

Chia Huang

#25475733

Fernando De Paz

#81962579

1. I got 'conflict' and 'interest' as usable tokens and ran a query to search for the word and their information. We got the pages of the occurrences of both 'conflict' and 'interest'. 'Of' token was left out since 'of' is part of the stop words.
2. Again we used sql query and ran the command 'GROUP BY word' and did an aggregate function of 'SUM (word)' and got a total of 50712 search results.
3. The estimate size of the file is around 6,656 mb but that is the table for the query. We won't able export the query table to a text file since the file size was big, the export took a lot of time. It was written to a .db file format