

CSCI 572 HOMEWORK 4

REPORT

ID – 7386071899

Website indexed: Reutersnews

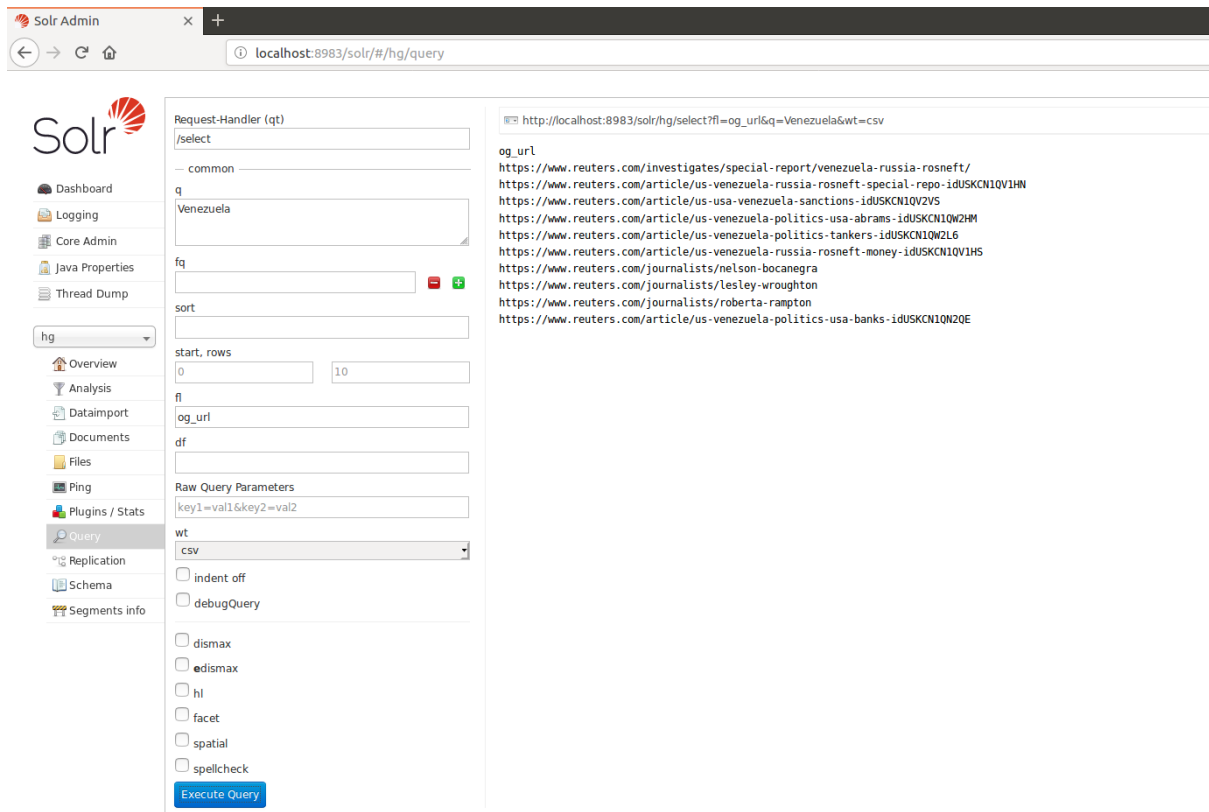
Steps followed to complete the assignment:

Step 1: Installation and Setup

- i. Downloaded Oracle's VirtualBox 5.2.26
- ii. Installed Ubuntu 16.04 on my Windows machine using Virtual Box
- iii. Downloaded Apache Solr 8.0.0.0 software
- iv. Downloaded files to be crawled (Reuters folder)

Step 2: Indexing HTML files using TIKA

- i. Started Solr server by typing command **bin/solr start** in solr directory
- ii. Created a new core "**hg**" using command **bin/solr create -c hg**
- iii. Updated managed-schema file in conf folder of the core –
Uncommented the line **<copyfield source="*" dest="_text_"/>**
- iv. Indexed the html pages by running the following command from solr's home directory-
bin/post -c hg -filetypes html crawl_data/reutersnews/
- v. Checked on Solr UI to verify that all the files were indexed properly
- vi. Made the following changes in solrconfig.xml
Updated the str element with name "df" - **<str name="df">_text_</str>**
- vii. Reloaded the core and tested by running queries



Step 3: Apache Web Server and PHP installation

- i. Installed XAMPP on my Ubuntu virtual machine. It helped me to install Apache & PHP.
- ii. Then cloned the github repository **solr-php-client**.
- iii. Added the files of solr-php-client project to the lamp/htdocs folder
- iv. Started all the servers and checked localhost to verify that web server was working correctly

Step 4: Page Rank Computation

- i. Wrote Java program to create edge_list.txt using JSoup library. In Java code, I created 2 hash maps "fileToUrl" and "urlToFile" using URLtoHTML_reuters_news.csv file
- ii. Wrote Python program to compute page rank using *edge_list.txt* file and NetworkX library
- iii. Page rank function in NetworkX library takes an edge list as input and calculates the page rank for each node. Here, nodes are web pages and edges are links connecting one web page to other.

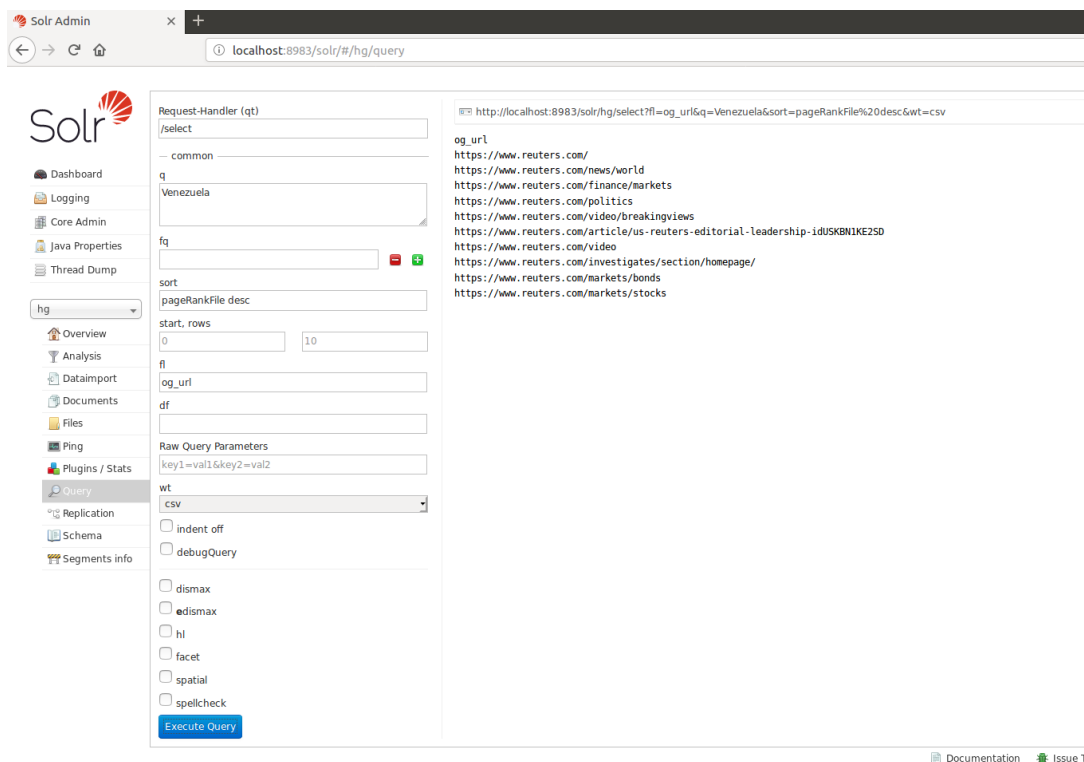
Parameters set: (for PageRank function)

*alpha=0.85, personalization=None, max_iter=30, tot=1e-06,
nstart=None, weight='weight', dangling= None*

- iv. The program stores computed page ranks in *external_pageRankFile.txt* in the format <document_id>=<page_rank_score>.

Step 5: Setting up PageRank in Solr

- i. Added the *external_pageRankFile.txt* in the data folder of core “hg”
- ii. Made necessary changes in managed-schema file by adding fieldtypes and field parameters.
- iii. Added listeners at the start of <query> tag in the solrconfig.xml file
- iv. Tested on Solr UI by placing “pageRankFile desc” in sort parameter of query



The screenshot shows the Solr Admin interface in a web browser. The URL in the address bar is `localhost:8983/solr/#/hg/query`. The left sidebar contains navigation links: Dashboard, Logging, Core Admin, Java Properties, Thread Dump, Overview, Analysis, Dataimport, Documents, Files, Ping, Plugins / Stats, Query (selected), Replication, Schema, and Segments info. The main panel is titled 'Request-Handler (qt)' and shows the following configuration:

- Request-Handler (qt): /select
- q: Venezuela
- fq: (empty)
- sort: pageRankFile desc
- start, rows: 0 to 10
- fl: og_url
- df: (empty)
- Raw Query Parameters: key1=val1&key2=val2
- wt: csv
- Indent off: ☐
- debugQuery: ☐
- dismax: ☐
- edismax: ☐
- hl: ☐
- facet: ☐
- spatial: ☐
- spellcheck: ☐

The 'Execute Query' button is at the bottom. The results section on the right shows the following URLs:

```
http://localhost:8983/solr/hg/select?fl=og_url&q=Venezuela&sort=pageRankFile%20desc&wt=csv
og_url
https://www.reuters.com/
https://www.reuters.com/news/world
https://www.reuters.com/finance/markets
https://www.reuters.com/politics
https://www.reuters.com/video/breakingviews
https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE25D
https://www.reuters.com/video
https://www.reuters.com/investigates/section/homepage/
https://www.reuters.com/markets/bonds
https://www.reuters.com/markets/stocks
```

Step 6: User Interface (PHP code)

- i. Created an interface for user to enter query and search for results.

- ii. Added a radio button for user to choose the algorithm that will be used to rank the search results - Lucene(Default) or PageRank
- iii. Extracted 4 fields from results – title, og_url, og_description and doc_id
- iv. If og_url is empty, used the “URLtoHTML_reuters_news.csv” to extract link of the page by its doc id

Why some pages have higher page rank than others?

Page rank of a page **A** is calculate using the formula

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

where d is a damping factor, T1, T2, T3 ... Tn are the pages that point to page A and C(x) is the number of links going out from page x.

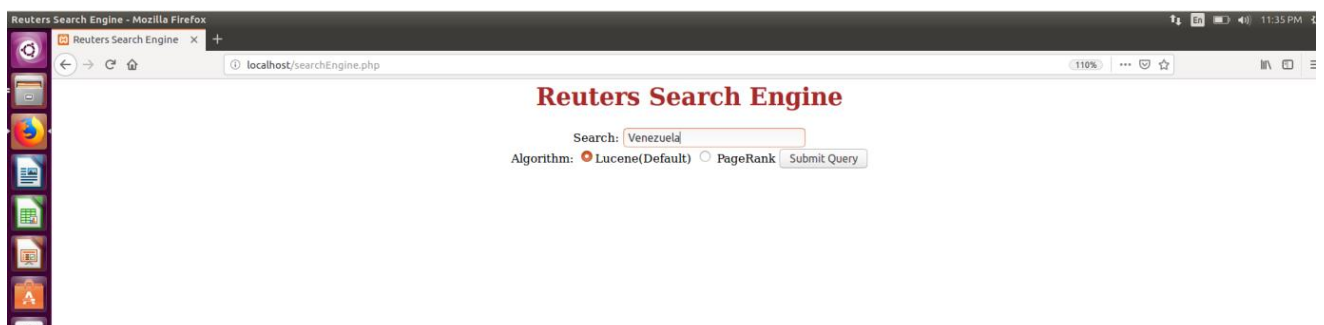
Thus, page rank of a page depends on the number of pages that point to it as well as the page rank of the pages that point to it.

As some pages are cited more often than others, they have higher page ranks.

A page with large number of incoming links especially from those pages that themselves have high page rank value leads to increase in its page rank value.

Flow for a Single Query (Screenshots)

1. Initial Page with Query box



2. Results for Lucene (Default) for Venezuela

Reuters Search Engine

Search: Venezuela

Algorithm: ☒ Lucene(Default) ☐ PageRank

Results 1 - 10 of 1709:

- Title: [How Russia sank billions of dollars into Venezuelan quicksand](#)

URL: <https://www.reuters.com/investigates/special-report/venezuela-russia-rosneft/>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/5ac186d9-6b06-4550-a2c1-3a83ff69cb84.html

Description: A Russian oil company spent a fortune in Venezuela despite suspecting it was being ripped off. Sources say the Kremlin wanted to help its South American ally.
- Title: [Special Report: How Russia sank billions of dollars into Venezuela quicksand | Reuters](#)

URL: <https://www.reuters.com/article/us-venezuela-russia-rosneft-special-repo-idUSKCN1OV1HN>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/2d57ef26-437b-4249-98be-1ff0d5b9d814.html

Description: At the end of 2015, managers at Rosneft, the Russian state-controlled oil firm, ...
- Title: [U.S. considers sanctions to restrict Visa, Mastercard in Venezuela: official | Reuters](#)

URL: <https://www.reuters.com/article/us-usa-venezuela-sanctions-idUSKCN1OV2VS>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/e49e96d0-9a07-4890-9c5d-302862160ed0.html

Description: The United States is considering imposing financial sanctions that could prohibi...
- Title: [U.S. envoy says Venezuela oil production dropping steadily | Reuters](#)

URL: <https://www.reuters.com/article/us-venezuela-politics-usa-sbrams-idUSKCN1OW2HM>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/c7253e2-eda2-47a6-97b2-8840c3ad3e89.html

Description: The U.S. special representative for Venezuela said on Friday that Venezuela'...
- Title: [Shipbroker McQuilling ends tanker contracts with Venezuela's PDVSA | Reuters](#)

URL: <https://www.reuters.com/article/us-venezuela-politics-tankers-idUSKCN1OWZL6>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/f543eabb-7882-4aa9-9962-83ed792b2306.html

Description: A second maritime firm has told Venezuela's state-run oil firm Petroleos de...
- Title: [Billions spent, little to show | Reuters](#)

URL: <https://www.reuters.com/article/us-venezuela-russia-rosneft-money-idUSKCN1OV1HS>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/8930da40-6b0b-4276-9c99-83217309951f.html

Description: Reuters estimates that after sinking more than \$9 billion into Venezuela in loan...
- Title: [Columns | Reuters.com](#)

URL: <https://www.reuters.com/journalists/nelson-bocanegra>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/c6803977-3c52-4814-993e-8870c797c68b.html

Description: Reuters.com brings you the latest news from around the world, covering breaking news in markets, business, politics, entertainment, technology, video and pictures.
- Title: [Columns | Reuters.com](#)

URL: <https://www.reuters.com/journalists/lesley-wroughton>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/57d346b9-8ffc-40f8-a9a8-194082e1d149.html

Description: Reuters.com brings you the latest news from around the world, covering breaking news in markets, business, politics, entertainment, technology, video and pictures.
- Title: [Columns | Reuters.com](#)

URL: <https://www.reuters.com/journalists/roberta-ramplog>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/d12d761f-755c-41d2-a68d-5658009dcf2.html

Description: Reuters.com brings you the latest news from around the world, covering breaking news in markets, business, politics, entertainment, technology, video and pictures.
- Title: [U.S. set to punish foreign banks who deal with Venezuela's Maduro: official | Reuters](#)

URL: <https://www.reuters.com/article/us-venezuela-politics-usa-banks-idUSKCN1QN2OE>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/67c990f1062-4cae-adfa-6d61d8278d11.html

Description: The United States has identified efforts by Venezuelan President Nicolas Maduro ...

3. Results for Page Rank for Venezuela

Reuters Search Engine

Search: Venezuela

Algorithm: ☐ Lucene(Default) ☒ PageRank

Results 1 - 10 of 1709:

- Title: [Business & Financial News, U.S & International Breaking News | Reuters](#)

URL: <https://www.reuters.com/>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/5ae3e0ec-c598-48d4-b3a3-ab29d3b3e778.html

Description: Reuters.com brings you the latest news from around the world, covering breaking news in markets, business, politics, entertainment, technology, video and pictures.
- Title: [World News | Reuters.com](#)

URL: <https://www.reuters.com/news/world>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/eece5caa-8d4a-4cb7-b670-f042b665123a.html

Description: Find late breaking world news from every corner of the globe at Reuters.com, your online source for international news coverage.
- Title: [Markets & Finance News | Reuters.com](#)

URL: <https://www.reuters.com/finance/markets>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/dcca6198-104a-469f-9b07-e324bcee4275.html

Description: Connect to the most current information on stocks and bonds on Reuters.com. Finance and investing news on NYSE, Dow Jones, S&P, and Nasdaq at Reuters.com.
- Title: [Politics & Political News | Reuters](#)

URL: <https://www.reuters.com/politics>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/b3557cce-8ba0-43d4-a477-8334f685745a.html

Description: Reuters.com is your politics news source for up to the minute, impartial political news coverage on the leaders, policies and agendas that shape our world.
- Title: [Breakingviews Video - Reuters.com](#)

URL: <https://www.reuters.com/video/breakingviews>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/19bb0d07-8fba-496c-ad7c-860833545f8b.html

Description: After scandal enveloped a member of South Korean boy band Big Bang, shares in the industry's talent agencies plummeted this week. Robyn Mak and Jeffrey Goldfarb discuss how YouTube and Spotify should be able to help the popular music export pick up...
- Title: [Reuters Editorial Leadership | Reuters](#)

URL: <https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/b7e1d031-d9dc-4acb-a788-5b5adda4f566.html

Description: Stephen J. Adler Editor-in-Chief
- Title: [Top News & Business Videos - Reuters.com](#)

URL: <https://www.reuters.com/video>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/ca69ffcd-c30f-45bb-bd16-100a7ee7ee14.html

Description: Reuters has learned that when he was growing up in El Paso, Texas, the kid who would grow up to be Democratic presidential hopeful Beto O'Rourke was a member of one of the earliest - and most influential - hacker collectives.
- Title: [Special Reports from Reuters journalists around the world](#)

URL: <https://www.reuters.com/investigates/section/homepage/>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/59797e04-f360-4c9e-9cce-b68ab6e3b0b3.html

Description: The latest Reuters investigations and other in-depth stories, plus a full archive of the past five years
- Title: [Rates & Bonds](#)

URL: <https://www.reuters.com/markets/bonds>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/0202b4c0-a6c6-49b5-9b99-95b9137a1533.html

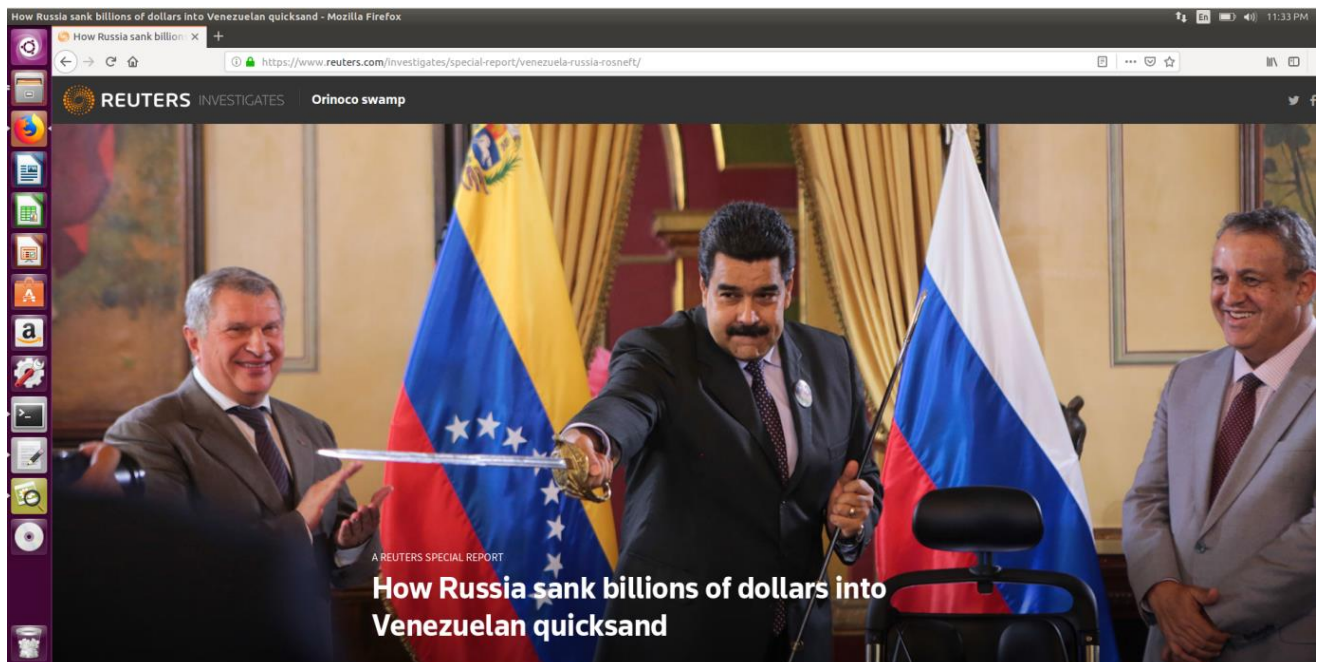
Description: Bond Yields, U.S. Debt, The Federal Reserve, and more...
- Title: [Stocks](#)

URL: <https://www.reuters.com/markets/stocks>

Doc Id: /home/shakshi/Desktop/solr-8.0.0/crawl_data/reutersnews/83931643-c4fd-4a6e-b137-26d551e12faf.html

Description: Reuters.com for the latest indices and markets news.

4. Actual Web Page by clicking on one of result's link



Query Results

1. Venezuela

LUCENE(DEFAULT)	PAGERANK
https://www.reuters.com/investigates/special-report/venezuela-russia-rosneft/	https://www.reuters.com/
https://www.reuters.com/article/us-venezuela-russia-rosneft-special-repo-idUSKCN1QV1HN	https://www.reuters.com/news/world
https://www.reuters.com/article/us-usa-venezuela-sanctions-idUSKCN1QV2VS	https://www.reuters.com/finance/markets
https://www.reuters.com/article/us-venezuela-politics-usa-abrams-idUSKCN1QW2HM	https://www.reuters.com/politics
https://www.reuters.com/article/us-venezuela-politics-tankers-idUSKCN1QW2L6	https://www.reuters.com/video/breakingviews
https://www.reuters.com/article/us-venezuela-russia-rosneft-money-idUSKCN1QV1HS	https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD
https://www.reuters.com/journalists/nelson-bocanegra	https://www.reuters.com/video
https://www.reuters.com/journalists/lesley-wroughton	https://www.reuters.com/investigates/section/homepage/
https://www.reuters.com/journalists/roberta-rampton	https://www.reuters.com/markets/bonds
https://www.reuters.com/article/us-venezuela-politics-usa-banks-idUSKCN1QN2QE	https://www.reuters.com/markets/stocks

2. Senate

LUCENE(DEFAULT)	PAGERANK
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8	https://www.reuters.com/politics
https://www.reuters.com/video/2019/03/04/senate-has-votes-to-block-border-emergen?videoid=521958334&videoChannel=-10465	https://www.reuters.com/video/breakingviews
https://www.reuters.com/video/2019/03/05/mcconnell-senate-has-votes-to-block-bord?videoid=521939215&videoChannel=-10465	https://www.reuters.com/video
https://www.reuters.com/article/us-usa-politics-beto-orourke-special-rep-idUSKCN1QW26N	https://www.reuters.com/news/pictures
https://www.reuters.com/article/us-usa-politics-beto-orourke-special-rep-idUSKCN1QW26N	https://www.reuters.com/subjects/sustainable-business
https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H	https://www.reuters.com/news/media
https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H	https://www.reuters.com/places/mexico
https://www.reuters.com/politics	https://www.reuters.com/video/2019/03/14/orourke-hits-the-ground-running-in-iowa?videoid=525923095&videoChannel=1003
https://www.reuters.com/article/us-congo-politics-idUSKCN1QW28R	https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK
https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV00E	https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H

3. Democrats

LUCENE(DEFAULT)	PAGERANK
https://www.reuters.com/article/us-usa-election-medicare-idUSKBN1QU189	https://www.reuters.com/
https://www.reuters.com/article/us-usa-election-medicare-idUSKBN1QU189	https://www.reuters.com/politics
https://www.reuters.com/article/us-usa-congress-whitaker-idUSKCN1PX1LD	https://www.reuters.com/subjects/supreme-court
https://www.reuters.com/video/2019/02/28/republicans-bash-cohen-democrats-at-cpac?videoId=520293110&videoChannel=13976	https://www.reuters.com/breakingviews
https://www.reuters.com/article/us-usa-election-biden-idUSKCN1QH1EX	https://www.reuters.com/news/archive/politicsNews
https://www.reuters.com/article/us-usa-trump-budget-idUSKBN1QS12D	https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK
https://www.reuters.com/video/2019/03/05/the-witch-hunt-continues-trump-on-democr?videoId=522293719&videoChannel=13976	https://www.reuters.com/article/us-usa-election-sanders-idUSKCN1QW2ZJ
https://www.reuters.com/article/us-usa-trump-russia-congress-idUSKCN1QV23E	https://www.reuters.com/article/us-usa-politics-beto-short-exclusive-idUSKCN1QW28H
https://www.reuters.com/politics	https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/article/us-usa-election-booker-idUSKCN1QW2P4

4. Republicans

LUCENE(DEFAULT)	PAGERANK
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8	https://www.reuters.com/politics
https://www.reuters.com/video/2019/02/28/republicans-bash-cohen-democrats-at-cpac?videoId=520293110&videoChannel=13976	https://www.reuters.com/news/us
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/news/archive/politicsNews
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C
https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY	https://www.reuters.com/article/us-usa-immigration-wall-insight-idUSKCN1QW1GR
https://www.reuters.com/article/us-usa-tax-survey-idUSKCN1QW1BY	https://www.reuters.com/article/us-usa-trump-russia-congress-idUSKCN1QV23E
https://www.reuters.com/article/us-california-fires-trump-idUSKCN1P31ND	https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8
https://www.reuters.com/article/us-usa-politics-trump-idUSKCN1QH29L	https://www.reuters.com/article/us-usa-court-census-idUSKCN1QW2MC

5. Patriot Movement

LUCENE(DEFAULT)	PAGERANK
https://www.reuters.com/article/us-usa-trump-turkey-idUSKCN1QV328	https://www.reuters.com/markets/bonds
https://www.reuters.com/article/us-france-protests-idUSKCN1Q325A	https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C
https://www.reuters.com/article/us-britain-climatechange-education-idUSKCN1QW2QU	https://www.reuters.com/article/canada-forex-idUSL1N2121OB
https://www.reuters.com/article/uk-people-rkelly-metoo-idUKKCN1QC0V7	https://www.reuters.com/article/us-climate-change-youth-usa-idUSKCN1QW1BE
https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S	https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S
https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S	https://www.reuters.com/article/us-slovakia-election-president-idUSKCN1QW2Z3
https://www.reuters.com/article/us-climate-change-youth-idUSKCN1QW01S	https://www.reuters.com/article/us-climate-change-youth-usa-idUSKCN1QW1BE
https://www.reuters.com/article/us-italy-china-analysis-idUSKCN1QW1E2	https://www.reuters.com/investigates/special-report/myanmar-suukyi-history/
https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C	https://www.reuters.com/article/us-auction-christies-georgemichael-art-idUSKCN1QP1EY
https://www.reuters.com/article/us-usa-election-orourke-idUSKCN1QV13C	https://www.reuters.com/article/us-italy-china-analysis-idUSKCN1QW1E2

6. Oscar 2019

LUCENE(DEFAULT)	PAGERANK
https://www.reuters.com/article/us-awards-oscars-women-idUSKCN1Q7130	https://www.reuters.com/
https://www.reuters.com/article/us-awards-oscars-diversity-idUSKCN1QE0N4	https://www.reuters.com/tools/mobile
https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT	https://www.reuters.com/news/world
https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT	https://www.reuters.com/finance
https://www.reuters.com/article/us-aviancaholdings-fleet-idUSKCN1QW2XT	https://www.reuters.com/finance/markets
https://www.reuters.com/journalists/pei-li	https://www.reuters.com/politics
https://www.reuters.com/article/us-britain-lgbt-celebrities-idUSKCN1QW2KY	https://www.reuters.com/video/breakingviews
https://www.reuters.com/theWire	https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD
https://www.reuters.com/article/us-britain-lgbt-rights-idUSKCN1QW2QI	https://www.reuters.com/video
https://www.reuters.com/article/us-britain-lgbt-rights-idUSKCN1QW2QI	https://www.reuters.com/investigates/section/homepage/

7. Channel

LUCENE(DEFAULT)	PAGERANK
https://www.reuters.com/video/2019/03/14/new-fines-for-insulting-russian-state-on?videoId=525822169&videoChannel=-10167	https://www.reuters.com/politics
https://www.reuters.com/video/2019/03/15/global-gloom-forces-japan-central-bank-t?videoId=526297395&videoChannel=-10464	https://www.reuters.com/video/breakingviews
https://www.reuters.com/video/2019/03/14/klopp-delighted-with-champions-league-wi?videoId=525657525&videoChannel=79	https://www.reuters.com/article/us-reuters-editorial-leadership-idUSKBN1KE2SD
https://www.reuters.com/video/2019/03/13/teenagers-killed-in-sao-paolo-school-sho?videoId=525625827&videoChannel=-10459	https://www.reuters.com/video
https://www.reuters.com/video/2019/03/12/nissan-renault-promise-equal-footing-aft?videoId=525085603&videoChannel=5	https://www.reuters.com/markets/bonds
https://www.reuters.com/video/2019/03/10/brexit-in-peril-if-pm-mays-deal-is-rejec?videoId=524245642&videoChannel=118261	https://www.reuters.com/places/china
https://www.reuters.com/video/2019/03/07/in-mexican-heartland-bad-guys-still-hold?videoId=523112542&videoChannel=-10465	https://www.reuters.com/news/lifestyle
https://www.reuters.com/video/2019/03/14/facebook-back-online-after-huge-outage?videoId=525957436&videoChannel=6	https://www.reuters.com/news/science
https://www.reuters.com/video/2019/03/15/at-kenya-summit-a-47bln-pledge-to-africa?videoId=526324412&videoChannel=73	https://www.reuters.com/subjects/sustainable-business
https://www.reuters.com/video/2019/03/13/liverpool-result-at-bayern-wont-affect-p?videoId=525185681&videoChannel=79	https://www.reuters.com/news/us

8. Wall

LUCENE(DEFAULT)	PAGERANK
http://www.reuters.com/investigates/section/wall-streets-way/	https://www.reuters.com/
https://www.reuters.com/article/us-usa-trump-budget-wall-exclusive-idUSKBN1QR0CW	https://www.reuters.com/finance
https://www.reuters.com/article/us-usa-immigration-wall-insight-idUSKCN1QW1GR	https://www.reuters.com/finance/markets
https://www.reuters.com/politics	https://www.reuters.com/politics
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/video/breakingviews
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/video
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/investigates/section/homepage/
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QW2AK	https://www.reuters.com/finance/markets/dividends
https://www.reuters.com/article/us-usa-trump-congress-emergency-idUSKCN1QV1A8	https://www.reuters.com/places/china
https://www.reuters.com/journalists/noel-randewich	https://www.reuters.com/finance/EarningsUS

OVERLAP GRAPH

Q.No	Query	#Overlaps
1	Venezuela	0
2	Senate	2
3	Democrats	2
4	Repulicans	3
5	Patriot Movement	3
6	Oscar 2019	0
7	Channel	0
8	Wall	1

