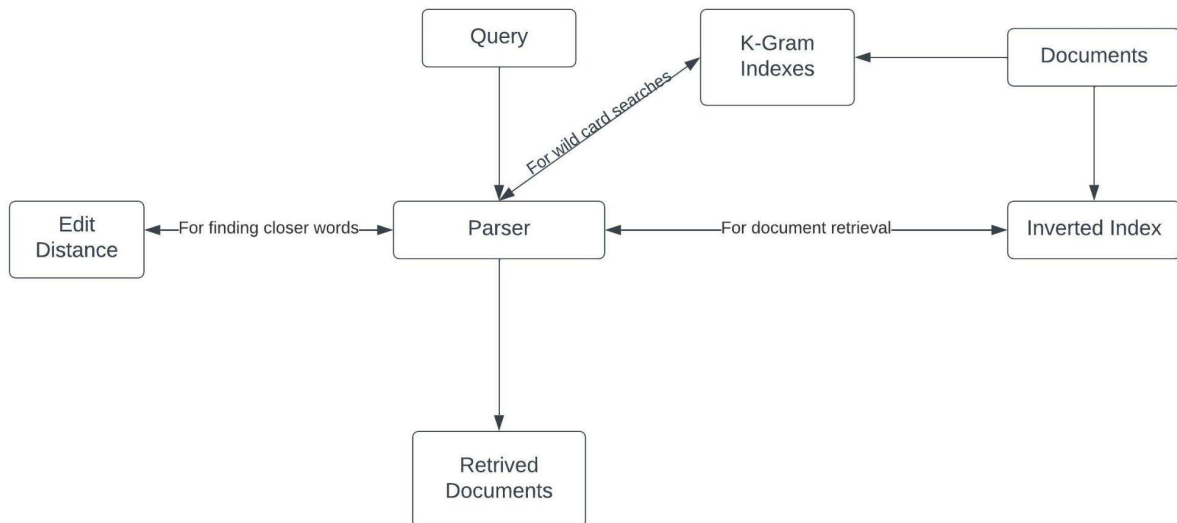


Design Document

The architecture of the application:

Both Inverted Index and Kgram index use a dictionary with LinkedList as values as a Data Structure, The linkedlist in the dictionary has skiplist implemented to reduce the time complexity to $O(\sqrt{n})$.

Inverted index has all words of all documents as keys and linked list of doc IDs as values, and k-gram index has 3-grams as keys and linked list of corresponding words as values.



Running Times:

Pre-processing & Building Index:

IRSystem succesfully built in 7.608170986175537 s

Search/Retrieval:

Give your query to get results/press EXIT to end the loop :
BRUTUS AND CAESAR AND NOT CALPHURNIA

5 Documents found relevant for your search :

1. antony-and-cleopatra_TXT_FolgerShakespeare.txt
2. hamlet_TXT_FolgerShakespeare.txt
3. henry-vi-part-2_TXT_FolgerShakespeare.txt
4. henry-v_TXT_FolgerShakespeare.txt
5. titus-andronicus_TXT_FolgerShakespeare.txt

Info retrieved in 0.005975246429443359 sec