Twitter Data

Gilad Mishne - @gilad

Twitter Search

Info 290 - Analyzing Big Data With Twitter

UC Berkeley Information School

August 2012
About Twitter

• The fastest, simplest way to communicate

• More than 140M active users
  • Majority (also) mobile; 60% out of U.S.

• More than 400M twitter.com visitors

• More than 400M tweets/day (peak: 25K/sec)

• 1,000 employees (majority in San Francisco)
  • 50% engineers
Twitter data: text
Twitter data: social graph

Credit: @isaach
Twitter data: time series
Twitter data: interest graph

Credit: @psychemedia
Combined: the pulse of the world

Paul Ryan: How the Conversation Took Off
Follow @gov for more about government & politics on Twitter.
What we do with large data
Scale
Search

Twitter search results for "SFGiants"

Top people:
- San Francisco Giants @SFGiants
  *Official Twitter home of the 2010 World Cham...*

Tweets:
- America's Cup @AmericasCup
  Do you think we are going to see any capsizes today? #AmericasCup
  pic.twitter.com/unbquPsl
  Promoted by America's Cup

- MLB Fan Cave @MLBFanCave
  Followers to Retweet are eligible to win a baseball signed by @SFGiants great Juan Marichal. #MLBFC

- San Francisco Giants @SFGiants
  #SFGiants now lead 8-2 in the 8th inning.

Recommendations

Who to follow · Refresh · View all

Longreads @longreads
Followed by Florian Leibert and
Follow

Adam Sharp @AdamS
Followed by Ana-Maria Popescu
Follow

George Takei @GeorgeTakei
Followed by Startup L. Jackson
Follow

Browse categories · Find friends

Tweets
What's happening now, tailored for you.

The Economist
@TheEconomist
Following

Indiana has seen a quiet whirlwind of education reform econ.st/O67qts

Extreme couponing

IN THE summer of 2011 a 16-year-old girl called Dayana Vazquez-Bucuer arrived at the reception desk of Roncalli High School, a nice private school in the south side...

25 RETWEETS 11 FAVORITES

10:06 PM - 22 Aug 12 via SocialFlow · Details

Curiosity Rover
@MarsCuriosity

Follow
Anti-Spam

Jan Jones
@JanJones1727

get free sample viagra - EXTRA LOW PRICES.
viagra > Save Your Money
bit.ly/OUaEzM

Reply Retweet Favorite

5:07 PM - 22 Aug 12 via twitterfeed · Embed this Tweet
The Speakers

(Plus, some Twitter features)
Twitter Overview

Raffi Krikorian
@raffi
Director of @twittereng’s Platform Services. I breed in San Francisco, California. http://t.co/eNmM6q

Othman Laraki
@othman
growth @ twitter by profession / kiter by addiction. Here a entrepreneur by passion / casablanca native by origin. San Francisco, CA.

Expanded Tweets

Raffi Krikorian
@raffi
was told to go on vacation. I am pretty stressed out by this. (@ Kahului Airport (OGG) w/ 5 others) 4sq.com/NNoQg6

A check-in at Kahului Airport (OGG)
Airport in Kahului, HI
foursquare @foursquare · Follow

Othman Laraki
@othman
More people in the world own cell phones than tooth brushes. I think we need to rethink Maslow’s hierarchy of needs.
mindjumpers.com/blog/2012/01/s...

Embedded links
Hadoop/Pig

Replies/conversations
Trends/Streaming

Protected accounts

Support for 30+ languages
Accidentally created an unwanted subdirectory named '~'. See if you can guess where this is going next...
Graphs, Recommendations, Relevance

Retweets (in the timeline)
Security, anomaly detection

Photos
Scalding

Geotagging, hashtags
Goals

- Work with real data, on real problems
- Learn how it is to work in a place like Twitter
- Build something useful
- Have a good time!
Questions?

Follow me: @gilad