

The Ninth International Conference on Advances in Databases, Knowledge, and Data Applications, Mai 21 - 25, 2017 - Barcelona, Spain

**Tutorial: Data Manipulation and Data Transformation using the Shell
Andreas Schmidt, Steffen Scholz**

Exercise I

Tasks:

1. Download the book „The Adventures of Tom Sawyer“ from <http://www.gutenberg.org/ebooks/74> (utf-8 format) and store it locally.
2. For cygwin users only: Convert file to Unix Format with command:
`dos2unix.exe <file>`
3. Browse through the first pages of the book using `less` (for help type `man less`)
 - a. Go to line 1234 of the file. What ist the third word?
 - b. Search for the first apperance of Alfred (what is his last name?)
 - c. Go tot he end of the document
 - d. Scroll backward 5 pages
 - e. Show only lines containing the word „Tom“
 - f. Quit less
4. How many chapters has the book (hint use `grep` with `-a` option)?
5. Count the number of empty lines in the file
6. Execute `grep Tom <file>` and `grep -o Tom <file>`. What is the difference?
7. How often does the names „Tom“ and „Huck“ appears in the book?