

CSCI 572 Information Retrieval

Assignment 4

Paras Sibal

6121-6253-78

Steps followed to complete this assignment:

1. Since, I am using a mac OS, there is no need for me to install virtual box and ubuntu.
2. Setting up of the system
 - a. Downloaded Xampp 7.4.30 64 bit for the Apache server to run my php code.
 - b. Also, with Xampp, PHP version used is 7.4.30.
 - c. Java (already present on the system)
3. Downloaded and extracted the source html files for LATIMES from the google drive and placed them in the /server/solr directory.
4. **Installation of Solr**
 - a. Downloaded Solr version 7.7.3.
 - b. Opened Terminal and changed the current directory to the solr-7.7.3 and executed bin/solr start to start the solr
 - c. Created the core on solr named, “**hw4**”, using the command **bin/solr create -c hw4**
 - d. Modified the “managed-schema” file under the conf folder to populate the _text_ field.
 - e. Modified Solr so as to query the default from _text_ by changing the df field inside the solrconfig.xml file.
 - f. Ran the command **bin/post -c hw4 -filetypes html**
/Users/paras/Desktop/solr-7.7.3/server/solr/LATIMES/latimes/ to index the html files. It took 10 minutes to parse 19193 documents and the index entries for 19193 documents were created.
 - g. Checked the Solr UI and see if the files have gotten indexed. Open the browser and go to **http://localhost:8983/solr/**. Selected the core “**hw4**” from the dropdown.

The screenshot displays the Solr Admin UI for the core 'hw4'. The left sidebar contains navigation links: Dashboard, Logging, Core Admin, Java Properties, Thread Dump, Overview (selected), Analysis, Dataimport, Documents, Files, Ping, Plugins / Stats, Query, Replication, Schema, and Segments info. The main content area is divided into three sections: Statistics, Instance, and Replication (Master). The Statistics section shows: Last Modified: 7 minutes ago, Num Docs: 19193, Max Doc: 19193, Heap Memory: -1, Usage: Deleted Docs: 0, Version: 248, Segment Count: 14, and Current: (green checkmark). The Instance section shows: CWD: /Users/paras/Desktop/solr-7.7.3/server, Instance: /Users/paras/Desktop/solr-7.7.3/server/solr/hw4, Data: /Users/paras/Desktop/solr-7.7.3/server/solr/hw4/data, Index: /Users/paras/Desktop/solr-7.7.3/server/solr/hw4/data/index, and Impl: org.apache.solr.core.NRTCachingDirectoryFactory. The Replication (Master) section shows a table with columns Version, Gen, and Size. The table has two rows: Master (Searching) with Version 1668656907508, Gen 40, and Size 290.05 MB; and Master (Replicable) with Version 1668656907508, Gen 40, and Size -. A Healthcheck section at the bottom right states: Ping request handler is not configured with a healthcheck file. At the bottom of the page, there are links for Documentation, Issue Tracker, IRC Channel, Community forum, and Solr Query Syntax.

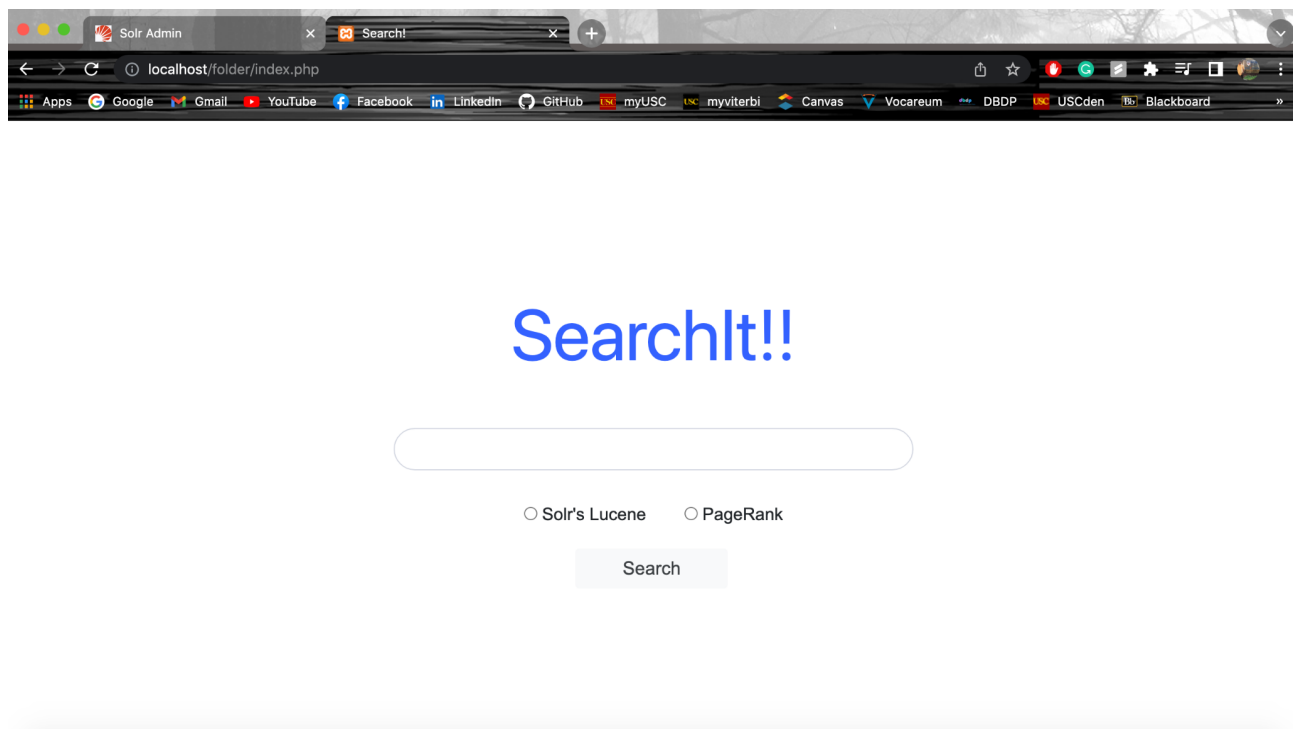
Version	Gen	Size
1668656907508	40	290.05 MB
1668656907508	40	-

5. Creation of EdgeList and PageRank

- Use the **Jsoup library** and **ExternalLinks.java file**(java program) to get the outgoing links and created the EdgeList text file, "**edgelist.txt**".
- Installed NetworkX, which is a python library to get graphs made out of the edge list. The file "**external_pageRankFile.txt**" was created using the python code written with the help of "**NetworkX**" library to get the scores of the document_ids.
- Page Rank was calculated with the following configuration:
alpha=0.85, personalization=None, max-iter=30, tol=10⁻⁶, nstart=None, weight="weights", dangling=None
- The "**external_pageRankFile.txt**" was placed under the data folder present inside my solr core folder.
- Modified the "managed-schema" file by adding in the following field entry.
<fieldType name="external" keyField="id" defVal="0" class="solr.ExternalFileField"/>
<field name="pageRankFile" type="external" stored="false" indexed="false"/>
- Modified the "solrconfig.xml" file by adding in the Listeners to reload the external file when either a searcher is loaded or when a new searcher is started.
<listener event="newSearcher" class="org.apache.solr.schema.ExternalFileFieldReloader"/>
<listener event="firstSearcher" class="org.apache.solr.schema.ExternalFileFieldReloader"/>
- Finally, reloaded the solr core by using the reload button present on Solr UI interface

6. Website index.php

- Designed a solr search engine interface using **solr-php-client, PHP, HTML and CSS**



7. Query Search on Solr UI

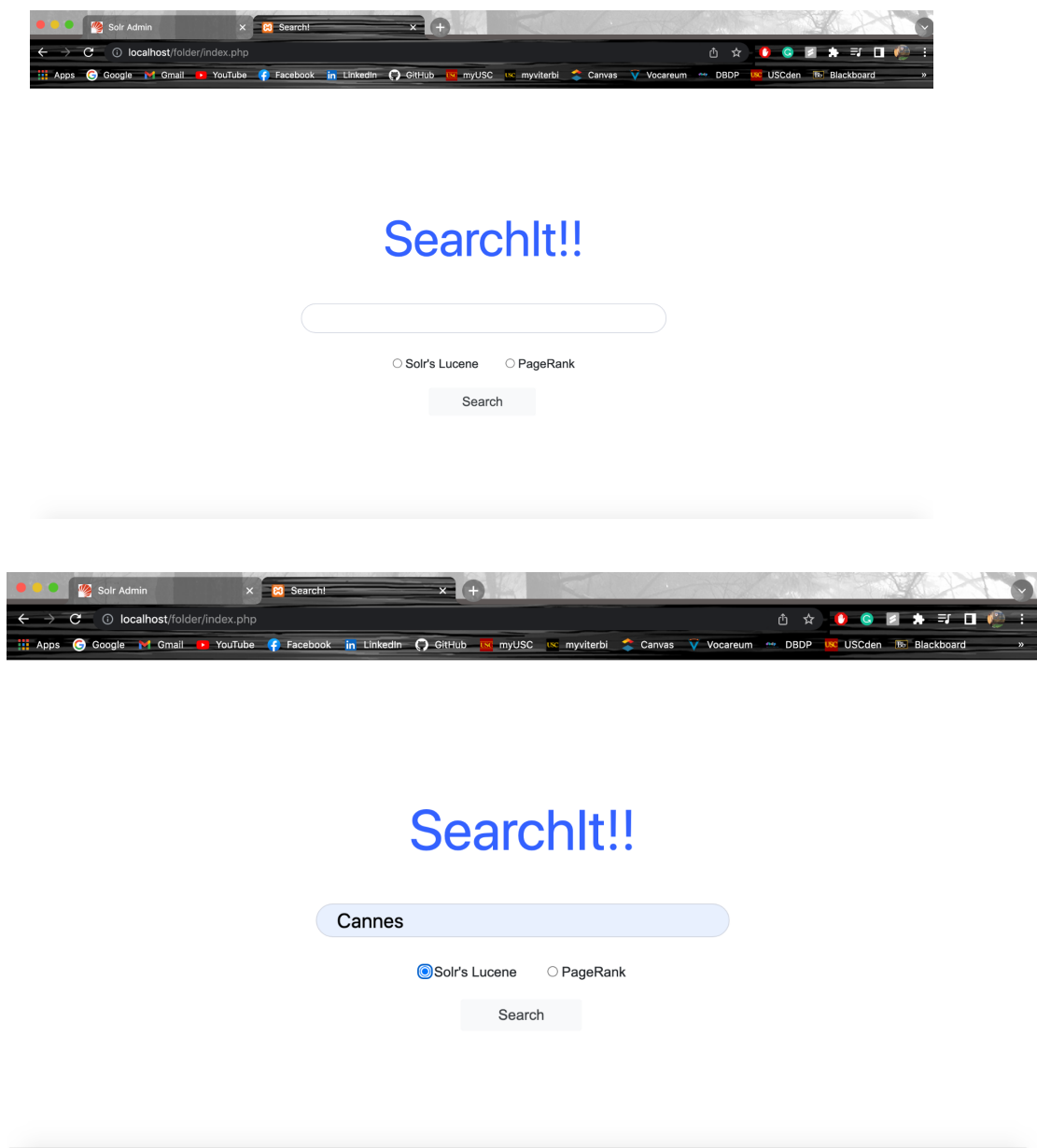
[illegible][illegible]

Explanation why some pages have higher PageRank values (1 para)

PageRank is used to rank web pages and analyze which web pages are more important than the others. It is mainly based on the number of external links that point to the webpage. The more in-links a webpage has, the higher the PageRank and the more important the web page is. Thus, any web page with a larger number of external links that point to that page will have a higher PageRank value. Due to these reasons, some pages have higher PageRank and lucene algorithm shows more relevant pages compared to the PageRank algorithm. Also, we can see that most of the links are irrelevant to the query executed but since they have a higher PageRank, they are being displayed.

Document the flow of any single query using screenshots.

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



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

Search!!!

Search Query: Cannes Algorithm: lucene

Showing 10 out of 144 results

<https://www.latimes.com/topic/cannes-film-festival>
Cannes Film Festival - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$70006ea-4463-474b-8982-01c1d84c96dd.html
N/A

<https://www.latimes.com/entertainment/movies/la-et-mn-cannes-wsnp-critics-conversation-20190528-story.html>
Cannes: Times critics on everything you need to know about this year's festival - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$26ca05f5-0756-432a-e059-646aa22054a2.html
The 72nd Festival de Cannes featured new films by Quentin Tarantino and Terrence Malick, opened with Jim Jarmusch's "The Dead Don't Die" and ended with the Palme d'Or going to "Parasite," a thriller directed by South Korean filmmaker Bong Joon-ho.

<https://www.latimes.com/entertainment/movies/la-et-mn-cannes-2015-winners-20190524-story.html>
Cannes 2015: The full list of winners - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$0340307-f1ca-4100-b001-dabdd09814.html
Director Jacques Audiard's film "Dheepan," a drama about three Tamil exiles trying to reconstruct their lives in France, has won the top prize at the Cannes Film Festival.

<https://www.latimes.com/entertainment/movies/la-et-mn-cannes-film-festival-award-winners-20190525-story.html>
Bong Joon-ho's 'Parasite' and Mati Diop's 'Atlantics' make history at Cannes Film Festival - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$4b1801c0-6891-4878-b308-15a67c3a.html
History was made and history repeated itself, albeit with a wicked twist, when Bong Joon-ho's "Parasite" won the Festival de Cannes' Palme d'Or on Saturday night.

<https://www.latimes.com/entertainment/movies/la-et-mn-cannes-parasite-atlantico-awards-20190526-story.html>
One thing the Cannes Film Festival jury got right: Bong Joon-ho's 'Parasite' victory - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$4c20a05-0911-b6c7-946a-2a71ed7c0e42.html
Times critic Justin Chang is filing dispatches from the 72nd Cannes Film Festival, which ran May 14-25 in France.

<https://www.latimes.com/entertainment/movies/la-et-mn-cannes-once-upon-a-time-in-hollywood-20190522-story.html>
Quentin Tarantino gets testy at Cannes over question about women's roles in his films - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$4348a12-ca31-4106-959-6539d9a75e61.html
There were plenty of hard questions Quentin Tarantino could have faced as he greeted reporters for the first time following the Cannes premiere of "Once Upon a Time ... in Hollywood."

<https://www.latimes.com/entertainment/movies/la-ca-mn-cannes-genre-parasite-wild-goose-lake-tarantino-20190522-story.html>
At Cannes, genre filmmaking from 'Parasite' to 'The Wild Goose Lake' turns heads - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$7536e-e381-4f6d-bad7-d6b75ea30761.html
Times critic Justin Chang is filing regular dispatches from the 72nd annual Cannes Film Festival, which runs through Saturday in France.

<https://www.latimes.com/entertainment/movies/la-et-mn-cannes-mekoub-my-love-intermezzo-abdelhalil-kechiche-20190523-story.html>
Review: Abdelhalil Kechiche's 'Mekoub, My Love: Intermezzo' is the disaster of the Cannes Film Festival - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$fd8562-9a2c-4ab1-9c39-0702f6ca3b3a.html
Times critic Justin Chang is filing regular dispatches from the 72nd Cannes Film Festival, which runs May 14-25 in France.

<https://www.latimes.com/entertainment/movies/la-et-mn-cannes-quentin-tarantino-once-upon-a-time-in-hollywood-20190521-story.html>
Quentin Tarantino shocks Cannes again with epic 'Once Upon a Time ... in Hollywood' - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$2420501-6c025-4a8f-b0b7-702f6ca07fcd.html
Times critic Justin Chang is filing regular dispatches from the 72nd annual Cannes Film Festival, which runs May 14-25 in France.

<https://www.latimes.com/entertainment/movies/la-et-mn-cannes-robert-eggers-the-lighthouse-robert-pattinson-20190524-story.html>
Robert Eggers, director of the smart and scary 'The Lighthouse,' sees comfort in horror - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$02b0a5e-e08e-456f-ee6f-0326812383a1.html
"The Witch" director Robert Eggers' serie new movie "The Lighthouse," with Robert Pattinson and Willem Dafoe, caused a sensation at the Cannes Film Festival.

⇒ Screenshot of page which shows the results for PageRank.

Search!!!

Search Query: Cannes Algorithm: pagerank

Showing 10 out of 144 results

<https://www.latimes.com/sitemap>
HTML Sitemap - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$99ba09-d8d1-43c2-bc9b-4f85eacae077.html
N/A

<https://www.latimes.com/entertainment/arts/story/2019-10-04/por-fin-joker-ajul-un-resumen-de-todo-el-drama-que-ha-provocado>
Por fin llega 'Joker'. Aquí un resumen de todo el drama que ha provocado - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$102893c-2b92-4757-9db5-72c48b76c54a.html
"Joker", la nueva historia de origen villano de Warner Bros. protagonizada por Joaquin Phoenix, ha suscitado mucho interés y controversia.

<https://www.latimes.com/entertainment-arts/movies/story/2019-11-08/documentaries-apollo-11-the-cave-sea-of-shadows-honeyland>
Four documentaries to watch as the race to the Oscars begins - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$8311c796-626a-4763-9120-77a0307aa702.html
"The Cave," "Apollo 11," "Honeyland" and "Sea of Shadows" explore vastly different worlds with equal intensity

<https://www.latimes.com/lifestyle/gallery/chole-sevigny-five-favorite-frocks>
Chloë Sevigny's five favorite frocks - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$9739033-80ca-40cf-ba43-94f2c0c04a07.html
Actor-director Chloë Sevigny recaps some of her most memorable red-carpet moments.

<https://www.latimes.com/topic/film-festivals>
Film Festivals - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$d1198af-5e51-4edf-af23-054a9f402616.html
Los Angeles Times coverage of film festivals, including Toronto, Telluride, Sundance, Cannes and more.

<https://www.latimes.com/topic/cannes-film-festival>
Cannes Film Festival - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$70006ea-4463-474b-8982-01c1d84c96dd.html
N/A

<https://www.latimes.com/people/gemara-molina>
Gemara Molina - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$d3e4894-711d-420a-8ea4-c0505bc907c1.html
N/A

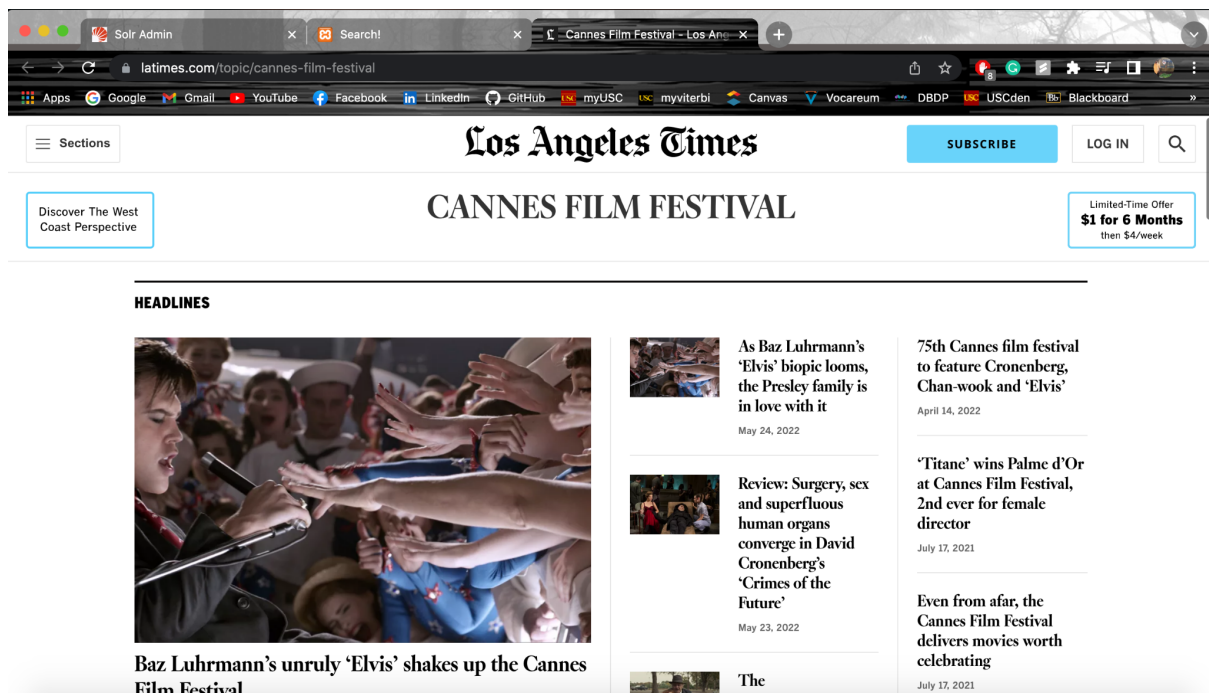
<https://www.latimes.com/entertainment-arts/music/story/2019-11-06/whitney-houston-lesbian-relationship-robyn-crawford>
Whitney Houston's best friend reveals their lesbian affair - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$02a32af-14a1-4a81-a602-e4c0699d6a6d.html
Robyn Crawford, Whitney Houston's best friend for years, says in a new memoir that they had a physical relationship — until the late singer got a record deal.

<https://www.latimes.com/entertainment-arts/movies/story/2019-10-31/apple-disney-hbo-max-streaming-movies>
As Apple, Disney and HBO start streaming, how will movies change? - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$55395e-a87a-418f-aa62-bca7d8b1c303.html
As new players like Apple, Disney and HBO bring their firepower to the streaming wars, many are wondering what implications will be for movies.

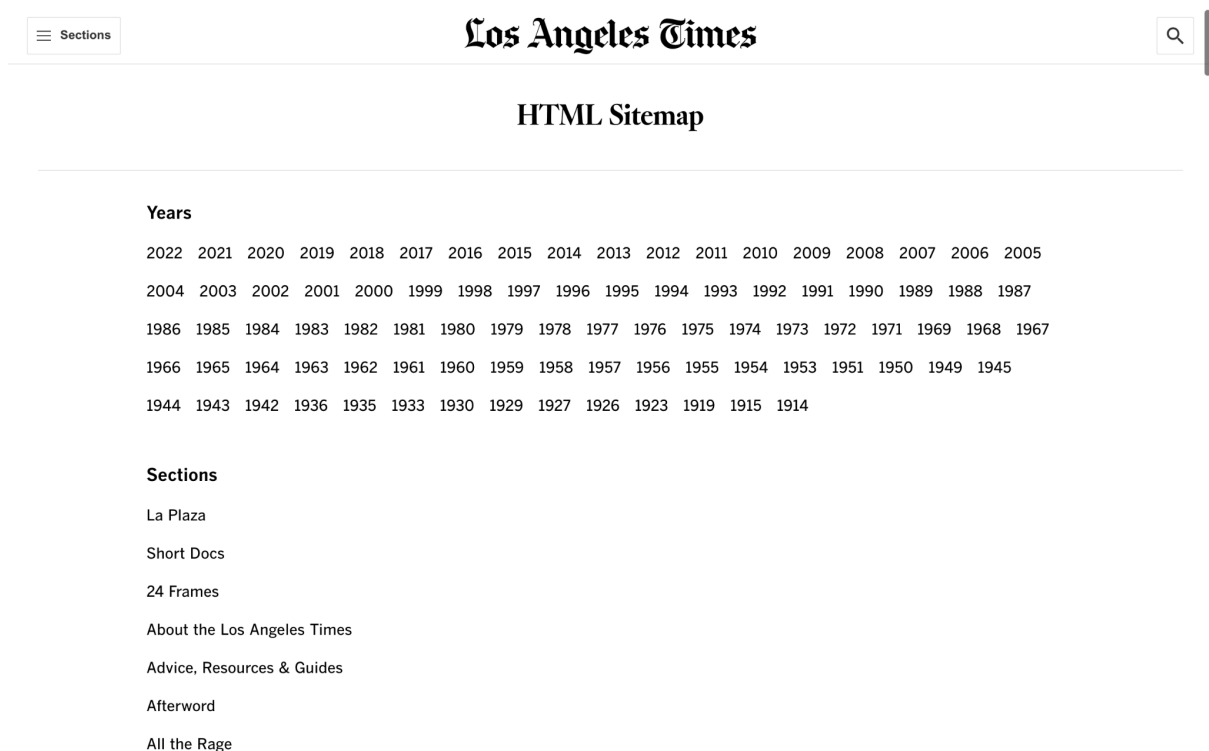
<https://www.latimes.com/topics/story/2019-09-17/patt-morrison-dana-thomas-fast-fashion-environment>
Column: Fast fashion is hurting the planet, and you can stop it - Los Angeles Times
/Users/parsa/Desktop/koib-7.7.3/serve/koib/LATIME\$latimes\$46460c76-e48f-4374-96d5-7449270502f1.html
Did you know it takes 5,000 gallons of water to make a T-shirt and jeans? Patt Morrison talks with Dana Thomas about her book "Fashionopolis."

⇒ Screenshot of page clicking on one of the result's link.

Cannes Lucene First Link:



Cannes Page Rank first link:



Top 10 Links to queries from both the ranking methods:

1. Cannes

Lucene	PageRank
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-film-festival-award-winners-20190525-story.html	https://www.latimes.com/lifestyle/gallery/chloe-sevignys-five-favorite-frocks
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-parasite-atlantics-awards-20190526-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-goose-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

2. Congress

Lucene	PageRank
https://www.latimes.com/opinion/enterthefray/la-ol-new-york-rhode-island-congress-marijuana-20190116-story.html	https://www.latimes.com/opinion

https://www.latimes.com/politics/story/2019-10-28/jerry-brown-to-testify-to-congress-rebutting-trumps-criticism-of-california	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

3. Democrats

Lucene	PageRank
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/2019-11-03/iowa-democrats-candidate-beat-trump	https://www.latimes.com/politics/story/2019-11-03/iowa-democrats-candidate-beat-trump

0-17/democrats-quick-impeachment-timing-complicated	1-08/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

4. Patriot Movement

Lucene	PageRank
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

5. Republicans

Lucene	PageRank
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/la-letter-to-the-editor-htmlstory.html
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

6. Senate

Lucene	PageRank
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/la-letter-to-the-editor-htmlstory.html
https://www.latimes.com/nation/la-pol-sctus-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
---	---

7. Olympics 2020

Lucene	PageRank
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

8. Stock

Lucene	PageRank
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-tesa-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-landr-y-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-ip-o-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/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/gap-ceo-was-done-in-by-fashion-missteps-and-fading-brands

9. Virus

Lucene	PageRank
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-htmlstory.html	https://www.latimes.com/espanol/
https://www.latimes.com/espanol/internacional	https://www.latimes.com/science

l/articulo/2019-11-09/monos-se-volvieron-inmunes-tras-inocularles-virus-del-ebola-con-mutacion	
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

The overlaps URLs have been highlighted in blue in the tables present above

Query Results Overlap between Lucene and PageRank

Query	Overlaps
Cannes	1
Congress	0
Democrats	0
Patriot Movement	0
Republicans	1
Senate	0
Olympics 2020	0
Stock	2
Virus	2

The graph for overlap is shown below.

