* should compute the ids of all records containing all words from a given query, as explained in the lecture.

For each method, write a unit test for a non-trivial input (with > 2 query words for *processQuery*).

Write a main program *SearchMain* that uses your extended *InvertedIndex* class to build an inverted index from a given text file (as shown in the lecture), and then answers arbitrary keyword queries entered by the user. For each query, output at most 5 records containing the query words. If you want, highlight the query words in the output, e.g. using ANSI escape codes.

You find a code design suggestion + some more implementation advice for this exercise on the Wiki. These suggestions are language-independent, and the corresponding files end in *.H*. You should consider these suggestions, but you do not have to follow them. Think twice when you deviate from them though. There is usually quite a bit of thought and a lot of programming experience behind our suggestions.

Exercise 3 (5 points)

Check out a working copy of your folder in the SVN repository of the course (see URL on your Daphne page), add your code to a new sub-directory *exercise-sheet-01*, and commit it. Make sure that everything runs through without errors on Jenkins (compiler, unit tests, checkstyle).

Also commit, in that sub-directory, a text file *experiences.txt* where you briefly describe your experiences with the first exercise sheet and the corresponding lecture. As a minimum, say how much time you invested and if you had major problems, and if yes, where.