Social Media & Text Analysis

lecture 2 - Twitter and Twitter API Tutorial

Instructor: Wei Xu

Website: socialmedia-class.org

Twitter API Tutorial: socialmedia-class.org

Social Media & Text Analytics | Syllabus | Twitter API Tutorial | Homework Assignments

Twitter's 404 error page -- the Fail Whale

Twitter API tutorial

by Wei Xu (July 1, 2015)

Follow @cocoweixu

1. Getting Twitter API keys

To start with, you will need to have a Twitter account and obtain credentia
Twitter developer site to access the Twitter API, following these steps:

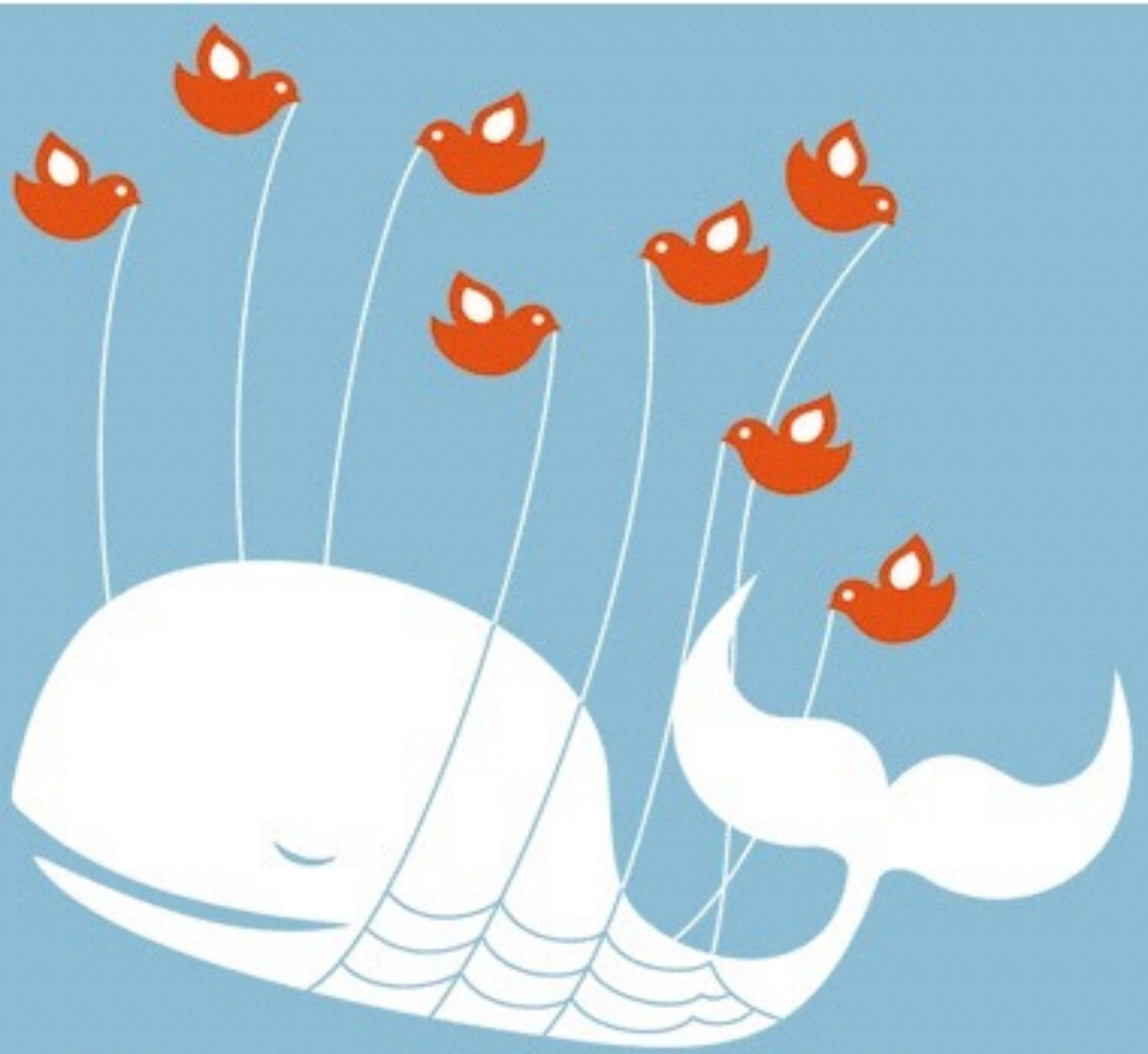
- Create a Twitter user account if you do not already have one.
- Go to https://apps.twitter.com/ and log in with your Twitter user acc
  your user account.
- Click "Create New App"

twitter

Home ›

Twitter is over capacity.

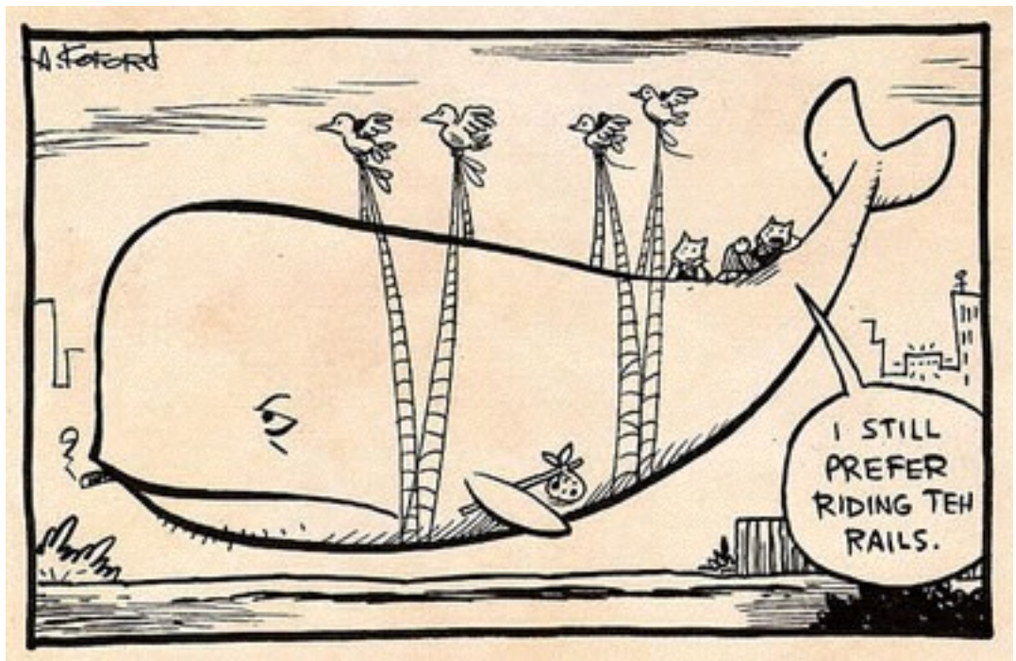
Please wait a moment and try again. For more information, check out Twitter Status »



English Deutsch Español Français Italiano 日本語

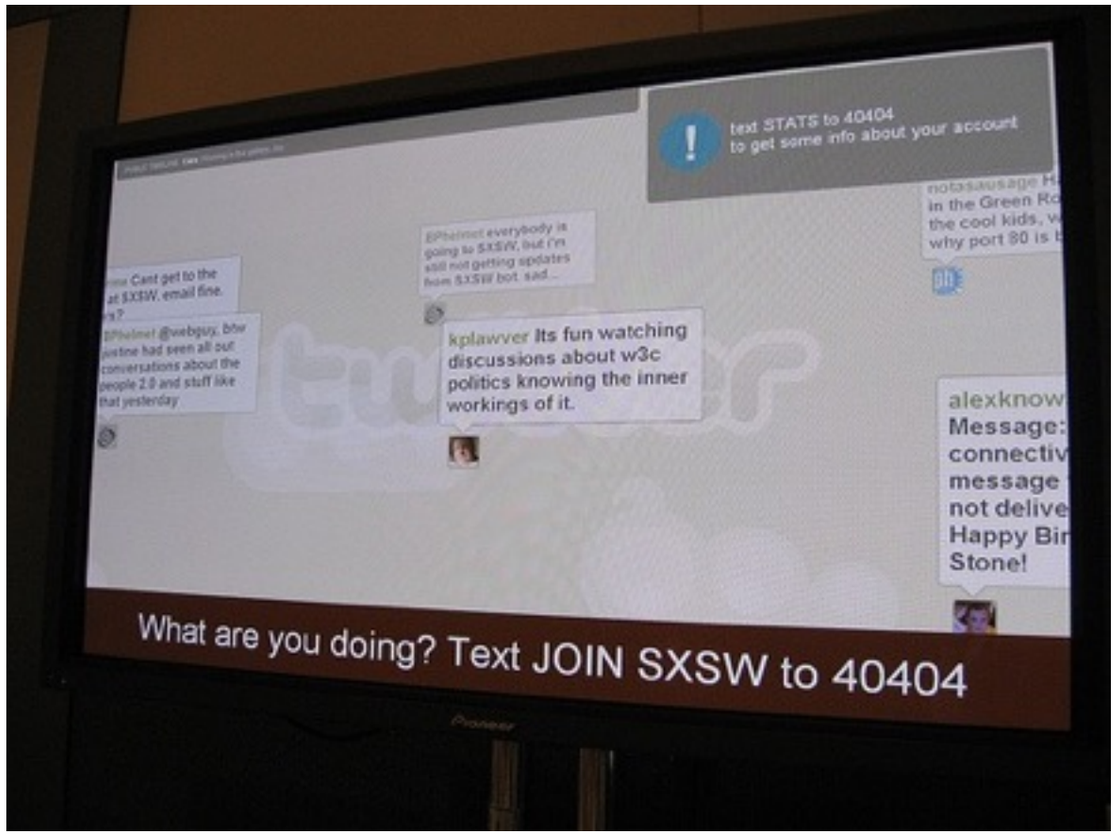
© 2011 Twitter About Us Contact Blog Status API Help Jobs TOS Privacy

todder
Todder had a long night.
Too many drinks! Please wait for the Tylenol to kick in and try again.


I STILL PREFER RIDING TEH RAILS.

Twitter History

- Jack Dorsey's idea
  (a NYU undergraduate then)
- 1st tweet on March 21, 2006
- exploded at SXSW 2007
  (20k→60k tweets/day)
- 100m tweets/quarter in 2008,
  50m tweets/day in 2010,
  400m tweets/day in 2013
- Huge API usage was
  unexpected as was the rise of
  the @ sign for replies



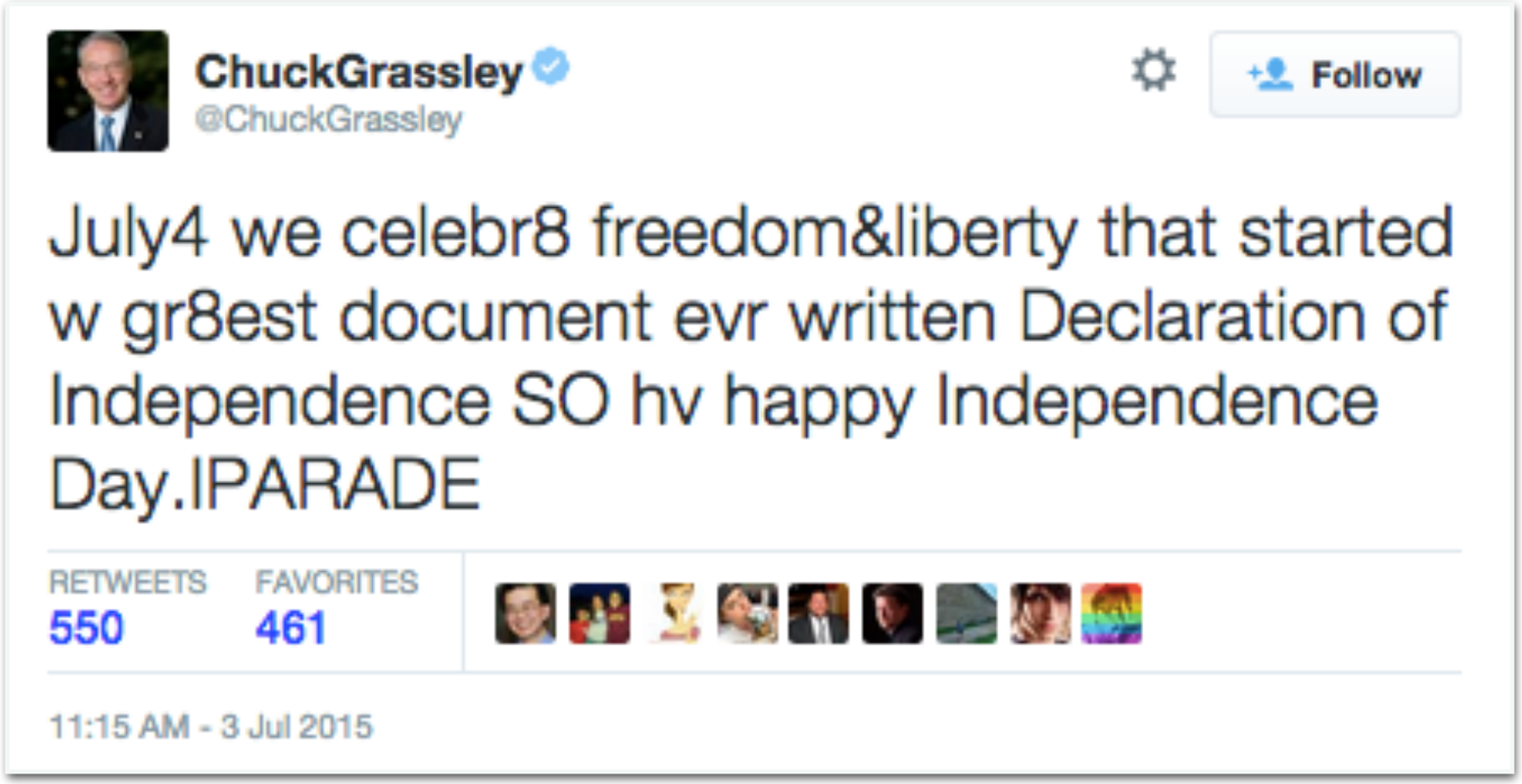
Twitter staff received the festival's Web Award prize with the remark "we'd like to thank you in 140 characters or less. And we just did!"

Twitter History

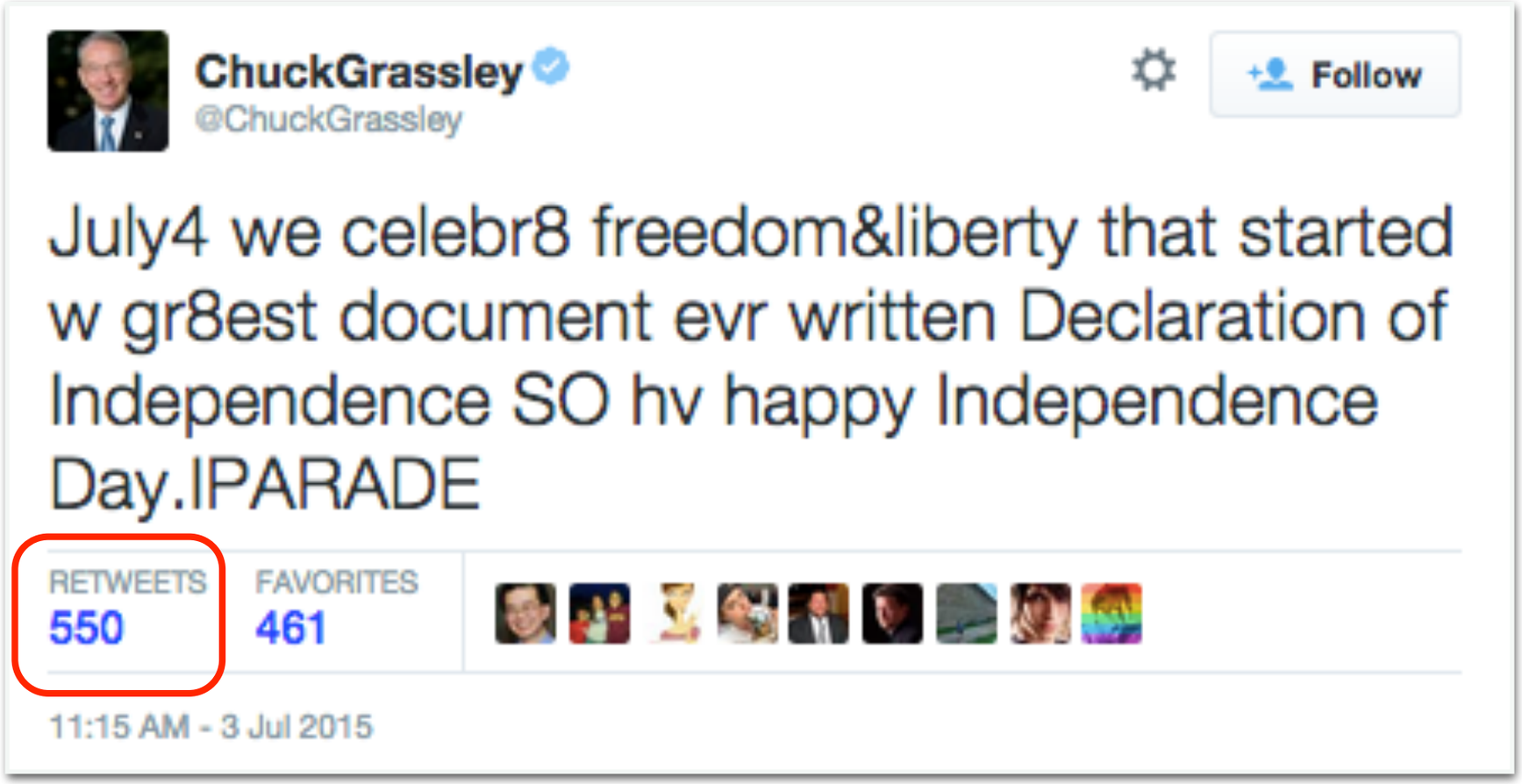
- IPO in 2013 Q4
- market value $24b, revenue $435m, net loss $162m in 2015 Q1
- CEO Dick Costolo resigned July 1st, 2015

Twitter Data

Tweets


ChuckGrassley
@ChuckGrassley
Follow
July4 we celebr8 freedom&liberty that started w gr8est document evr written Declaration of Independence SO hv happy Independence Day.IPARADE
RETWEETS
550
FAVORITES
461
11:15 AM - 3 Jul 2015

ReTweets

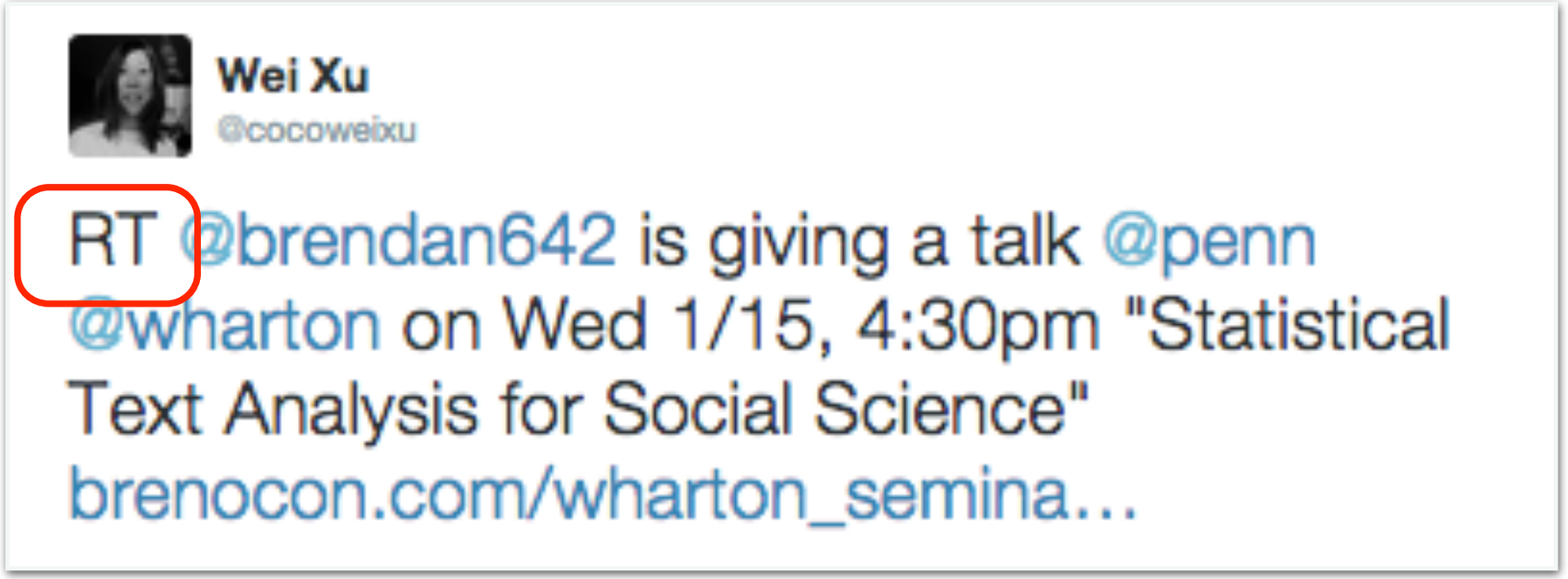


a re-posting of someone else's Tweet

Disclaimer:

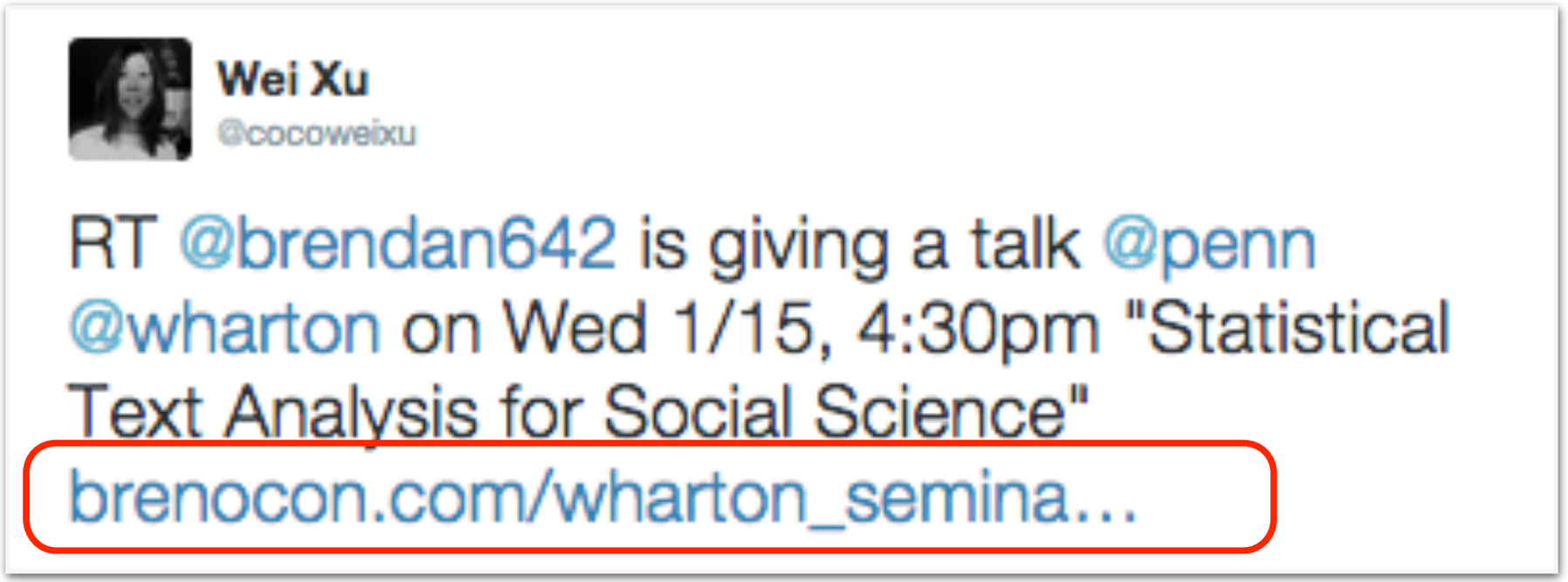
I am going to use many of my own tweets as examples in the slides …

ReTweets



- not an official Twitter feature
- often signifies quoting another user
- sometimes creates problems for data analytics

Embedded Links



- shortened for display

Embedded Links

- can provide extra external information for text processing

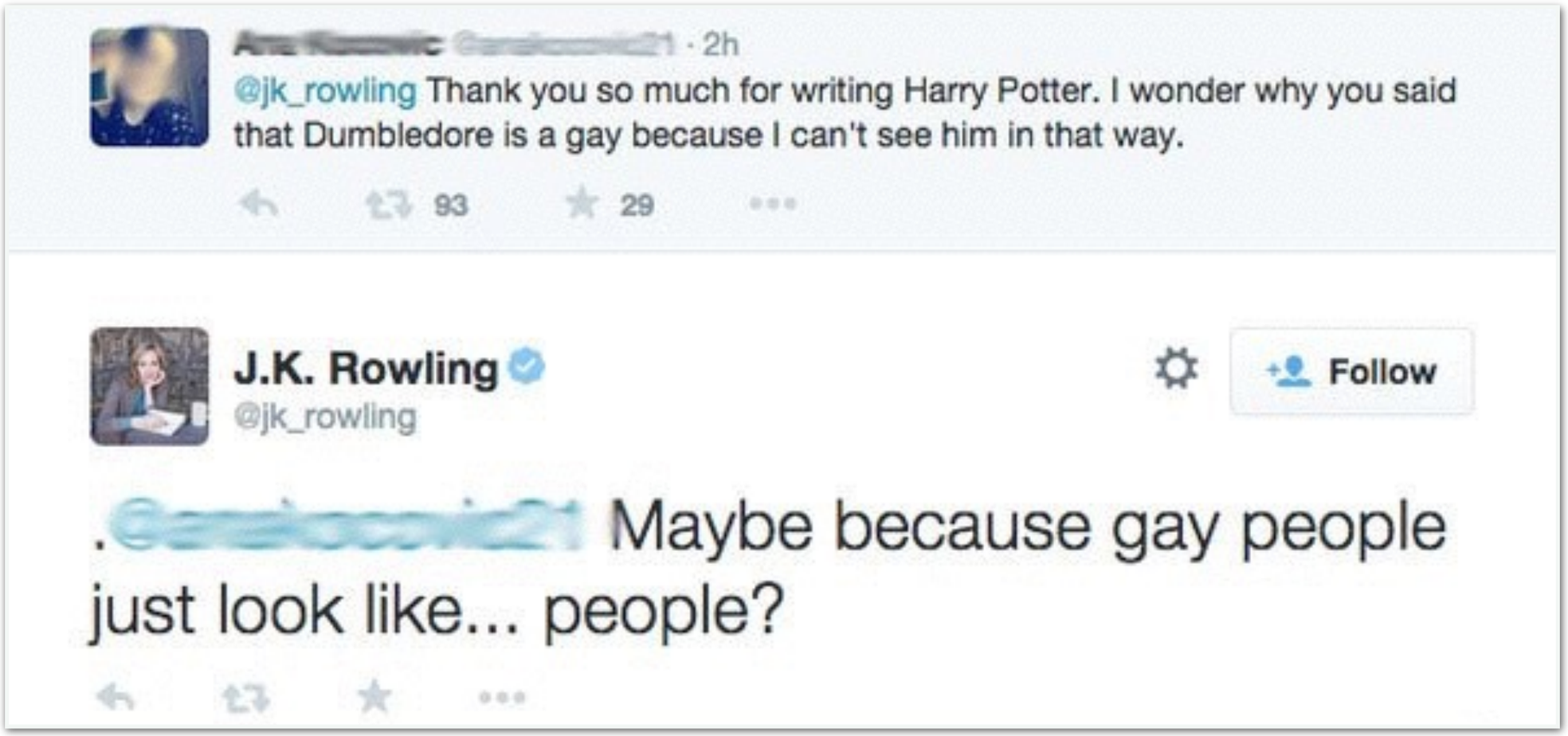
Mentions

Wei Xu
@cocoweixu

RT @brendan642 is giving a talk @penn @wharton on Wed 1/15, 4:30pm "Statistical Text Analysis for Social Science" brenocon.com/wharton_semina…

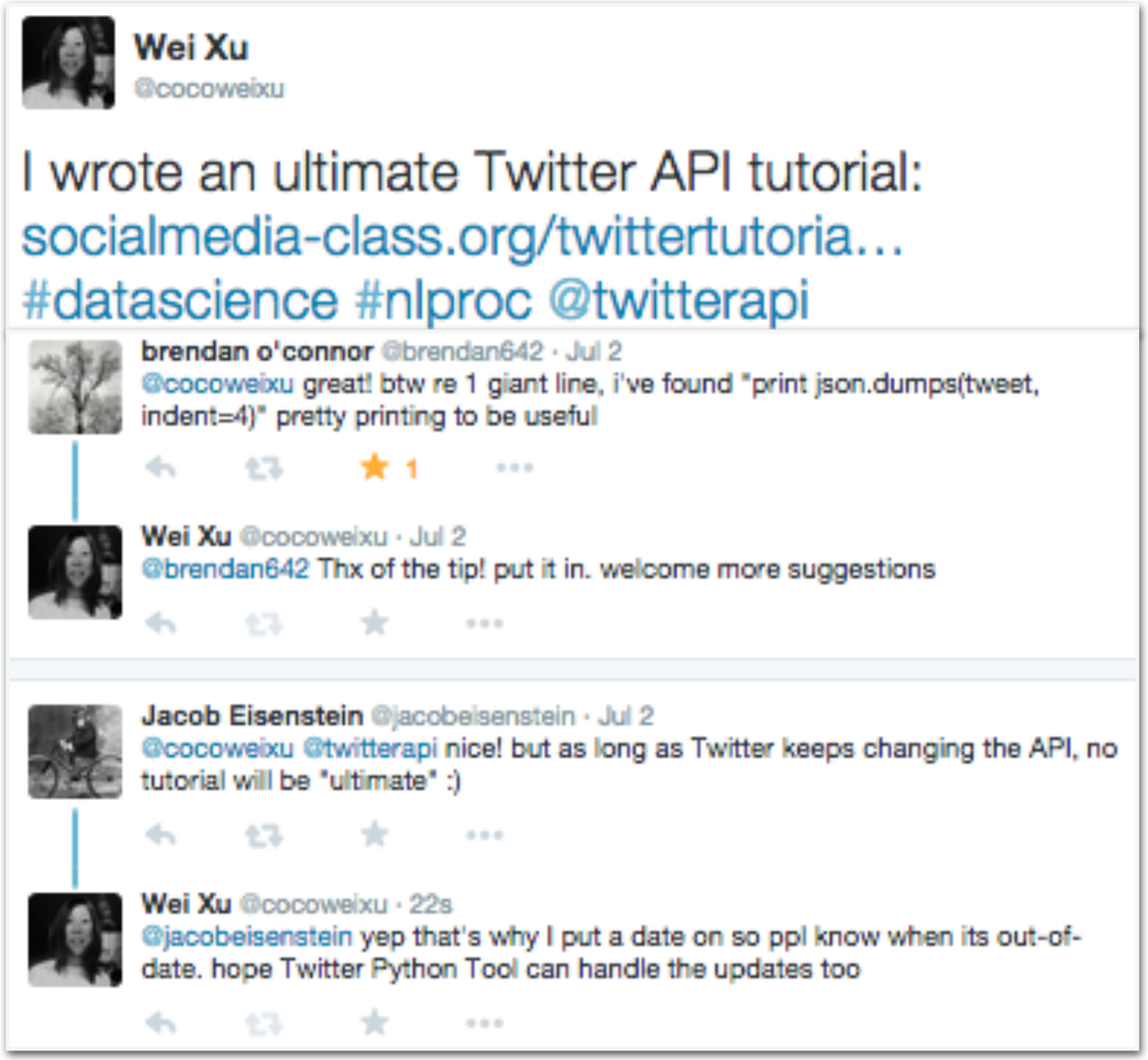
- user's @username anywhere in the body of the Tweet

Replies/Conversations



- Tweet starts with a @username

Replies/Conversations



- can have multi-round conversations

Images

Wei Xu
@cocoweixu
I wrote an ultimate Twitter API tutorial:
socialmedia-class.org/twittertutoria…
#datascience #nlproc @twitterapi
View translation
Social Media & Text Analytics
Syllabus
Twitter API Tutorial
Homework Assignments
Twitter's 404 error page -- the Fail Whale
Twitter API tutorial
by Wei Xu (July 1, 2015)
Follow @cocoweixu
1. Getting Twitter API keys
To start with, you will need to have a Twitter account and obtain credentia
Twitter developer site to access the Twitter API, following these steps:
Create a Twitter user account if you do not already have one.
Go to https://apps.twitter.com/ and log in with your Twitter user acc
your user account.
Click "Create New App"
RETWEETS
27
FAVORITES
53
11:55 AM - 2 Jul 2015

Hashtags

Wei Xu
@cocoweixu

I wrote an ultimate Twitter API tutorial:
socialmedia-class.org/twittertutoria…
#datascience #nlproc @twitterapi

View translation

Social Media & Text Analytics | Syllabus | Twitter API Tutorial | Homework Assignments

Twitter's 404 error page -- the Fail Whale

Twitter API tutorial

by Wei Xu (July 1, 2015) Follow @cocoweixu

1. Getting Twitter API keys

To start with, you will need to have a Twitter account and obtain credentia
Twitter developer site to access the Twitter API, following these steps:

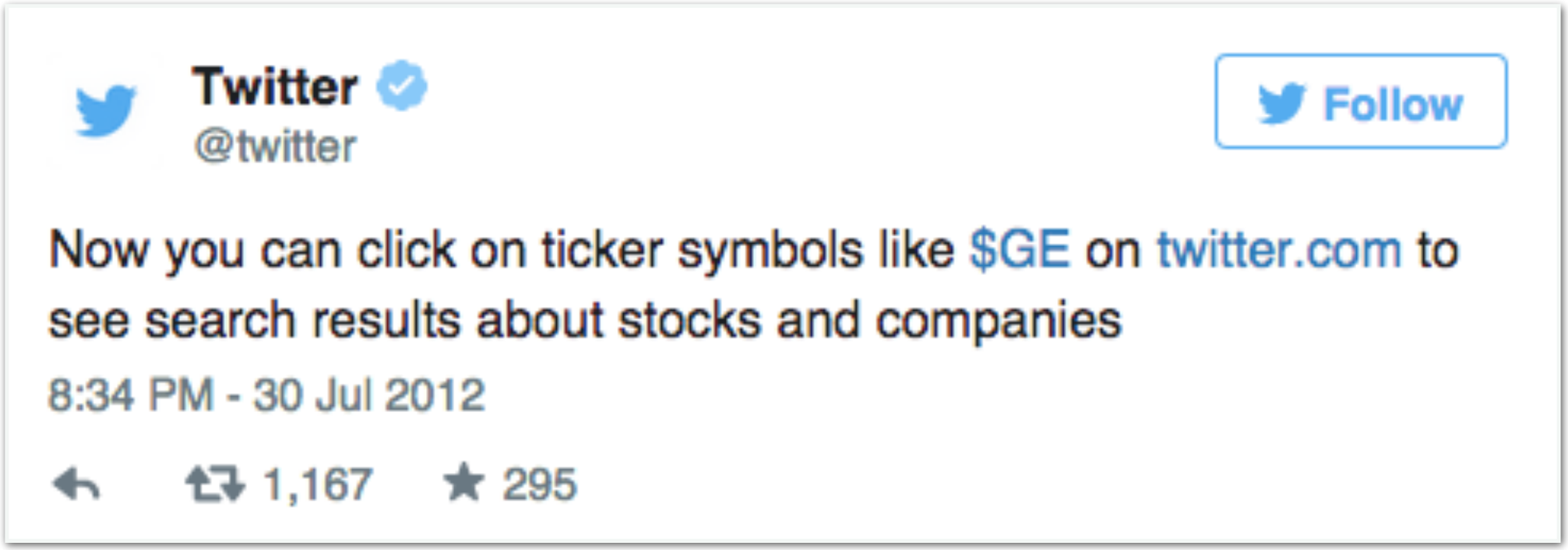
- Create a Twitter user account if you do not already have one.
- Go to https://apps.twitter.com/ and log in with your Twitter user acc
  your user account.
- Click "Create New App"

RETWEETS 27 FAVORITES 53

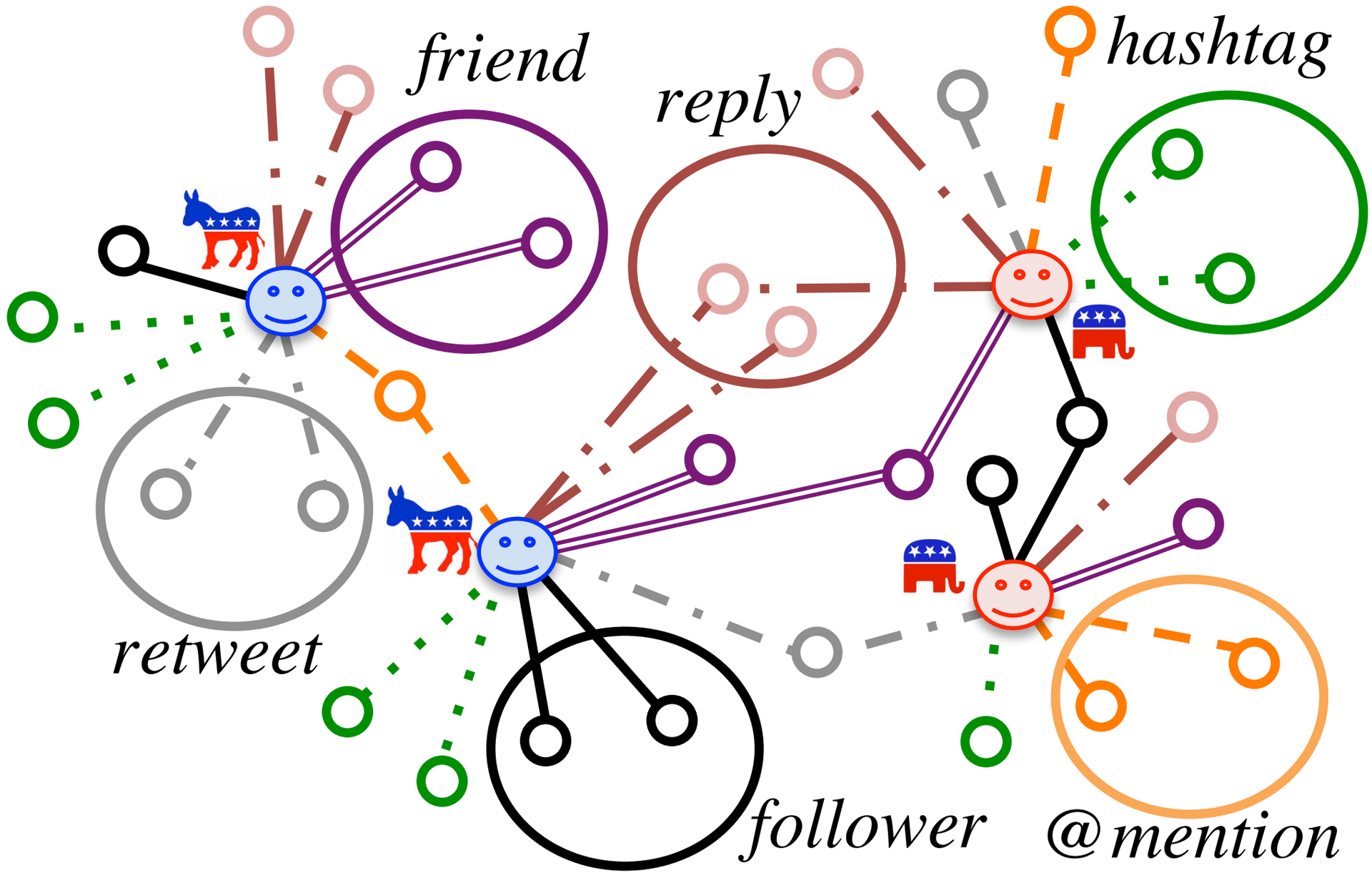
11:55 AM - 2 Jul 2015

hashtags are powerful

Cashtags


Twitter
@twitter
Follow
Now you can click on ticker symbols like $GE on twitter.com to see search results about stocks and companies
8:34 PM - 30 Jul 2012
1,167
295

Twitter's Social Graph



Source: Volkova, Van Durme, Yarowsky, Bachrach
"Tutorial on Social Media Predictive Analytics" NAACL 2015

Twitter API

What is an API?

Application **P**rogramming **I**nterface

API is a set of protocols that specify how software programs comment with each other.

What is an API?

> **Without API:**
> An app finds the current weather in London by opening http://www.weather.com/ and reading the webpage like a human does, interpreting the content.

> **With API:**
> An app finds the current weather in London by sending a message to the weather.com API (in a structured format like XML). The weather.com API then replies with a structured response.

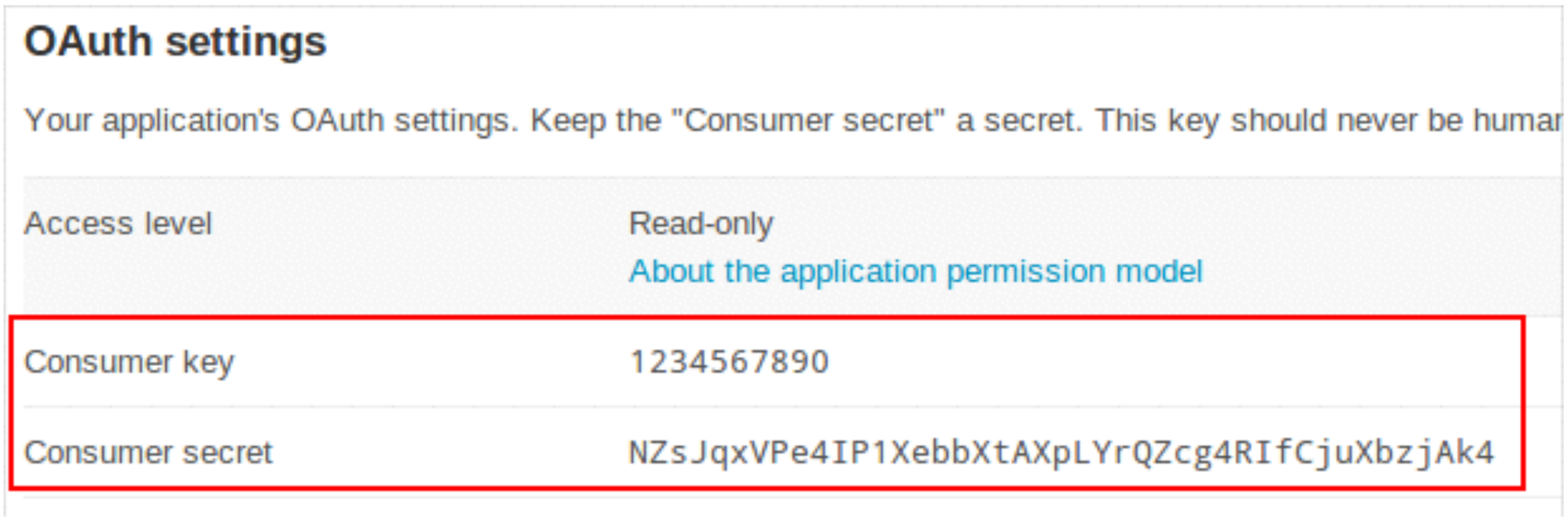
Source: Chris Beach @ Quora

Two Most Popular APIs

Streaming API	**REST API**
a sample of public tweets and events as they published on Twitter (can specify search terms or users)	- search - trends - read author profile and follower data - post / modify
only real-time data	historical data up to a week
continuous net connection	one-time request
no limit	rate limit (varies for different requests)

OAuth

- Twitter uses OAuth to provide authorized access to its API.
- which means, to start with needs:
  - a Twitter account
  - OAuth access tokens from apps.twitter.com



Python Twitter Tools

ndation [US] https://pypi.python.org/pypi/twitter

» Package Index > twitter > 1.17.0

twitter 1.17.0

An API and command-line toolset for Twitter (twitter.com)

Downloads ↓

Python Twitter Tools
==========

[![Build Status]https://travis-ci.org/sixohsix/twitter.svg]https://travis-ci.org/sixohsix/twitter [!
[Coverage Status]https://coveralls.io/repos/sixohsix/twitter/badge.png?
branch=master]https://coveralls.io/r/sixohsix/twitter?branch=master

The Minimalist Twitter API for Python is a Python API for Twitter,
everyone's favorite Web 2.0 Facebook-style status updater for people
on the go.

Streaming API

```
# Import the necessary package to process data in JSON format
try:
    import json
except ImportError:
    import simplejson as json

# Import the necessary methods from "twitter" library
from twitter import Twitter, OAuth, TwitterHTTPError, TwitterStream

# Variables that contains the user credentials to access Twitter API
ACCESS_TOKEN = 'YOUR ACCESS TOKEN"'
ACCESS_SECRET = 'YOUR ACCESS TOKEN SECRET'
CONSUMER_KEY = 'YOUR API KEY'
CONSUMER_SECRET = 'ENTER YOUR API SECRET'

oauth = OAuth(ACCESS_TOKEN, ACCESS_SECRET, CONSUMER_KEY, CONSUMER_SECRET)

# Initiate the connection to Twitter Streaming API
twitter_stream = TwitterStream(auth=oauth)

# Get a sample of the public data following through Twitter
iterator = twitter_stream.statuses.sample()
```

OAuth →

connection →

JSON

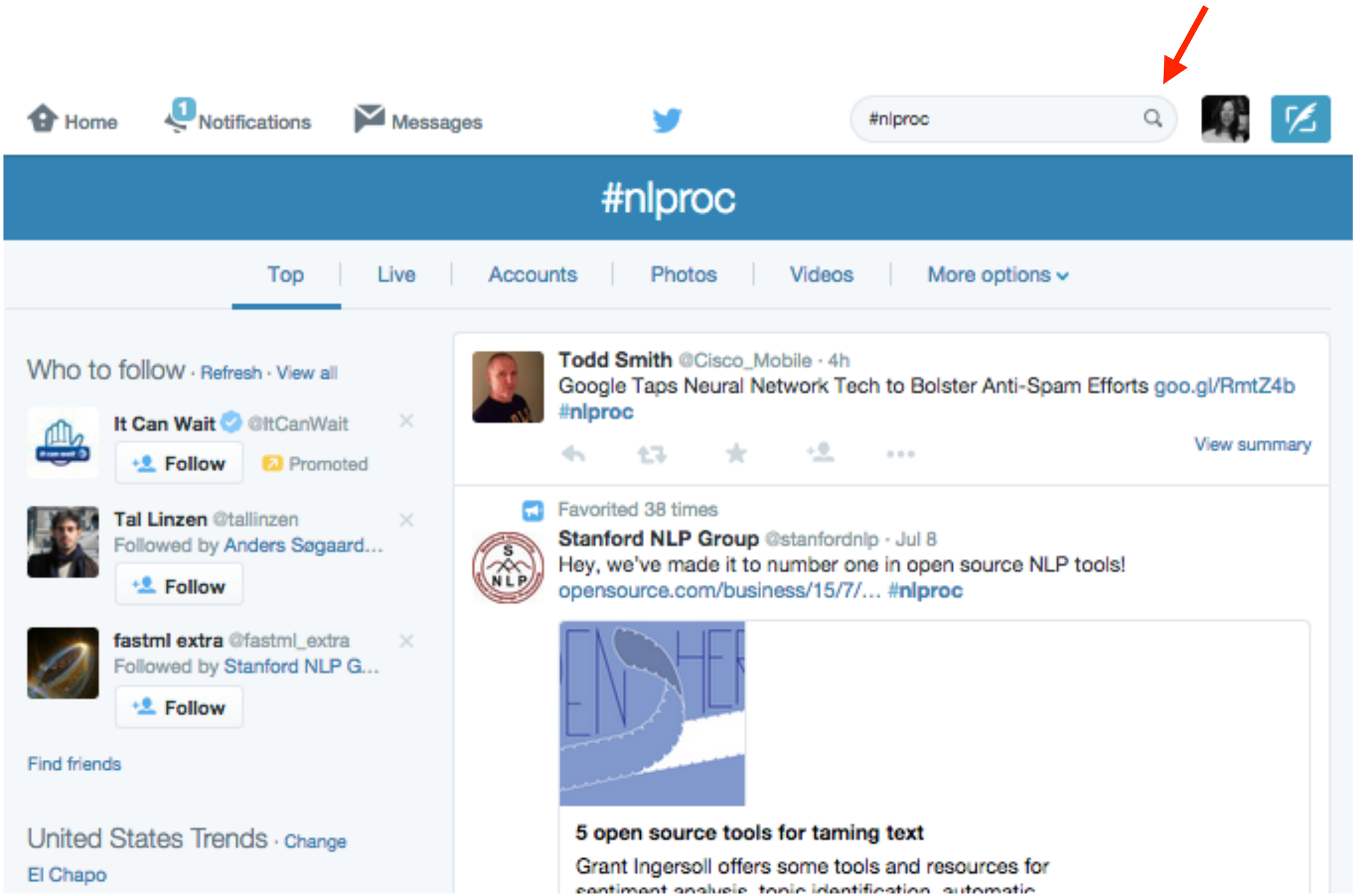
JavaScript Object Notation

JSON is a minimal, readable format for structuring data.

A Tweet in JSON

```
{
   "favorited": false,
   "contributors": null,
   "truncated": false,
   "text": "#CFP Workshop on Noisy User-generated Text at ACL - Beijing 31 July 2015. Papers due: 11 May 2015. http://t.co/rcygyEowqH   #NLProc #WNUT15",
   "possibly_sensitive": false,
   "in_reply_to_status_id": null,
   "user": {
      "follow_request_sent": null,
      "profile_use_background_image": true,
      "default_profile_image": false,
      "id": 237918251,
      "verified": false,
      "profile_image_url_https": "https://pbs.twimg.com/profile_images/527088456967544832/Dn
```

Search


Home
Notifications
Messages
#nlproc
#nlproc
Top
Live
Accounts
Photos
Videos
More options
Who to follow · Refresh · View all
It Can Wait @ItCanWait
Follow
Promoted
Tal Linzen @tallinzen
Followed by Anders Søgaard...
Follow
fastml extra @fastml_extra
Followed by Stanford NLP G...
Follow
Find friends
United States Trends · Change
El Chapo
Todd Smith @Cisco_Mobile · 4h
Google Taps Neural Network Tech to Bolster Anti-Spam Efforts goo.gl/RmtZ4b #nlproc
View summary
Favorited 38 times
Stanford NLP Group @stanfordnlp · Jul 8
Hey, we've made it to number one in open source NLP tools!
opensource.com/business/15/7/... #nlproc
5 open source tools for taming text
Grant Ingersoll offers some tools and resources for

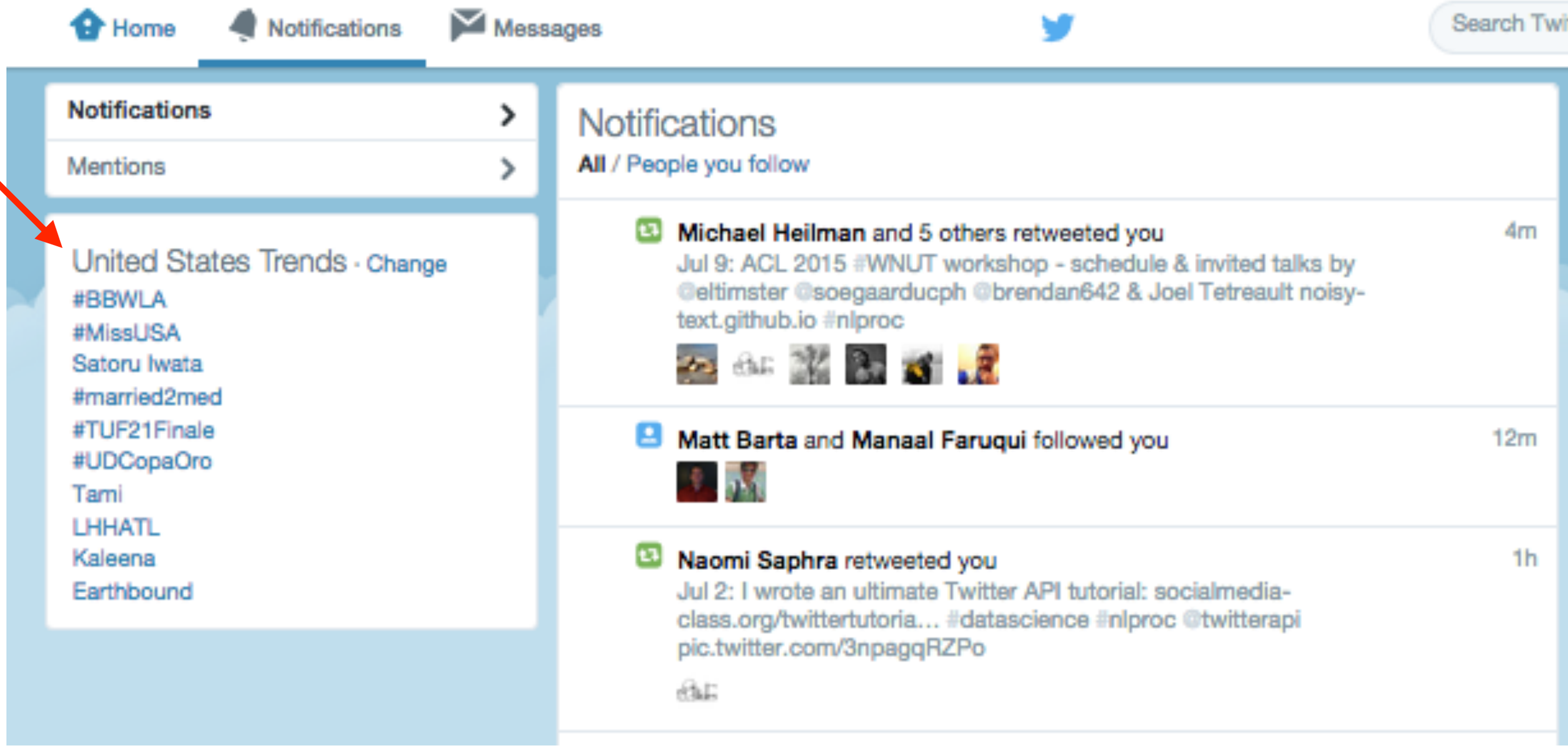
Search API

```
# Initiate the connection to Twitter REST API
twitter = Twitter(auth=oauth)
```

```
# Search for latest tweets about "#nlproc"
twitter.search.tweets(q='#nlproc')
```

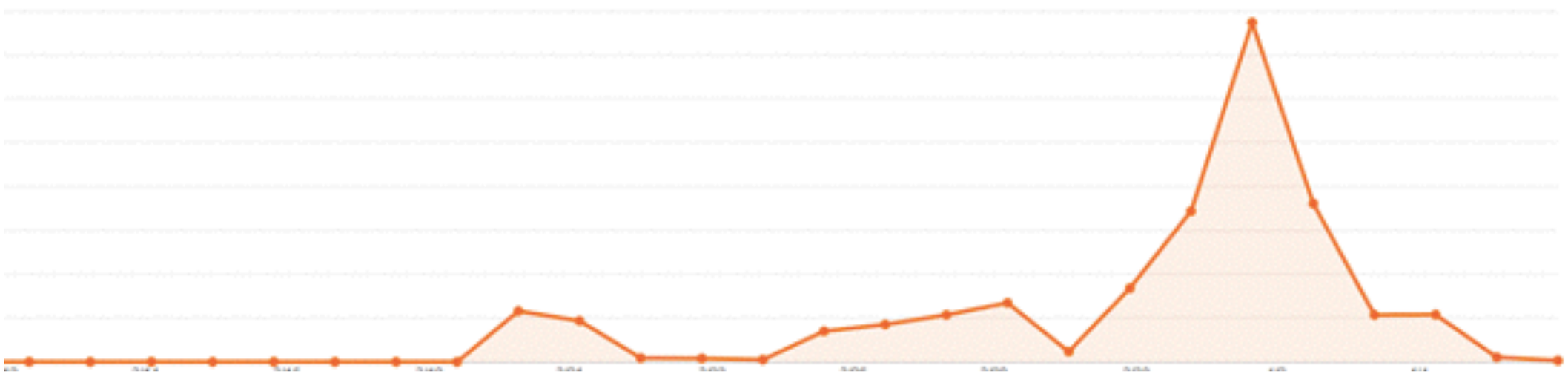
```
twitter.search.tweets(q='#nlproc', result_type='recent', lang='en', count=10)
```

Trends


Home
Notifications
Messages
Search Twi
Notifications
Mentions
United States Trends · Change
#BBWLA
#MissUSA
Satoru Iwata
#married2med
#TUF21Finale
#UDCopaOro
Tami
LHHATL
Kaleena
Earthbound
Notifications
All / People you follow
Michael Heilman and 5 others retweeted you
4m
Jul 9: ACL 2015 #WNUT workshop - schedule & invited talks by @eltimster @soegaarducph @brendan642 & Joel Tetreault noisy-text.github.io #nlproc
Matt Barta and Manaal Faruqui followed you
12m
Naomi Saphra retweeted you
1h
Jul 2: I wrote an ultimate Twitter API tutorial: socialmedia-class.org/twittertutoria… #datascience #nlproc @twitterapi pic.twitter.com/3npagqRZPo

Trends

trending topics are determined by an unpublished algorithm, which finds words, phrases and hashtags that have had a sharp increase in popularity, as opposed to overall volume.



Trends API

Where On Earth ID

```
# Get all (it's always 10) trending topics in San Francisco (its WOEID is 2487956)
sfo_trends = twitter.trends.place(_id = 2487956)
```


```
{
   "created_at":"2015-07-01T22:09:55Z",
   "trends":[
      {
         "url":"http://twitter.com/search?q=%23LiesIveToldMyParents",
         "query":"%23LiesIveToldMyParents",
         "name":"#LiesIveToldMyParents",
         "promoted_content":null
      },
      {
         "url":"http://twitter.com/search?q=%22Kevin+Love%22",
         "query":"%22Kevin+Love%22",
         "name":"Kevin Love",
         "promoted_content":null
      },

      ... [and another 8 trends omitted here to save space]
```

Homework #1

- Become a Twitter User (if you haven't)
- Learn to program in Python (if you haven't)
- Learn to use Twitter API

Social Media & Text Analytics | Syllabus | Twitter API Tutorial | Homework Assignments

Twitter's 404 error page -- the Fail Whale

Twitter API tutorial

by Wei Xu (July 1, 2015) Follow @cocoweixu

1. Getting Twitter API keys

To start with, you will need to have a Twitter account and obtain credentia Twitter developer site to access the Twitter API, following these steps:

- Create a Twitter user account if you do not already have one.
- Go to https://apps.twitter.com/ and log in with your Twitter user acc your user account.
- Click "Create New App"

Thank You!

Instructor: Wei Xu

www.cis.upenn.edu/~xwe/

Twitter API Tutorial: socialmedia-class.org