

Our program uses the tokenizer from assignment 1, and the sortest-list from assignment 2. The tokenzier is tweaked however, to take in files instead of strings. It also uses all characters that aren't letters or numbers as tokens instead of tokens the user can input. We also created an indexer.c that reads in a file or directory and calls the sorted-list, inserting 1 word at a time from the files. If the file is valid, it tokenizes it, inserts it ointo a list, then outputs into a file. We used linked lists and hashmaps.

Analysis:

Iterate over m input files -  $O(m)$

Tokenize file of length in into k tokens -  $O(n)$

Write the token into the file -  $O(1)$

Final:  $O(mn)$