

Homework 5 Report

Name: Jiayuan Shi

USC ID: 2659142134

1.Steps to complete assignment

Step 1: Use Norvig' s program to implement spelling correction

I downloaded Norvig' s spelling program writing in PHP and run on my server. Then I created my own "big.txt" using Apache Tika. Also I enhanced my spelling correction program with set of terms I am responsible for.

Step 2: Implement the autocomplete function

In order to do this, I used the FuzzyLookupFactory feature of Solr/Lucene. I used jQuery and AJAX to create asynchronous request. Besides, I wrote another PHP script called "solr.php" as a proxy to solve the cross domain issue.

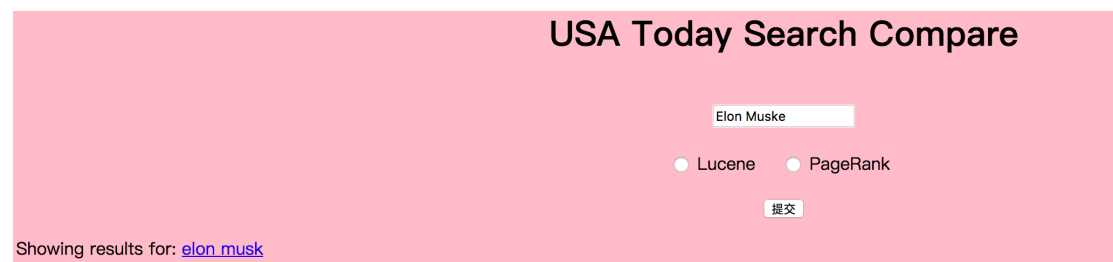
Step 3: Produce snippets of each result

I produced snippets to each search result by regular expression matching of the query term.

2. Analysis of the results

(1) misspelled terms – spelling correction:

Elon Muske → elon musk



Northi → north

USA Today Search Compare

Northi

Lucene

PageRank

提交

Showing results for: [north](#)

Hurricaane → hurricane

USA Today Search Compare

Hurricaane

Lucene

PageRank

提交

Showing results for: [hurricane](#)

iPhonee → iphone

USA Today Search Compare

iPhonee

Lucene

PageRank

提交

Showing results for: [iphone](#)

Climmate → climate

USA Today Search Compare

Climmate

Lucene

PageRank

提交

Showing results for: [climate](#)

(2) autocomplete

USA Today Search Compare

Clima

Lucene

PageRank

climate

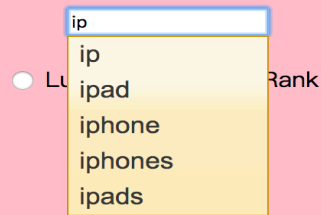
climbing

climb

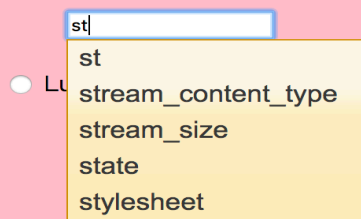
climbs

climbed

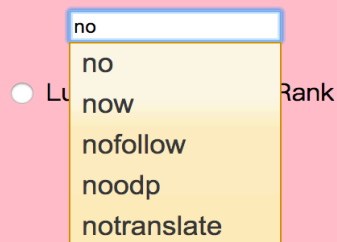
USA Today Search Compare



USA Today Search Compare



USA Today Search Compare



USA Today Search Compare

