Search Engine Project

Name: Soumil Nitn Shah
Id: - 1031686
Project: Search Engine

Screen Shots:
Terminal Screen Shots

```
/Users/soumilshah/CLionProjects/SearchEngine/cmake-build-debug/SearchEngine
1 : Crawling on : /Users/soumilshah/CLionProjects/Assingment3/www.cplusplus.com.txt


2 : Crawling on : /Users/soumilshah/CLionProjects/Assingment3/www.bridgeport.edu.txt


3 : Crawling on : /Users/soumilshah/CLionProjects/Assingment3/www.olympic.org.txt




Crawling Finish


Enter the Search Word
Bridgeport
Word is Present /Users/soumilshah/CLionProjects/Assingment3/www.bridgeport.edu.txt
Enter the Search Word
Q

Process finished with exit code 0
```
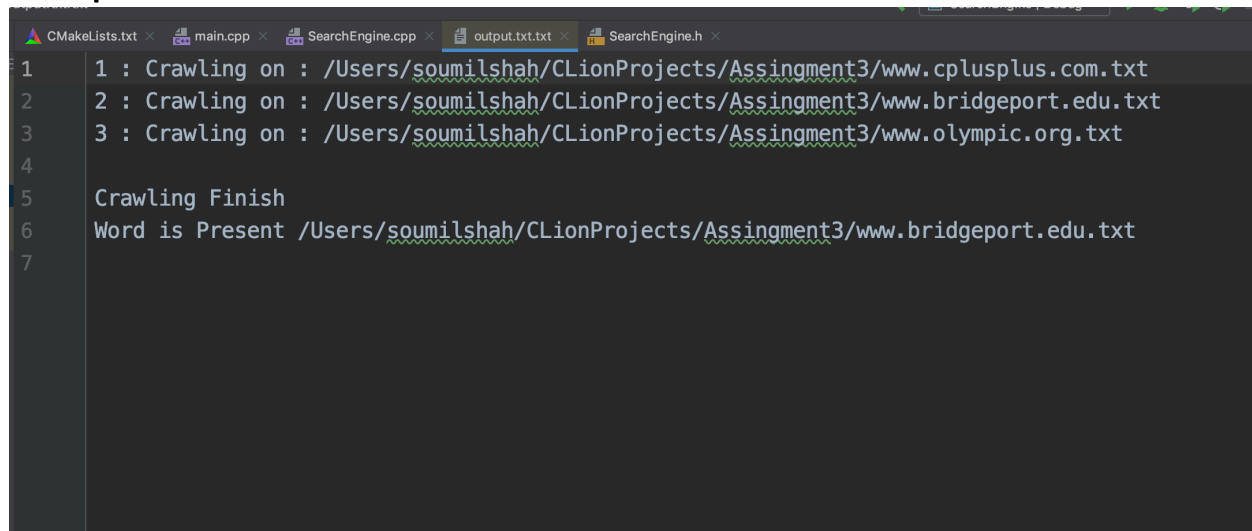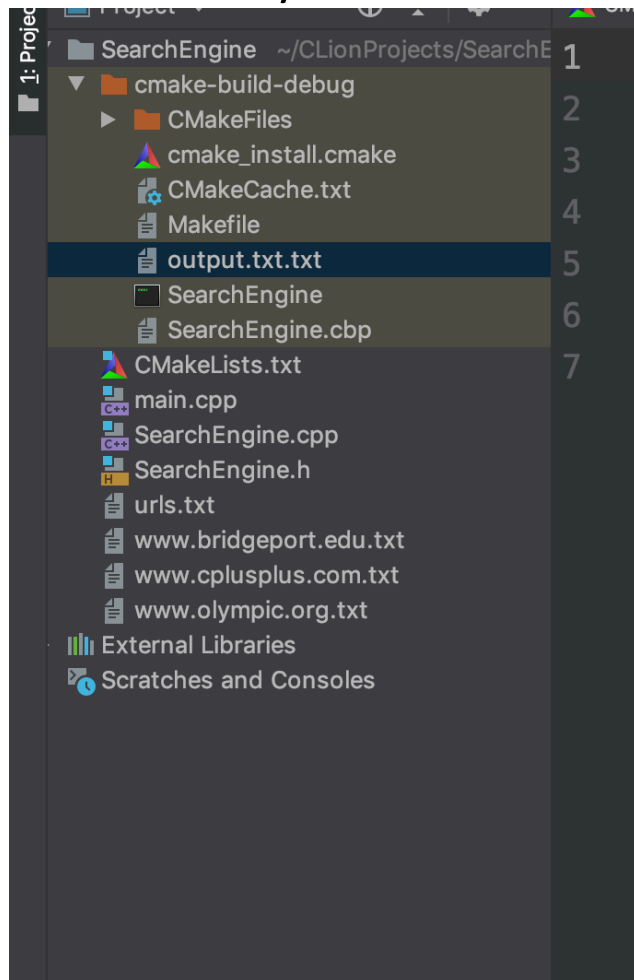
Output File

CMakeLists.txt	main.cpp	SearchEngine.cpp	output.txt.txt	SearchEngine.h

```
1     1 : Crawling on : /Users/soumilshah/CLionProjects/Assingment3/www.cplusplus.com.txt
2     2 : Crawling on : /Users/soumilshah/CLionProjects/Assingment3/www.bridgeport.edu.txt
3     3 : Crawling on : /Users/soumilshah/CLionProjects/Assingment3/www.olympic.org.txt
4
5     Crawling Finish
6     Word is Present /Users/soumilshah/CLionProjects/Assingment3/www.bridgeport.edu.txt
7
```

File Directory

SearchEngine ~/CLionProjects/SearchE
▼ cmake-build-debug
  ▶ CMakeFiles
    cmake_install.cmake
    CMakeCache.txt
    Makefile
    output.txt.txt
    SearchEngine
    SearchEngine.cbp
  CMakeLists.txt
  main.cpp
  SearchEngine.cpp
  SearchEngine.h
  urls.txt
  www.bridgeport.edu.txt
  www.cplusplus.com.txt
  www.olympic.org.txt
External Libraries
Scratches and Consoles

```cpp
/*
 * main.CPP File
 * SearchEngine
 * Author : Soumil Nitin SHah
 * Version 0.0.1
 * Email soushah@my.bridgeport.edu
 * Student id : 1031686
 * Github:
 * Last Modified : 6 october 2019
 *
 * Operating System: MAC os
 * Configured with: --
prefix=/Applications/Xcode.app/Contents/Developer/us
r --with-gxx-include-
dir=/Applications/Xcode.app/Contents/Developer/Platf
orms/MacOSX.platform/Developer/SDKs/MacOSX10.14.sdk/
usr/include/c++/4.2.
 * Apple LLVM version 10.0.1 (clang-1001.0.46.3)
 * Target: x86_64-apple-darwin18.7.0
 * InstalledDir:
/Applications/Xcode.app/Contents/Developer/Toolchain
s/XcodeDefault.xctoolchain/usr/bin
 */


#include <iostream>
#include <map>
#include <set>
#include <iterator>
#include <string>
#include <iomanip>
#include <sstream>
#include <fstream>
#include <vector>
#include "SearchEngine.h"

using namespace std;
int main()
```

```cpp
{
    SearchEngine G00gle;
    string start =
"/Users/soumilshah/CLionProjects/Assingment3/urls.tx
t";
    G00gle.startCrawler(start);

    string sterm = "";

    while (sterm!= "q" || sterm !="Q")
    {

        cout << "Enter the Search Word" << endl;
        cin >> sterm;
        if (sterm == "Q")
        {
            break;
        }else if(sterm == "q")
        {
            break;
        }
        G00gle.search(sterm);
    }

    return 0;
}
```

```cpp
/*
 * SearchEngine.CPP File
 * SearchEngine
 * Author : Soumil Nitin SHah
 * Version 0.0.1
 * Email soushah@my.bridgeport.edu
 * Student id : 1031686
 * Github:
 * Last Modified : 6 october 2019
 *
 * Operating System: MAC os
 * Configured with: --
prefix=/Applications/Xcode.app/Contents/Develop
er/usr --with-gxx-include-
dir=/Applications/Xcode.app/Contents/Developer/
Platforms/MacOSX.platform/Developer/SDKs/MacOSX
10.14.sdk/usr/include/c++/4.2.
 * Apple LLVM version 10.0.1 (clang-
1001.0.46.3)
 * Target: x86_64-apple-darwin18.7.0
 * InstalledDir:
/Applications/Xcode.app/Contents/Developer/Tool
chains/XcodeDefault.xctoolchain/usr/bin
 */

#include "SearchEngine.h"
#include <iostream>
#include <map>
#include <set>
#include <iterator>
#include <string>
#include <iomanip>
#include <sstream>
```

```cpp
#include <fstream>
#include <vector>

using namespace std;



ofstream fout("output.txt.txt");

SearchEngine::SearchEngine()
{

}

SearchEngine::~SearchEngine()
{

}

string SearchEngine::clean_word(std::string
word)
{
    string result {};
    for(char c : word)
    {
        if (isalpha(c))
        {
            result += c;
        }
        else
        {
            continue;
        }
```

```cpp
    }
    return result;
}

void SearchEngine::startCrawler(std::string
path)
{
    std::ifstream in_file (path);
    while (std::getline(in_file, line))
    {
        std::stringstream ss(line);      // Each
URL
        FilePath.push_back(Tem+line);    // Get
all Path
        //cout << Tem+line << endl;
    }
    auto it = FilePath.begin();
    int Row = 0;


    while (it != FilePath.end())
    {
        Row = Row + 1;
        cout << Row<<" : Crawling on : " << *it
<< endl;cout << "\n";
        fout << Row<<" : Crawling on : " << *it
<< endl;cout << "\n";

        ifstream in_file (*it);
// Create a file Object
        while (std::getline(in_file, line))
// Iterate Over First File
        {
            // cout << line << endl;
```

```cpp
// Each Sentences
            std::stringstream ss(line);
            while (ss >> word)
// Each Word
            {
                cleaned_word =
clean_word(word);    // clean Word
                P[Row][cleaned_word]++;
            }
        }cout <<"\n";
        it++;
    }cout<<"\n";fout<<"\n";
    cout << "Crawling Finish" <<
endl;cout<<"\n";
    fout << "Crawling Finish" <<
endl;cout<<"\n";
}

void SearchEngine::search(std::string
user_input)
{
    for (int i =1; i<=FilePath.size(); i++)
    {


        if(P[i][user_input])
        {
            // cout << "Found" << i << endl;
            cout << "Word is Present " <<
FilePath.at(i-1) << endl;
            fout << "Word is Present " <<
FilePath.at(i-1) << endl;
        }
```

```
        }
}
```

```
/*
 * SearchEngine.h File
 * SearchEngine
 * Author : Soumil Nitin SHah
 * Version 0.0.1
 * Email soushah@my.bridgeport.edu
 * Student id : 1031686
 * Github:
 * Last Modified : 6 october 2019
 *
 * Operating System: MAC os
 * Configured with: --
prefix=/Applications/Xcode.app/Conte
nts/Developer/usr --with-gxx-
include-
dir=/Applications/Xcode.app/Contents
/Developer/Platforms/MacOSX.platform
/Developer/SDKs/MacOSX10.14.sdk/usr/
include/c++/4.2.
 * Apple LLVM version 10.0.1 (clang-
1001.0.46.3)
 * Target: x86_64-apple-darwin18.7.0
 * InstalledDir:
/Applications/Xcode.app/Contents/Dev
eloper/Toolchains/XcodeDefault.xctoo
lchain/usr/bin
```

```cpp
 */



#include <iostream>
#include <map>
#include <set>
#include <iterator>
#include <string>
#include <iomanip>
#include <sstream>
#include <fstream>
#include <vector>

using namespace std;

#ifndef SEARCHENGINE_SEARCHENGINE_H
#define SEARCHENGINE_SEARCHENGINE_H


class SearchEngine {


public:
    /*
     * Declaring variables and
```

```cpp
Attributes
    */
    std::map<std::string, int>
words;
    std::string line;
    std::string word;
    std::string cleaned_word ;

    std::string user_input;
    vector <string> FilePath;
    string Tem =
"/Users/soumilshah/CLionProjects/Ass
ingment3/";
    map <int, map<string, int>> P;

public:
    /*
     * Constructor and Destructor
     */
    SearchEngine();

    ~SearchEngine();

public:
    void startCrawler(std::string
path);
```

```cpp
    string clean_word(std::string
word);
    void search(std::string
user_input);
};


#endif //SEARCHENGINE_SEARCHENGINE_H
```