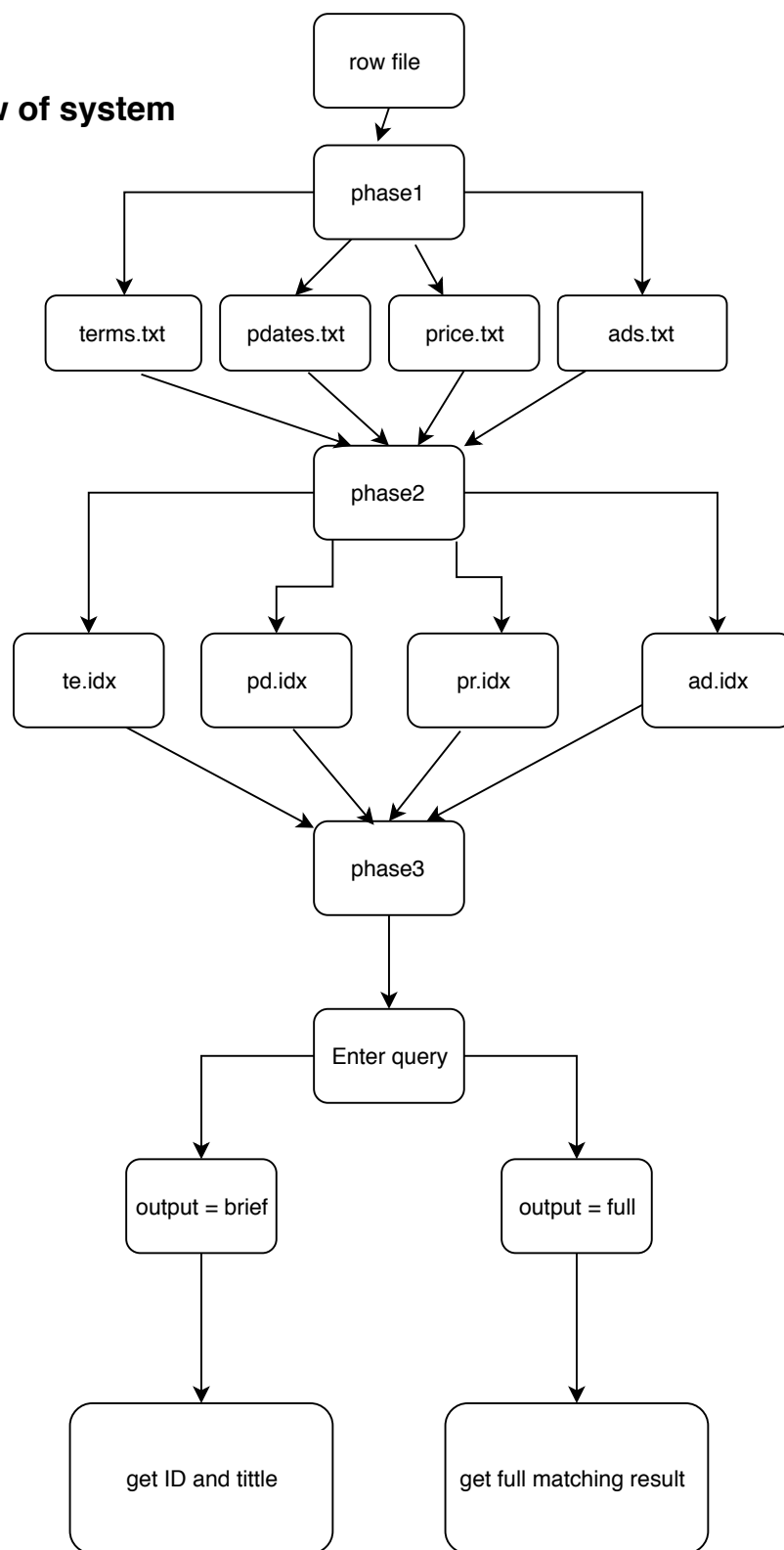


## General overview of system



# Report

## Small user guide

At the beginning of our program, we run phase1.py first and user need to input task file which provide the path. Then the system will create four txt files that is terms.txt, pdates.txt, prices.txt and ads.txt. Then run phase2.py, the four txt files which generate in phase1 will be sorted and 4 index file, ad.idx, te.idx, pd.idx, pr.idx. In phase3, user need to enter conditions that select information from idx. The conditions satisfied by title, description, post date, location, category and the price. User can select brief or full matching results. For each brief result, system only return ID and tittle of matching result. For each full result, system return will return id, tittle, description, category, price,price, location and date of matching results. If the conditions that user entered is not exist, system will return 0.

## Description of algorithm for efficiently evaluating queries

For each function, we build an aid empty list. When the conditions that user input match the information in database, system will return all matching ID and gather all matching id in aid list. Then make use of aid list in brief or full function to find all matching results.

## Testing strategy

Our testing strategy in phase1 is compare the 4 txt files with the file that provided by professor. We test phase3 by enter conditions include title, description, post date, location, category and the price. We have to check if system can print correct output if user enter different kinds of input. For example, ignore big digit or small digit of key words, ignore space between key words and get the valuable information, % is wild card symbol to search key words. If there is multiple conditions of input, whether the output get all matching output. To check whether we can get

the correct output, we use output = full that can see if all matching output is correct. To make sure we get all matching information, we create database and write query to search in database and compare the result with the output we get.

## **Group work break-down strategy**

My group made up of three members: Qi Yang, Ziyuan Liu, Jianfei Ma. Jianfei wrote the search.py and dateup.py to help Ziyuan better write phase3. Ziyuan wrote the main function in phase3. I wrote Phase1, report and write phase2 with Ziyuan. We spend around 3 to 4 days to finish this project. I wrote phase1. At first I didn't think about the condition of underscore '\_' and dashed '-' characters in terms file. Then get remind by Jiabin Fan, so I add this situation. Ziyuan at first didn't get the correct output when have multiple conditions. For example, when enter location = edmonton data=2018/11/07, he got the output of location=edmonton and the output of data=2018/11/07 rather than common part of these two conditions. Jianfei at first don't know how to get the correct date time if adding some dates on current date. After discussion, he changed the string into datetime format and then change it back. The method to get good collaboration is we take this whole apart and each member is responsible for own part. When meet problems, we discuss together and help each other to check.