**Executing SetRank – Documentation**

* clone github repository into arbitrary folder on local environment
* open Git BASH in folder SetRank and change directory into folder
* enter the master branch and follow implementation instructions in the provided README.rmd
* retrieve Elasticsearch 5.4.0 from

https://www.elastic.co/downloads/past-releases/elasticsearch-5-4-0

and save locally to change configurations

* open the elasticsearch/config/elasticsearch.yml and add the following two commands: (Where exactly add the lines? In the end of document? With # or not? Done without #)
  + script.inline: true
  + script.ingest: true
* start the elasticsearch engine in the bin folder with “elasticsearch” (In terminal 🡪 requires installing the shellcheck package in editor to execute 🡪 further requires to install JDK or set JAVA\_HOME on Windows System
* take necessary pre-steps to run files
* install requirements.txt (or elasticsearch 5.4.0 and textblob 0.13.0 locally in editor) open project in editor and add code fixes to files as follows:
  + #/code/SetRank/autoSetRank\_TREC.py

#Line 343 🡪 When I open file, it is Line 342

- rankings = all\_docno\_rankings[query\_id]

+ rankings = all\_docno\_rankings[int(query\_id)-1]

* + #Line 403

- if args.mode == "query": # save results only for query level aggregation

+ if args.agglevel == "query": # save results only for query level aggregation

* + #/code/SetRank/setRank\_ESR.py

#Line 40 and 45

- if len(line) != 3:

+ if len(line) < 1 or len(line) > 3:

- type = line[2]

+ type = line[2] if len(line) > 2 else "."

* + #/code/SetRank/setRank\_TREC.py

#Which Line??? Line 383 for me

- parser.add\_argument('-output', required=False, default="../results/trec/setrank.run",

+

* mklink /D TREC\_BIO TREC-BIO for Windows(cmd als Admin ausführen!) in command line or ln -s TREC\_BIO TREC-BIO for Linux (anders herum! zuerst <link> dann <target>
* add the downloaded datasets to the correct folders (takes quite some time as the data provided is huge) 🡪 WinRar needed to unzip tar.gz file in Windows
* querying is only provided to reproduce Table 3 from the original paper, therefore, query differences must be executed separately for each algorithm (4 times baseline algorithms + SetRank)
* creating indexes for baseline algorithms is also necessary manually in elasticsearch
  + Remark: not necessarily python3 (depends on version that is used) 🡪 for JL simply “python”
* Create .run files 🡪 does not work as proposed in README, have to call .py files directly in PyCharm (editor) and set parameters locally to execute and create .run files in results folder
* Execute pytrec\_eval for setRank\_tuned.run files and retrieve evaluation values