

Keyword search for "government"

TFIDF (TF)(IDF)	TF (occurrence/document)	IDF - bing	URI
0.067	135/15208 = 0.00887	$\log_2(16.8B/89.4M)$ =7.61495031333	http://www.truthdispatch.com
$0.001386870954 * 7.6$ =0.010	21/15142 = 0.001386870954	$\log_2(16.8B/89.4M)$ =7.61495031333	http://www.risingkashmir.com
$0.00528634361 * 7.6$ = 0.04	12/2270 = 0.00528634361	$\log_2(16.8B/89.4M)$ =7.61495031333	http://www.studiothick.com
$0.0011146201 * 7.6$ = 0.00848777667	12/10766 = 0.0011146201	$\log_2(16.8B/89.4M)$ =7.61495031333	http://2016candidatesdeclarations.blogspot.com/
$0.00084817642 * 7.6$ = 0.00645882129	12/14148 = 0.00084817642	$\log_2(16.8B/89.4M)$ =7.61495031333	http://www.healthcareinfosecurity.com
$0.00059817556 * 7.6$ = 0.00454613425	4/6687 = 0.00059817556	$\log_2(16.8B/89.4M)$ =7.61495031333	http://www.youngbloodstudios.com
$0.00081015392 * 7.6$ = 0.00615716979	6/7406 = 0.00081015392	$\log_2(16.8B/89.4M)$ =7.61495031333	http://youranonnews.tumblr.com/
$0.00024945654 * 7.6$ = 0.0018958697	7/28061 = 0.00024945654	$\log_2(16.8B/89.4M)$ =7.61495031333	http://www.ohsogeeky.com/
$0.00085719183 * 7.6$ = 0.0065146579	5/5833 = 0.00085719183	$\log_2(16.8B/89.4M)$ =7.61495031333	http://www.egham-airtrack.co.uk
$0.0016594756 * 7.6$ = 0.01261201456	5/3013 = 0.0016594756	$\log_2(16.8B/89.4M)$ =7.61495031333	http://miningmediagroup.com/

Ten random URI page ranking from Twitter API URI list. All had a positive keyword search for “government.”

Pagerank - Google	URI
0.0	http://www.truthdispatch.com
0.0	http://2016candidatesdeclarations.blogspot.com/
0.0	http://miningmediagroup.com/
0.0	http://www.ohsogeeky.com/
0.1	http://www.egham-airtrack.co.uk
0.2	http://www.youngbloodstudios.com
0.3	http://www.studiothick.com
0.4	http://youranonnews.tumblr.com/
0.5	http://www.healthcareinfosecurity.com
0.6	http://www.risingkashmir.com

Page ranking does not provide much information. It would appear that there is no correlation between page ranking and TFIDF.

Extra Credit

Kendall tau Rank Correlation	
Kendall tau	-0.453539431095123
2-sided p-value	0.0951447412371635