Applied Textmining

First Assignment

Checkout the course project from *https://github.com/dimalabs/applied-textmining* and follow the instructions in *INSTALL.txt* to set it up on your local computer.

1. Keyword extraction

Have a look at *de.tuberlin.dima.textmining.assignment1.KeywordFinder*. This class is fed with a text and must extract a number of keywords from it.

Run our testcase *de.tuberlin.dima.textmining.RunAssignment1* to process the script of the first *Star Wars* movie and find *10* keywords from that. The keywords should be meaningful and ordered descendingly by their relevance for the text. Be creative :)

We will test your solution on a different text and show exceptional examples in the next lecture (unless you tell us not to).

Deadline

Please upload your solution as a patch file to the *ISIS* system until 16th of October 2011.