

Indexing the Web Using Solr

By Akshay Bhatia; email: akshaybh@usc.edu; usc ID: 4983495776

Site Used

Reuters News

Steps followed to complete this assignment

I completed the assignment on Ubuntu

as follows:

1. Installed Solr on Ubuntu
2. Created a new core, "myexample" in Solr for the assignment
3. Used SimplePostTool to create Solr's index using the HTML files provided
4. Wrote a Java program on Eclipse to create an edge list, which works as follows:
 - a. Create two HashMaps, `fileUrlMap` and `urlFileMap`, which map the filenames to the urls inside them and vice versa using the provided CSV file
 - b. For each file in the crawl data folder, use JSoup to parse all outgoing links in the HTML file.
 - c. Find the corresponding filename for the link from the original page using the map
 - d. Add lines to a text file, "edgelist.txt" in the format "<original page> <page linked to>"
5. Wrote a Python program to calculate PageRank using the NetworkX library, which works as follows:
 - a. Creates a DiGraph using the edge list generated as output of Eclipse java program

```
G = nx.read_edgelist("edgeList.txt", create_using=nx.DiGraph())
```

- b. Uses NetworkX's `pagerank` function to compute PageRank for the above graph, using the following parameters for its initial configuration
- ```
pr = nx.pagerank(G, alpha=0.85, personalization=None, max_iter=30, tol=1e-06, nstart=None, weight='weight', dangling=None)
```
- c. Creates an output file, "external\_pageRankFile.txt"
  - d. Writes the PageRank to a text file, "external\_pageRankFile.txt" in the <docId>=<pageRankValue> format as instructed. The directory path ("/home/akshay/shared/reutersnews/") is appended to the beginning of the page filename, which becomes the key part, and the value of pageRank is the value part. These key=value pairs are written to the text file.
6. Configured Solr to allow sorting on the basis of PageRank by editing the managed-schema and solrconfig.xml files as specified.
  7. Created a PHP program as follows:
    - a. Used solr-php-client to allow accessing Solr and querying it.
    - b. Create an associative array mapping file names to URLs, to fetch URLs from in case they weren't present in Solr output.

- c. Wrote PHP code to specify the sorting algorithms, giving a choice between Lucene (score desc) and PageRank (pageRankFile desc), and perform a search on Solr based on the query parameters input in the text box provided.
  - d. Displayed output by displaying each URL's title, url, id and description (if one exists), where each result's title and url links you to its page
  - e. Used Bootstrap 4 to style the page to look similar to how Google displays search results.
8. Used Apache2 & php7 to host a web server on Ubuntu for the created PHP file.

## Why do some pages have higher PageRank than others?

The PageRank (PR) of each page is a measure of how 'important' a page is. It depends on PR of other pages which are pointing to it. It is calculated as follows:

$$PR(A) = (1-d) + d (PR(T1)/C(T1) + ... + PR(Tn)/C(Tn))$$

Where d is a damping factor that lies between 0 and 1, and C(Ti) is the number of outgoing links from page Ti.

As one may notice from the calculation formula, a page with a large number of incoming links will have a high pageRank, while one with a few will have a smaller pageRank. Thus, some pages, like the Reuters News home page will have a very high page-rank since all pages on the Reuters News web site point to it.

## Flow Screenshots

Screenshot of initial page where you enter the query in input box



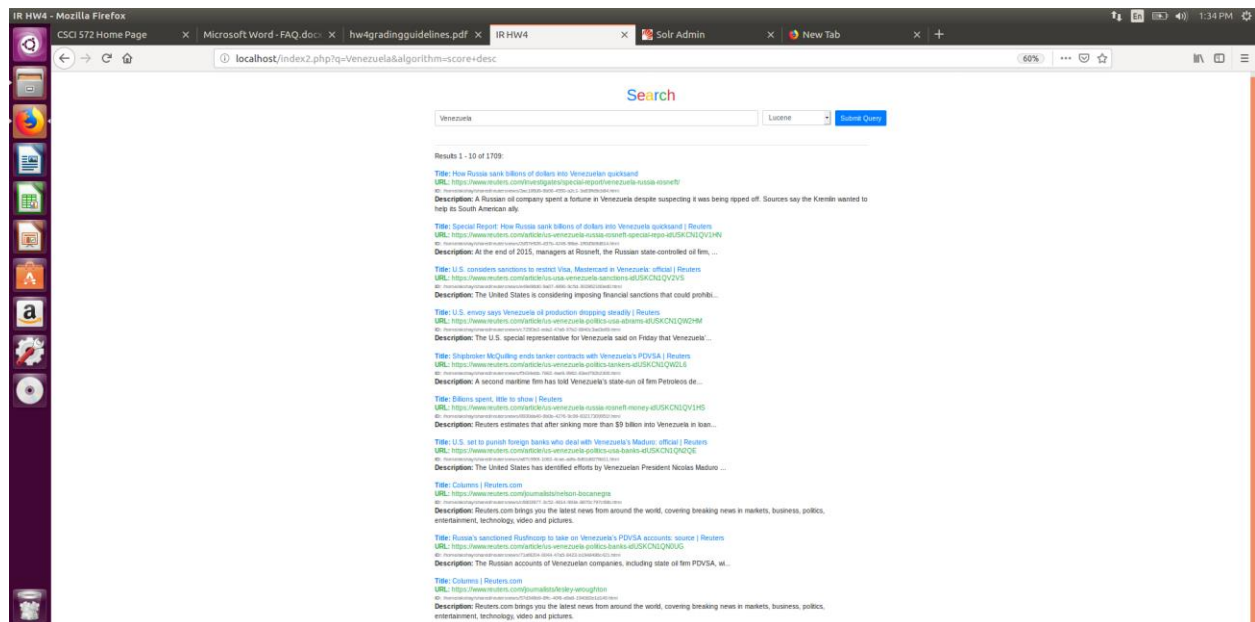
Search

Enter search term here

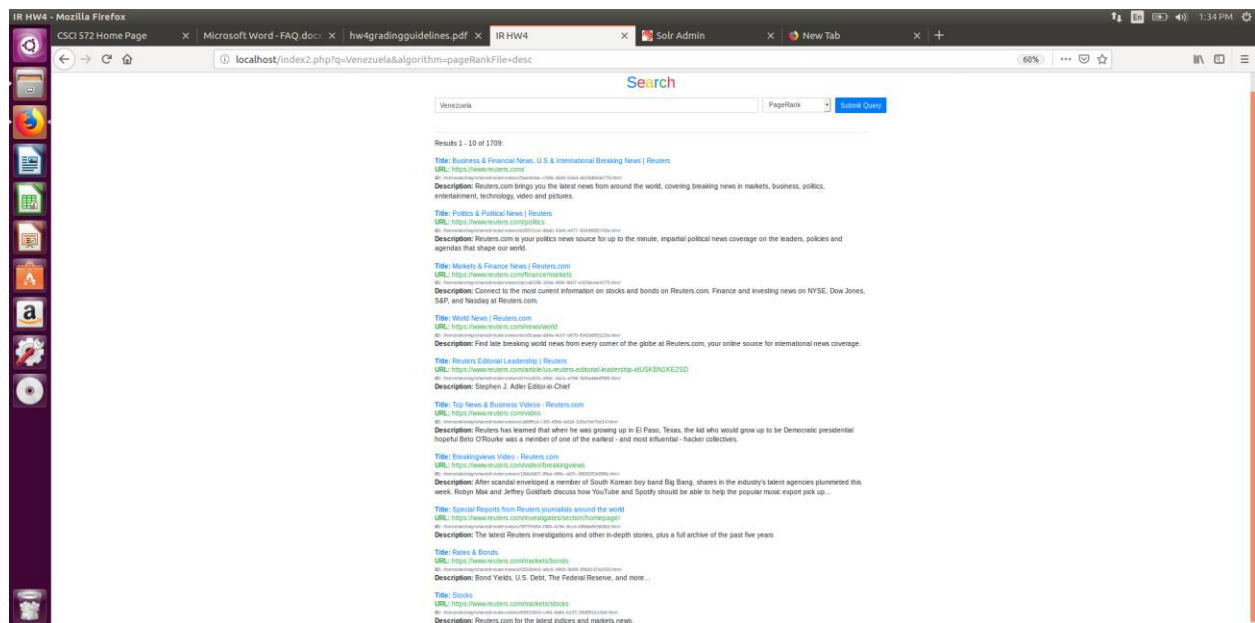
Lucene

Submit

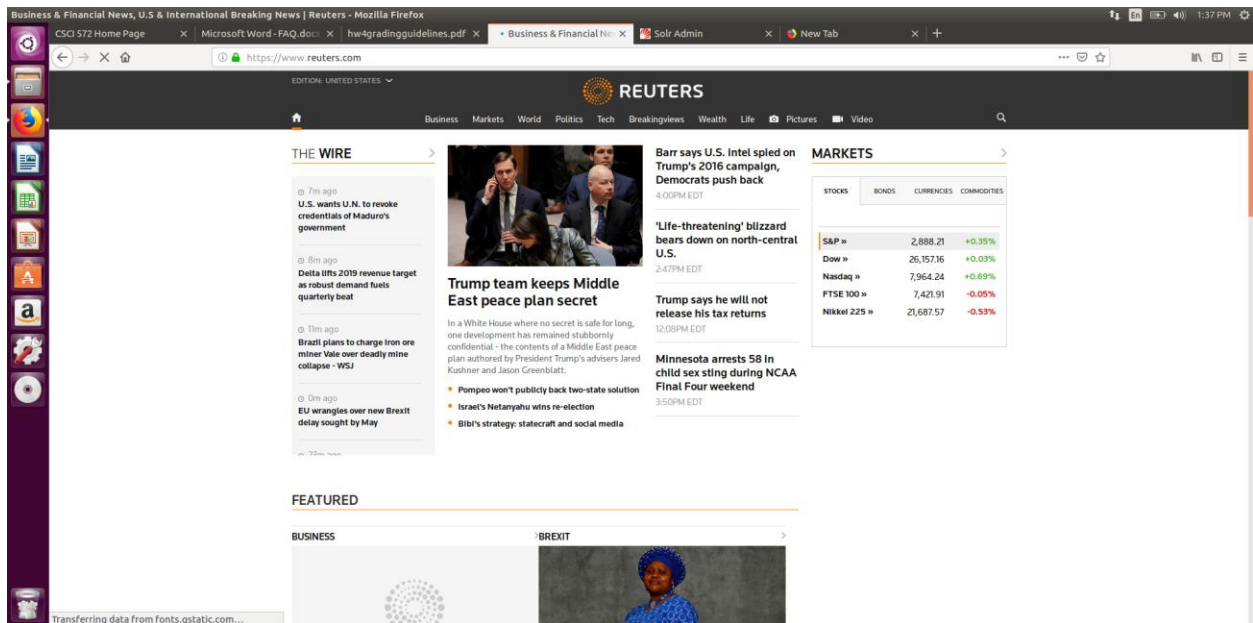
## Screenshot of page which shows the results for Lucene (Default)



## Screenshot of page which shows the results for PageRank



Screenshot of the actual webpage that opens when clicking on one of the result's link



## Table of URLs for each query

Overlapping results highlighted.

| Query     | Lucene                                                                                                                                                                                      | PageRank                                                                                                                                                                  |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Venezuela | <a href="https://www.reuters.com/investigates/special-report/venezuela-russia-rosneft/">https://www.reuters.com/investigates/special-report/venezuela-russia-rosneft/</a>                   | <a href="https://www.reuters.com/">https://www.reuters.com/</a>                                                                                                           |
|           | <a href="https://www.reuters.com/article/us-venezuela-russia-rosneft-special-repo-idUSKCN1QV1HN">https://www.reuters.com/article/us-venezuela-russia-rosneft-special-repo-idUSKCN1QV1HN</a> | <a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a>                                                                                           |
|           | <a href="https://www.reuters.com/article/us-usa-venezuela-sanctions-idUSKCN1QV2VS">https://www.reuters.com/article/us-usa-venezuela-sanctions-idUSKCN1QV2VS</a>                             | <a href="https://www.reuters.com/finance/markets">https://www.reuters.com/finance/markets</a>                                                                             |
|           | <a href="https://www.reuters.com/article/us-venezuela-politics-usa-abrams-idUSKCN1QW2HM">https://www.reuters.com/article/us-venezuela-politics-usa-abrams-idUSKCN1QW2HM</a>                 | <a href="https://www.reuters.com/news/world">https://www.reuters.com/news/world</a>                                                                                       |
|           | <a href="https://www.reuters.com/article/us-venezuela-politics-tankers-idUSKCN1QW2L6">https://www.reuters.com/article/us-venezuela-politics-tankers-idUSKCN1QW2L6</a>                       | <a href="https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD">https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD</a> |
|           | <a href="https://www.reuters.com/article/us-venezuela-russia-rosneft-money-idUSKCN1QV1HS">https://www.reuters.com/article/us-venezuela-russia-rosneft-money-idUSKCN1QV1HS</a>               | <a href="https://www.reuters.com/video">https://www.reuters.com/video</a>                                                                                                 |
|           | <a href="https://www.reuters.com/article/us-venezuela-politics-usa-banks-idUSKCN1QN2QE">https://www.reuters.com/article/us-venezuela-politics-usa-banks-idUSKCN1QN2QE</a>                   | <a href="https://www.reuters.com/video/breakingviews">https://www.reuters.com/video/breakingviews</a>                                                                     |
|           | <a href="https://www.reuters.com/journalists/nelson-bocanegra">https://www.reuters.com/journalists/nelson-bocanegra</a>                                                                     | <a href="https://www.reuters.com/investigates/section/homepage/">https://www.reuters.com/investigates/section/homepage/</a>                                               |
|           | <a href="https://www.reuters.com/article/us-venezuela-politics-banks-idUSKCN1QN0UG">https://www.reuters.com/article/us-venezuela-politics-banks-idUSKCN1QN0UG</a>                           | <a href="https://www.reuters.com/markets/bonds">https://www.reuters.com/markets/bonds</a>                                                                                 |
|           | <a href="https://www.reuters.com/journalists/lesley-wroughton">https://www.reuters.com/journalists/lesley-wroughton</a>                                                                     | <a href="https://www.reuters.com/markets/stocks">https://www.reuters.com/markets/stocks</a>                                                                               |
| Senate    | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8</a>                   | <a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a>                                                                                           |

| Query      | Lucene                                                                                                                                                                                                                                                                | PageRank                                                                                                                                                                                                                                                        |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|            | <a href="https://www.reuters.com/video/2019/03/04/senate-has-votes-to-block-border-emergen?videoid=521958334&amp;videoChannel=-10465">https://www.reuters.com/video/2019/03/04/senate-has-votes-to-block-border-emergen?videoid=521958334&amp;videoChannel=-10465</a> | <a href="https://www.reuters.com/video">https://www.reuters.com/video</a>                                                                                                                                                                                       |
|            | <a href="https://www.reuters.com/video/2019/03/05/mcconnell-senate-has-votes-to-block-bord?videoid=521939215&amp;videoChannel=-10465">https://www.reuters.com/video/2019/03/05/mcconnell-senate-has-votes-to-block-bord?videoid=521939215&amp;videoChannel=-10465</a> | <a href="https://www.reuters.com/video/breakingviews">https://www.reuters.com/video/breakingviews</a>                                                                                                                                                           |
|            | <a href="https://www.reuters.com/article/us-usa-politics-beto-orourke-special-rep-idUSKCN1QW26N">https://www.reuters.com/article/us-usa-politics-beto-orourke-special-rep-idUSKCN1QW26N</a>                                                                           | <a href="https://www.reuters.com/news/pictures">https://www.reuters.com/news/pictures</a>                                                                                                                                                                       |
|            | <a href="https://www.reuters.com/article/us-usa-politics-beto-orourke-special-rep-idUSKCN1QW26N">https://www.reuters.com/article/us-usa-politics-beto-orourke-special-rep-idUSKCN1QW26N</a>                                                                           | <a href="https://www.reuters.com/news/media">https://www.reuters.com/news/media</a>                                                                                                                                                                             |
|            | <a href="https://www.reuters.com/article/us-congo-politics-idUSKCN1QW28R">https://www.reuters.com/article/us-congo-politics-idUSKCN1QW28R</a>                                                                                                                         | <a href="https://www.reuters.com/subjects/sustainable-business">https://www.reuters.com/subjects/sustainable-business</a>                                                                                                                                       |
|            | <a href="https://www.reuters.com/article/us-usa-privacy-idUSKBN1QS23L">https://www.reuters.com/article/us-usa-privacy-idUSKBN1QS23L</a>                                                                                                                               | <a href="https://www.reuters.com/places/mexico">https://www.reuters.com/places/mexico</a>                                                                                                                                                                       |
|            | <a href="https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H">https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H</a>                                                                                   | <a href="https://www.reuters.com/video/2019/03/14/orourke-hits-the-ground-running-in-iowa?videoid=525923095&amp;videoChannel=1003">https://www.reuters.com/video/2019/03/14/orourke-hits-the-ground-running-in-iowa?videoid=525923095&amp;videoChannel=1003</a> |
|            | <a href="https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H">https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H</a>                                                                                   | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK</a>                                                                                       |
|            | <a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a>                                                                                                                                                                                       | <a href="https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H">https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H</a>                                                                             |
| Democrats  | <a href="https://www.reuters.com/article/us-usa-election-medicare-idUSKBN1QU189">https://www.reuters.com/article/us-usa-election-medicare-idUSKBN1QU189</a>                                                                                                           | <a href="https://www.reuters.com/">https://www.reuters.com/</a>                                                                                                                                                                                                 |
|            | <a href="https://www.reuters.com/article/us-usa-election-medicare-idUSKBN1QU189">https://www.reuters.com/article/us-usa-election-medicare-idUSKBN1QU189</a>                                                                                                           | <a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a>                                                                                                                                                                                 |
|            | <a href="https://www.reuters.com/article/us-usa-congress-whitaker-idUSKCN1PX1LD">https://www.reuters.com/article/us-usa-congress-whitaker-idUSKCN1PX1LD</a>                                                                                                           | <a href="https://www.reuters.com/breakingviews">https://www.reuters.com/breakingviews</a>                                                                                                                                                                       |
|            | <a href="https://www.reuters.com/video/2019/02/28/republicans-bash-cohen-democrats-at-cpac?videoid=520293110&amp;videoChannel=13976">https://www.reuters.com/video/2019/02/28/republicans-bash-cohen-democrats-at-cpac?videoid=520293110&amp;videoChannel=13976</a>   | <a href="https://www.reuters.com/subjects/supreme-court">https://www.reuters.com/subjects/supreme-court</a>                                                                                                                                                     |
|            | <a href="https://www.reuters.com/article/us-usa-election-biden-idUSKCN1QH1EX">https://www.reuters.com/article/us-usa-election-biden-idUSKCN1QH1EX</a>                                                                                                                 | <a href="https://www.reuters.com/news/archive/politicsNews">https://www.reuters.com/news/archive/politicsNews</a>                                                                                                                                               |
|            | <a href="https://www.reuters.com/article/us-usa-trump-budget-idUSKBN1QS12D">https://www.reuters.com/article/us-usa-trump-budget-idUSKBN1QS12D</a>                                                                                                                     | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK</a>                                                                                       |
|            | <a href="https://www.reuters.com/video/2019/03/05/the-witch-hunt-continues-trump-on-democr?videoid=522293719&amp;videoChannel=13976">https://www.reuters.com/video/2019/03/05/the-witch-hunt-continues-trump-on-democr?videoid=522293719&amp;videoChannel=13976</a>   | <a href="https://www.reuters.com/article/us-usa-election-sanders-idUSKCN1QW2ZJ">https://www.reuters.com/article/us-usa-election-sanders-idUSKCN1QW2ZJ</a>                                                                                                       |
|            | <a href="https://www.reuters.com/article/us-usa-trump-russia-congress-idUSKCN1QV23E">https://www.reuters.com/article/us-usa-trump-russia-congress-idUSKCN1QV23E</a>                                                                                                   | <a href="https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H">https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H</a>                                                                             |
|            | <a href="https://www.reuters.com/article/us-usa-election-booker-idUSKCN1QW2P4">https://www.reuters.com/article/us-usa-election-booker-idUSKCN1QW2P4</a>                                                                                                               | <a href="https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY">https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY</a>                                                                                                                   |
|            | <a href="https://www.reuters.com/article/us-usa-election-booker-idUSKCN1QW2P4">https://www.reuters.com/article/us-usa-election-booker-idUSKCN1QW2P4</a>                                                                                                               | <a href="https://www.reuters.com/article/us-usa-election-booker-idUSKCN1QW2P4">https://www.reuters.com/article/us-usa-election-booker-idUSKCN1QW2P4</a>                                                                                                         |
| Republican | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8</a>                                                                                             | <a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a>                                                                                                                                                                                 |
|            | <a href="https://www.reuters.com/video/2019/02/28/republicans-bash-cohen-democrats-at-cpac?videoid=520293110&amp;videoChannel=13976">https://www.reuters.com/video/2019/02/28/republicans-bash-cohen-democrats-at-cpac?videoid=520293110&amp;videoChannel=13976</a>   | <a href="https://www.reuters.com/news/us">https://www.reuters.com/news/us</a>                                                                                                                                                                                   |
|            | <a href="https://www.reuters.com/article/us-california-fires-trump-idUSKCN1P31ND">https://www.reuters.com/article/us-california-fires-trump-idUSKCN1P31ND</a>                                                                                                         | <a href="https://www.reuters.com/news/archive/politicsNews">https://www.reuters.com/news/archive/politicsNews</a>                                                                                                                                               |
|            | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK</a>                                                                                             | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK</a>                                                                                       |
|            | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK</a>                                                                                             | <a href="https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY">https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY</a>                                                                                                                   |
|            | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK</a>                                                                                             | <a href="https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C">https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C</a>                                                                                                       |
|            | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK</a>                                                                                             | <a href="https://www.reuters.com/article/us-usa-immigration-wall-insight-idUSKCN1QW1GR">https://www.reuters.com/article/us-usa-immigration-wall-insight-idUSKCN1QW1GR</a>                                                                                       |
|            | <a href="https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY">https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY</a>                                                                                                                         | <a href="https://www.reuters.com/article/us-usa-trump-russia-congress-idUSKCN1QV23E">https://www.reuters.com/article/us-usa-trump-russia-congress-idUSKCN1QV23E</a>                                                                                             |
|            | <a href="https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY">https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY</a>                                                                                                                         | <a href="https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8">https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8</a>                                                                                       |
|            | <a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a>                                                                                                                                                                                       | <a href="https://www.reuters.com/article/us-usa-court-census-idUSKCN1QW2MC">https://www.reuters.com/article/us-usa-court-census-idUSKCN1QW2MC</a>                                                                                                               |

| Query            | Lucene                                                                                                                                                                                                                                                                    | PageRank                                                                                                                                                                                |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Patriot Movement | <a href="https://www.reuters.com/article/us-usa-trump-turkey-idUSKCN1QV328">https://www.reuters.com/article/us-usa-trump-turkey-idUSKCN1QV328</a>                                                                                                                         | <a href="https://www.reuters.com/markets/bonds">https://www.reuters.com/markets/bonds</a>                                                                                               |
|                  | <a href="https://www.reuters.com/article/us-france-protests-idUSKCN1Q325A">https://www.reuters.com/article/us-france-protests-idUSKCN1Q325A</a>                                                                                                                           | <a href="https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C">https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C</a>                               |
|                  | <a href="https://www.reuters.com/article/us-britain-climatechange-education-idUSKCN1QW2QU">https://www.reuters.com/article/us-britain-climatechange-education-idUSKCN1QW2QU</a>                                                                                           | <a href="https://www.reuters.com/article/canada-forex-idUSL1N2121OB">https://www.reuters.com/article/canada-forex-idUSL1N2121OB</a>                                                     |
|                  | <a href="https://www.reuters.com/article/uk-people-rkelly-metoo-idUKKCN1QC0V7">https://www.reuters.com/article/uk-people-rkelly-metoo-idUKKCN1QC0V7</a>                                                                                                                   | <a href="https://www.reuters.com/article/us-slovakia-election-president-idUSKCN1QW2Z3">https://www.reuters.com/article/us-slovakia-election-president-idUSKCN1QW2Z3</a>                 |
|                  | <a href="https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S">https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S</a>                                                                                                                 | <a href="https://www.reuters.com/article/us-climate-change-youth-usa-idUSKCN1QW1BE">https://www.reuters.com/article/us-climate-change-youth-usa-idUSKCN1QW1BE</a>                       |
|                  | <a href="https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S">https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S</a>                                                                                                                 | <a href="https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S">https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S</a>                               |
|                  | <a href="https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S">https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S</a>                                                                                                                 | <a href="https://www.reuters.com/article/us-climate-change-youth-usa-idUSKCN1QW1BE">https://www.reuters.com/article/us-climate-change-youth-usa-idUSKCN1QW1BE</a>                       |
|                  | <a href="https://www.reuters.com/article/us-italy-china-analysis-idUSKCN1QW1E2">https://www.reuters.com/article/us-italy-china-analysis-idUSKCN1QW1E2</a>                                                                                                                 | <a href="https://www.reuters.com/investigates/special-report/myanmar-suukyi-history/">https://www.reuters.com/investigates/special-report/myanmar-suukyi-history/</a>                   |
|                  | <a href="https://www.reuters.com/article/us-italy-china-analysis-idUSKCN1QW1E2">https://www.reuters.com/article/us-italy-china-analysis-idUSKCN1QW1E2</a>                                                                                                                 | <a href="https://www.reuters.com/article/us-auction-christies-georgemichael-art-idUSKCN1QP1EY">https://www.reuters.com/article/us-auction-christies-georgemichael-art-idUSKCN1QP1EY</a> |
|                  | <a href="https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C">https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C</a>                                                                                                                 | <a href="https://www.reuters.com/article/us-italy-china-analysis-idUSKCN1QW1E2">https://www.reuters.com/article/us-italy-china-analysis-idUSKCN1QW1E2</a>                               |
|                  |                                                                                                                                                                                                                                                                           |                                                                                                                                                                                         |
|                  |                                                                                                                                                                                                                                                                           |                                                                                                                                                                                         |
| Oscar 2019       | <a href="https://www.reuters.com/article/us-awards-oscar-diversity-idUSKCN1QE0N4">https://www.reuters.com/article/us-awards-oscar-diversity-idUSKCN1QE0N4</a>                                                                                                             | <a href="https://www.reuters.com/">https://www.reuters.com/</a>                                                                                                                         |
|                  | <a href="https://www.reuters.com/article/us-awards-oscar-women-idUSKCN1Q7130">https://www.reuters.com/article/us-awards-oscar-women-idUSKCN1Q7130</a>                                                                                                                     | <a href="https://www.reuters.com/tools/mobile">https://www.reuters.com/tools/mobile</a>                                                                                                 |
|                  | <a href="https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT">https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT</a>                                                                                                               | <a href="https://www.reuters.com/finance">https://www.reuters.com/finance</a>                                                                                                           |
|                  | <a href="https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT">https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT</a>                                                                                                               | <a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a>                                                                                                         |
|                  | <a href="https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT">https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT</a>                                                                                                               | <a href="https://www.reuters.com/finance/markets">https://www.reuters.com/finance/markets</a>                                                                                           |
|                  | <a href="https://www.reuters.com/journalists/pei-li">https://www.reuters.com/journalists/pei-li</a>                                                                                                                                                                       | <a href="https://www.reuters.com/news/world">https://www.reuters.com/news/world</a>                                                                                                     |
|                  | <a href="https://www.reuters.com/theWire">https://www.reuters.com/theWire</a>                                                                                                                                                                                             | <a href="https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD">https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD</a>               |
|                  | <a href="https://www.reuters.com/article/us-britain-lgbt-rights-idUSKCN1QW2QI">https://www.reuters.com/article/us-britain-lgbt-rights-idUSKCN1QW2QI</a>                                                                                                                   | <a href="https://www.reuters.com/video">https://www.reuters.com/video</a>                                                                                                               |
|                  | <a href="https://www.reuters.com/article/us-britain-lgbt-celebrities-idUSKCN1QW2KY">https://www.reuters.com/article/us-britain-lgbt-celebrities-idUSKCN1QW2KY</a>                                                                                                         | <a href="https://www.reuters.com/video/breakingviews">https://www.reuters.com/video/breakingviews</a>                                                                                   |
|                  | <a href="https://www.reuters.com/article/us-britain-lgbt-rights-idUSKCN1QW2QI">https://www.reuters.com/article/us-britain-lgbt-rights-idUSKCN1QW2QI</a>                                                                                                                   | <a href="https://www.reuters.com/investigates/section/homepage/">https://www.reuters.com/investigates/section/homepage/</a>                                                             |
|                  |                                                                                                                                                                                                                                                                           |                                                                                                                                                                                         |
|                  |                                                                                                                                                                                                                                                                           |                                                                                                                                                                                         |
| Channel          | <a href="https://www.reuters.com/video/2019/03/11/exclusive-trump-to-ask-for-86-billion-fo?videoid=524332535&amp;videoChannel=-10465">https://www.reuters.com/video/2019/03/11/exclusive-trump-to-ask-for-86-billion-fo?videoid=524332535&amp;videoChannel=-10465</a>     | <a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a>                                                                                                         |
|                  | <a href="https://www.reuters.com/video/2019/01/24/pelosi-calls-off-trump-speech-in-house-c?videoid=506936337&amp;videoChannel=13976">https://www.reuters.com/video/2019/01/24/pelosi-calls-off-trump-speech-in-house-c?videoid=506936337&amp;videoChannel=13976</a>       | <a href="https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD">https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD</a>               |
|                  | <a href="https://www.reuters.com/video/2019/02/28/krk-chinas-tencent-eyeing-bids-for-unive?videoid=520241085&amp;videoChannel=5">https://www.reuters.com/video/2019/02/28/krk-chinas-tencent-eyeing-bids-for-unive?videoid=520241085&amp;videoChannel=5</a>               | <a href="https://www.reuters.com/video">https://www.reuters.com/video</a>                                                                                                               |
|                  | <a href="https://www.reuters.com/video/2019/01/16/the-shifting-sands-of-gulf-energy?videoid=504277328&amp;videoChannel=118285">https://www.reuters.com/video/2019/01/16/the-shifting-sands-of-gulf-energy?videoid=504277328&amp;videoChannel=118285</a>                   | <a href="https://www.reuters.com/video/breakingviews">https://www.reuters.com/video/breakingviews</a>                                                                                   |
|                  | <a href="https://www.reuters.com/video/2019/03/12/daniel-ricciardo-realistic-about-renault?videoid=524995984&amp;videoChannel=79">https://www.reuters.com/video/2019/03/12/daniel-ricciardo-realistic-about-renault?videoid=524995984&amp;videoChannel=79</a>             | <a href="https://www.reuters.com/markets/bonds">https://www.reuters.com/markets/bonds</a>                                                                                               |
|                  | <a href="https://www.reuters.com/video/2019/03/15/zucks-right-hand-man-quits-facebook?videoid=526038800&amp;videoChannel=5">https://www.reuters.com/video/2019/03/15/zucks-right-hand-man-quits-facebook?videoid=526038800&amp;videoChannel=5</a>                         | <a href="https://www.reuters.com/places/china">https://www.reuters.com/places/china</a>                                                                                                 |
|                  | <a href="https://www.reuters.com/video/2019/02/01/jailed-reuters-reporters-appeal-to-myanmar?videoid=510223083&amp;videoChannel=117760">https://www.reuters.com/video/2019/02/01/jailed-reuters-reporters-appeal-to-myanmar?videoid=510223083&amp;videoChannel=117760</a> | <a href="https://www.reuters.com/subjects/supreme-court">https://www.reuters.com/subjects/supreme-court</a>                                                                             |
|                  | <a href="https://www.reuters.com/video/2019/03/12/nissan-renault-promise-equal-footing-aft?videoid=525085603&amp;videoChannel=-10464">https://www.reuters.com/video/2019/03/12/nissan-renault-promise-equal-footing-aft?videoid=525085603&amp;videoChannel=-10464</a>     | <a href="https://www.reuters.com/news/media">https://www.reuters.com/news/media</a>                                                                                                     |
|                  | <a href="https://www.reuters.com/video/2019/01/15/ban-ki-moon-backs-tech-solutions-to-clim?videoid=503833588&amp;videoChannel=118285">https://www.reuters.com/video/2019/01/15/ban-ki-moon-backs-tech-solutions-to-clim?videoid=503833588&amp;videoChannel=118285</a>     | <a href="https://www.reuters.com/news/lifestyle">https://www.reuters.com/news/lifestyle</a>                                                                                             |
|                  |                                                                                                                                                                                                                                                                           |                                                                                                                                                                                         |
|                  |                                                                                                                                                                                                                                                                           |                                                                                                                                                                                         |

|              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Query</b> | <b>Lucene</b><br>https://www.reuters.com/video/2019/03/10/malawi-woman-looks-to-change-girls-educat?videoid=523504071&videoChannel=73                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | <b>PageRank</b><br>https://www.reuters.com/news/entertainment/arts                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Wall</b>  | http://www.reuters.com/investigates/section/wall-streets-way/<br>https://www.reuters.com/article/us-usa-trump-budget-wall-exclusive-idUSKBN1QR0CW<br>https://www.reuters.com/article/us-usa-immigration-wall-insight-idUSKCN1QW1GR<br>https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK<br>https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK<br>https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK<br>https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK<br><a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a><br>https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8<br>https://www.reuters.com/video/2019/03/15/tech-stocks-lift-wall-street?videoid=526390533&videoChannel=303 | https://www.reuters.com/<br>https://www.reuters.com/finance<br><a href="https://www.reuters.com/politics">https://www.reuters.com/politics</a><br>https://www.reuters.com/finance/markets<br>https://www.reuters.com/video<br>https://www.reuters.com/video/breakingviews<br>https://www.reuters.com/investigates/section/homepage/<br>https://www.reuters.com/places/china<br>https://www.reuters.com/finance/funds<br>https://www.reuters.com/finance/markets/europe |

## Overlap

All queries except “Venezuela, Oscar 2019 and Channel” produced overlaps. The overlaps have been highlighted in the table. Here is a graph for overlaps.

## Overlap Graph

