

CSCI-572 Homework 4 Report

Name : Shwetanshu Rohatgi

Website Indexed : Latimes

Steps followed to complete HW4:

Step 1: VM Installation with Ubuntu:

- Installed Ubuntu 16.04 on my Windows laptop inside Oracle VM Virtual Box.
- Setup a Shared folder named Shared-ubuntu

Step 2: Solr Installation:

- I downloaded solr 7.7.0 from the mentioned website and unzipped the tar.
- Started the Apache solr server using `bin/solr start`.
- Created a core using the command: `bin/solr create -c solr_core`, as suggested in the document and named it 'solr_core'.
- Solr was launched in the web browser at port 8983: `localhost:8983/solr/#`

Step 3: Indexing the HTML files:

- Started the solr server on 8983 port using `bin/solr start`.
- Edited the managed-schema file in the conf folder so that all the text content can be mapped properly.
- Refreshed Solr admin and Core page and submitted a few Queries through Solr UI to query solr_core html pages.

Step 4: Apache Web Server, PHP installation:

- I installed both using the below commands:

```
sudo apt-get install php7 php7.0-cli  
sudo apt-get install apache2  
sudo apt-get install libapache2-mod-php
```
- Copied the Apache folder and PHP files in `var/www/html` and tested for basic html/php rendering.

Step 5: Page Rank algorithm:

- Downloaded Jsoup jars and wrote a Java file ExtractLinks.java to find and save edgelist for the html files in latimes folder.
- Constructed various data structures such as: urlFileMap and fileUrlMap to construct the edgelist.
- All the outputs written into a separate file called *networkEdgeList.txt*.
- Summary:
 - a) Extract outgoing links
 - b) Find the file name for each outgoing link using the provided map* files
 - c) Create edges between files that have links connecting them.

Step 6: Python code and NetworkX Configuration to compute PageRanks:

- Downloaded and installed the NetworkX python library in Anaconda from <https://pypi.python.org/pypi/networkx/>
 - `G = nx.read_edgelist("networkEdgeList.txt", create_using=nx.DiGraph())`
 - Pagerank Algorithm Configuration:
`alpha=0.85, personalization=None, max_iter=30, tol=1e-06, nstart=None, weight='weight', dangling=None`
`pRank = nx.pagerank(G, alpha=0.85, personalization=None, max_iter=30, tol=1e-06, nstart=None, weight='weight', dangling=None)`

Step 7: Added following lines to managed-schema to handle PageRank external file:

- `<fieldType name="external" keyField="id" defVal="0" class="solr.ExternalFileField" />`
- `<field name="pageRankFile" type="external" stored="false" indexed="false" />`

Step 8: Added following lines to solrconfig.xml:

- `<listener event="newSearcher" class="org.apache.solr.schema.ExternalFileFieldReloader" />`
- `<listener event="firstSearcher" class="org.apache.solr.schema.ExternalFileFieldReloader" />`

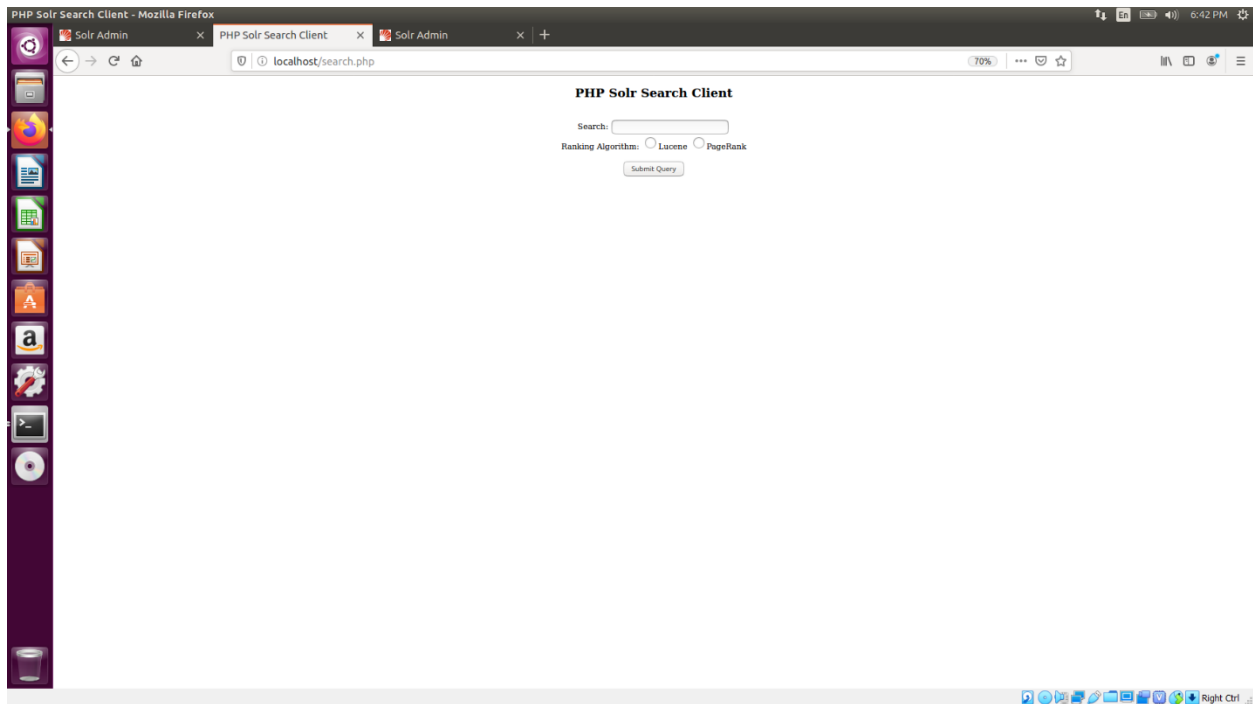
Step 9: Finalizing Code for Solr-PHP Client:

- I placed the calculated external_pageRankFile.txt file in data folder of the created solr_core. and reloaded to take care of new changes.
- Developed client-code in PHP to handle HTTP request/response with Solr to fetch search results from both Lucene and PageRank algorithms.

Explanation regarding why some pages have higher PageRank values than others:

PageRank works by counting the number of in-links to a page and quality of those in-links to determine a rough estimate of how important the website is. Some pages have higher PageRank because the algorithm calculates score for each *.html file in latimes crawl_data folder by computing the number of incoming links to each webpage. If a webpage has lots of in-links then chances are it will have higher PageRank score. In our case, this data is being fed from external_edgelist.txt file and using NetworkX Library we compute PageRank scores. The underlying assumption is that more important websites are likely to receive more links from other websites and hence a higher PageRank Score.

Documenting flow for any single query using screenshots:



Initial Page with Query Box

PHP Solr Search Client - Mozilla Firefox

Search: Ranking Algorithm: ☐ Lucene ☒ PageRank Submit Query

Results 1 - 10 of 1624:

1.	ID /Shared-ubuntu/solr-7.7.0/tatimes/79940e4d-eaef-4874-9045-5862ae8f73a.html Title Faced with the most damning impeachment testimony yet, Republicans on info full clown mode - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-10-23/rep-republicans-storm-hearing-impeachment Description It's easier to bluster about the impeachment inquiry process than deny William Taylor's account of the quid pro quo President Trump demanded of Ukraine.
2.	ID /Shared-ubuntu/solr-7.7.0/tatimes/be7b0464-b44c-4774-b2d9-906a7f3bc5ff.html Title Democrats' hopes of winning over Republicans for impeachment faded - Los Angeles Times URL https://www.latimes.com/politics/story/2019-10-20/democrats-hoped-david-will-over-republicans-on-impeachment-but-its-not-looking-that-way-as-yet Description Nancy Pelosi once said impeachment would only work with bipartisan support, but it's not turning out that way so far.
3.	ID /Shared-ubuntu/solr-7.7.0/tatimes/433b58bc-64ae-441e-e91f-76b50edecac.html Title Opinion: What else will Trump do before GOP supports impeachment? - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-11-08/impeachment-trump-republicans-tourney Description Is there a line Trump could cross that would lead Republicans to put national interest ahead of party? That question looms larger than his abuses.
4.	ID /Shared-ubuntu/solr-7.7.0/tatimes/90ad4af3-304f-404e-9c2b-90563c0ad170.html Title Impeachment prospect worries Senate Republicans from swing states, and could endanger GOP majority - Los Angeles Times URL https://www.latimes.com/politics/story/2019-10-24/vulnerable-senate-republicans-impeachment Description Some Democrats have worried that political backlash to the drive to impeach President Trump could cost them control of the House. But the unpredictable issue could pose a bigger threat to the Republican majority in the Senate.
5.	ID /Shared-ubuntu/solr-7.7.0/tatimes/accfb584-4666-409f-8431-7e7e7021250b.html Title Column: Republicans blame Democratic leadership for California's problems. They have a point, but ... - Los Angeles Times URL https://www.latimes.com/california/story/2019-10-16/california-republicans-democrats-wildfires-homelessness Description It's true Democrats have had a firm grip on California politics recently, but what have Republicans offered?
6.	ID /Shared-ubuntu/solr-7.7.0/tatimes/9f9b44e67-3d24-4544-9366-a1a06e6e0c1f.html Title What's really happening inside the impeachment inquiry room, stormed by Republicans - Los Angeles Times URL https://www.latimes.com/politics/story/2019-10-23/impeachment-deposition-room-stormed-by-republicans Description Dozens of GOP lawmakers, including some of Trump's most loyal supporters, have been attending the closed-door impeachment hearings — given equal time to ask questions.
7.	ID /Shared-ubuntu/solr-7.7.0/tatimes/3a070ed-5548-4036-ba63-85306e3e4248.html Title Opinion: Rich Californians sure are eager to help fund Trump's attacks on the Golden State - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-09-17/california-republicans-trump-fundraise-president Description It's disheartening that fellow Californians are so eager to support a president whose retrograde positions on the environment threaten the very things that make the Golden State great.
8.	ID /Shared-ubuntu/solr-7.7.0/tatimes/62c9e-904-9974-474043d309d9d.html Title Is Trump an obstacle to Republicans winning back Katie Hill's seat? - Los Angeles Times URL https://www.latimes.com/politics/story/2019-10-28/katie-hill-resignation-trump-obstacle-for-republicans Description Candidate scramble to run for Katie Hill's seat
9.	ID /Shared-ubuntu/solr-7.7.0/tatimes/c0b2e2c-e84e-49c5-b1ba-39f6e3717c64.html Title Column: Don't believe anything you read about the 2020 election URL https://www.latimes.com/opinion/story/2019-11-06/trump-republicans-democrats-predictions-nothing-2020 Description If you pay attention to the pundits and politicos, your head will never stop spinning
10.	ID /Shared-ubuntu/solr-7.7.0/tatimes/441402c1-1362-402a-a053-59a922c9476d.html Title Republicans see warning signs after election losses - Los Angeles Times URL https://www.latimes.com/politics/story/2019-11-06/republicans-election-trump-candidate-loses-kentucky Description Republican losses in Kentucky, Virginia and Pennsylvania showed that suburban voters increasingly are moving away from Trump and the GOP as the president heads into his reelection race.

Results for 'Republicans' Query (Lucene)

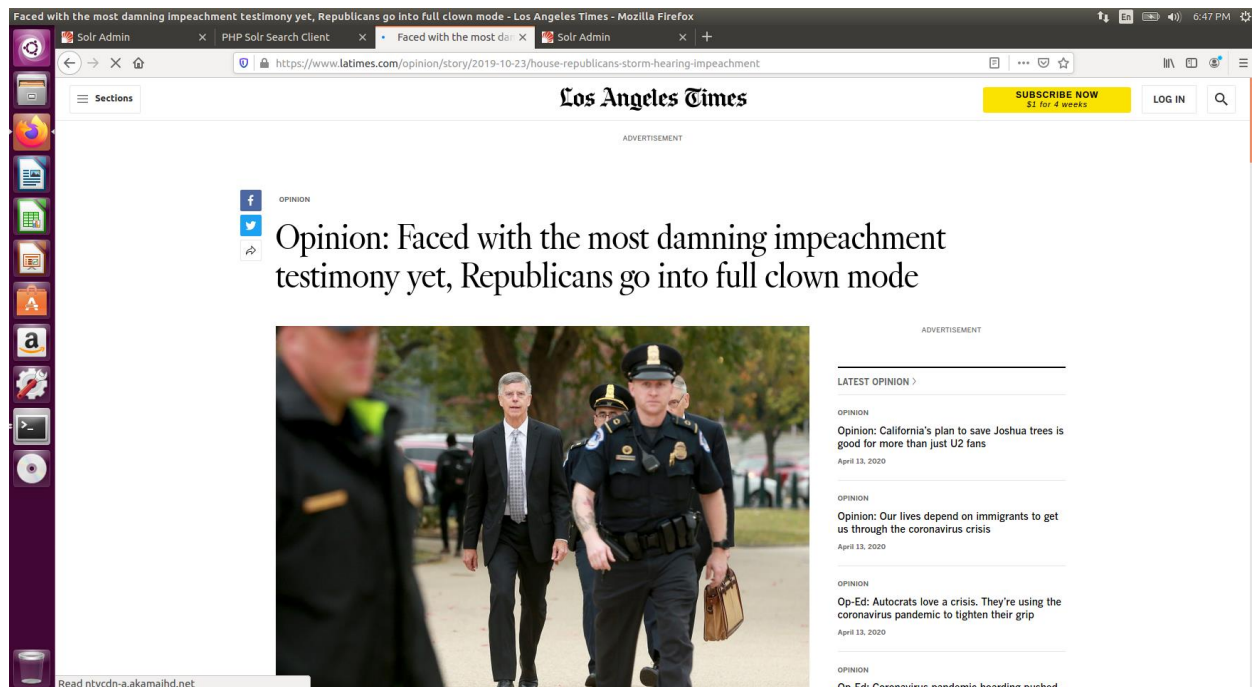
PHP Solr Search Client - Mozilla Firefox

Search: Ranking Algorithm: ☐ Lucene ☒ PageRank Submit Query

Results 1 - 10 of 1624:

1.	ID /Shared-ubuntu/solr-7.7.0/tatimes/576336-4e-644b-4b09-ab4c-555448c06841.html Title News from California, the nation and world - Los Angeles Times URL https://www.latimes.com/ Description N/A
2.	ID /Shared-ubuntu/solr-7.7.0/tatimes/cab28094-610b-446f-93d1-1b9838f5ade9.html Title Opinion - Los Angeles Times URL https://www.latimes.com/opinion Description Opinion, op-eds, letters to the editor, analysis and policy ideas from the Los Angeles Times.
3.	ID /Shared-ubuntu/solr-7.7.0/tatimes/2e9b42e-1a65-4304-9aea-63eadf9124c.html Title Submit a Letter to the Editor - Los Angeles Times URL https://www.latimes.com/opinion/a/letter-to-the-editor.html/story.html Description N/A
4.	ID /Shared-ubuntu/solr-7.7.0/tatimes/be5c4742-8647-4153-8c4d-985899de2da5.html Title Opinion: Lies and conspiracy theories drive Facebook profits - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-11-09/facebook-buttler-political-ads-lies Description As long as the financial returns of a social platform align with outrage, the site will be optimized for distributing disinformation.
5.	ID /Shared-ubuntu/solr-7.7.0/tatimes/820c3021-7034-48e7-a5cc-a91713e898c.html Title U.S. learned wrong lessons from Berlin Wall, it out as Trump - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-11-08/berlin-wall-30th-anniversary-cold-war-donald-trump Description Three decades into the post-Cold War era, we can assess what U.S. primacy has yielded. U.S. global leadership has been irrelevant or counterproductive.
6.	ID /Shared-ubuntu/solr-7.7.0/tatimes/2c0b5e5b-1996-4484-55ee-0244211b052c.html Title Elizabeth Warren took down an Obama nominee. Was it for nothing? - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-11-09/elizabeth-warren-used-hate-knuckle-facties-to-take-down-an-obama-nominee Description In 2014, Sen. Elizabeth Warren demonstrated her ability to galvanize liberals. But as it turns out, the fight she won was mostly symbolic.
7.	ID /Shared-ubuntu/solr-7.7.0/tatimes/afbb21ab-1b7b-4da3-8363-80215bb263ed.html Title Elizabeth Warren is nothing like Donald Trump - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-11-09/elizabeth-warren-trump Description The argument that Elizabeth Warren is like Trump because she makes bold promises ignores her background in law, government and advocacy.
8.	ID /Shared-ubuntu/solr-7.7.0/tatimes/56074e45-e9b3-4d21-a3e6-1b5844a4f8bf.html Title Opinion: Devin Nunes wants impeachment focus on 2016, not Trump - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-11-09/devin-nunes-impeachment-williams-hal Description The GOP seeks to call impeachment witnesses who will back their theory that Ukraine meddled in the 2016 election and Burisma was corrupt.
9.	ID /Shared-ubuntu/solr-7.7.0/tatimes/05ccf320-91d7-44c7-ba81-3d693cbe9d6e.html Title U.C. should not dumb itself down by dropping the SAT - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-11-09/u-c-college-admissions Description Studying really hard can earn you high grades, but it might not do much for you on the SAT. Colleges need to know who is both smart and studious.
10.	ID /Shared-ubuntu/solr-7.7.0/tatimes/e7134d6c-e1e5-4ed6-8552-df52d346c18.html Title Opinion: What else will Trump do before GOP supports impeachment? - Los Angeles Times URL https://www.latimes.com/opinion/story/2019-11-08/impeachment-trump-republicans-tourney Description Is there a line Trump could cross that would lead Republicans to put national interest ahead of party? That question looms larger than his abuses.

Results for 'Republicans' Query (PageRank)



Webpage opened on First Result (Lucene) Click

Table containing list of URLs of the top ten results produced by both ranking methods

Cannes: 1 Overlap

Lucene Results	PageRank Results
https://www.latimes.com/topic/cannes-film-festival	https://www.latimes.com/sitemap
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-wrap-critics-conversation-20190526-story.html	https://www.latimes.com/espanol/entretenimiento/articulo/2019-10-04/por-fin-llega-joker-aqui-un-resumen-de-todo-el-drama-que-ha-provocado
https://www.latimes.com/entertainment/movies/moviesnow/la-et-mn-cannes-2015-winners-20150524-story.html	https://www.latimes.com/entertainment-arts/movies/story/2019-11-08/documentaries-apollo-11-the-cave-sea-of-shadows-honeyland
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-parasite-atlantics-awards-20190526-story.html	https://www.latimes.com/lifestyle/gallery/chloe-sevignys-five-favorite-frocks
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-film-festival-award-winners-20190525-story.html	https://www.latimes.com/topic/film-festivals
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-once-upon-a-time-in-hollywood-20190522-story.html	https://www.latimes.com/topic/cannes-film-festival
https://www.latimes.com/entertainment/movies/la-ca-mn-cannes-genre-parasite-wild-geese-lake-tarantino-20190522-story.html	https://www.latimes.com/people/genaro-molina

https://www.latimes.com/entertainment/movies/la-et-mn-cannes-mektoub-my-love-intermezzo-abdellatif-kechiche-20190523-story.html	https://www.latimes.com/entertainment-arts/music/story/2019-11-06/whitney-houston-lesbian-relationship-robyn-crawford
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-quentin-tarantino-once-upon-a-time-in-hollywood-20190521-story.html	https://www.latimes.com/entertainment-arts/movies/story/2019-10-31/apple-disney-hbomax-streaming-movies
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-robert-eggers-the-lighthouse-robert-pattinson-20190524-story.html	https://www.latimes.com/opinion/story/2019-09-17/patt-morrison-dana-thomas-fast-fashion-environment

Congress: 0 Overlaps

Lucene Results	PageRank Results
https://www.latimes.com/politics/story/2019-10-28/jerry-brown-to-testify-to-congress-rebutting-trumps-criticism-of-california	https://www.latimes.com/opinion
https://www.latimes.com/opinion/enterthefray/la-ol-new-york-rhode-island-congress-marijuana-20190116-story.html	https://www.latimes.com/espanol/politica
https://www.latimes.com/politics/la-na-pol-congress-sexual-harassment-20181212-story.html	https://www.latimes.com/politics/story/2019-11-08/michael-bloomberg-files-papers-for-2020-democratic-presidential-primary
https://www.latimes.com/politics/la-pol-ca-richest-california-lawmakers-20180305-story.html	https://www.latimes.com/opinion/story/2019-11-09/facebook-twitter-political-ads-lies
https://www.latimes.com/politics/story/2019-09-09/congress-gun-control-government-shutdown	https://www.latimes.com/politics/story/2019-11-08/elizabeth-warren-used-bare-knuckle-tactics-to-take-down-an-obama-nominee
https://www.latimes.com/opinion/story/2019-10-16/democratic-debate-trump-congress-executive-action	https://www.latimes.com/opinion/story/2019-11-09/devin-nunes-impeachment-witness-list
https://www.latimes.com/politics/story/2019-10-14/trump-ukraine-aid-congress-impeachment	https://www.latimes.com/environment/story/2019-11-08/eastern-sierra-towns-recreation-tourism-forest-service
https://www.latimes.com/politics/la-na-pol-congress-bioweapons-detection-system-20190414-story.html	https://www.latimes.com/politics/story/2019-11-08/steve-bannon-roger-stone-trump-campaign-wikileaks
https://www.latimes.com/politics/la-na-pol-congress-mueller-pressure-impeachment-20190529-story.html	https://www.latimes.com/opinion/story/2019-11-08/impeachment-trump-republicans-toomey
https://www.latimes.com/politics/la-pol-ca-richest-in-congress-darrell-issa-story.html	https://www.latimes.com/politics/story/2019-11-08/alexander-vindman-fiona-hill-impeachment-transcripts

Democrats: 0 Overlaps

Lucene Results	PageRank Results
https://www.latimes.com/politics/story/2019-11-07/house-democrats-subpoena-mick-mulvaney-in-impeachment-probe	https://www.latimes.com/california

https://www.latimes.com/opinion/story/2019-11-05/impeachment-donald-trump-democrats-political-cost-elections	https://www.latimes.com/opinion
https://www.latimes.com/politics/story/2019-11-03/iowa-democrats-candidate-beat-trump	https://www.latimes.com/politics
https://www.latimes.com/politics/story/2019-10-17/democrats-quick-impeachment-timing-complicated	https://www.latimes.com/politics/story/2019-11-08/michael-bloomberg-files-papers-for-2020-democratic-presidential-primary
https://www.latimes.com/politics/story/2019-10-08/impeachment-trump-democrats-whistleblower-identity-testimony	https://www.latimes.com/politics/story/2019-11-08/democratic-primary-debate-moves-from-ucla-to-loyola-marymount-university
https://www.latimes.com/politics/story/2019-09-24/democrats-and-pelosi-appear-close-to-tipping-point-on-impeachment	https://www.latimes.com/politics/story/2019-11-08/elizabeth-warren-used-bare-knuckle-tactics-to-take-down-an-obama-nominee
https://www.latimes.com/politics/story/2019-09-29/trump-allies-and-democrats-reveal-the-deep-divisions-over-impeachment-inquiry	https://www.latimes.com/opinion/story/2019-11-09/elizabeth-warren-trump
https://www.latimes.com/politics/story/2019-08-01/assignment-post-debate-health-care-analysis-new-story	https://www.latimes.com/opinion/story/2019-11-09/devin-nunes-impeachment-witness-list
https://www.latimes.com/nation/ct-democrats-economic-plan-20170824-story.html	https://www.latimes.com/politics/story/2019-11-08/steve-bannon-roger-stone-trump-campaign-wikileaks
https://www.latimes.com/california/story/2019-10-21/california-independent-voters-can-cast-ballots-for-democrats-not-trump-march-2020-primary	https://www.latimes.com/opinion/story/2019-11-08/impeachment-trump-republicans-toomey

Patriot Movement: 0 overlaps

Lucene Results	PageRank Results
https://www.latimes.com/archives/la-xpm-2012-oct-23-la-me-russell-means-20121023-story.html	https://www.latimes.com/environment
https://www.latimes.com/staff/megan-garvey	https://www.latimes.com/science
https://www.latimes.com/archives/la-xpm-2001-jun-10-mn-8792-story.html	https://www.latimes.com/california/story/2019-11-09/prop-187-anniversary-commemoration
https://www.latimes.com/entertainment-arts/books/story/2019-10-25/shadowlands-anthony-mccann-oregon-standoff	https://www.latimes.com/entertainment-arts/story/2019-11-09/classical-music-things-to-do-in-la-this-week-nov-10-17-magic-flute-la-opera
https://www.latimes.com/sports/highschool/story/2019-11-06/girls-tennis-southern-section-playoff-results-and-updated-pairings	https://www.latimes.com/entertainment-arts/story/2019-11-09/theater-things-to-do-in-la-this-week-nov-10-17-dr-seuss-how-the-grinch-stole-christmas-old-globe-key-largo-andy-garcia-geffen-playhouse
https://www.latimes.com/politics/story/2019-09-20/bernie-sanders-muslim-voters-2020	https://www.latimes.com/food/story/2019-11-07/kung-pao-chicken-history-recipe-gong-bao
https://www.latimes.com/sports/nba/la-sp-nba-best-game-ever-20181222-story.html	https://www.latimes.com/obituaries/story/2019-11-04/james-stern-obit-black-activist-who-led-neo-nazi-group-dies-amid-bid-to-destroy-it
https://www.latimes.com/sports/highschool/story/2019-11-06/girls-tennis-southern-section-playoff-results-and-updated-pairings	https://www.latimes.com/topic/column-one

https://www.latimes.com/entertainment/movies/la-et-mn-july-4-box-office-history-20180629-story.html	https://www.latimes.com/topic/mexico-americas
https://www.latimes.com/entertainment/movies/la-et-mn-july-4-box-office-history-20180629-story.html	https://www.latimes.com/topic/museums

Republicans: 1 Overlap

Lucene Results	PageRank Results
https://www.latimes.com/opinion/story/2019-10-23/house-republicans-storm-hearing-impeachment	https://www.latimes.com/
https://www.latimes.com/politics/story/2019-10-30/democrats-hoped-theyd-win-over-republicans-on-impeachment-but-its-not-looking-that-way-so-far	https://www.latimes.com/opinion
https://www.latimes.com/opinion/story/2019-11-08/impeachment-trump-republicans-toomey	https://www.latimes.com/opinion/submit-letter-to-the-editor
https://www.latimes.com/politics/story/2019-10-04/vulnerable-senate-republicans-impeachment	https://www.latimes.com/opinion/story/2019-11-09/facebook-twitter-political-ads-lies
https://www.latimes.com/california/story/2019-10-16/california-republicans-democrats-wildfires-homelessness	https://www.latimes.com/opinion/story/2019-11-08/berlin-wall-30th-anniversary-cold-war-donald-trump
https://www.latimes.com/politics/story/2019-10-23/impeachment-deposition-room-stormed-by-republicans	https://www.latimes.com/politics/story/2019-11-08/elizabeth-warren-used-bare-knuckle-tactics-to-take-down-an-obama-nominee
https://www.latimes.com/opinion/story/2019-09-17/california-republicans-trump-fundraising-president	https://www.latimes.com/opinion/story/2019-11-09/elizabeth-warren-trump
https://www.latimes.com/politics/story/2019-10-28/katie-hill-resignation-trump-obstacle-for-republicans	https://www.latimes.com/opinion/story/2019-11-09/devin-nunes-impeachment-witness-list
https://www.latimes.com/opinion/story/2019-11-06/trump-republicans-democrats-predictions-polling-2020	https://www.latimes.com/opinion/story/2019-11-09/sat-uc-college-admissions
https://www.latimes.com/politics/story/2019-11-06/republicans-election-trump-candidate-loses-kentucky	https://www.latimes.com/opinion/story/2019-11-08/impeachment-trump-republicans-toomey

Senate: 0 Overlaps

Lucene Results	PageRank Results
https://www.latimes.com/politics/story/2019-11-07/gop-is-already-thinking-about-how-to-turn-a-senate-impeachment-trial-to-trumps-advantage	https://www.latimes.com/
https://www.latimes.com/politics/story/2019-10-30/california-donors-spend-millions-on-2020-senate-races-across-the-country	https://www.latimes.com/opinion
https://www.latimes.com/politics/story/2019-10-04/vulnerable-senate-republicans-impeachment	https://www.latimes.com/california/story/2019-11-09/prop-187-anniversary-commemoration
https://www.latimes.com/politics/la-na-pol-kavanaugh-hearing-20180927-story.html	https://www.latimes.com/opinion/submit-letter-to-the-editor

https://www.latimes.com/nation/la-pol-scotus-confirmation-votes-over-the-years-20181005-htmlstory.html	https://www.latimes.com/opinion/story/2019-11-09/facebook-twitter-political-ads-lies
https://www.latimes.com/politics/story/2019-08-14/2020-senate-control-presidential-race	https://www.latimes.com/opinion/story/2019-11-08/berlin-wall-30th-anniversary-cold-war-donald-trump
https://www.latimes.com/politics/story/2019-08-14/john-hickenlooper-quits-presidential-race-for-senate-run	https://www.latimes.com/politics/story/2019-11-08/elizabeth-warren-used-bare-knuckle-tactics-to-take-down-an-obama-nominee
https://www.latimes.com/politics/story/2019-10-30/california-donors-spend-millions-on-2020-senate-races-across-the-country	https://www.latimes.com/opinion/story/2019-11-09/elizabeth-warren-trump
https://www.latimes.com/politics/la-na-pol-william-barr-senate-confirm-attorney-general-20190214-story.html	https://www.latimes.com/opinion/story/2019-11-09/devin-nunes-impeachment-witness-list
https://www.latimes.com/opinion/story/2019-10-15/wmcaleen-homeland-security-trump-nielsen-senate-acting-secretary	https://www.latimes.com/opinion/story/2019-11-09/sat-uc-college-admissions

Olympics 2020: 0 Overlaps

Lucene Results	PageRank Results
https://www.latimes.com/sports/olympics/story/2019-08-15/2020-tokyo-olympics-searching-for-answers-amid-heat-wave	https://www.latimes.com/
https://www.latimes.com/sports/olympics	https://www.latimes.com/sitemap
https://www.latimes.com/people/david-wharton	https://www.latimes.com/sports
https://www.latimes.com/people/david-wharton	https://www.latimes.com/opinion
https://www.latimes.com/sports/olympics/story/2019-10-16/2020-tokyo-olympics-marathon-sapporo-avoid-heat	https://www.latimes.com/business
https://www.latimes.com/sports/olympics/story/2019-07-31/deadly-heat-wave-2020-summer-olympics-tokyo	https://www.latimes.com/entertainment-arts
https://www.latimes.com/sports/story/2019-11-05/ioc-anti-doping-2020-tokyo-olympics	https://www.latimes.com/world-nation
https://www.latimes.com/sports/olympics/story/2019-10-10/naomi-osaka-chooses-japan-over-u-s-2020-tokyo-olympics	https://www.latimes.com/environment
https://www.latimes.com/sports/story/2019-07-23/tokyos-rough-road-to-2020-summer-olympics	https://www.latimes.com/newsroom-directory
https://www.latimes.com/people/david-wharton	https://www.latimes.com/politics

Stock: 1 Overlap

Lucene Results	PageRank Results
https://www.latimes.com/business/story/2019-11-06/uber-lock-up-period-ends-with-falling-stock-protests	https://www.latimes.com/
https://www.latimes.com/business/la-fi-lyft-stock-20190401-story.html	https://www.latimes.com/business

https://www.latimes.com/business/la-fi-hy-tesla-stock-20160518-snap-story.html	https://www.latimes.com/business/real-estate
https://www.latimes.com/business/story/2019-11-06/uber-lock-up-period-ends-with-falling-stock-protests	https://www.latimes.com/business/real-estate/story/2019-11-09/hulu-looking-for-alaska-landry-bender
https://www.latimes.com/business/story/2019-11-09/stock-managers-who-played-defense-in-2019-are-left-scrambling-to-make-up-ground	https://www.latimes.com/business/story/2019-11-08/trump-china-tariffs
https://www.latimes.com/business/la-fi-uber-ipo-stock-trading-price-20190510-story.html	https://www.latimes.com/business/real-estate/story/2019-11-09/hot-property-very-brady-sale-malibu
https://www.latimes.com/business/story/2019-09-09/at-t-stock-surges-elliott-reveals-3-2-billion-stake	https://www.latimes.com/politics/story/2019-11-08/elizabeth-warren-used-bare-knuckle-tactics-to-take-down-an-obama-nominee
https://www.latimes.com/business/la-fi-stock-market-status-update-20180209-story.html	https://www.latimes.com/business/real-estate/story/2019-11-09/hot-property-newsletter-rifleman-leads-the-posse-on-a-little-mosey-down-memory-lane
https://www.latimes.com/business/la-fi-stock-market-correction-20180212-story.html	https://www.latimes.com/business/story/2019-11-09/stock-managers-who-played-defense-in-2019-are-left-scrambling-to-make-up-ground
https://www.latimes.com/business/la-fi-stock-market-cpi-roundup-20180214-story.html	https://www.latimes.com/business/story/2019-11-08/gap-ceo-was-done-in-by-fashion-missteps-and-fading-brands

Virus: 1 Overlaps

Lucene Results	PageRank Results
https://www.latimes.com/espanol/internacional/articulo/2019-11-09/monos-se-volvieron-inmunes-tras-inocularles-virus-del-ebola-con-mutacion	https://www.latimes.com/terms-of-service
https://www.latimes.com/science/sciencenow/la-sci-science-of-zika-five-ways-20160223-htmllstory.html	https://www.latimes.com/espanol/
https://www.latimes.com/espanol/internacional/articulo/2019-11-09/monos-se-volvieron-inmunes-tras-inocularles-virus-del-ebola-con-mutacion	https://www.latimes.com/science
https://www.latimes.com/science/story/2019-10-21/virus-afm-illness-paralyzing-kids	https://www.latimes.com/espanol/internacional
https://www.latimes.com/science/story/2019-10-21/virus-afm-illness-paralyzing-kids	https://www.latimes.com/espanol/deportes/articulo/2019-11-09/como-ser-campeon-de-league-of-legends
https://www.latimes.com/socal/daily-pilot/news/tn-dpt-me-west-nile-20170707-story.html	https://www.latimes.com/business/story/2019-10-13/body-parts-harvesting-hinders-coroner-autopsies
https://www.latimes.com/science/sciencenow/la-sci-sn-acute-flaccid-myelitis-polio-20181017-story.html	https://www.latimes.com/espanol/mexico/articulo/2019-11-09/crean-en-mexico-plataforma-para-medir-impacto-ambiental-de-basura-electronica
https://www.latimes.com/local/lanow/la-me-ln-virulent-newcastle-disease-outbreak-in-southern-california-20190607-story.html	https://www.latimes.com/espanol/internacional/articulo/2019-11-09/el-nuevo-mitsubishi-mirage-pudiera-arribar-a-mediados-del-proximo-ano
https://www.latimes.com/world-nation/story/2019-09-04/anti-vaxxers-helping-polio-comeback-pakistan	https://www.latimes.com/espanol/internacional/articulo/2019-11-09/monos-se-volvieron-inmunes-tras-inocularles-virus-del-ebola-con-mutacion

<https://www.latimes.com/local/obituaries/la-me-deborah-asnis-20150921-story.html>

<https://www.latimes.com/espanol/internacional/articulo/2019-11-09/una-tortuga-tecnologica-para-detectar-microplasticos-en-los-oceanos>

Overlaps:

