CSCI 572 HOMEWORK 4 : COMPARING SEARCH ENGINE RANKING ALGORITHMS

Name: DEEPAK SINGH

USC ID : 5026958753

Website Indexed: BOSTON GLOBE

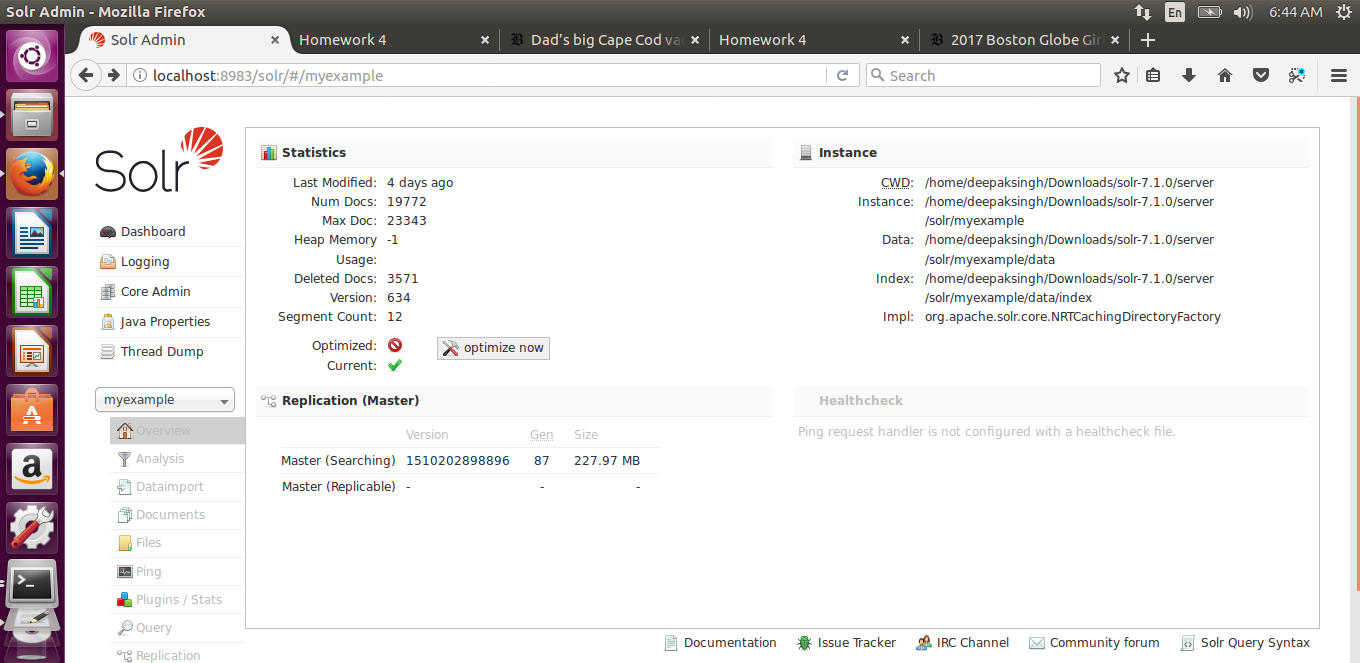
Step 1 : Solr Installation

* I downloaded solr and for the website as suggested in the guidelines.
* I started solr using **bin/solr start**
* I then created a core using **bin/solr create -c myexample**
* Solr was seen on the web browser at port 8983 – **localhost:8983/solr/#**

Step 2: Indexing the Web-Page

* I downloaded the data from the Google Drive for Boston\_Globe
* For indexing the managed-schema file where changed to remove the exceptions like NumberFormat, Invalid Date, etc.
* bin/post -c myexample -filetypes html ﻿/home/deepaksingh/Downloads/solr-7.1.0/BG/BG/BG/

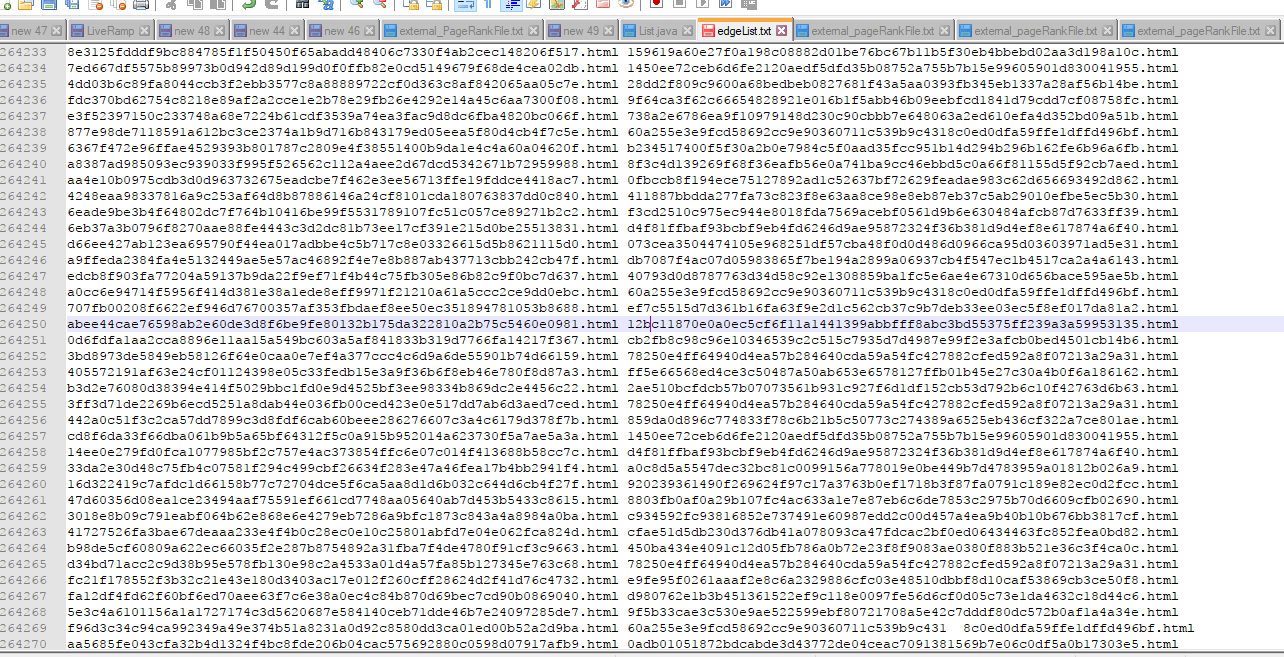
- did the indexing.

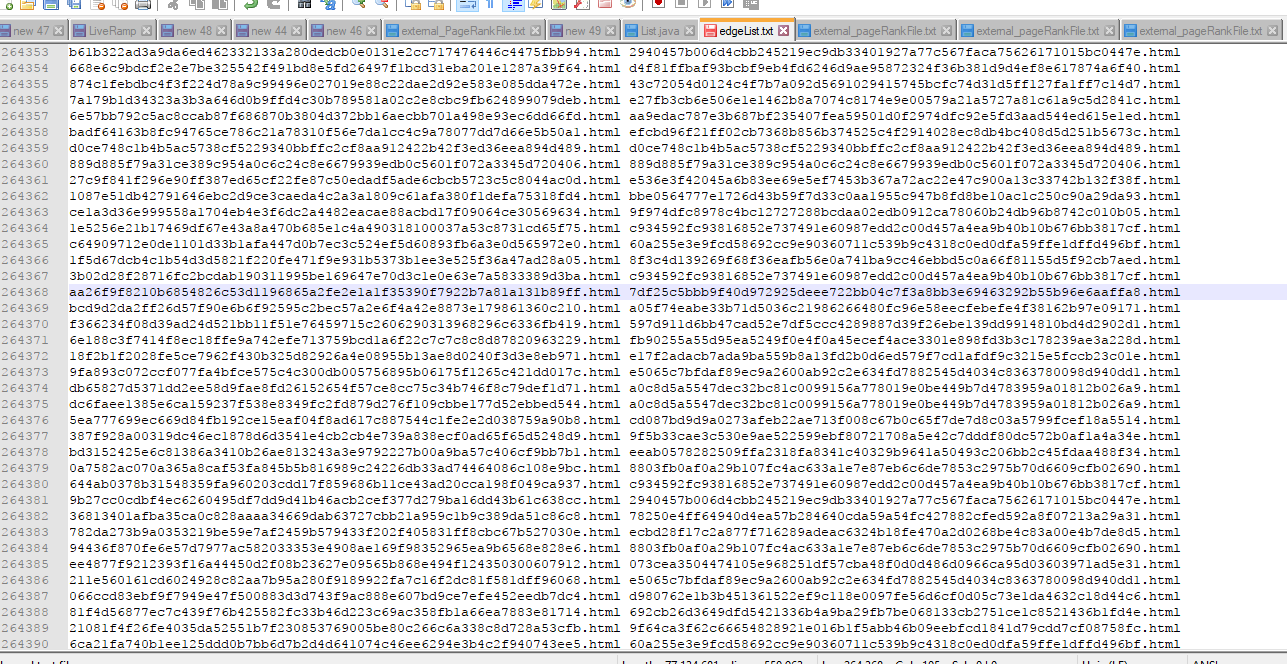


Step 3 : PageRank Algorithm:

* I downloaded Jsoup from the Web and wrote a Java file List.java to find the edgelist for the websites.
* JSOUP library was used to loop over each file in the input directory containing html files. jsoup is a Java library for working with real-world HTML. It provides a very convenient API for extracting and manipulating data, using the best of DOM, CSS, and jquery-like methods.
* A java program using JSOUP generated edge list of files which is used by used networkx to generate PageRank for each data. I called pageRank method of networkx to generate pageRank values for documents and wrote these values to ‘external\_PageRank.txt’ file using python code.

The following is a edgeList snippet :



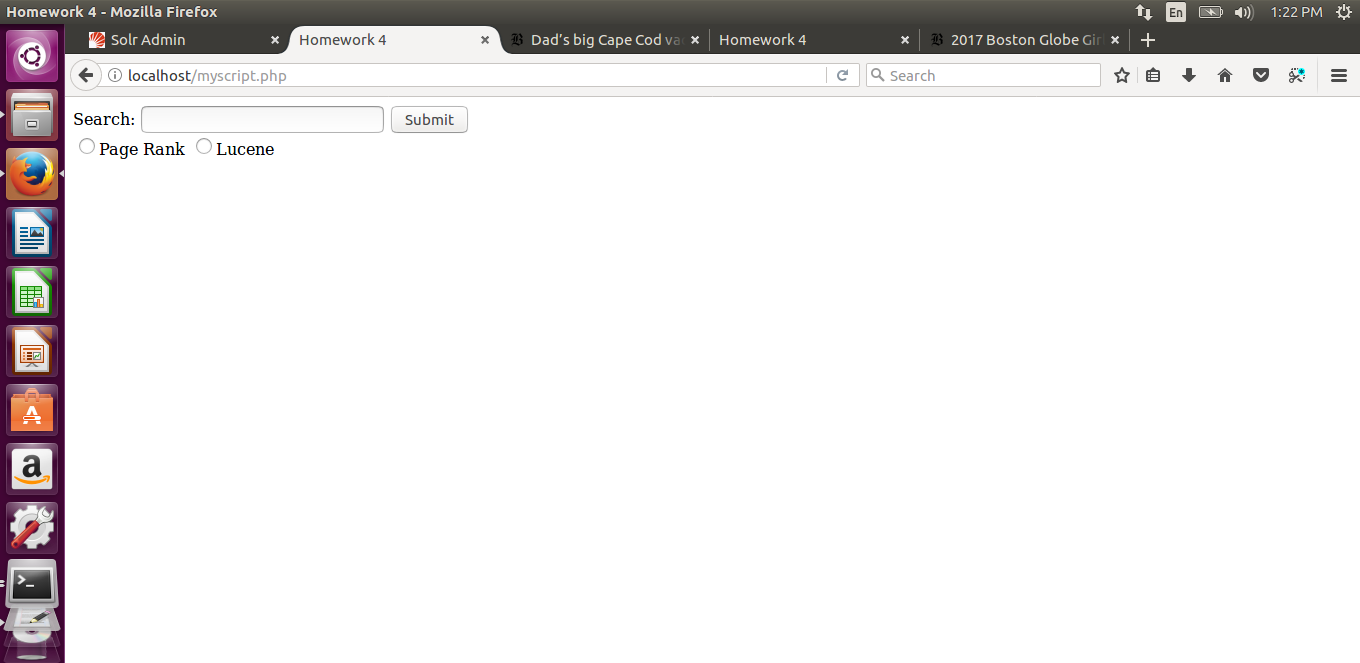


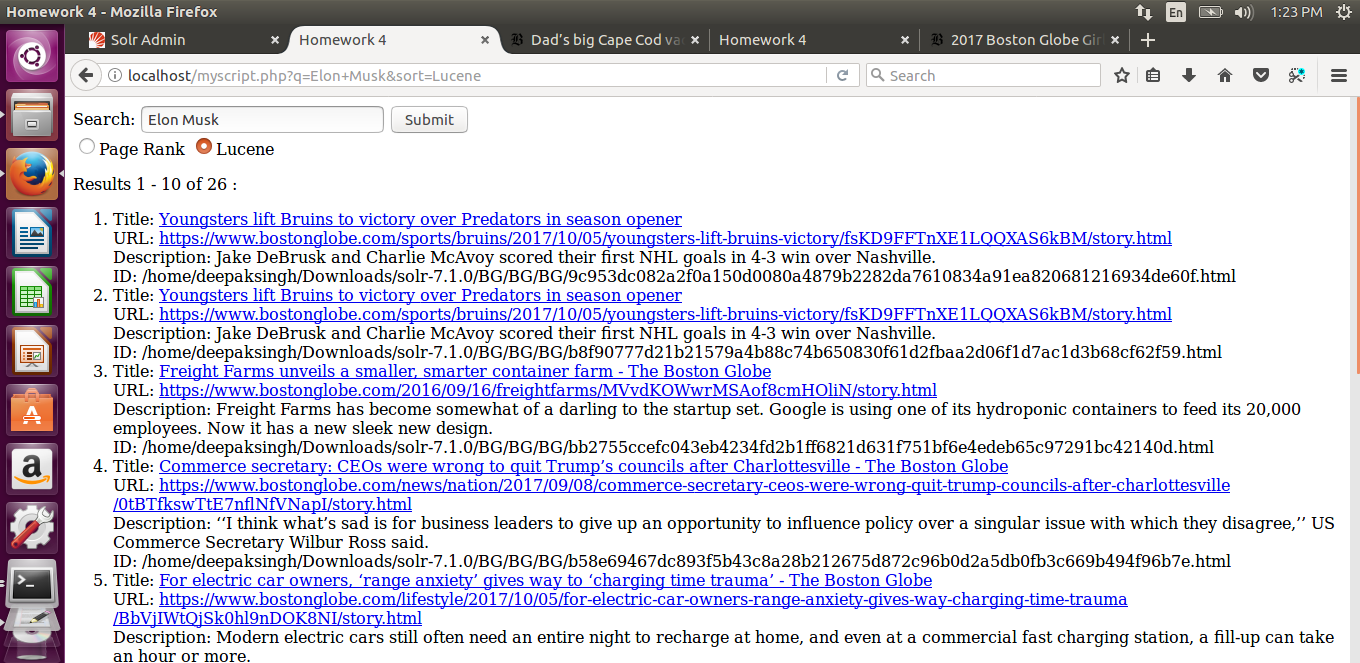
* The I used Networkx to read the graph and output a pagerank file name external\_pageRankFile.txt as given in the folder.
* PageRank Factors : alpha=0.85, personalization=None, max\_iter=30, tol=1e-06, nstart=None, weight='weight',dangling=None
* The IDEs used for java was Eclipse and PyDev for Python.
* Then I copied the file into the data folder of my core.

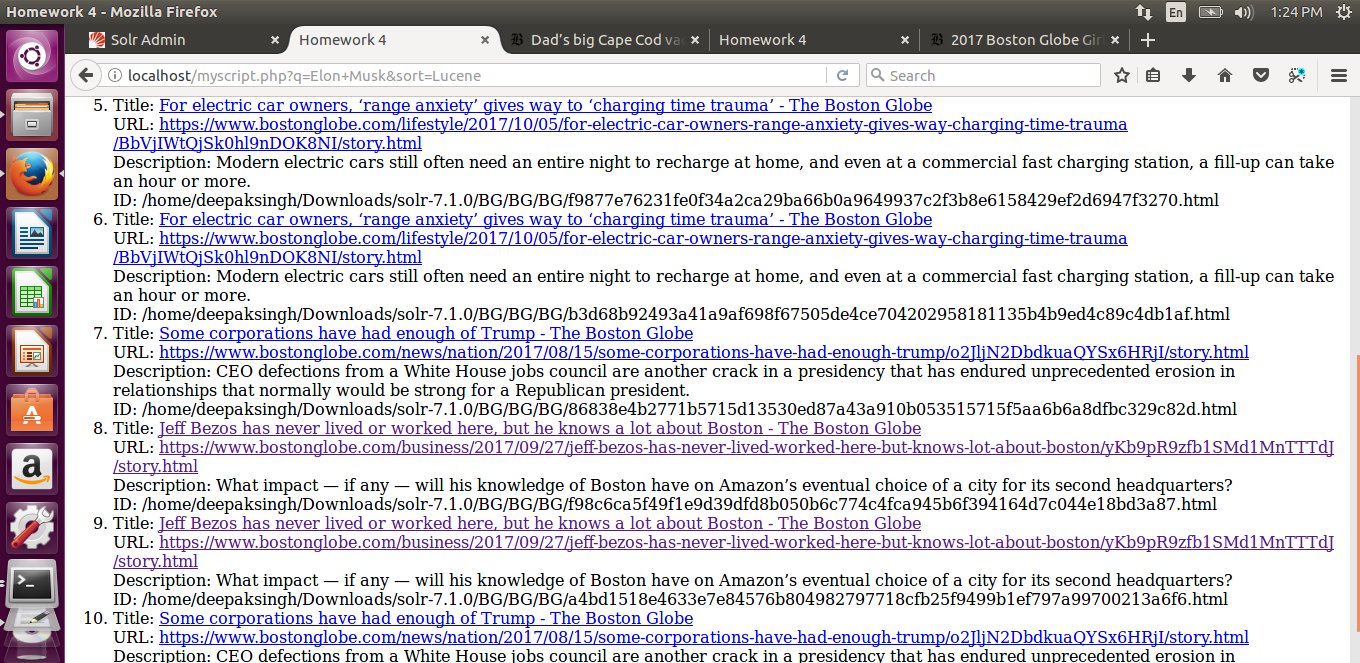
**The Following Snippet Shows a working model of my code :**

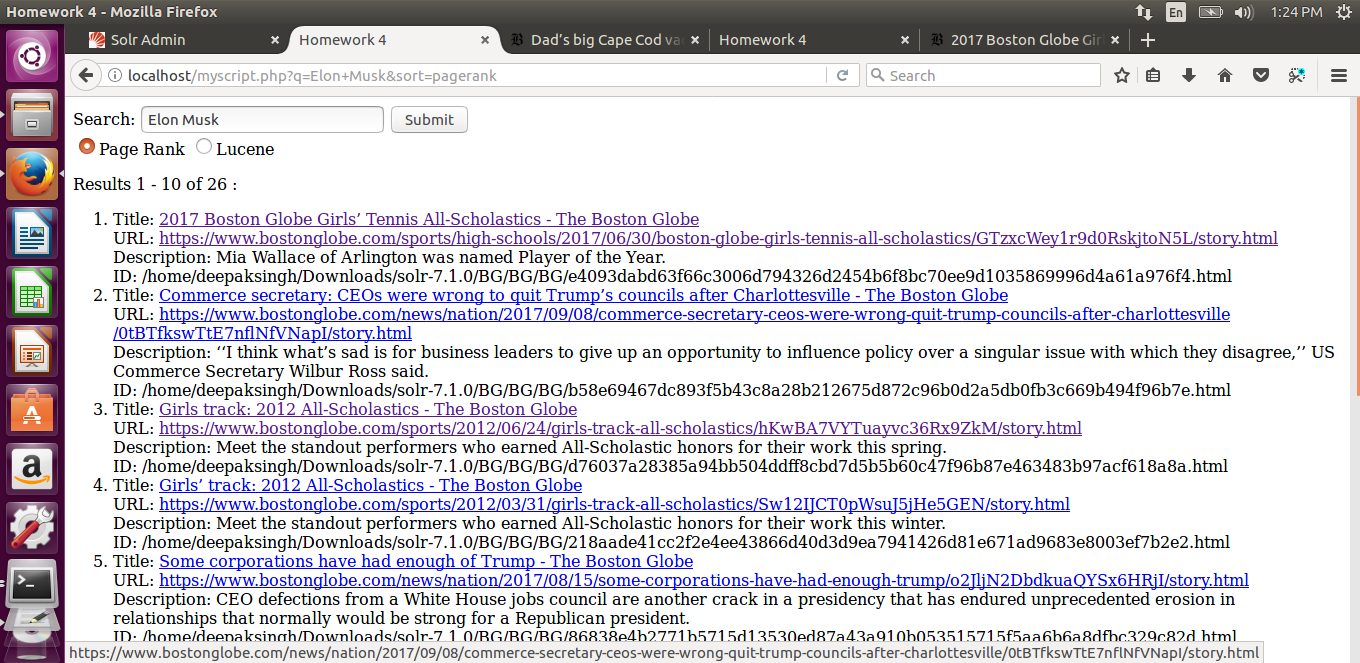
1. Steps in Query

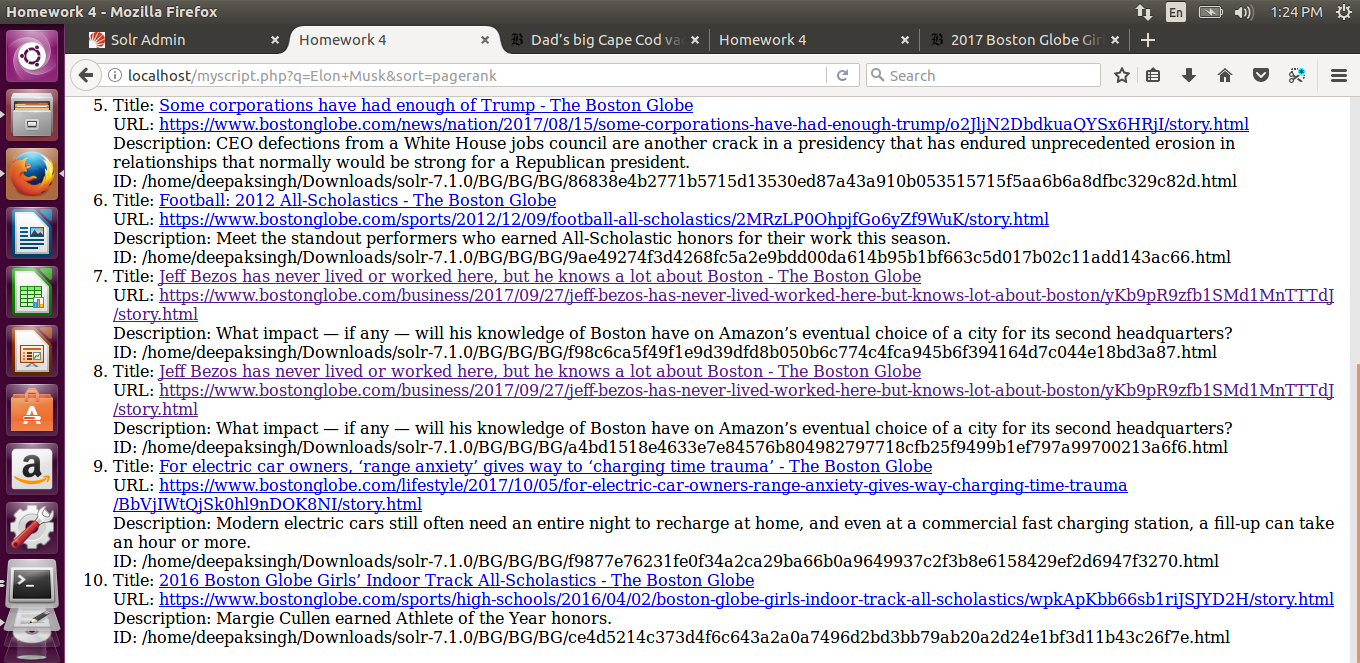
myscript.php gives the GUI on the localhost for querying solr. After querying, we can select the algorithm to be executed – Lucene or PageRank. The results are then displayed on the same page. These results are clickable links that can be routed to the destination link. The following snippets shows the process of Querying solr.

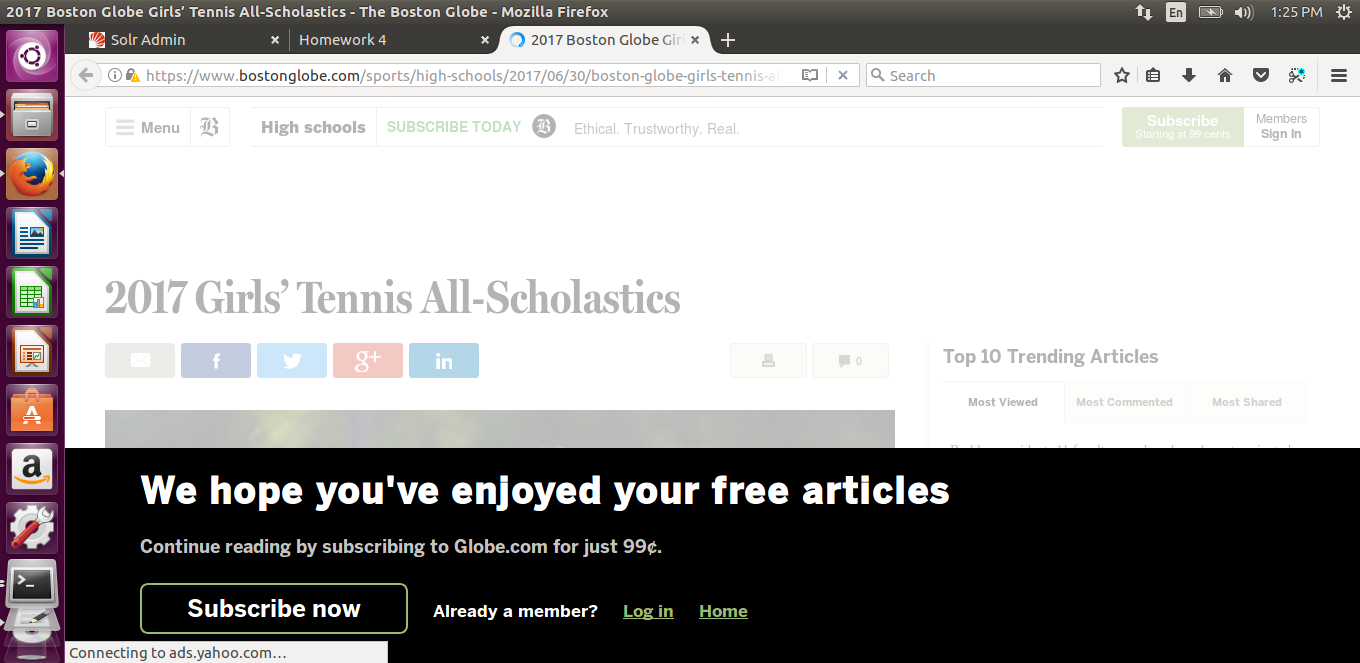












**1.Elon Musk**

|  |  |
| --- | --- |
| **PageRank** | **Lucene** |
| https://www.bostonglobe.com/sports/high-schools/2017/06/30/boston-globe-girls-tennis-all-scholastics/GTzxcWey1r9d0RskjtoN5L/story.html | https://www.bostonglobe.com/sports/bruins/2017/10/05/youngsters-lift-bruins-victory/fsKD9FFTnXE1LQQXAS6kBM/story.html |
| https://www.bostonglobe.com/news/nation/2017/09/08/commerce-secretary-ceos-were-wrong-quit-trump-councils-after-charlottesville/0tBTfkswTtE7nflNfVNapI/story.html | https://www.bostonglobe.com/sports/bruins/2017/10/05/youngsters-lift-bruins-victory/fsKD9FFTnXE1LQQXAS6kBM/story.html |
| https://www.bostonglobe.com/sports/2012/06/24/girls-track-all-scholastics/hKwBA7VYTuayvc36Rx9ZkM/story.html | https://www.bostonglobe.com/2016/09/16/freightfarms/MVvdKOWwrMSAof8cmHOliN/story.html |
| https://www.bostonglobe.com/sports/2012/03/31/girls-track-all-scholastics/Sw12IJCT0pWsuJ5jHe5GEN/story.html | https://www.bostonglobe.com/news/nation/2017/09/08/commerce-secretary-ceos-were-wrong-quit-trump-councils-after-charlottesville/0tBTfkswTtE7nflNfVNapI/story.html |
| https://www.bostonglobe.com/news/nation/2017/08/15/some-corporations-have-had-enough-trump/o2JljN2DbdkuaQYSx6HRjI/story.html | https://www.bostonglobe.com/lifestyle/2017/10/05/for-electric-car-owners-range-anxiety-gives-way-charging-time-trauma/BbVjIWtQjSk0hl9nDOK8NI/story.html |
| https://www.bostonglobe.com/sports/2012/12/09/football-all-scholastics/2MRzLP0OhpjfGo6yZf9WuK/story.html | https://www.bostonglobe.com/lifestyle/2017/10/05/for-electric-car-owners-range-anxiety-gives-way-charging-time-trauma/BbVjIWtQjSk0hl9nDOK8NI/story.html |
| https://www.bostonglobe.com/business/2017/09/27/jeff-bezos-has-never-lived-worked-here-but-knows-lot-about-boston/yKb9pR9zfb1SMd1MnTTTdJ/story.html | https://www.bostonglobe.com/news/nation/2017/08/15/some-corporations-have-had-enough-trump/o2JljN2DbdkuaQYSx6HRjI/story.html |
| https://www.bostonglobe.com/business/2017/09/27/jeff-bezos-has-never-lived-worked-here-but-knows-lot-about-boston/yKb9pR9zfb1SMd1MnTTTdJ/story.html | https://www.bostonglobe.com/business/2017/09/27/jeff-bezos-has-never-lived-worked-here-but-knows-lot-about-boston/yKb9pR9zfb1SMd1MnTTTdJ/story.html |
| https://www.bostonglobe.com/lifestyle/2017/10/05/for-electric-car-owners-range-anxiety-gives-way-charging-time-trauma/BbVjIWtQjSk0hl9nDOK8NI/story.html | https://www.bostonglobe.com/business/2017/09/27/jeff-bezos-has-never-lived-worked-here-but-knows-lot-about-boston/yKb9pR9zfb1SMd1MnTTTdJ/story.html |
| https://www.bostonglobe.com/sports/high-schools/2016/04/02/boston-globe-girls-indoor-track-all-scholastics/wpkApKbb66sb1riJSJYD2H/story.html | https://www.bostonglobe.com/news/nation/2017/08/15/some-corporations-have-had-enough-trump/o2JljN2DbdkuaQYSx6HRjI/story.html |

1. **Star Wars**

|  |  |
| --- | --- |
| **PageRank** | **Lucene** |
| https://www.bostonglobe.com/sports/high-schools/2017/06/30/boston-globe-girls-tennis-all-scholastics/GTzxcWey1r9d0RskjtoN5L/story.html | https://www.bostonglobe.com/arts/movies/2015/12/15/the-force-with-star-wars-the-force-awakens/KyOQ1W4zeMyVDncrIirp5M/story.html |
| https://www.bostonglobe.com/news/nation/2017/09/08/commerce-secretary-ceos-were-wrong-quit-trump-councils-after-charlottesville/0tBTfkswTtE7nflNfVNapI/story.html | https://www.bostonglobe.com/arts/movies/2015/12/15/the-force-with-star-wars-the-force-awakens/KyOQ1W4zeMyVDncrIirp5M/story.html |
| https://www.bostonglobe.com/sports/2012/06/24/girls-track-all-scholastics/hKwBA7VYTuayvc36Rx9ZkM/story.html | https://www.bostonglobe.com/metro/2015/12/17/first-line-for-the-force-fan-stakes-out-spot-for-latest-star-wars-film/QBYM4R1QwbCoA7c7cBs8wO/story.html |
| https://www.bostonglobe.com/sports/2012/03/31/girls-track-all-scholastics/Sw12IJCT0pWsuJ5jHe5GEN/story.html | https://www.bostonglobe.com/lifestyle/style/2015/12/18/for-star-wars-die-hards-hour-marathon-daunting-but-enjoyable-task/W08Q3d2TeXWCQzDalZaZsI/story.html |
| https://www.bostonglobe.com/news/nation/2017/08/15/some-corporations-have-had-enough-trump/o2JljN2DbdkuaQYSx6HRjI/story.html | https://www.bostonglobe.com/arts/2015/12/12/starwarsgenerations/PLDMidc6USG536AfUaPA7N/story.html |
| https://www.bostonglobe.com/sports/2012/12/09/football-all-scholastics/2MRzLP0OhpjfGo6yZf9WuK/story.html | https://www.bostonglobe.com/lifestyle/style/2015/12/12/star-wars-sneakers-not-just-for-kids/XnHidQOeClvVNY94HfptQJ/story.html |
| https://www.bostonglobe.com/business/2017/09/27/jeff-bezos-has-never-lived-worked-here-but-knows-lot-about-boston/yKb9pR9zfb1SMd1MnTTTdJ/story.html | https://www.bostonglobe.com/metro/2015/12/17/charlie-baker-not-fan-new-star-wars-flicks/UX6q3i6wY5F6AIuydftuTI/story.html |
| https://www.bostonglobe.com/business/2017/09/27/jeff-bezos-has-never-lived-worked-here-but-knows-lot-about-boston/yKb9pR9zfb1SMd1MnTTTdJ/story.html | https://www.bostonglobe.com/arts/movies/2015/12/12/getting-latest-scoop-scoops-star-wars/1CDNVgBUSzNeG9FPW50yJP/story.html |
| https://www.bostonglobe.com/lifestyle/2017/10/05/for-electric-car-owners-range-anxiety-gives-way-charging-time-trauma/BbVjIWtQjSk0hl9nDOK8NI/story.html | https://www.bostonglobe.com/arts/movies/2015/12/12/getting-latest-scoop-scoops-star-wars/1CDNVgBUSzNeG9FPW50yJP/story.html |
| https://www.bostonglobe.com/sports/high-schools/2016/04/02/boston-globe-girls-indoor-track-all-scholastics/wpkApKbb66sb1riJSJYD2H/story.html | https://www.bostonglobe.com/arts/movies/2015/12/12/oscar-isaac-talks-about-acting-new-star-wars-movie/Z85IkTprnFbJz9wT7PoygM/story.html |

1. **North Korea**

|  |  |
| --- | --- |
| **PageRank** | **Lucene** |
| N/A | https://www.bostonglobe.com/news/world/2017/09/08/korea-braces-for-another-possible-korea-missile-test/wrfYUFe7CEgGvKbcg41ZbN/story.html |
| https://www.bostonglobe.com/metro/regionals/north/2017/09/06/fines-levied-asphalt-plant-site-westford/oXHMCX7XzDsIBEkigXmQUJ/story.html | https://www.bostonglobe.com/news/world/2017/08/09/trump-threatens-fire-and-fury-against-north-korea-endangers/sSaFxbsoTVeQ3syE8gJvgO/story.html |
| N/A | https://www.bostonglobe.com/news/world/2017/08/09/trump-threatens-fire-and-fury-against-north-korea-endangers/sSaFxbsoTVeQ3syE8gJvgO/story.html |
| N/A | https://www.bostonglobe.com/opinion/2017/08/09/plan-for-north-korea-before-this-spirals-out-control/Q23bbp1BbDPlTxdvYxv3LK/story.html |
| N/A | https://www.bostonglobe.com/opinion/2017/08/09/plan-for-north-korea-before-this-spirals-out-control/Q23bbp1BbDPlTxdvYxv3LK/story.html |
| https://www.bostonglobe.com/metro/2017/03/25/massachusetts-failing-mental-health-care-system-even-lucky-ones-aren-lucky/EYWOJ46DAJZ78tifZWBWOK/story.html | https://www.bostonglobe.com/news/world/2017/09/25/north-korea-minister-says-can-shoot-down-bombers/8TtjEtXdimYcIXEArVHu5K/story.html |
| https://www.bostonglobe.com/sports/2012/12/09/girls-soccer-all-scholastics/mSimga4QrjSDaUIwWgBCbK/story.html | https://www.bostonglobe.com/news/politics/2017/09/26/north-korea-seeks-help-from-republican-analysts-what-with-trump/Iro675CGFxHP3wqY4ZbzhK/story.html |
| https://www.bostonglobe.com/news/bigpicture/2017/08/21/there-goes-sun-total-solar-eclipse/txO5gaCugLPTHKdQYhUhSO/story.html | https://www.bostonglobe.com/news/politics/2017/09/26/north-korea-seeks-help-from-republican-analysts-what-with-trump/Iro675CGFxHP3wqY4ZbzhK/story.html |
| https://www.bostonglobe.com/magazine/2017/05/19/five-things-may/hZW68EP3ZdxOn3cZ8DKtUK/story.html | https://www.bostonglobe.com/news/nation/2017/09/03/trump-says-north-korea-very-hostile-and-dangerous/SoR0T7tl6i3MdNpwHonu9M/story.html |
| https://www.bostonglobe.com/opinion/editorials/2017/06/24/bristol-county-deserves-better-sheriff-than-hodgson/ipneUr85oZ9R6qPlrSZ4zO/story.html | https://www.bostonglobe.com/news/nation/2017/09/03/trump-says-north-korea-very-hostile-and-dangerous/SoR0T7tl6i3MdNpwHonu9M/story.html |

1. **LA Dodgers**

|  |  |
| --- | --- |
| **PageRank** | **Lucene** |
| https://www.bostonglobe.com/lifestyle/food-dining/2017/08/07/instagram-ruining-food-and-might-only-one-who-cares/GlZVsoSzmMKRpzrtenjfQN/story.html | https://www.bostonglobe.com/sports/2017/10/07/mlb-boxscore-dodgers-diamondbacks/TAdQkAqY6OqncGtJjZyG8O/story.html |
| https://www.bostonglobe.com/lifestyle/names | https://www.bostonglobe.com/sports/2012/08/26/gonzo/Avzuwi4ZrRfOBdQKd1UFGP/story.html |
| N/A | https://www.bostonglobe.com/lifestyle/names/2017/09/28/carl-crawford-boston-carried-hate-for-that-city-for-long-time/SKvbZfN6QwarHEy24f965H/story.html |
| https://www.bostonglobe.com/lifestyle/2014/08/07/bostonians-accept-ice-bucket-challenge/DjJuksqazGGN7wNGTvBLZK/story.html | https://www.bostonglobe.com/lifestyle/names/2017/09/28/carl-crawford-boston-carried-hate-for-that-city-for-long-time/SKvbZfN6QwarHEy24f965H/story.html |
| N/A | N/A |
| https://www.bostonglobe.com/ideas/2014/10/31/text-aye-matey-the-pirate-party-push-for-direct-democracy/X8dl2dKD73HiXGTXd1yuDP/story.html | N/A |
| https://www.bostonglobe.com/sports/2015/06/13/for-red-sox-one-inning-microcosm-misery/pi6vCRDKVo7qs3G8TfWCoO/story.html | N/A |
| https://www.bostonglobe.com/arts/theater-art/2016/03/03/the-ticket-dance-picks/lMHNCoNmQzxi5ilPBsstIN/story.html | https://www.bostonglobe.com/sports/2017/09/30/baseball-standings/G6K9mjYu6AokJmy36PkhcL/story.html |
| https://www.bostonglobe.com/metro/2015/07/27/for-olympic-organizers-many-conflicting-constituencies-appease/QeZsPEEzLTohF9KNv8ADdK/story.html | https://www.bostonglobe.com/sports/redsox/2017/10/04/diamondbacks-advance-after-wild-one-over-rockies/ZXhz3EJeboHWpmrFA4UcRP/story.html |
| https://www.bostonglobe.com/arts/television/2017/10/03/pbs-the-collection-dressed-for-excess/bsTTLZdpXROxIeZBOFKRHL/story.html | https://www.bostonglobe.com/sports/redsox/2017/10/04/diamondbacks-advance-after-wild-one-over-rockies/ZXhz3EJeboHWpmrFA4UcRP/story.html |

**5 . Puerto Rico**

|  |  |
| --- | --- |
| **PageRank** | **Lucene** |
| https://www.bostonglobe.com/lifestyle/names | https://www.bostonglobe.com/news/politics/2017/10/03/six-things-trump-probably-shouldn-have-said-about-puerto-rico-while-puerto-rico/z2A6pU5IozLPkJdMzSY3HP/story.html |
| N/A | https://www.bostonglobe.com/news/politics/2017/10/03/six-things-trump-probably-shouldn-have-said-about-puerto-rico-while-puerto-rico/z2A6pU5IozLPkJdMzSY3HP/story.html |
| https://www.bostonglobe.com/sports/2017/09/30/bob-ryan-the-state-college-sports-has-become-national-embarrassment/G6iDZAj3qHj7F5PprsMOfK/story.html | https://www.bostonglobe.com/news/politics/2017/10/03/six-things-trump-probably-shouldn-have-said-about-puerto-rico-while-puerto-rico/z2A6pU5IozLPkJdMzSY3HP/story.html |
| https://www.bostonglobe.com/opinion/2017/10/05/how-much-longer-will-republicans-set-aside-their-integrity/nWN3ZIhQPjj3GQYj5uwsFP/story.html | https://www.bostonglobe.com/news/politics/2017/10/03/six-things-trump-probably-shouldn-have-said-about-puerto-rico-while-puerto-rico/z2A6pU5IozLPkJdMzSY3HP/story.html |
| https://www.bostonglobe.com/metro/2017/09/09/bay-staters-step-again-help-southern-neighbors-during-hurricane-irma/qu6OHqSHs6PU7F9ox9vjzN/story.html | https://www.bostonglobe.com/news/politics/2017/10/03/six-things-trump-probably-shouldn-have-said-about-puerto-rico-while-puerto-rico/z2A6pU5IozLPkJdMzSY3HP/story.html |
| https://www.bostonglobe.com/staff/horowitz | https://www.bostonglobe.com/news/politics/2017/10/03/six-things-trump-probably-shouldn-have-said-about-puerto-rico-while-puerto-rico/z2A6pU5IozLPkJdMzSY3HP/story.html |
| https://www.bostonglobe.com/magazine/2017/05/05/blind-date-was-raised-with-great-manners-props-his-mom-and-dad/G5QsKLvyE8HtoToCSxD9jO/story.html | https://www.bostonglobe.com/opinion/editorials/2017/09/26/where-urgency-puerto-rico-response/oFxAJlj9duwxiaxVGK0LIL/story.html |
| https://www.bostonglobe.com/arts/music/2017/09/29/beyonce-sings-spanish-for-relief-efforts-gente-remix/I9Lv5IJq1QS9bFqXURh4VP/story.html | https://www.bostonglobe.com/opinion/editorials/2017/09/26/where-urgency-puerto-rico-response/oFxAJlj9duwxiaxVGK0LIL/story.html |
| N/A | https://www.bostonglobe.com/opinion/editorials/2017/09/23/don-let-puerto-rico-languish/k7iPpkMGa54CDFd65Iwh4H/story.html |
| https://www.bostonglobe.com/business/2017/10/06/hiring-falls-after-hurricanes/TwNKvEJrikFWq62p4JsTfM/story.html | https://www.bostonglobe.com/opinion/editorials/2017/09/23/don-let-puerto-rico-languish/k7iPpkMGa54CDFd65Iwh4H/story.html |

1. **Hurricane Harvey**

|  |  |
| --- | --- |
| **PageRank** | **Lucene** |
| https://www.bostonglobe.com/magazine/2017/05/02/dad-big-cape-cod-vacation-gamble/xuxOIIi4Qsi54f8WTkaErJ/story.html | https://www.bostonglobe.com/news/nation/2017/09/08/house-seals-deal-hurricane-relief/mgMI2SyD8kUIb4ewoFSBII/story.html |
| https://www.bostonglobe.com/metro/2014/09/13/secretary-state-galvin-faces-criticism-for-keeping-government-secrets/e70pa1N2jfwTzKQyrLMWSL/story.html | https://www.bostonglobe.com/news/bigpicture/2017/08/28/hurricane-harvey-devastates-southeast-texas/ZWm7Ln0xffqQqpglzffYxL/story.html |
| https://www.bostonglobe.com/business/2017/09/26/biogen-best-hope-alzheimer/iYq5Yiz7kTRmbWQZeO8TEK/story.html | https://www.bostonglobe.com/news/bigpicture/2017/08/28/hurricane-harvey-devastates-southeast-texas/ZWm7Ln0xffqQqpglzffYxL/story.html |
| https://www.bostonglobe.com/metro/2014/03/22/remy/xFRaOQqrnZ1S1pfLa2eKgK/story.html | https://www.bostonglobe.com/lifestyle/names/2017/09/28/carl-crawford-boston-carried-hate-for-that-city-for-long-time/SKvbZfN6QwarHEy24f965H/story.html |
| N/A | https://www.bostonglobe.com/news/bigpicture/2017/08/28/hurricane-harvey-devastates-southeast-texas/ZWm7Ln0xffqQqpglzffYxL/story.html |
| https://www.bostonglobe.com/metro/regionals/north/2017/09/06/fines-levied-asphalt-plant-site-westford/oXHMCX7XzDsIBEkigXmQUJ/story.html | https://www.bostonglobe.com/news/bigpicture/2017/08/28/hurricane-harvey-devastates-southeast-texas/ZWm7Ln0xffqQqpglzffYxL/story.html |
| https://www.bostonglobe.com/sports/2017/07/22/halloffame/3Lqz1v7TaLebyhdZEVIE4H/story.html | https://www.bostonglobe.com/news/bigpicture/2017/08/28/hurricane-harvey-devastates-southeast-texas/ZWm7Ln0xffqQqpglzffYxL/story.html |
| https://www.bostonglobe.com/lifestyle/food-dining/2017/08/07/instagram-ruining-food-and-might-only-one-who-cares/GlZVsoSzmMKRpzrtenjfQN/story.html | https://www.bostonglobe.com/news/bigpicture/2017/08/28/hurricane-harvey-devastates-southeast-texas/ZWm7Ln0xffqQqpglzffYxL/story.html |
| https://www.bostonglobe.com/business/2014/09/21/child-poverty-continues-climb-massachusetts/cz4Df6tzBt2Rxdj1nM9d5J/story.html | https://www.bostonglobe.com/news/bigpicture/2017/08/28/hurricane-harvey-devastates-southeast-texas/ZWm7Ln0xffqQqpglzffYxL/story.html |
| https://www.bostonglobe.com/ideas/2014/10/18/vote-all-you-want-the-secret-government-won-change/jVSkXrENQlu8vNcBfMn9sL/story.html | https://www.bostonglobe.com/news/bigpicture/2017/08/28/hurricane-harvey-devastates-southeast-texas/ZWm7Ln0xffqQqpglzffYxL/story.html |

1. **iPhone X**

|  |  |
| --- | --- |
| **PageRank** | **Lucene** |
| https://www.bostonglobe.com/magazine/2017/05/02/dad-big-cape-cod-vacation-gamble/xuxOIIi4Qsi54f8WTkaErJ/story.html | https://www.bostonglobe.com/business/2017/09/21/should-you-queue-for-new-iphone-wait-for/MZLL85K3mhI2iGVYuy8ofN/story.html |
| https://www.bostonglobe.com/metro/2014/09/13/secretary-state-galvin-faces-criticism-for-keeping-government-secrets/e70pa1N2jfwTzKQyrLMWSL/story.html | https://www.bostonglobe.com/business/2017/09/21/should-you-queue-for-new-iphone-wait-for/MZLL85K3mhI2iGVYuy8ofN/story.html |
| N/A | https://www.bostonglobe.com/ideas/2017/09/12/apple-world-iphone-doesn-need-look-like-iphone/C8EFHm2BFe6eV9oeCudAKO/story.html |
| https://www.bostonglobe.com/business/2017/09/26/biogen-best-hope-alzheimer/iYq5Yiz7kTRmbWQZeO8TEK/story.html | https://www.bostonglobe.com/ideas/2017/09/12/apple-world-iphone-doesn-need-look-like-iphone/C8EFHm2BFe6eV9oeCudAKO/story.html |
| https://www.bostonglobe.com/metro/2014/03/22/remy/xFRaOQqrnZ1S1pfLa2eKgK/story.html | https://www.bostonglobe.com/ideas/2017/09/12/apple-world-iphone-doesn-need-look-like-iphone/C8EFHm2BFe6eV9oeCudAKO/story.html |
| N/A | https://www.bostonglobe.com/business/2017/09/12/super-new-iphone-may-not-quite-super-enough-for-price/cVm5OkTOwWHIRtNtC0RVDJ/story.html |
| N/A | https://www.bostonglobe.com/business/2017/09/12/super-new-iphone-may-not-quite-super-enough-for-price/cVm5OkTOwWHIRtNtC0RVDJ/story.html |
| https://www.bostonglobe.com/metro/regionals/north/2017/09/06/fines-levied-asphalt-plant-site-westford/oXHMCX7XzDsIBEkigXmQUJ/story.html | https://www.bostonglobe.com/business/2017/09/12/super-new-iphone-may-not-quite-super-enough-for-price/cVm5OkTOwWHIRtNtC0RVDJ/story.html |
| https://www.bostonglobe.com/sports/2017/07/22/halloffame/3Lqz1v7TaLebyhdZEVIE4H/story.html | https://www.bostonglobe.com/business/2017/09/12/super-new-iphone-may-not-quite-super-enough-for-price/cVm5OkTOwWHIRtNtC0RVDJ/story.html |
| N/A | NA |

1. **Paris Climate Deal**

|  |  |
| --- | --- |
| **PageRank** | **Lucene** |
| https://www.bostonglobe.com/magazine/2017/05/02/dad-big-cape-cod-vacation-gamble/xuxOIIi4Qsi54f8WTkaErJ/story.html | N/A |
| https://www.bostonglobe.com/metro/2014/03/22/remy/xFRaOQqrnZ1S1pfLa2eKgK/story.html | https://www.bostonglobe.com/news/politics/2017/06/02/john-kerry-calls-trump-climate-change-announcement-day-craven-ignorance-and-cynicism/GWBwSaQwm24lsxfX5jHmWM/story.html |
| N/A | https://www.bostonglobe.com/news/politics/2017/06/02/john-kerry-calls-trump-climate-change-announcement-day-craven-ignorance-and-cynicism/GWBwSaQwm24lsxfX5jHmWM/story.html |
| N/A | https://www.bostonglobe.com/magazine/2015/07/22/energy-secretary-ernie-moniz-cabinet-star-born/PsqmP1TairoTYQ1OSbt3uJ/story.html |
| N/A | https://www.bostonglobe.com/business/2017/06/14/stata-slowing-down-hardly/xBhdDp7fFMfsxpjCzGunQJ/story.html |
| https://www.bostonglobe.com/sports/2017/07/31/woman-says-pete-rose-had-sex-with-her-while-she-was-minor/D9ZIMvvmmreOcV8aBFzPdN/story.html | https://www.bostonglobe.com/metro/2017/06/02/climate/qELM7JPNpKnORMMqkzi2XJ/story.html |
| N/A | https://www.bostonglobe.com/metro/2017/06/02/climate/qELM7JPNpKnORMMqkzi2XJ/story.html |
| https://www.bostonglobe.com/opinion/2017/10/09/cancer-chemo-and-road-self-compassion/27h4g3idTBd4qF9OLUjZ1L/story.html | https://www.bostonglobe.com/metro/2017/06/02/climate/qELM7JPNpKnORMMqkzi2XJ/story.html |
| https://www.bostonglobe.com/business/2017/06/12/greenway-deal-could-clear-way-for-winthrop-square-tower/3mcqxcBQZ1HQf92LVp9a5L/story.html | https://www.bostonglobe.com/metro/2017/07/04/united-church-christ-approves-emergency-resolution-climate-change/pgzKhYbm3UEyJcxA21GPAI/story.html |
| https://www.bostonglobe.com/news/bigpicture | https://www.bostonglobe.com/metro/2017/07/04/united-church-christ-approves-emergency-resolution-climate-change/pgzKhYbm3UEyJcxA21GPAI/story.html |

**Why Some Pages Have Higher Page Rank Values**

• The PR of each page depends on the PR of the pages pointing to it and the outgoing links the page has. What that means to us is that we can just go ahead and calculate a page’s PR without knowing the final value of the PR of the other pages. That seems strange but, basically, each time we run the calculation we’re getting a closer estimate of the final value. Hence, In the assignment we create a graph that decides the list of edges pointing the page.   
• As we add more internal links in a site, the PR will be spread out more evenly between the pages. Important, high-quality sites receive a higher PageRank.  
• From the results it is observed that the PR algorithm returns pages with high rank but are not so relevant to the search query.  
• It’s probably better to get lots (perhaps thousands) of links from sites with small PR than to spend any time or money desperately trying to get just the one link from a high PR page. Of course, important pages mean nothing to you if they don’t match your query.

* Thus the major reason that pagerank has higher page rank value is the fact that it analyses the weight of page on the number of links that point it.

**The Graph of OverLap:**

|  |  |
| --- | --- |
| **Query** | **Overlap** |
| Elon Musk | 5 |
| Star Wars | 0 |
| North Korea | 0 |
| LA Dodgers | 0 |
| Puerto Rico | 0 |
| Hurricane Harvey | 0 |
| Iphone X | 0 |
| Paris Climate Deal | 0 |

**Conclusion:** Thus, we implemented Solr with Lucene and PageRank and compared the results.