REPORT FOR C

1. **Function Design:**

The program was designed with modularity in mind and was hence implements 4 separate but complimentary functions

Get\_documents takes in the filename as input and return a dictionary of which the keys are number from 1 to n and the values are the individual documents contained within the input file

Get\_index takes the documents dictionary and return a new dictionary called index, which has as keys individual words from all of the documents texts and as values set that contains the documents indices where the word appeard

Find\_documents returns the set of documents that contain all of the words queried by the user

Main contains the body of the program and the user interface. It hands the call to all previous functions

1. **Data structures**

Two data structures that were heavily relied on in this project: set and dictionaries

Dictionaires are used both to index the dictionaries and create the words index, they provide easy key value pairing.

Sets are used to store the documents where a certain word appeared, this choice is motivated for two key reasons:

1. The guaranteed uniqueness of each item inside the set makes it easy to avoid duplicates while creating the index
2. The intersection method makes it easy to find the key documents that cite all the queried words
3. **Error handling:**

Since the user does not choose the filename themselves, the error handling in our program is limited to invalid inputs at the interface level, which is handled with simple if-else statement

1. **The difficulties encountered:**

No significant difficulty was faced during the development of this program