

Adding Spell Checking, AutoComplete and Snippets to Your Search Engine

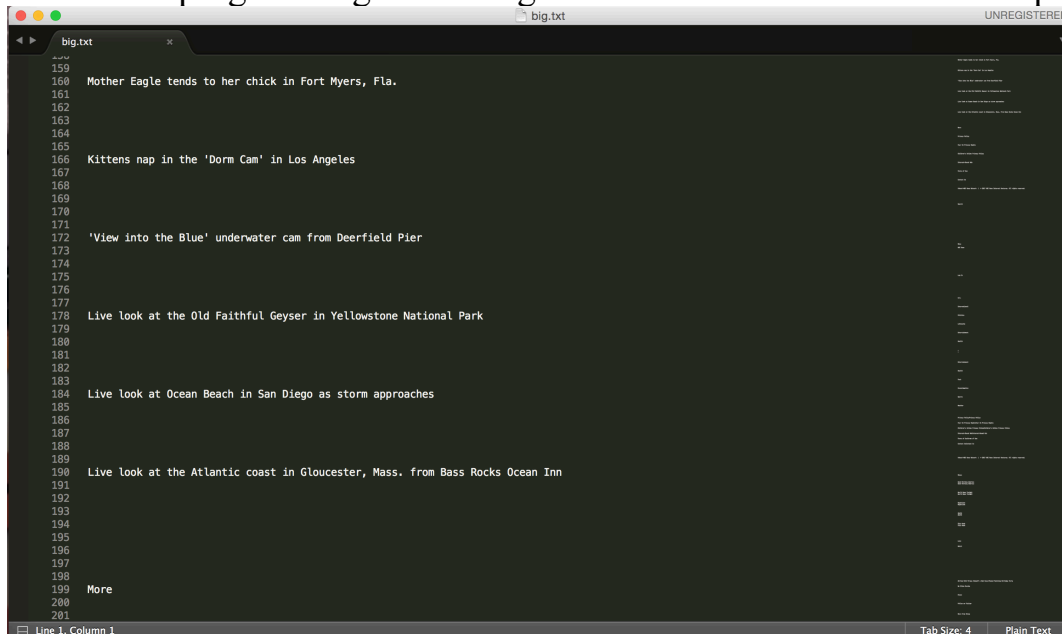
1.Steps I followed to complete this assignment

a First change the solrconfig.xml and add the suggest to Solr.

```
1018 <searchComponent class="solr.SuggestComponent" name="suggest">
1019   <lst name="suggester">
1020     <str name="name">suggest</str>
1021     <str name="lookupimpl">FuzzyLookupFactory</str>
1022     <str name="field">_text_</str>
1023     <str name="suggestAnalyzerFieldType">string</str>
1024     <str name="exactMatchFirst">true</str>
1025     <str name="preserveSep">true</str>
1026     <str name="preservePositionIncrements">true</str>
1027     <str name="nonFuzzyPrefix">10</str>
1028   </lst>
1029 </searchComponent>
1030
1031 <requestHandler class="solr.SearchHandler" name="/suggest">
1032   <lst name="defaults">
1033     <str name="suggest">true</str>
1034     <str name="suggest.count">5</str>
1035     <str name="suggest.dictionary">suggest</str>
1036   </lst>
1037   <arr name="components">
1038     <str>suggest</str>
1039   </arr>
1040 </requestHandler>
1041
1042
```

b Then reload the core

c Use Tika program to generate big.txt from content of all html files provided.



d write a php including JQuery ui library, and the library of implementing snippet to finish the autocomplete feature and spell corrector and snippet feature. I also

download the SpellCorrector.php from phpclass.org to transform the big.txt to series id.txt.

```
hw5.php
49 </head>
50 <body>
51 <form accept-charset="utf-8" method="get">
52   <label for="q">Search:</label>
53   <input id="q" name="q" type="text" value="<?php echo htmlspecialchars($query,
54     ENT_QUOTES, 'utf-8'); ?>" />
55   <input type="submit" />
56   <input type="checkbox" name="pagerank">Use Page Rank Algorhthm<br>
57 </form>
58 <?php if ($results): ?>
59   <?php
60     $total = (int)$results->response->numFound;
61     $start = min(1, $total);
62     $end = min($limit, $total);
63   ?>
64   <?php if ($corrector != ""): ?>
65     <p>Do you want to search: <a href="http://localhost/572hw4/hw5.php?q=<?php
66       echo $corrector; ?>"><?php echo $corrector; ?></a></p>
67     <?php
68       $results = $solr->search($corrector, 0, $limit, $param);
69       $total = (int)$results->response->numFound;
70       $start = min(1, $total);
71       $end = min($limit, $total);
72     ?>
73   <?php endif; ?>
74   <?php
75     $load = false;
76     if (!$load)
77     {
78       if (($handle = fopen("mapABCNewsDataFile.csv", "r")) !== FALSE) {
79         while (($data = fgetcsv($handle)) !== FALSE) {
80           $array[$data[0]] = $data[1];
81         }
82         fclose($handle);
83       }
84       $load = true;
85     }
86   ?>
```

2. Misspelled term:

a “snpchat”, it shows “snapchat”

localhost/572hw4/hw5.php?q=snpchat

Search:

☐ Use Page Rank Algorhthm

Do you want to search: [snpchat](#)

Results 1 - 10 of 95:

[View the Detail Websit](#) Snapchat Videos at ABC News Video Ar

Author: None | Size: 48462

Home> Topics> News> Technology> Snapchat Snapchat Photos

[View the Detail Websit](#) Snapchat Videos at ABC News Video Ar

Author: None | Size: 50405

Home> Topics> News> Technology> Snapchat Snapchat News

b “russal”, it shows “russel”

localhost/572hw4/hw5.php?q=russal

Search:

☐ Use Page Rank Algorhthm

Do you want to search: [russel](#)

Results 1 - 10 of 17:

[View the Detail Websit](#) John Russel Houser Videos at ABC

Author: None | Size: 50630

[View the Detail Websit](#) 6 die in gas leak at South African na

Author: ABC News | Size: 23292

... people Friday, including three members of the military, au

for paramedic companv ER24. In a statement. ER24 describ

c “nasad”, it shows “nasa”

Search:
☐ Use Page Rank Algorithm

Do you want to search: [nasa](#)

Results 1 - 10 of 4937:

[View the Detail Websit](#) NASA

Author: None | Size: 47764

Home> Topics> News> NASA NASA News Tw

[View the Detail Websit](#) Incredible Images Captu

Author: ABC News | Size: 227987

...Live Live Digital coverage of today's events D
GPS 10 Mission Mother Earth tends to be a chie

e “nasdaq”, it shows “nasdaq”

Search:
☐ Use Page Rank Algorithm

Do you want to search: [nasdaq](#)

Results 1 - 10 of 67:

[View the Detail Websit](#) NASDAQ Composite Index Videos at A

Author: None | Size: 43866

Home> Topics> Business> Indexes> NASDAQ Composite Ind

[View the Detail Websit](#) NASDAQ 100 Index Videos at ABC Ne

Author: None | Size: 50615

Home> Topics> Business> Indexes> NASDAQ 100 Index NAS

[View the Detail Websit](#) NASDAO Videos at ABC News Video .

d “brexits”, it shows “brexit”

Search:
☐ Use Page Rank Algorithm

Do you want to search: [brexit](#)

Results 1 - 10 of 132:

[View the Detail Websit](#) Brexit European Union Videos

Author: None | Size: 46271

Home> Topics> News> World> Brexit Brexit Photos I

[View the Detail Websit](#) Brexit European Union Videos

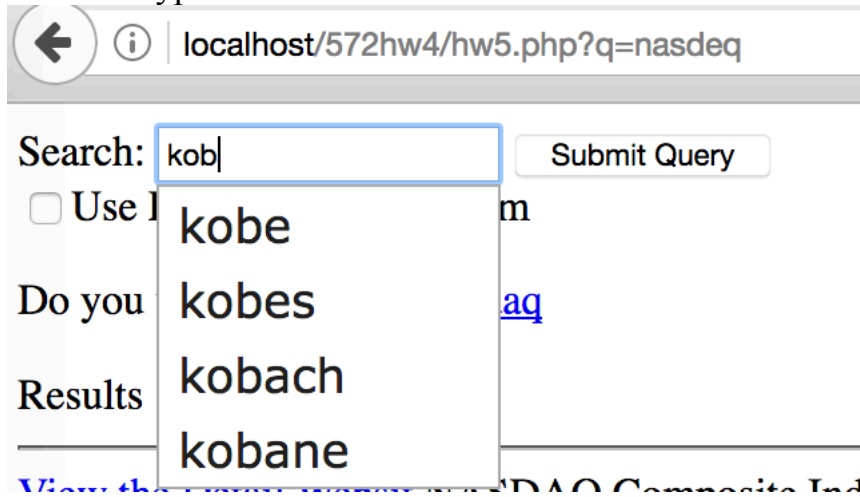
Author: None | Size: 51693

Home> Topics> News> World> Brexit Brexit Videos I

[View the Detail Websit](#) Tony Blair's new mission is to

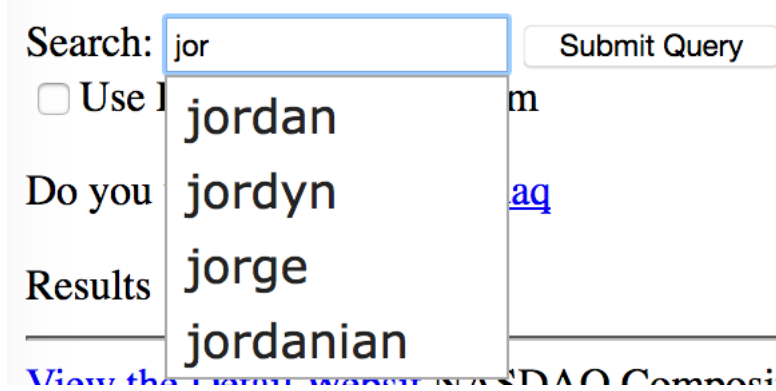
3. Auto-completion term:

a when I type “kob”



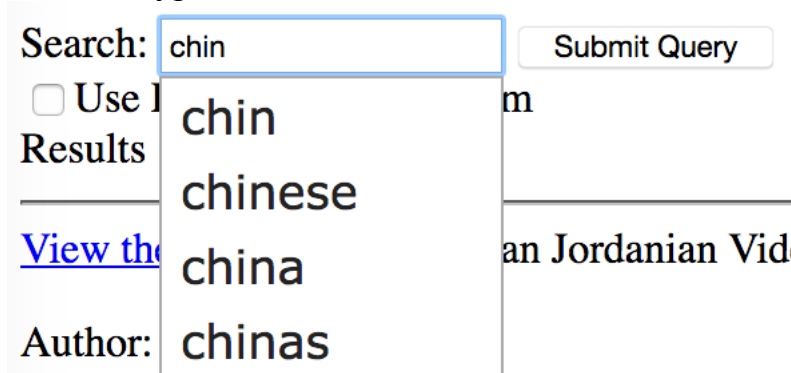
The screenshot shows a web browser address bar with the URL `localhost/572hw4/hw5.php?q=nasdeq`. Below the address bar is a search form with the label "Search:" and a text input field containing "kob". To the right of the input field is a button labeled "Submit Query". Below the input field, a dropdown menu is open, displaying four suggestions: "kobe", "kobes", "kobach", and "kobane". To the left of the input field, there is a checkbox labeled "Use I" and some partially visible text "Do you" and "Results".

b when I type “jor”



The screenshot shows the same web browser interface as above, but the search input field now contains "jor". The dropdown menu displays four suggestions: "jordan", "jordyn", "jorge", and "jordanian". The "Submit Query" button and the "Use I" checkbox are still visible.

c when I type “chin”



The screenshot shows the same web browser interface, but the search input field now contains "chin". The dropdown menu displays four suggestions: "chin", "chinese", "china", and "chinas". The "Submit Query" button and the "Use I" checkbox are still visible.

d when I type “c”

Search: c

Submit Query

☐ Use I

Results

[View the](#)

Author:

charset

content

crawldata

canonical

m

an Jordanian Vide

e when I type “lea”

Search: lea

Submit Query

☐ Use I

Results

[View the](#)

Author:

lea

leaps

lead

leader

m

an Jordanian Vide