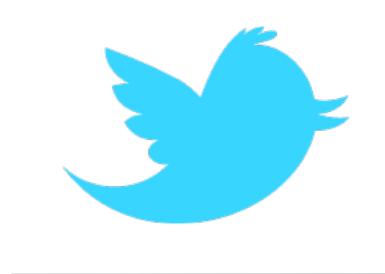
TwitNews: Ranking Tweets Using Structured Social Data



Chris Moghbel Lei Jin



Introduction

The Birth of Twitter

- Introduced publicly July 15, 2006
- Originally designed as a means of sharing small tidbits of personal information
- Motto: "What are you doing?"
- > Trends introduced in April 2009



The Rise of Twitter and Microblogging

- Invaluable tool in recent protests and revolutions in Iran, Egypt, and Tunisia
- Current record is 7,196 tweets per second during the 2011 FIFA Women's World Cup
- Used to spread information about various breaking news such as Michael Jackson's death



Twitter Today

- Grown from 400,000 tweets per quarter in 2007 to 200 million tweets a day in 2011
- New Motto: "What's happening?"
- ➤ November 2010, co-founder Biz Stone mentioned, for the first time, the idea of a Twitter news service

Tweets · Top ▼ Refine results »

20 new tweets



MsDionne83 Dionnell or Dee

True story...RT @Str8ChillenN804: #inthe90s u watched all da "house party" movies....lol

1 minute ago



stefanoalgieri stefano algieri

#inthe90s I was wheeling around in nothing but my spiderman undies wish I could do tha now and not get frowned upon #idkilltobelittleagain

1 minute ago



LiveLikeRowena Rowena Domingo
#inthe90s our social network was actually going outside to play



nlabella nicolina labella #inthe90s I was oblivious to the corruption in the world 1 minute ago



Siddyboo sidney beasley
#inthe90s things for meh weree wayyy easier #MissMyChildhood
1 minute ago



Uwe_x007 Y.B
#inthe90s I made a valentines day for my first crush..sprayed it with
my mother's perfume too

1 minute ago



1 minute ago



JET_LIFE_KING TheJETking

#inthe90s the realest generation was born RT if you think so

Breaking News via Twitter

I want to harness the speed at which news propagates on Twitter to stay up to date on the latest breaking news...

...but how do I find the highest quality results from this massive influx of Tweets?

Tweets · Top ▼ Refine results ×

20 new tweets



maddoghall maddog Hall

Dennis Ritchie, Creator of the "C" language and co-xreator of Unix is dead at age 70.....all programmers owe him a moment of silence.

17 hours ago



Retweeted 100+ times



Yardboy Yardboy

The first few comments of the Wired article about **Dennis Ritchie**'s passing are pretty good. vurl.me/GVYI

5 minutes ago



washingtonpost The Washington Post

Creator of C programming language and Unix dies at 70: wapo.st/qraRTx

10 minutes ago



sfranklin1717 Skip Franklin

Goodbye, Mr. Ritchie, you will be missed. But the C legacy will live on for a good long time. sgp.cm/6d3e75

13 minutes ago



pdp7 Drew Fustini

Unix & C were products of pure research, started w/ equipment bought on a promise that they'd develop a word processor wired.com/wiredenterpris...

37 minutes ago



nitashatiku Nitasha Tiku

The comments written in C on **Dennis Ritchie**'s obituary are pretty adorable wired.com/wiredenterpris...

39 minutes ago



rixstepnews Rixstep News

#dmr created C not out of curiosity but because B hung the 1st Unix system. arstechnica.com/business/news/...

39 minutes ago



auartara Famanda Omentara

Problem: Finding Relevant Information

There's a lot of noise on Twitter...



Is there some way we can automatically find breaking news and filter out the most relevant content?



- Related Work
- Framework
- Ranking Algorithms
- Results
- Conclusion



- Related Work
- Framework
- Ranking Algorithms
- Results
- Conclusion

TrustRank

$$TR(p_i) = d \cdot \sum_{j} TR(p_j) / c_j + b_i$$

$$b_i = \begin{cases} (1-d)/N_T & \text{if } p_i \text{ is trusted} \\ 0 & \text{otherwise} \end{cases}$$

- Generalization of PageRank formula developed by Page and Brin
- Solves problems such as crawler trap by only randomly jumping to pages within a trusted set

TwitRank: Extracting and Ranking Relevant Tweets For News Events

- Similar project regarding finding breaking news from within Twitter data
- Implementation differs utilizes machine learning and human involvement in solution
- Determined that work was inconclusive

Topical Semantics of Twitter Links

- Retweets provide a better indication of topical relevance than follower counts
- Retweets can be considered as votes of relevance

An Empirical Study on Learning to Rank of Tweets

URL links are good indications of topical relevance



- >> Related Work
- Framework
- Ranking Algorithms
- Results
- Conclusion

Tweet Speak: RT @ #!

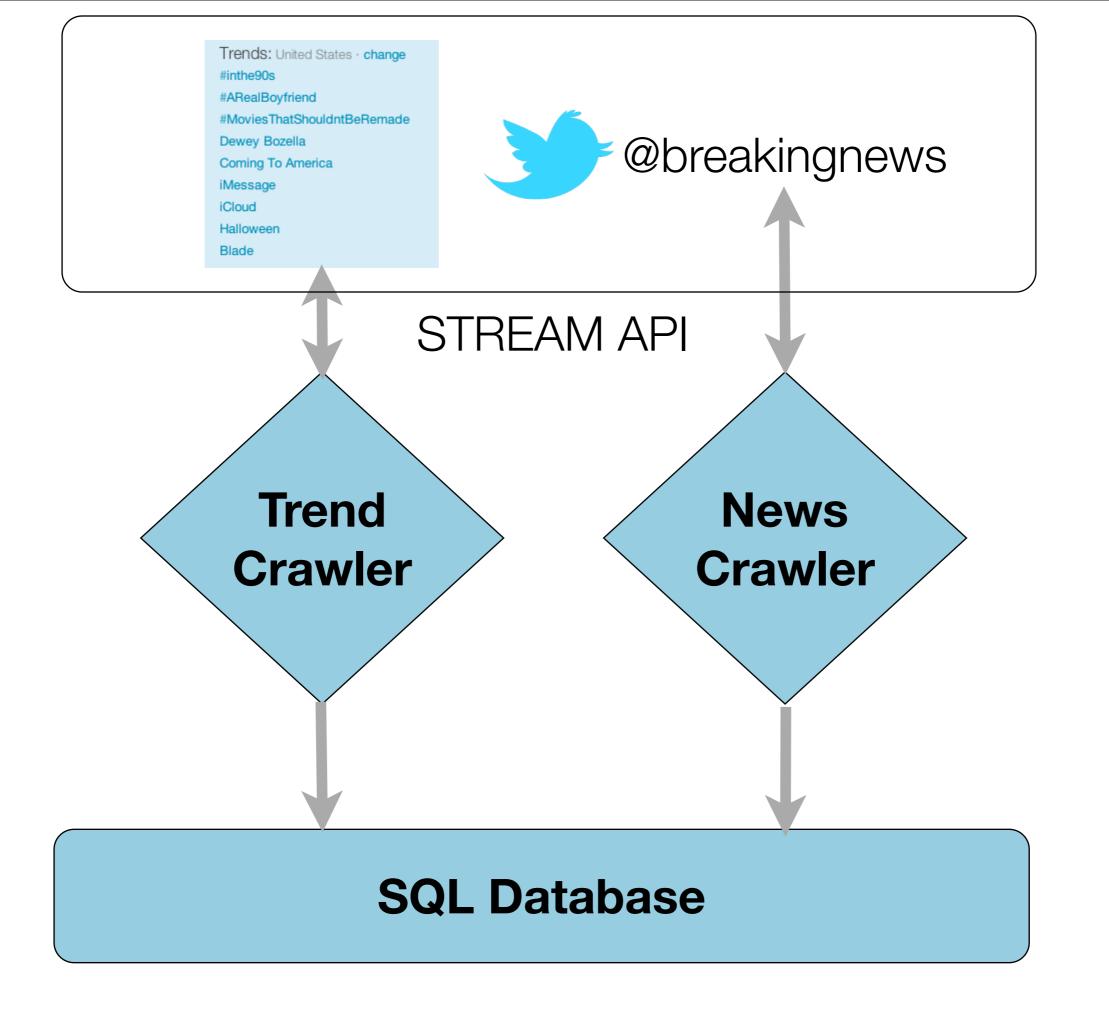
- A Tweet is a 140 character limited snippet that allows users to express comments to the public and exchange information. Also known as a microblog.
- "RT" or "Retweeting" a Tweet allows users to share that Tweet with their followers.
- "#" or "Hashtags" are symbol initiated keywords that mark specific topics in a Tweet.
- Verified users are Twitter users who have had their accounts manually verified by Twitter in order to certify their accounts as authentic and trustworthy.

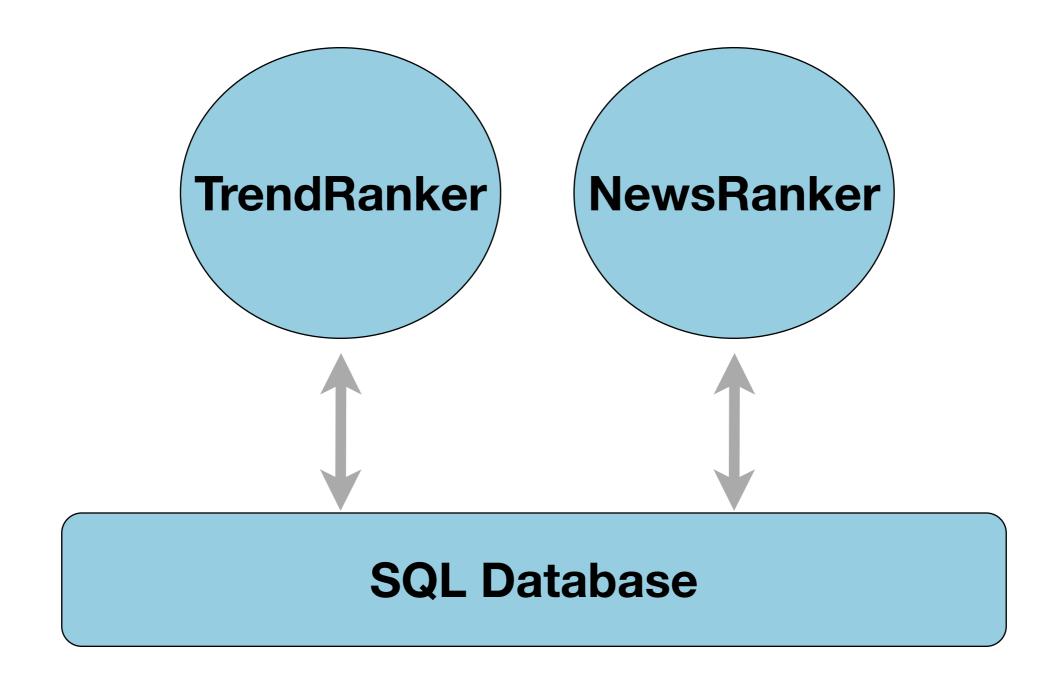
Twitter Trends

Trends: United States · change
#inthe90s
#ARealBoyfriend
#MoviesThatShouldntBeRemade
Dewey Bozella
Coming To America
iMessage
iCloud
Halloween
Blade

Twitter surfaces 10 "Trends", which correspond to the current most popular keywords being used on Twitter.

Trends are often used to aggregate user submissions on a given topic.







- >> Related Work
- >> Framework
- Ranking Algorithms
- Results
- Conclusion

A Reminder...

$$TR(p_i) = d \cdot \sum_{j} TR(p_j) / c_j + b_i$$

$$b_i = \begin{cases} (1-d)/N_T & \text{if } p_i \text{ is trusted} \\ 0 & \text{otherwise} \end{cases}$$

TrustRank

TrustRank --> Naive TwitNews Ranking

$$RANK(T_i) = \alpha \cdot \log\left(\frac{\sum rt_i}{RT}\right) + \beta \cdot v_i$$

- >> Verified Users are equivalent to a trusted source
- Retweets are considered votes of confidence, and we equate the sum of retweets to sum of TrustRank from all in-links in the TrustRank

TwitNews Ranking Algorithm

$$RANK(T_i) = \alpha \cdot \log\left(\frac{\sum rt_i}{RT}\right) + \beta \cdot v_i + \gamma \cdot l_i + \delta \cdot \log\left(\frac{f_i}{F}\right)$$

 $rt_i = A$ retweet of T_i

 f_i = Number of followers of the tweeter, at the time of the tweet

RT = The maximum number of retweets, limited to 101 by the API

F = The number of followers of the most followed user on Twitter

$$\alpha + \beta + \gamma + \delta = 1$$

Integrating Our Algorithm with Existing IR Models

- Created simple Boolean model for tweets based on keyword matching.
- Defined keywords as any proper nouns parsed from the Tweet text.
- Combined simple Boolean model with existing TwitRank formula to create a more generalized ranking.
- Existing IR model matches Tweets with a given query, and TwitRank refines the results by ordering on relevance.

Spam Detection and Other Improvements

- Penalize rank based on number of tokens matching a list of "bad words" such as adult language
- Penalize rank of Tweets with an inordinate number of Trend names included within Tweet
- Eliminate "stop words" before parsing keywords to improve Boolean matching model.



- >> Related Work
- >> Framework
- Ranking Algorithms
- Results
- Conclusion

Experiment 1: Ranking Tweets for Trends

- Crawled current trends from Twitter.
- Crawled tweets for each trend downloaded.
- Applied ranking algorithm plus spam detection to rank tweets for each trend.
- Data set of > 1.5 Million Tweets

enoitules ... Advanced search Рпузсу About Google Search Help Give us feedback C000000000816 Next VEW IOI ISTIED & OD do a barrel roll gif do a barrel roll ytmnd bnuoz lior lamed a ob Eqm llor lansd a ob do a barrel roll collegehumor do a barrel roll song breodbruoe lor lemed a ob Searches related to do a barrel roll Mewsweek and The Daily Beast shared this ... fart — that — tecently posted by his daughter — that ... Nov 3, 2011 - "It looks worse than it is." — Texas family law judge William Adams, newsweek.tumblr.com/.../go-to-google-com-and-search-do-a-barrel-r... Newsweek (Go to Google.com and search "Do a barrel roll") sirt benefic omeM striog Points Terr tech bloggers discovered that typing in the query "do a barrel roll" ... Nov 3, 2011 - Google search was furned upside down quite literally on Thursday whenfgeslab.falkingpointsmemo.com/.../google-explains-barrel-roll-search... Google Explains Barrel Roll Search Command | TPM Idea Lab ... 8 mort gring the flightpath. Starting from 8 ... to do a barrel roll from the pilot's perspective, it may best be thought of as a roll of as a roll of the decided and the decided of the forms of the decided of the forms of the decided en.wikipedia.org/wiki/Barrel_roll +1 Barrel roll - Wikipedia, the free encyclopedia Nov 4, 2011 - Google's latest Easter egg turned into an Internet meme. T+ qse.00, r888865, r185, 0\2910\instruction \text{(00.asp. 11)} Like Google's 'Do a Barrel Roll'? You'll Love These Other Easter ... Google still finds time to make our lives more fun. In a week when its latest. 4 days ago – It takes care of more than 1 billion searches worldwide every day but Do a barrel roll: Google's best Easter eggs | Metro.co.uk Defeat Alystazor in the Firelands without any raid member getting hit by the following 1101-lamed-s-ob\E182=Inemevairios\moo.bsedwow.www Do a Barrel Roll! - Achievement - World of Warcraft "Do a Barrel Roll" search trick, which went viral on Twitter and brightened a ... s hanathorid has rattiut no lesiv tnew dainly Not de got a feet on Thursday with the highly popular by Chris Taylor - In 12,676 Google+ circles Add to circles

mashable.com/2011/11/03/google-easter-eggs-2/ Beyond Barrel Roll: 10 Hidden Google Tricks

is known for being whimsical, but "do a barrel roll" might just take the cake. by Tim Stenovec - In 838 Google+ circles Add to circles Nov 3, 2011 – With its interactive doodles and incerner is known for Nov 3, 2011 - With its interactive doodles and insanely large pumpkins, Google/google-do-a-barrel-roll.../mos.feonptoniffud.www

Do a Barrel Roll!

User	Verified?	Followers	Tweet	Link Domain
dmbmeg	No	440	just realized "do a barrel roll" is trend- ing and I just tweeted about Justin Bieber. Please don't unfollow me. I'll do better next time.	None
brunnoleonardo	No	724	Do a Barrel Roll! ?	None
jonalammers	No	131	RT @lisajorinde: RT @NinaSylvie: hahahaa typ in bij google do a barrel roll , je schrikt je dood	None
xddlovatosmile	No	331	Ho digitato "do a barrel roll" su Google,google se ne va di testa ogni tanto.AHAHAHAHAHAH	None
vadro	No	23	RT @Nesttor77: Entren a Google y pongan do a barrel roll! Jajajaja mola mucho xD	None
JeromeParadis	No	2669	Nice Easter egg! Re: Do A Barrel Roll! In Google! Right Now! via @techcrunch http://t.co/mrgOAgMh	techcrunch.c om
DoweyBaothma n	No	117	RT @MishaelSaad: THINGS YOU SHOULDNT DO: Drugs.	None
chellyybean	No	224	umm wtfi want my google to do a barrel roll. WHY WON'T IT WORK?!	None
LizaCRVG	No	118	@JonataanW_OFC - Abra a pagina inicial do google, e procure: do a barrel roll e me diz o que achou =)	None
GlitterMonseterx	No	831	Going on a laptop JUST to search up " Do a Barrel Roll "	None

Top 10 Tweets for trend "DO A BARREL ROLL" using the Timestamp method.

User	Verified?	Followers	Tweet	Link Domain
mashable	Yes	2,568,170	"Do a Barrel Roll" on Google, and You Won't Be Disappointed - http://t.co/pqaDJ4Fi	mashable.co m
Alyssa_Milano	Yes	1,828,268	Go to http://t.co/5QlfhQwX & type "do a barrel roll".The results will entertain you. / via @BuzzFeed	google.com
Kotaku	Yes	100,767	Google "Do a Barrel Roll" Right Now http://t.co/OJMZZMyj	kotaku.com
jennydeluxe	Yes	420,292	Everyone do this NOW rt @ryeclifton Google "do a barrel roll" via @lindseyweber	None
threadless	Yes	1,672,228	Type "do a barrel roll" into http://t.co/ AGwbxxEv. When your screen stops spin- ning, request a reprint of this - http://t.co/ FsISaFwT	google.com
toptweets_es	Yes	277,021	RT @jbrownridge: 1) Go to http://t.co/ 4fy4lVeH 2) Search for "do a barrel roll"	google.com
toptweets_es	Yes	277,020	RT @abrmorales: 1. Entra a http://t.co/ ANAGfAOy 2. Escribe: do a barrel roll 3. Pulsa enter.	google.es
Scrobleizer	Yes	215,917	RT @mashable: "Do a Barrel Roll" on Google, and You Won't Be Disappointed - http://t.co/pqaDJ4Fi	mashable.co m
arrington	Yes	99,944	RT @Techmeme: Do A Barrel Roll! In Google! Right Now! (@mjburnsy / TechCrunch) http://t.co/b2suSooM http://t.co/WZyncXwx	techcrunch.c om
Econsultancy	Yes	74,588	A little fun! RT @lakey: The mother of all Google tricks try entering the search query: "do a barrel roll" http://t.co/6Sr5LQMw	google.com

Top 10 Tweets for trend "DO A BARREL ROLL" using the TwitNews ranking algorithm.



User	Verified?	Followers	Tweet	Link Domain
KayBee0311	No	58	"@BreakingNews: Producers say new James Bond film will be called 'Skyfall,' star Daniel Craig, Javier Bardem - AP" Yay can't wait!!	None
MalcomIngram	No	11,653	James Bond is 007 in SKYFALL?!?! Worst title ever Better one off the top of my head James Bond is 007 in TIME EQUALS DEATH #BetterBondTitles	None
Diianamedez	No	59	Novo filme do James Bond :) Aiii Daniel deli- iiciiiaaa :P	None
CajunBrooke	No	357	New James Bond movie coming out, but not until November 2012. Can I buy tickets now?	None
NorbertKuehn	No	52	RT @HuffingtonPost: Breaking updates: All the new James Bond film details as they happen! http://t.co/6ViOiQ2K	huffingtonpos t.com
digitalspybrk	No	3,887	Skyfall' plot sees MI6 "under attack", James Bond loyalty to M tested http://t.co/dITRYmth	digitalspy.co m
digitalspyent	No	5,013	Skyfall' plot sees MI6 "under attack", James Bond loyalty to M tested http://t.co/ GSCFgT1R	digitalspy.co m
lincingcom	No	56	New James Bond Film is Officially Ti- tled 'Skyfall'; Mendes and Producers Up- date Casting and Plot Details http://t.co/ TugXWONC	showbiz.lincin g.com
Phildaddy007	No	50	Well now Daniel Craig is now a trending topic over James Bond That works, I'll go with that he is the #BestBond	None
Khusomir	No	130	Enough with James Bond, such a boring series	None

Top 10 Tweets for trend "James Bond" using the Timestamp method.

User	Verified?	Followers	Tweet	Link Domain
BBCWorld	Yes	1,182,858	Next James Bond film will be titled Skyfall, producers announce. More soon http://t.co/ jTltwN70	bbc.co.uk
AP	Yes	593,789	James Bond returns next year in "Skyfall," starring Daniel Craig and Javier Bardem. It's 007's 23rd movie: http://t.co/5nKhsM1C -EF	hosted.ap.org
BreakingNews	Yes	3,225,044	Google Producers say new James Bond film will be called 'Skyfall,' star Daniel Craig, Javier Bardem - AP	None
HuffingtonPost	Yes	1,345,737	Breaking updates: All the new James Bond film details as they happen! http://t.co/ 6ViOiQ2K	huffingtonpos t.com
BBCNews	Yes	379,179	Next James Bond film will be titled Skyfall, producers announce. More soon http://t.co/ S4kmCo2G	bbc.co.uk
holacom	Yes	157,398	Comienza el rodaje de 'Skyfall' el nuevo filme de James Bond, con Daniel Craig y Javier Bardem como protagonistas: http://t. co/g1LW82bl	noticias.hola. com
ParisMatch	Yes	53,248	James Bond tombera du ciel: Lors d une confÃl'rence de presse, le titre des 23e aven- tures du cÃl'IÃ lbre agent secret http://t.co/ DXGY0ZTB	parismatch.c om
NoticiasMVS	Yes	34,532	La nueva pelà cula de James Bond se lla- marÃą 'Skyfall' http://t.co/XXMDD0Jm	noticiasmvs.c om
MetroUK	Yes	18,863	James Bond film title announced as Skyfall at launch with Daniel Craig http://t.co/ EL9rePnJ	metro.co.uk
YahooOmgUk	Yes	6,555	The James Bond/Daniel Craig news WE'D been waiting for http://t.co/koAiVBFw	uk.omg.yaho o.com

Top 10 Tweets for trend "James Bond" using the TwitNews ranking algorithm.

Experiment 2: Finding and Ranking Tweets Related to News

- Crawled latest tweets from @breakingnews
- Used these tweets as example queries
- Crawled tweets with matching keywords from Twitter
- Applied Boolean model combined with TwitNews ranking algorithm to find tweets for a given "query" and rank them.
- Data set of > 2 Million Tweets.

User	Verified?	Followers	Tweet	Link Domain
mdulian	No	786	ESPN is reporting that it looks like Joe Paterno will be gone! RT @Elliott Sadler: What is Penn State goin (cont) http://t.co/HRXHgBOI	None
1DSheeranBiebs	No	1251	@AwwMusic when I have credit and I'm with Joe next I'll ring you and you can talk to him hahahhaaa	None
LordFestusIII	No	52	RT @speakz: RIP Joe Fraizer.Ali called you an uncle tom & a paper champion in build up to your 1st fightYou responded by beating his	None
NewsDangler	No	52	I'm at Nyack Beach State Park (698 N Broadway, Upper Nyack) w/ 2 others [pic]: http://t.co/ HXEJ5Dho	https:// foursquare.com
jOleary20	No	170	Hope everything works out for Penn State @ItsMeSnitchez10	None
gonzalezmam	No	2128	Penn State's Paterno cancels briefing amid scandal http://t.co/ajjYsbZ9	http://reuters.com
freehan11	No	542	This scandal at Penn State is truly disgust- ing. No one ever learns that coverups al- ways come uncovered. God Bless the vic- tims.	None
avalam00	No	21	RT @mark may: The next AD at Penn State should be Matt Millen	None
Bojandles	No	34	Penn State should get the death penalty Sick!	None
DatDudeNickT93	No	48	Guys, Real Joe Pa has been dead now for some time. They simply are propping his corpse up in the booth on GameDays #WeekendAtBernies Style	None

Top 10 Tweets for query "Penn State said to be planning football coach Joe Paterno's departure, sources say - @NYTimes http://t.co/1KAsWAZu" using the Timestamp method.

User	Verified?	Followers	Tweet	Link Domain
kimtee	No	103	"@BreakingNews: Penn State said to be planning football coach Joe Paterno's de- parture, sources say - @NYTimes http://t. co/sOCyq3pg WHAT.	nytimes.com
JustRizz	No	165	RT @SportsCenter: Breaking News: The New York Times: Penn State is planning head coach Joe Paterno's exit http://t.co/ K5AfLhqQ	nytimes.com
MsMoo_Pow	No	141	RT @SportsCenter: Breaking News: The New York Times: Penn State is planning head coach Joe Paterno's exit http://t.co/ K5AfLhqQ	nytimes.com
BrwnEyeGemini	No	151	@2 shay my nigga The New York Times: Penn State is planning head coach Joe Pa- terno's exit http://t.co/9DigGXjm ;; that ish right there	nytimes.com
WrnrG	No	26	Report: Penn State planning Joe Pa- ternos exit amid sexual-abuse scan- dal - College Football News — FOX Sports on http://msn.foxsports.com/ collegefootball/ story/Penn-State-planning- Joe-Paterno-exit-amid-sexual-abuse - MSN: http://t.co/IQSzSoBm	msn.foxsports.com
nthartness6	No	605	RT @SportsCenter: Breaking News: The New York Times: Penn State is planning head coach Joe Paterno's exit http://t.co/ K5AfLhqQ	nytimes.com
jcr4522	No	60	SportsCenter: Breaking News: The New York Times: Penn State is planning head coach Joe Paterno's exit http://t.co/rn7Yfz8M	nytimes.com
DHarguth	No	155	RT @SportsCenter: Breaking News: The New York Times: Penn State is plan- ning head coach Joe Paterno's exit url- http://t.co/K5AfLhqQ	nytimes.com
D2Lo	No	41	RT @SportsCenter: Breaking News: The New York Times: Penn State is planning head coach Joe Paterno's exit http://t.co/ K5AfLhqQ	nytimes.com
final118	No	64	RT @USATODAY: Penn State cancels Joe Paterno's scheduled news conference http://t.co/QuOhFXr8	content.usatoday.com

Top 10 Tweets for query "Penn State said to be planning football coach Joe Paterno's departure, sources say - @NYTimes http://t.co/1KAsWAZu" using the Boolean model.

User	Verified?	Followers	Tweet	Link Domain
USATODAY	Yes	189,806	Report: Penn State planning Joe Paterno's exit http://t.co/zVEhlm3s	content.usatody.com
USATODAY	Yes	189,804	Penn State cancels Joe Paterno's scheduled news conference http://t.co/QuOhFXr8	content.usatody.com
michaelombardi	Yes	133,191	According to the NY Times Penn State is planning for Joe Paterno's exit. http://t. co/gT3volwC These is a horrendusly sad story	nytimes.com
YahooSports	Yes	40,747	Paper: Penn State planning Joe Paterno's exit http://t.co/KmBM17y2	rivals.yahoo.com
ESPNNewYork	Yes	33,045	Penn State Nittany Lions call off Joe Paterno's news conference http://t.co/ Z0sVeYRT	espn.go.com
vleach44	Yes	16,125	RT @SportsCenter: Breaking News: The New York Times: Penn State is planning head coach Joe Paterno's exit http://t.co/ DGsPU6zT	nytimes.com
CBSi	Yes	10,712	Report: Penn State planning Joe Paterno's exit - http://t.co/BRkBG5WD	eye-on- collegefootbal.blogs. cbssports.com
OpieRadio	No	105,719	YAWN! knew this 18hrs ago! BELIEVE RT @BreakingNews Penn State planning football coach Joe Paterno's departure @NYTimes http://t.co/9VaXtLXk	nytimes.com
evertus	No	50,069	Penn State president cancels Joe Paterno's press conference (morningcall): Share With Friends: — Top News http://t.co/7KbKSxFy	mcall.com
hackhype	No	29,468	Report: Penn State planning coach Pa- terno's exit: Joe Paterno's tenure as head coach of Nittany Lions is likely http://t.co/lyXotQWN	nfl.com

Top 10 Tweets for query "Penn State said to be planning football coach Joe Paterno's departure, sources say - @NYTimes http://t.co/1KAsWAZu" using the combined Boolean and TwitNews ranking algorithm.



- >> Related Work
- >> Framework
- Ranking Algorithms
- > Results
- Conclusion

Recap

- Developed ranking algorithm for Tweets that takes into account structured social data
- Performed much better than timestamp or simple IR methods
- Can be combined with existing IR methods to improve search in microblogging systems

Thanks!

Q&A