

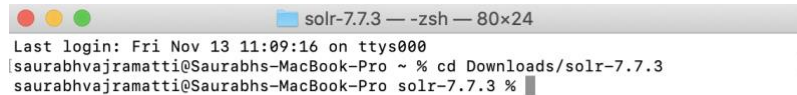
REPORT

Name: Saurabh Vajramatti

USC ID: 5761344286

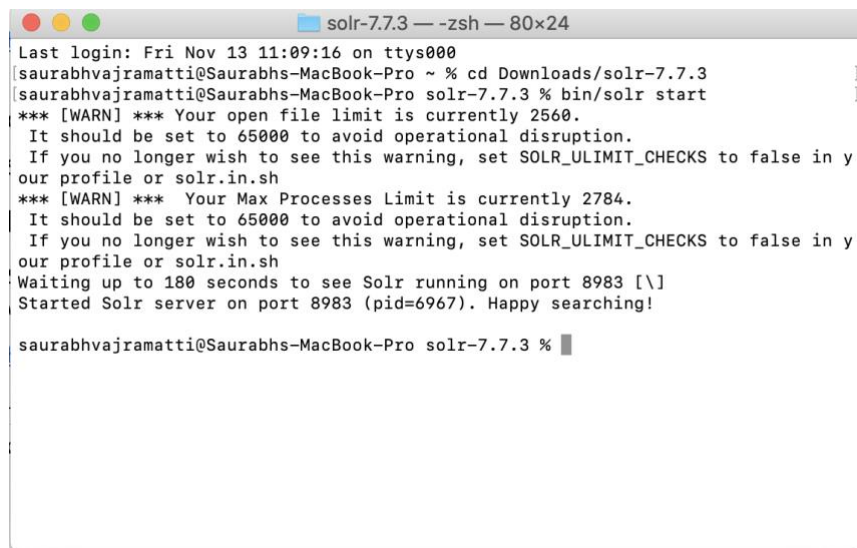
1. Downloaded the solr-7.7.3. Extracted the compressed file.
2. Moved to directory /Users/saurabhvajramatti/Downloads/solr-7.7.3 by using the command:

`cd Users/saurabhvajramatti/Downloads/solr-7.7.3`

A terminal window titled 'solr-7.7.3 — -zsh — 80x24'. The prompt is 'Last login: Fri Nov 13 11:09:16 on ttys000'. The user is 'saurabhvajramatti@Saurabhs-MacBook-Pro'. The command 'cd Downloads/solr-7.7.3' is entered and executed. The prompt changes to 'saurabhvajramatti@Saurabhs-MacBook-Pro solr-7.7.3 %'.

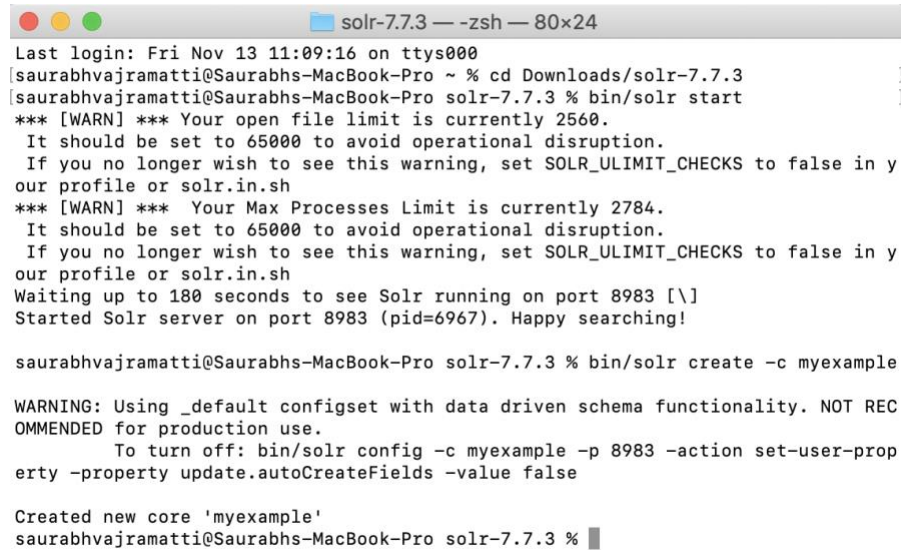
3. Started the Solr server by running this command:

`bin/solr start`

A terminal window titled 'solr-7.7.3 — -zsh — 80x24'. The prompt is 'Last login: Fri Nov 13 11:09:16 on ttys000'. The user is 'saurabhvajramatti@Saurabhs-MacBook-Pro'. The command 'cd Downloads/solr-7.7.3' is entered and executed. The prompt changes to 'saurabhvajramatti@Saurabhs-MacBook-Pro solr-7.7.3 %'. The command 'bin/solr start' is entered and executed. The output shows warnings about open file limit and max processes limit, followed by 'Waiting up to 180 seconds to see Solr running on port 8983 [\]' and 'Started Solr server on port 8983 (pid=6967). Happy searching!'. The prompt changes to 'saurabhvajramatti@Saurabhs-MacBook-Pro solr-7.7.3 %'.

4. Created a new core by running command:

```
bin/solr create -c myexample
```

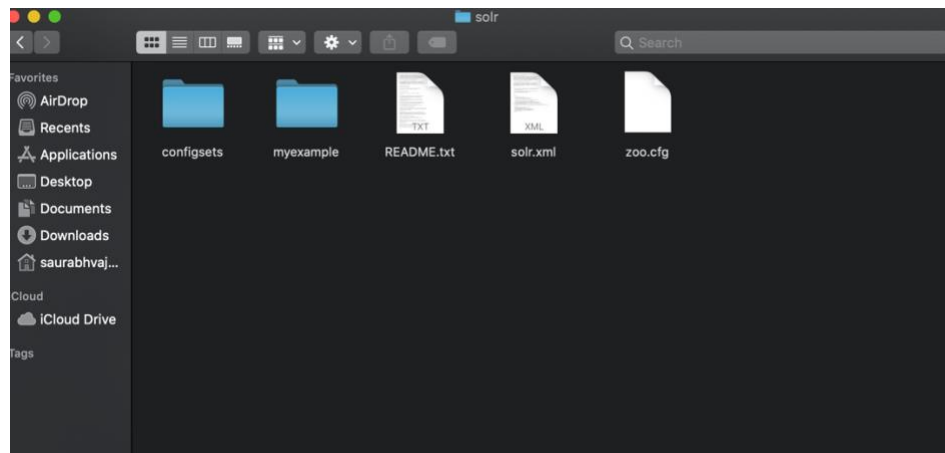


```
solr-7.7.3 — zsh — 80x24
Last login: Fri Nov 13 11:09:16 on ttys000
[saurabhvajramatti@Saurabhs-MacBook-Pro ~ % cd Downloads/solr-7.7.3
[saurabhvajramatti@Saurabhs-MacBook-Pro solr-7.7.3 % bin/solr start
*** [WARN] *** Your open file limit is currently 2560.
It should be set to 65000 to avoid operational disruption.
If you no longer wish to see this warning, set SOLR_ULIMIT_CHECKS to false in y
our profile or solr.in.sh
*** [WARN] *** Your Max Processes Limit is currently 2784.
It should be set to 65000 to avoid operational disruption.
If you no longer wish to see this warning, set SOLR_ULIMIT_CHECKS to false in y
our profile or solr.in.sh
Waiting up to 180 seconds to see Solr running on port 8983 []
Started Solr server on port 8983 (pid=6967). Happy searching!

saurabhvajramatti@Saurabhs-MacBook-Pro solr-7.7.3 % bin/solr create -c myexample

WARNING: Using _default configset with data driven schema functionality. NOT REC
OMMENDED for production use.
To turn off: bin/solr config -c myexample -p 8983 -action set-user-prop
erty -property update.autoCreateFields -value false

Created new core 'myexample'
saurabhvajramatti@Saurabhs-MacBook-Pro solr-7.7.3 %
```



5. Made changes in managed_schema file by Uncommenting the line - <!-- <copyField source="*" dest="_text_"/> -->, so that it looks like <copyField source="*" dest="_text_"/>

```
conf — vi managed-schema — 80x24
<!-- docValues are enabled by default for long type so we don't need to index
the version field -->
<field name="_version_" type="plong" indexed="false" stored="false"/>
<field name="_root_" type="string" indexed="true" stored="false" docValues="
false" />
<field name="_text_" type="text_general" indexed="true" stored="false" multi
Valued="true"/>

<!-- This can be enabled, in case the client does not know what fields may be
e searched. It isn't enabled by default
because it's very expensive to index everything twice. -->
<copyField source="*" dest="_text_"/>

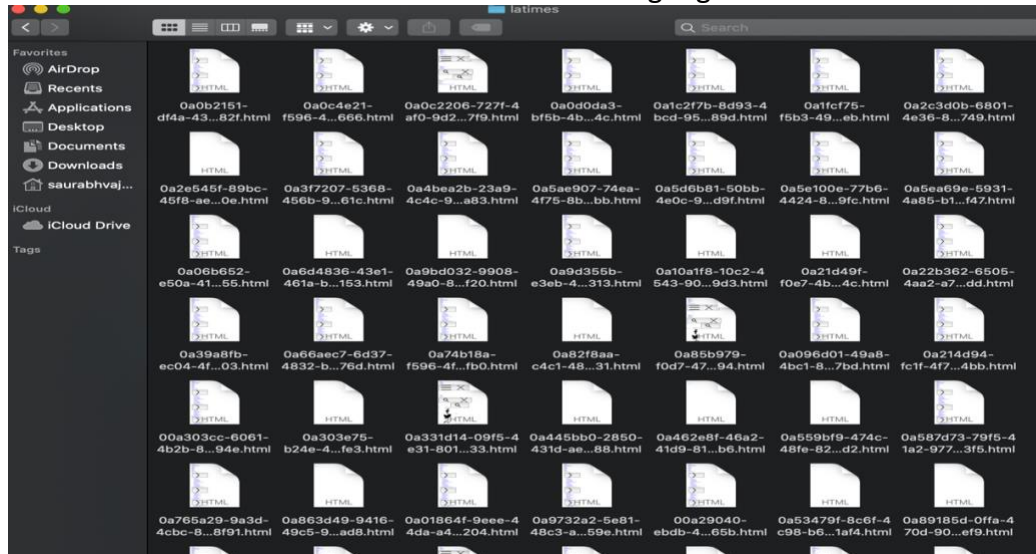
<!-- Dynamic field definitions allow using convention over configuration
for fields via the specification of patterns to match field names.
EXAMPLE: name="*_i" will match any field ending in _i (like myid_i, z_i)
RESTRICTION: the glob-like pattern in the name attribute must have a "*"
only at the start or the end. -->

<dynamicField name="*_i" type="pint" indexed="true" stored="true"/>
<dynamicField name="*_is" type="pints" indexed="true" stored="true"/>
<dynamicField name="*_s" type="string" indexed="true" stored="true" />
<dynamicField name="*_ss" type="strings" indexed="true" stored="true"/>
-- INSERT --
```

6. Made changes in solrconfig.xml file by uncommenting a line and changing text to _text_ to get this: <str name="df">_text_</str>

```
solrconfig.xml
Welcome  solrconfig.xml x
Users > saurabhvajramatti > Downloads > solr-7.7.3 > server > solr > myexample > conf > solrconfig.xml
675 | not be initialized until the first request that uses it.
676 |
677 | -->
678 | <!-- SearchHandler
679 |
680 | http://wiki.apache.org/solr/SearchHandler
681 |
682 | For processing Search Queries, the primary Request Handler
683 | provided with Solr is "SearchHandler" It delegates to a sequent
684 | of SearchComponents (see below) and supports distributed
685 | queries across multiple shards
686 | -->
687 | <requestHandler name="/select" class="solr.SearchHandler">
688 | <!-- default values for query parameters can be specified, these
689 | will be overridden by parameters in the request
690 | -->
691 | <lst name="defaults">
692 | <str name="echoParams">explicit</str>
693 | <int name="rows">10</int>
694 | <!-- Default search field -->
695 | <str name="df">_text_</str>
696 | <!-- Change from JSON to XML format (the default prior to Solr 7.0)
697 | <str name="wt">xml</str>
698 | -->
699 | </lst>
700 | <!-- In addition to defaults, "appends" params can be specified
701 | to identify values which should be appended to the list of
702 | multi-val params from the query (or the existing "defaults").
703 | -->
704 | <!-- In this example, the param "fq=instock:true" would be appended to
705 | any query time fq params the user may specify, as a mechanism for
706 | partitioning the index, independent of any user selected filtering
707 | that may also be desired (perhaps as a result of faceted searching).
708 |
709 | NOTE: there is *absolutely* nothing a client can do to prevent these
710 | "appends" values from being used, so don't use this mechanism
711 | unless you are sure you always want it.
712 | -->
```

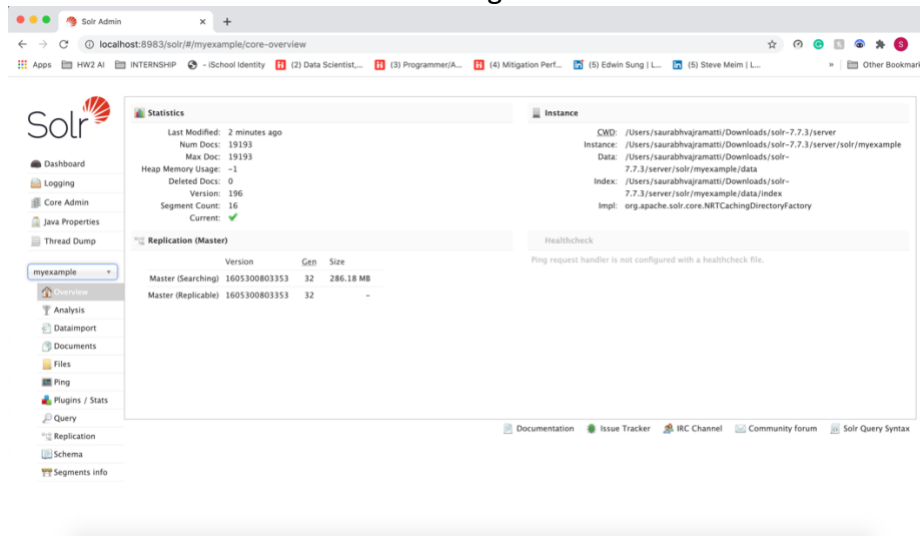
7. Downloaded the documents in LATIMES folder of the google drive.



8. Performed indexing of only HTML files present in the folder latimes. I performed the indexing to my core 'myexample'. Used the command:
bin/post -c myexample -filetypes html /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes

```
solr-7.7.3 — -zsh — 80x24
POSTing file 5d642fcf-a9d3-4448-bb36-d67a5556b78f.html (text/html) to [base]/ext
ract
POSTing file f0c2fe36-1dc0-420b-81c6-8bdfbf94f542.html (text/html) to [base]/ext
ract
POSTing file 6b363ac7-6879-452e-b705-c3d21c435551.html (text/html) to [base]/ext
ract
POSTing file 00eaa0a8-67b9-4503-8ab3-77ac94f53656.html (text/html) to [base]/ext
ract
POSTing file cd73f4f7-7c28-4a62-81d1-4c711c564dd8.html (text/html) to [base]/ext
ract
POSTing file be77472a-cddf-4964-8294-4bf480845f52.html (text/html) to [base]/ext
ract
POSTing file fdb09540-56ff-4ebb-9ac4-4e6c831ca557.html (text/html) to [base]/ext
ract
POSTing file 69ee8a9e-3f32-40ba-8872-9d2a448ff2ec.html (text/html) to [base]/ext
ract
POSTing file 762c6313-2bb4-4622-baa5-74f2d8ed6ba1.html (text/html) to [base]/ext
ract
POSTing file b26eb8f3-d3da-42f3-b2d2-e4de0c799288.html (text/html) to [base]/ext
ract
19193 files indexed.
COMMITting Solr index changes to http://localhost:8983/solr/myexample/update...
Time spent: 0:07:36.260
saurabhvajramatti@Saurabhs-MacBook-Pro solr-7.7.3 %
```

9. Had a look at the core statistics after reloading the core.



10. Now we go on to create the edgelist file which will contain information on which pages have a link between them. I wrote a code in Java using Jsoup library. I used Eclipse as the development IDE.

11. Next we need to create the external_pageRankFile which will contain the scores of the webpages based on the number of links (in-links to be specific) associated with a webpage which represents a node in the graph that results from the edges in edgelist file.

- a. Took the edgelist.txt file as input, which contains the information about all the edges which represent links between webpages.
- b. The networkx library to generate the graph from the information in the edgelist.txt file.
- c. Made use of the function read_edgelist with "edgelist.txt" and create_using=nx.DiGraph() as its parameters.
- d. Next I called the function pagerank with an initial configuration as follows:
 - i. alpha=0.85 :the damping parameter
 - ii. personalization=None :Uniform distribution is used
 - iii. max_iter=30 : It signifies the maximum number of iterations
 - iv. tol=1e-06 : It is the error tolerance
 - v. nstart=None : Starting value for page rank iteration
 - vi. weight='weight' : data to use as weight
 - vii. dangling=None : Used to specify nodes which do not have any outgoing edges.
- e. The final resulting external_pageRankFile will have data in the format page_id=score.
- f. So the basic idea is that we are considering the webpages as the nodes and the links between them as edges. We first identify the network and generate the edgelist file which will contain all the edges. Next we use this file as input to generate the graph which will be used to compute the PageRank. So the score

(page rank) of a webpage (graph node) depends on the number of in-links along with the importance of the nodes pointing to it. Then we use it to generate the PageRank file.

12. Pasted the external_pageRankFile.txt in my code file
13. Made changes to managed_schema: `<fieldType name="external" keyField="id" defVal="0" class="solr.ExternalFileField"/>`
`<field name="pageRankFile" type="external" stored="false" indexed="false"/>`
14. Added these lines to solrconfig.xml: `<listener event="newSearcher" class="org.apache.solr.schema.ExternalFieldReloader">`
`<listener event="firstSearcher" class="org.apache.solr.schema.ExternalFieldReloader">`
15. Reloaded the solr

The screenshot shows the Solr Admin web interface in a browser. The URL is `localhost:8983/solr/#/myexample/core-overview`. The interface includes a sidebar with 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: 2 minutes ago, Num Docs: 19193, Max Doc: 19193, Heap Memory Usage: -1, Deleted Docs: 0, Version: 196, Segment Count: 16, and Current: ✓. The Instance section shows: CWD: /Users/saurabhvajramatti/Downloads/solr-7.7.3/server, Instance: /Users/saurabhvajramatti/Downloads/solr-7.7.3/server/solr/myexample, Data: /Users/saurabhvajramatti/Downloads/solr-7.7.3/server/solr/myexample/data, Index: /Users/saurabhvajramatti/Downloads/solr-7.7.3/server/solr/myexample/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 1605300803353, Gen 32, and Size 286.18 MB; and Master (Replicable) with Version 1605300803353, Gen 32, and Size -. The Healthcheck section shows: Ping request handler is not configured with a healthcheck file. At the bottom, there are links for Documentation, Issue Tracker, IRC Channel, Community forum, and Solr Query Syntax.

	Version	Gen	Size
Master (Searching)	1605300803353	32	286.18 MB
Master (Replicable)	1605300803353	32	-

16. Got results for sample query “Cannes” by default and PageRank method

The screenshot shows the Solr Admin web interface. On the left is a sidebar with navigation links like Dashboard, Logging, Core Admin, etc. The main area is divided into two panels. The left panel contains form fields for a query: Request-Handler (qt) is set to /select, q is 'Cannes', fq is empty, sort is 'pageRankFile desc', start is 0, and rows is 10. The right panel displays the JSON response from the query, showing a list of document IDs and their corresponding page rank values.

This screenshot shows the Solr Admin interface with a different query. The q field is still 'Cannes', but the sort field is now 'key1'. The JSON response on the right shows a different set of document IDs, reflecting the change in sorting criteria.

Cannes-Overlap 0

Lucene	Custom Page Rank
https://www.latimes.com/topic/cannes-film-festival	https://www.latimes.com/entertainment-arts/movies/story/2019-09-13/toronto-film-festival-best-of-wrap-parasite
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-wrap-critics-conversation-20190526-story.html	https://www.latimes.com/entertainment-arts/story/2019-10-01/la-et-mn-cannes-wrap-critics-conversation-20190526-story.html

	mg-prince-harry-meghan-suing-mail-associated-newspapers
https://www.latimes.com/entertainment/movies/moviesnow/la-et-mn-cannes-2015-winners-20150524-story.html	https://www.latimes.com/entertainment-arts/movies/story/2019-09-09/jojo-rabbit-knives-out-toronto-film-festival
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-film-festival-award-winners-20190525-story.html	https://www.latimes.com/entertainment/movies/la-ca-mn-cannes-taron-egerton-rocketman-20190523-story.html
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-parasite-atlantics-awards-20190526-story.html	https://www.latimes.com/books/jacketcopy/la-et-jc-carrie-fisher-book-sales-20161228-story.html
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-once-upon-a-time-in-hollywood-20190522-story.html	https://www.latimes.com/local/lanow/la-me-ln-hadid-sentencing-20170720-story.html
https://www.latimes.com/entertainment/movies/la-ca-mn-cannes-genre-parasite-wild-goose-lake-tarantino-20190522-story.html	https://www.latimes.com/sitemap/2013/2
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-mektoub-my-love-intermezzo-abdellatif-kechiche-20190523-story.html	https://www.latimes.com/sitemap/2017/1
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/la-et-carrie-fisher-over-the-edge-and-back-19870731-story.html
https://www.latimes.com/entertainment/movies/la-et-mn-cannes-robert-eggers-the-lighthouse-robert-pattinson-20190524-story.html	https://www.latimes.com/people/gennaro-molina

Congress- Overlap 1 (this was repeated: <https://www.latimes.com/politics/la-pol-ca-richest-in-congress-darrell-issa-story.html>)

Lucene	Custom Page Rank
https://www.latimes.com/opinion/enterthefray/la-ol-new-york-rhode-island-congress-marijuana-20190116-story.html	https://www.latimes.com/local/lanow/la-me-ln-essential-california-20161114-story.html
https://www.latimes.com/politics/story/2019-10-28/jerry-brown-to-testify-to-congress-rebutting-trumps-criticism-of-california	https://www.latimes.com/california/story/2019-10-01/asylum-9thcircuit-trump

https://www.latimes.com/politics/la-na-pol-congress-sexual-harassment-20181212-story.html	https://www.latimes.com/travel/story/2019-10-25/can-you-carry-cbd-plane-across-state-lines
https://www.latimes.com/politics/la-pol-ca-richest-california-lawmakers-20180305-story.html	https://www.latimes.com/politics/la-pol-ca-richest-in-congress-darrell-issa-story.html
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-healthcare-plan-too-liberal-to-win
https://www.latimes.com/opinion/story/2019-10-16/democratic-debate-trump-congress-executive-action	https://www.latimes.com/politics/la-na-pol-2020-democratic-presidential-candidates-2019-story.html
https://www.latimes.com/politics/story/2019-10-14/trump-ukraine-aid-congress-impeachment	https://www.latimes.com/sitemap/1993/4
https://www.latimes.com/politics/la-na-pol-congress-bioweapons-detection-system-20190414-story.html	https://www.latimes.com/politics/story/2019-10-17/gordon-sondland-ukraine-house-impeachment-testimony
https://www.latimes.com/politics/la-na-pol-congress-mueller-pressure-impeachment-20190529-story.html	https://www.latimes.com/politics/story/2019-09-25/trump-impeachment-nancy-pelosi-ukraine-biden-call
https://www.latimes.com/politics/la-pol-ca-richest-in-congress-darrell-issa-story.html	https://www.latimes.com/politics/story/2019-10-13/trump-scandal-beginning-to-smell-a-lot-like-watergate

Democrats -Overlap 0

Lucene	Custom Page Rank
https://www.latimes.com/politics/story/2019-11-07/house-democrats-subpoena-mick-mulvaney-in-impeachment-probe	https://www.latimes.com/local/lanow/la-me-ln-essential-california-20161114-story.html
https://www.latimes.com/opinion/story/2019-11-05/impeachment-donald-trump-democrats-political-cost-elections	https://www.latimes.com/politics/la-me-on-politics-column-20170323-story.html
https://www.latimes.com/politics/story/2019-11-03/iowa-democrats-candidate-beat-trump	https://www.latimes.com/people/seema-mehta
https://www.latimes.com/politics/story/2019-10-17/democrats-quick-impeachment-timing-complicated	https://www.latimes.com/politics/la-pol-ca-richest-in-congress-darrell-issa-story.html

https://www.latimes.com/politics/story/2019-10-08/impeachment-trump-democrats-whistleblower-identity-testimony	https://www.latimes.com/world-nation/story/2019-10-30/ukraine-trump-impeachment
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-healthcare-plan-too-liberal-to-win
https://www.latimes.com/politics/story/2019-09-29/trump-allies-and-democrats-reveal-the-deep-divisions-over-impeachment-inquiry	https://www.latimes.com/politics/la-na-pol-2020-democratic-presidential-candidates-2019-story.html
https://www.latimes.com/politics/story/2019-08-01/assignment-post-debate-health-care-analysis-new-story	https://www.latimes.com/politics/la-pol-ca-ceqa-homeless-shelter-20190515-story.html
https://www.latimes.com/nation/ct-democrats-economic-plan-20170824-story.html	https://www.latimes.com/politics/la-na-pol-julian-assange-indictment-wikileaks-20190523-story.html
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/politics/story/2019-10-17/gordon-sondland-ukraine-house-impeachment-testimony

Patriot Movement – Overlap 0

Lucene	Custom Page Rank
https://www.latimes.com/archives/la-xpm-2012-oct-23-la-me-russell-means-20121023-story.html	https://www.latimes.com/business/technology/story/2019-08-29/ab5-uber-lyft-newsom-lorena-gonzalez-ballot-tony-west
https://www.latimes.com/staff/megan-garvey	https://www.latimes.com/entertainment/arts/la-ca-cm-queer-threads-book-20170702-htmlstory.html
https://www.latimes.com/archives/la-xpm-2001-jun-10-mn-8792-story.html	https://www.latimes.com/travel/story/2019-10-25/can-you-carry-cbd-plane-across-state-lines
https://www.latimes.com/entertainment-arts/books/story/2019-10-25/shadowlands-anthony-mccann-oregon-standoff	https://www.latimes.com/socal/burbank-leader/news/story/2019-10-24/a-word-please-a-hyphen-rule-reversal-and-other-associated-press-twitter-proclamations
https://www.latimes.com/sports/highschool/story/2019-11-06/girls-tennis-southern-section-playoff-results-and-updated-pairings	https://www.latimes.com/california/story/2019-11-09/prop-187-anniversary-commemoration

https://www.latimes.com/politics/story/2019-09-20/bernie-sanders-muslim-voters-2020	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/nba/la-sp-nba-best-game-ever-20181222-story.html	https://www.latimes.com/entertainment-arts/music/story/2019-07-17/bad-bunny-ricky-martin-lin-manuel-miranda-want-puerto-rican-governor-out
https://www.latimes.com/sports/highschool/story/2019-11-06/girls-tennis-southern-section-playoff-results-and-updated-pairings	https://www.latimes.com/environment/story/2019-09-19/climate-change-youth-activism
https://www.latimes.com/entertainment/movies/la-et-mn-july-4-box-office-history-20180629-story.html	https://www.latimes.com/environment
https://www.latimes.com/entertainment/movies/la-et-mn-july-4-box-office-history-20180629-story.html	https://www.latimes.com/world-nation/story/2019-09-25/ukraine-president-trump-impeachment-inquiry

Republicans -Overlap 0

Lucene	Custom Page Rank
https://www.latimes.com/opinion/story/2019-10-23/house-republicans-storm-hearing-impeachment	https://www.latimes.com/opinion/story/2019-11-01/artificial-intelligence-ai-wall-street
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/politics/la-me-on-politics-column-20170323-story.html
https://www.latimes.com/opinion/story/2019-11-08/impeachment-trump-republicans-toomey	https://www.latimes.com/people/seema-mehta
https://www.latimes.com/politics/story/2019-10-04/vulnerable-senate-republicans-impeachment	https://www.latimes.com/politics/la-pol-ca-richest-in-congress-darrell-issa-story.html
https://www.latimes.com/california/story/2019-10-16/california-republicans-democrats-wildfires-homelessness	https://www.latimes.com/opinion/editorials/la-ed-migrants-dead-drowned-rio-grande-el-salvador-mexico-trump-20190626-story.html
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-healthcare-plan-too-liberal-to-win

https://www.latimes.com/opinion/story/2019-09-17/california-republicans-trump-fundraising-president	https://www.latimes.com/politics/story/2019-09-25/trump-impeachment-nancy-pelosi-ukraine-biden-call
https://www.latimes.com/politics/story/2019-10-28/katie-hill-resignation-trump-obstacle-for-republicans	https://www.latimes.com/opinion/editorials/la-ed-climate-change-global-warming-part-2-story.html
https://www.latimes.com/opinion/story/2019-11-06/trump-republicans-democrats-predictions-polling-2020	https://www.latimes.com/politics/story/2019-10-13/trump-scandal-beginning-to-smell-a-lot-like-watergate
https://www.latimes.com/politics/story/2019-11-06/republicans-election-trump-candidate-loses-kentucky	https://www.latimes.com/politics/story/2019-10-09/trumps-legal-strategy-against-impeachment-shaky-constitutional-grounds-scholars-say

Senate -Overlap 0

Lucene	Custom Page Rank
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/opinion/story/2019-11-01/artificial-intelligence-ai-wall-street
https://www.latimes.com/politics/story/2019-10-30/california-donors-spend-millions-on-2020-senate-races-across-the-country	https://www.latimes.com/politics/la-me-on-politics-column-20170323-story.html
https://www.latimes.com/politics/story/2019-10-04/vulnerable-senate-republicans-impeachment	https://www.latimes.com/people/seema-mehta
https://www.latimes.com/politics/la-na-pol-kavanaugh-hearing-20180927-story.html	https://www.latimes.com/opinion/editorials/la-ed-migrants-dead-drowned-rio-grande-el-salvador-mexico-trump-20190626-story.html
https://www.latimes.com/nation/la-pol-scotus-confirmation-votes-over-the-years-20181005-htmlstory.html	https://www.latimes.com/politics/story/2019-11-08/elizabeth-warren-healthcare-plan-too-liberal-to-win
https://www.latimes.com/politics/story/2019-08-14/2020-senate-control-presidential-race	https://www.latimes.com/politics/la-na-pol-2020-democratic-presidential-candidates-2019-story.html
https://www.latimes.com/politics/story/2019-08-14/john-hickenlooper-quits-presidential-race-for-senate-run	https://www.latimes.com/politics/la-pol-ca-ceqa-homeless-shelter-20190515-story.html

https://www.latimes.com/politics/story/2019-10-30/california-donors-spend-millions-on-2020-senate-races-across-the-country	https://www.latimes.com/california/story/2019-11-09/prop-187-anniversary-commemoration
https://www.latimes.com/politics/la-na-pol-william-barr-senate-confirm-attorney-general-20190214-story.html	https://www.latimes.com/entertainment-arts/tv/story/2019-09-29/netflix-the-politician-bette-midler-judith-light-ryan-murphy
https://www.latimes.com/opinion/story/2019-10-15/wmcaleen-homeland-security-trump-nielsen-senate-acting-secretary	https://www.latimes.com/sitemap/1993/4

Olympics 2020 -Overlap 0

Lucene	Custom Page Rank
https://www.latimes.com/sports/olympics/story/2019-08-15/2020-tokyo-olympics-searching-for-answers-amid-heat-wave	https://www.latimes.com/sports/clippers/story/2019-11-03/montrezl-harrell-clippers-made-the-right-decision-jazz-nba
https://www.latimes.com/sports/olympics	https://www.latimes.com/california/story/2019-10-10/pge-admits-equipment-may-have-sparked-several-fires-this-year
https://www.latimes.com/people/david-wharton	https://www.latimes.com/entertainment-arts/story/2019-10-01/did-stephen-miller-accidentally-blow-the-whistle-on-trumps-presidency
https://www.latimes.com/people/david-wharton	https://www.latimes.com/business/technology/story/2019-08-29/ab5-uber-lyft-newsom-lorena-gonzalez-ballot-tony-west
https://www.latimes.com/sports/olympics/story/2019-10-16/2020-tokyo-olympics-marathon-sapporo-avoid-heat	https://www.latimes.com/politics/la-me-on-politics-column-20170323-story.html
https://www.latimes.com/sports/olympics/story/2019-07-31/deadly-heat-wave-2020-summer-olympics-tokyo	https://www.latimes.com/entertainment/music/la-et-ms-kevin-bean-kroq-fm-20150414-story.html
https://www.latimes.com/sports/story/2019-11-05/ioc-anti-doping-2020-tokyo-olympics	https://www.latimes.com/sports/story/2019-11-06/diamond-league-drop-discus-triple-jump-steeplechase
https://www.latimes.com/sports/olympics/story/2019-10-10/naomi-osaka	https://www.latimes.com/people/seema-mehta

chooses-japan-over-u-s-2020-tokyo-olympics	
https://www.latimes.com/sports/story/2019-07-23/tokyos-rough-road-to-2020-summer-olympics	https://www.latimes.com/politics/la-pol-ca-richest-in-congress-darrell-issa-story.html
https://www.latimes.com/people/david-wharton	https://www.latimes.com/politics/story/2019-11-08/elizabeth-warren-healthcare-plan-too-liberal-to-win

Stock -Overlap 0

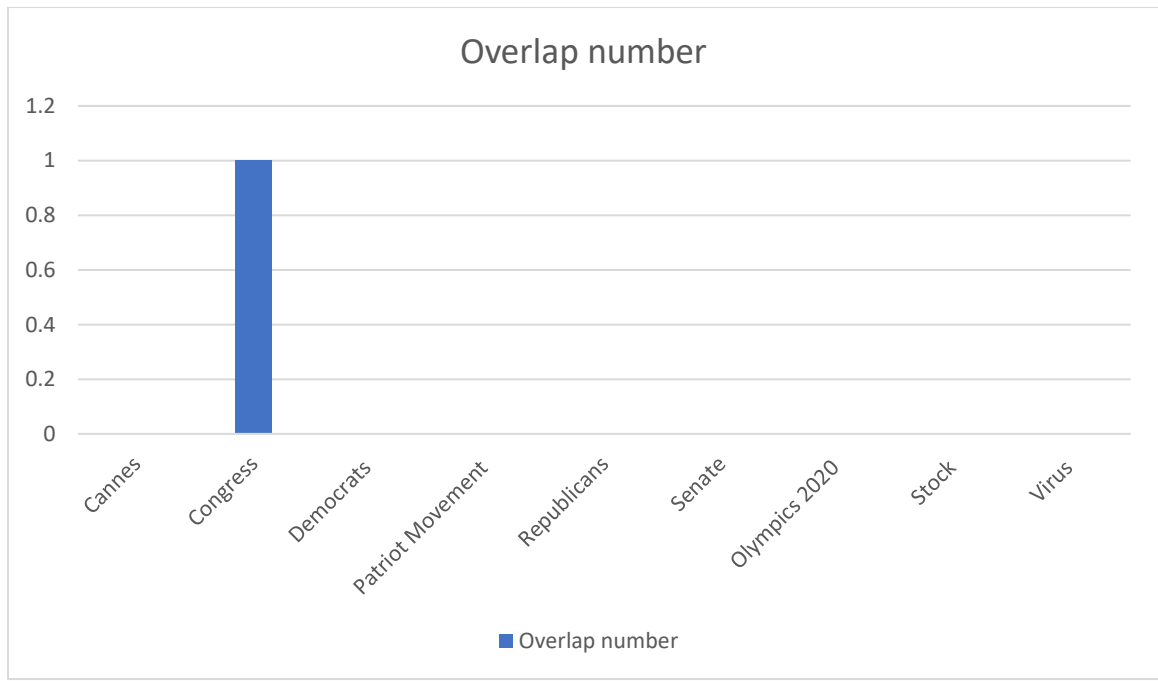
Lucene	Custom Page Rank
https://www.latimes.com/business/story/2019-11-06/uber-lock-up-period-ends-with-falling-stock-protests	https://www.latimes.com/people/dan-beucke
https://www.latimes.com/business/la-fi-lyft-stock-20190401-story.html	https://www.latimes.com/opinion/story/2019-11-01/artificial-intelligence-ai-wall-street
https://www.latimes.com/business/la-fi-hy-tesla-stock-20160518-snap-story.html	https://www.latimes.com/business/la-fi-edison-pge-20190422-story.html
https://www.latimes.com/business/story/2019-11-06/uber-lock-up-period-ends-with-falling-stock-protests	https://www.latimes.com/business/la-fi-tn-uber-board-20170930-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/real-estate/story/2019-09-20/producer-dana-brunetti-sells-toluca-lake-home-to-maroon-5s-james-valentine
https://www.latimes.com/business/la-fi-uber-ipo-stock-trading-price-20190510-story.html	https://www.latimes.com/business/la-fi-jobs-report-20190705-story.html
https://www.latimes.com/business/story/2019-09-09/at-t-stock-surges-elliott-reveals-3-2-billion-stake	https://www.latimes.com/business/la-xpm-2014-apr-11-la-fi-mo-minion-mayhem-universal-studios-20140411-story.html
https://www.latimes.com/business/la-fi-stock-market-status-update-20180209-story.html	https://www.latimes.com/business/story/2019-10-04/tesla-puts-driverless-cars-on-public-byways-controlled-by-humans-with-smartphones-abuse-is-rsafety-regulators-shrug

https://www.latimes.com/business/la-fi-stock-market-correction-20180212-story.html	https://www.latimes.com/business/real-estate/story/2019-10-18/architect-edward-killingsworth-house-sells-for-first-time
https://www.latimes.com/business/la-fi-stock-market-cpi-roundup-20180214-story.html	https://www.latimes.com/business/story/2019-09-09/elon-musk-says-he-was-involved-in-solarcity-discussions-prior-to-tesla-merger

Virus -Overlap 0

Lucene	Custom Page Rank
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/espanol/internacional/articulo/2019-11-02/mazda-muestra-en-el-auto-show-de-tokio-su-primer-vehiculo-electrico
https://www.latimes.com/science/sciencenow/la-sci-science-of-zika-five-ways-20160223-htmlstory.html	https://www.latimes.com/sitemap/2007/1
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/entertainment-arts/story/2019-10-25/movies-on-tv-this-week-sunday-oct-27-2019
https://www.latimes.com/science/story/2019-10-21/virus-afm-illness-paralyzing-kids	https://www.latimes.com/entertainment-arts/tv/story/2019-10-31/apple-tv-plus-shows-movies-price-launch-date
https://www.latimes.com/science/story/2019-10-21/virus-afm-illness-paralyzing-kids	https://www.latimes.com/local/lanow/la-me-ln-measles-at-lax-again-20190320-story.html
https://www.latimes.com/socal/daily-pilot/news/tn-dpt-me-west-nile-20170707-story.html	https://www.latimes.com/science/sciencenow/la-sci-sn-porter-ranch-methane-20160225-story.html
https://www.latimes.com/science/sciencenow/la-sci-sn-acute-flaccid-myelitis-polio-20181017-story.html	https://www.latimes.com/business/hiltzik/la-fi-hiltzik-drug-prices-20160601-snap-story.html
https://www.latimes.com/local/lanow/la-me-ln-virulent-newcastle-disease-outbreak-in-southern-california-20190607-story.html	https://www.latimes.com/espanol/vidayestilo/la-es-un-nuevo-estudio-de-la-epa-cuantifica-los-efectos-del-cambio-climatico-20190408-story.html
https://www.latimes.com/world-nation/story/2019-09-04/anti-vaxxers-helping-polio-comeback-pakistan	https://www.latimes.com/food/story/2019-10-25/plenty-vertical-farm-compton

https://www.latimes.com/local/obituaries/la-me-deborah-asnis-20150921-story.html	https://www.latimes.com/espanol/etiqueta/salud
---	---



17. I cloned the GitHub repository <https://github.com/PTCInc/solr-php-client.git> and started developing the php code. Also, I made sure to use the correct paths to the apache server and the csv files. Added radio buttons, search box and limited the results to 10. The screenshots are shown below.

PHP Solr Client Example

localhost:9000/page.php

Search :

☐ Lucene ☐ Page Rank

PHP Solr Client Example

localhost:9000/page.php?q=Virus&method=lucene

Search :

☒ Lucene ☐ Page Rank

Results 1 - 10 of 242:

Title : [Monos se volvieron inmunes tras inocularles virus del Ébola con mutación - Los Angeles Times](#)
URL : <https://www.latimes.com/espanol/internacional/articulo/2019-11-09/monos-se-volvieron-inmunes-tras-inocularles-virus-del-ebola-con-mutacion>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/93b7133f-b7f3-409e-bb23-c2c936725830.html
Description : Científicos han inoculado el virus del Ébola con una mutación en una proteína llamada VP3 a monos, que enfermaron y se volvieron inmunes, informó hoy la revista Cell Reports.

Title : [Five ways scientists are going after the Zika virus - Los Angeles Times](#)
URL : <https://www.latimes.com/science/sciencenow/la-sci-science-of-zika-five-ways-20160223.htmlstory.html>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/9ddf16c7-f3c6-484f-9276-44ed0f12dbb4.html
Description : To determine whether the Zika virus is indeed responsible for a surge in birth defects, public health officials are drawing on expertise from many disciplines.

Title : [Monos se volvieron inmunes tras inocularles virus del Ébola con mutación - Los Angeles Times](#)
URL : <https://www.latimes.com/espanol/internacional/articulo/2019-11-09/monos-se-volvieron-inmunes-tras-inocularles-virus-del-ebola-con-mutacion>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/ba09422e-fed8-49e0-9944-973c7e754bc1.html
Description : Científicos han inoculado el virus del Ébola con una mutación en una proteína llamada VP3 a monos, que enfermaron y se volvieron inmunes, informó hoy la revista Cell Reports.

Title : [Virus may be culprit in AFM, the paralyzing illness affecting kids - Los Angeles Times](#)
URL : <https://www.latimes.com/science/story/2019-10-21/virus-afm-illness-paralyzing-kids>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/733787a8-6c61-4d00-af2b-3e48668f9001.html
Description : Scientists have found the strongest evidence yet that a virus is to blame for AFM, a mysterious polio-like illness that can quickly paralyze children.

Title : [Virus may be culprit in AFM, the paralyzing illness affecting kids - Los Angeles Times](#)
URL : <https://www.latimes.com/science/story/2019-10-21/virus-afm-illness-paralyzing-kids>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/aa26c5e1-6da7-4f49-81dc-36a907202ff7.html
Description : Scientists have found the strongest evidence yet that a virus is to blame for AFM, a mysterious polio-like illness that can quickly paralyze children.

Title : [Laguna Beach woman tests positive for West Nile virus; first case in O.C. this year - Los Angeles Times](#)
URL : <https://www.latimes.com/socal/daily-pilot/news/tn-dpt-me-west-nile-20170707-story.html>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/a8b52a29-f6f2-4a58-8b4b-352a89aac45a.html
Description : A Laguna Beach woman is the first person to test positive for West Nile virus in Orange County this year, the county Health Care Agency said Friday.

Title : [Q&A: What is AFM? Everything you need to know about the polio-like virus suddenly affecting children across the U.S. - Los Angeles Times](#)
URL : <https://www.latimes.com/science/sciencenow/la-sci-sn-acute-flaccid-mylitis-polio-20181017-story.html>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/f5d77f1c-7186-46ec-8114-ab2a2d36dae7.html

PHP Solr Client Example

localhost:9000/page.php?q=Virus&method=pagerank

Search: Virus

☐ Lucene ☒ Page Rank

Results 1 - 10 of 242:

Title : [Mazda muestra en el auto show de Tokio su primer vehículo eléctrico - Los Angeles Times](#)
URL : <https://www.latimes.com/sspanol/internacional/articulo/2019-11-02/mazda-muestra-en-el-auto-show-de-tokio-su-primer-vehiculo-electrico>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/a274c249-7d37-4de7-a172-b6f977727e82.html
Description : Mazda asegura que el MX-30 es un modelo que ha sido desarrollado pensando en los conductores, que cuenta con unas sorprendentes puertas freestyle, que utiliza materiales ecológicos y monta una batería de tamaño adecuado que le proporciona una autonomía de aproximadamente 130 millas.

Title : [HTML Sitemap - January 2007 - Los Angeles Times](#)
URL : <https://www.latimes.com/sitemap/2007/1>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/2d9bb099-c4eb-4a09-96f5-93df29d9d682.html
Description :

Title : [Movies on TV this week: Sunday, Oct. 27, 2019 - Los Angeles Times](#)
URL : <https://www.latimes.com/entertainment-arts/story/2019-10-25/movies-on-tv-this-week-sunday-oct-27-2019>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/76b1112e-28a4-4dc6-916d-88f592903663.html
Description : Movies on TV this week: Sunday, Oct. 27, 2019

Title : [Apple TV+: A guide to its shows, movies, price and launch date - Los Angeles Times](#)
URL : <https://www.latimes.com/entertainment-arts/tv/story/2019-10-31/apple-tv-plus-shows-movies-price-launch-date>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/aSec262c-24ce-401c-8d9a-3893dd5c29b8.html
Description : Our guide to the original TV shows and movies, catalog titles, launch date, price and other features of Apple's new streaming service, Apple TV+.

Title : [Officials warn of measles exposure at LAX — again - Los Angeles Times](#)
URL : <https://www.latimes.com/local/lanow/la-me-ln-measles-at-lax-again-20190320-story.html>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/2cd348c1-5711-4a00-a689-d10761fd112f.html
Description : For the second time this month, Los Angeles County public health officials are warning travelers that a person with measles flew into Los Angeles International Airport.

Title : [Porter Ranch leak declared largest methane leak in U.S. history - Los Angeles Times](#)
URL : <https://www.latimes.com/science/sciencenow/la-sci-sn-porter-ranch-methane-20160225-story.html>
ID : /Users/saurabhvajramatti/Saurabh Vajramatti USC/IR/HW4/Info/latimes/bf29edbf-4443-487c-9463-9ac7f0ed3aac.html
Description : Scientists who flew an airplane equipped with sensors through the plume of natural gas leaking into the Porter Ranch area have found that at its peak, the nearly four-month leak released roughly 100,000 tons of methane — effectively doubling the methane emissions rate of the entire Los Angeles...

Title : [Here's why it's so difficult to solve soaring drug prices - Los Angeles Times](#)
URL : <https://www.latimes.com/business/la-fi-drug-prices-20190601-story.html>

Movies on TV this week: Sund...

latimes.com/entertainment-arts/story/2019-10-25/movies-on-tv-this-week-sunday-oct-27-2019

Sections

Los Angeles Times

Subscribe Now \$1 for 8 weeks

LOG IN

ADVERTISEMENT

National news. West Coast perspective.

Your support makes our journalism possible.

8 weeks for \$1

or

1 year for \$196 \$98

Cancel anytime.

SUBSCRIBE

Los Angeles Times

The state of what's next.

Already a subscriber? [Login](#)

Print subscriber? [Activate digital access](#)

Why some pages have higher PageRank values than others?

So, our custom page rank algorithm works by taking into consideration the in-links and out-links. The importance of a node depends on the number of in-links associated with the node. Also, the importance of a node also increases if it is being pointed to by important nodes, which in-turn are ones who have high in-link count. So, basically the nodes(webpages) with the highest score according to the scheme mentioned before would be the ones having higher PageRank than the others.