Social Media & Text Analysis

part 3 - Intro to Twitter API

Ohio State University

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)

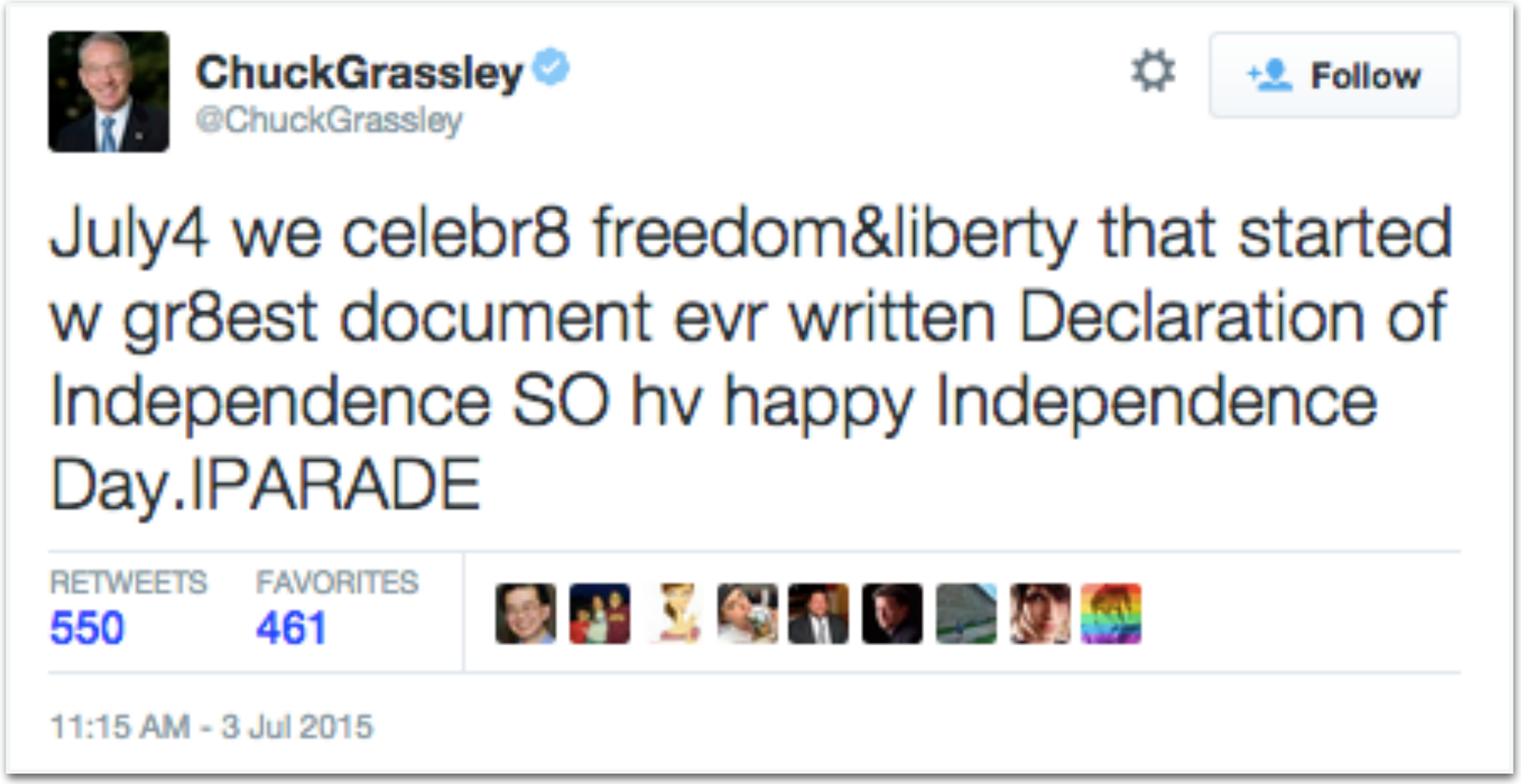
Follow @cocoweixu

1. Getting Twitter API keys

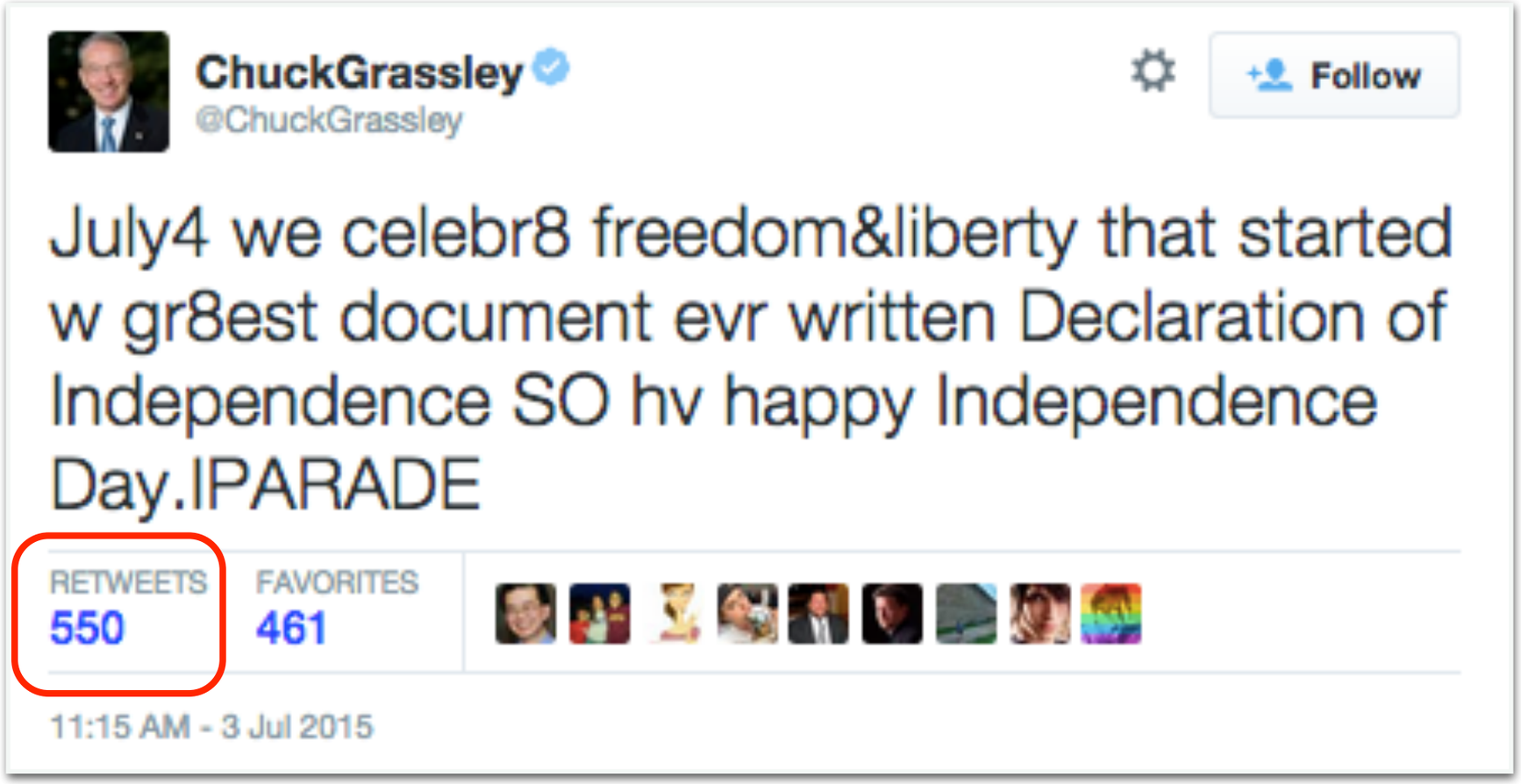
To start with, you will need to have a Twitter account and obtain credentia
Twitter developer site to access the Twitter API, following these steps:

- 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"

Tweets

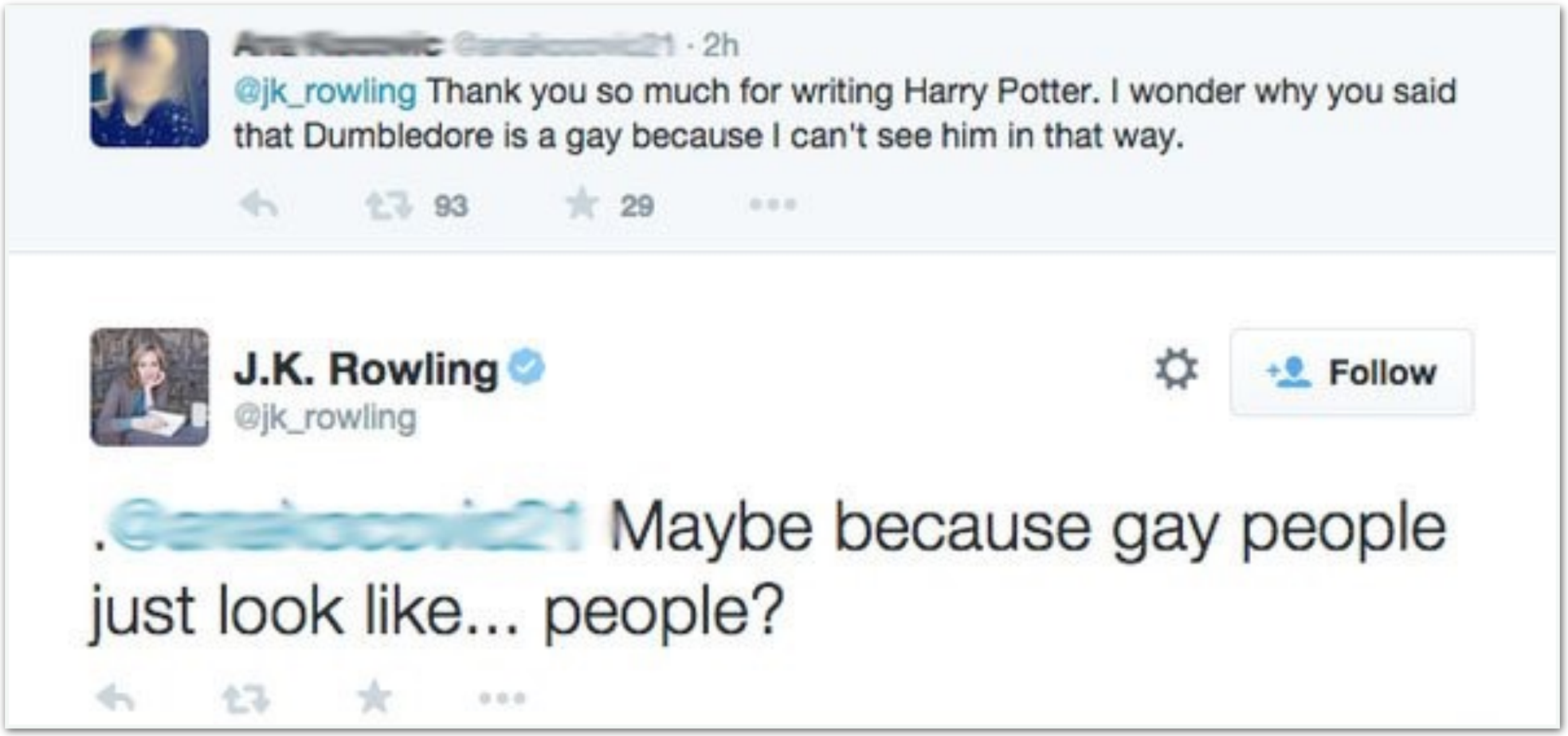

ChuckGrassley
@ChuckGrassley
Follow
July4 we celebr8 freedom&liberty that started w gr8est document evr written Declaration of Independence SO hv happy Independence Day.IPARADE
RETWEETS
550
FAVORITES
461
11:15 AM - 3 Jul 2015

ReTweets



a re-posting of someone else's Tweet

Replies/Conversations



- Tweet starts with a @username

Twitter API

What is an API?

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

API is a set of protocols that specify how software programs communicate 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

Twitter API

- Twitter is recognized for having one of the most open and powerful developer APIs of any major technology company.

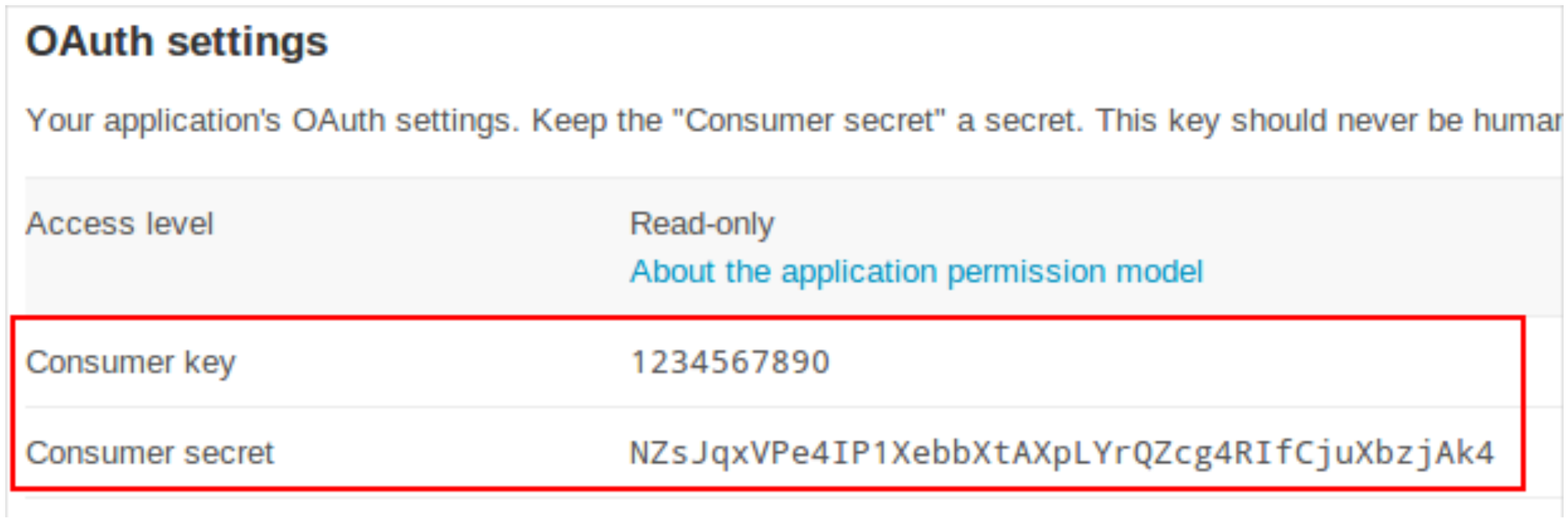
- The first version of its public API was released in September 2006.

Twitter API

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 Authentication

- 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

www.tweepy.org

OAuth Authentication

Twitter uses OAuth to provide authorized access to the API.

```
[ ]  # Import the tweepy library
     import tweepy
     from tweepy.streaming import StreamListener

     # 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'

     # Setup tweepy to authenticate with Twitter credentials:
     auth = tweepy.OAuthHandler(CONSUMER_KEY, CONSUMER_SECRET)
     auth.set_access_token(ACCESS_TOKEN, ACCESS_SECRET)
```

Streaming API

```
[ ]  # This is a basic listener that just prints received tweets to stdout.
     class StdOutListener(StreamListener):

         def on_data(self, data):
             print(json.loads(data))
             return True

         def on_error(self, status):
             if status_code == 420:
                 return False


     # tweepy.Stream.sample() will give a live stream (~1% sample) of all public tweets
     # Warning: it will continue to run indefinitely until you stop it.

     listener = StdOutListener()
     twitterStream = tweepy.Stream(auth, listener)
     twitterStream.sample()
```

JSON

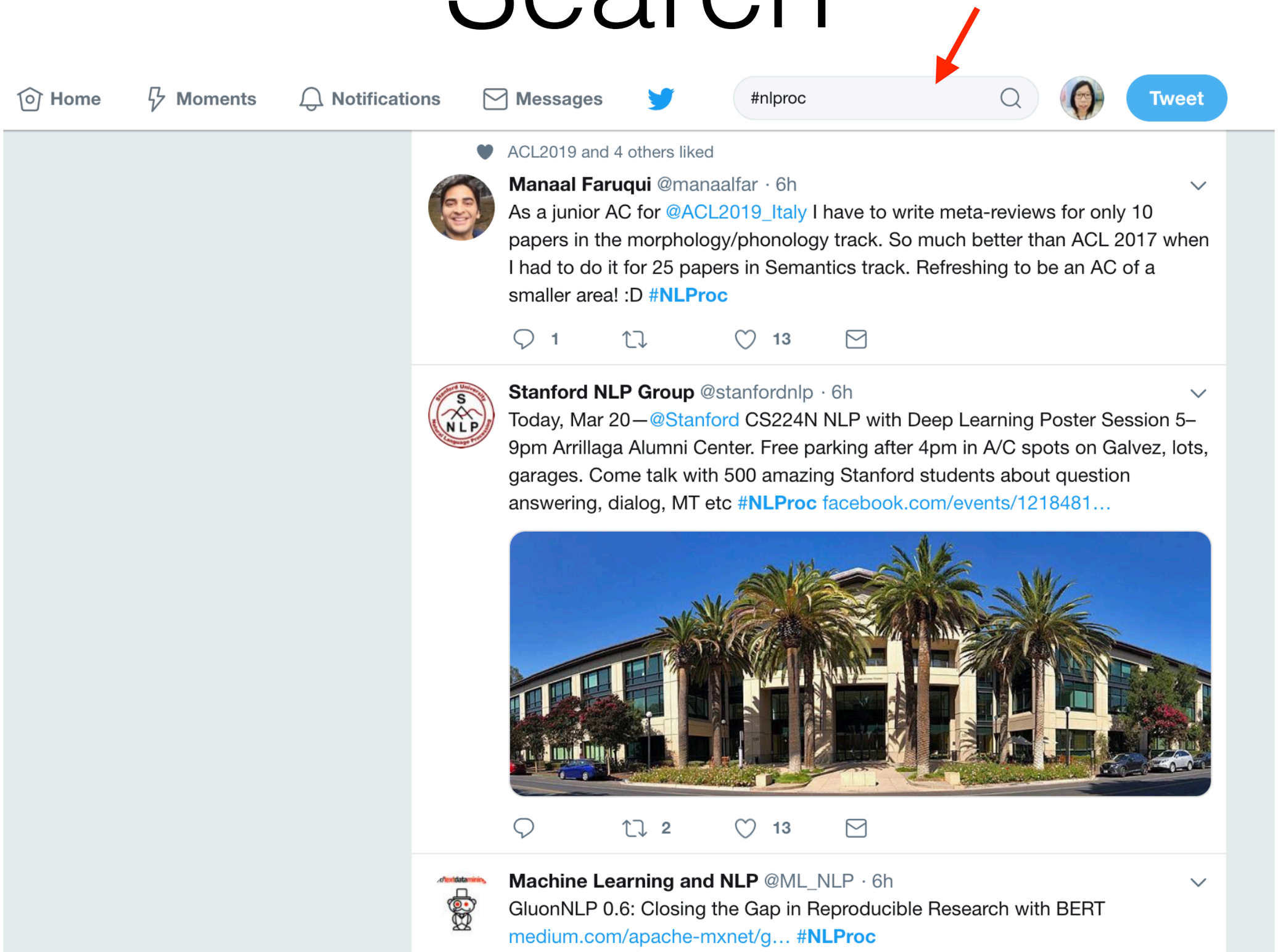
JavaScript Object Notation

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
Moments
Notifications
Messages
#nlproc
Tweet
ACL2019 and 4 others liked
Manaal Faruqui @manaalfar · 6h
As a junior AC for @ACL2019_Italy I have to write meta-reviews for only 10 papers in the morphology/phonology track. So much better than ACL 2017 when I had to do it for 25 papers in Semantics track. Refreshing to be an AC of a smaller area! :D #NLProc
1
13
Stanford NLP Group @stanfordnlp · 6h
Today, Mar 20—@Stanford CS224N NLP with Deep Learning Poster Session 5–9pm Arrillaga Alumni Center. Free parking after 4pm in A/C spots on Galvez, lots, garages. Come talk with 500 amazing Stanford students about question answering, dialog, MT etc #NLProc facebook.com/events/1218481…
2
13
Machine Learning and NLP @ML_NLP · 6h
GluonNLP 0.6: Closing the Gap in Reproducible Research with BERT
medium.com/apache-mxnet/g… #NLProc

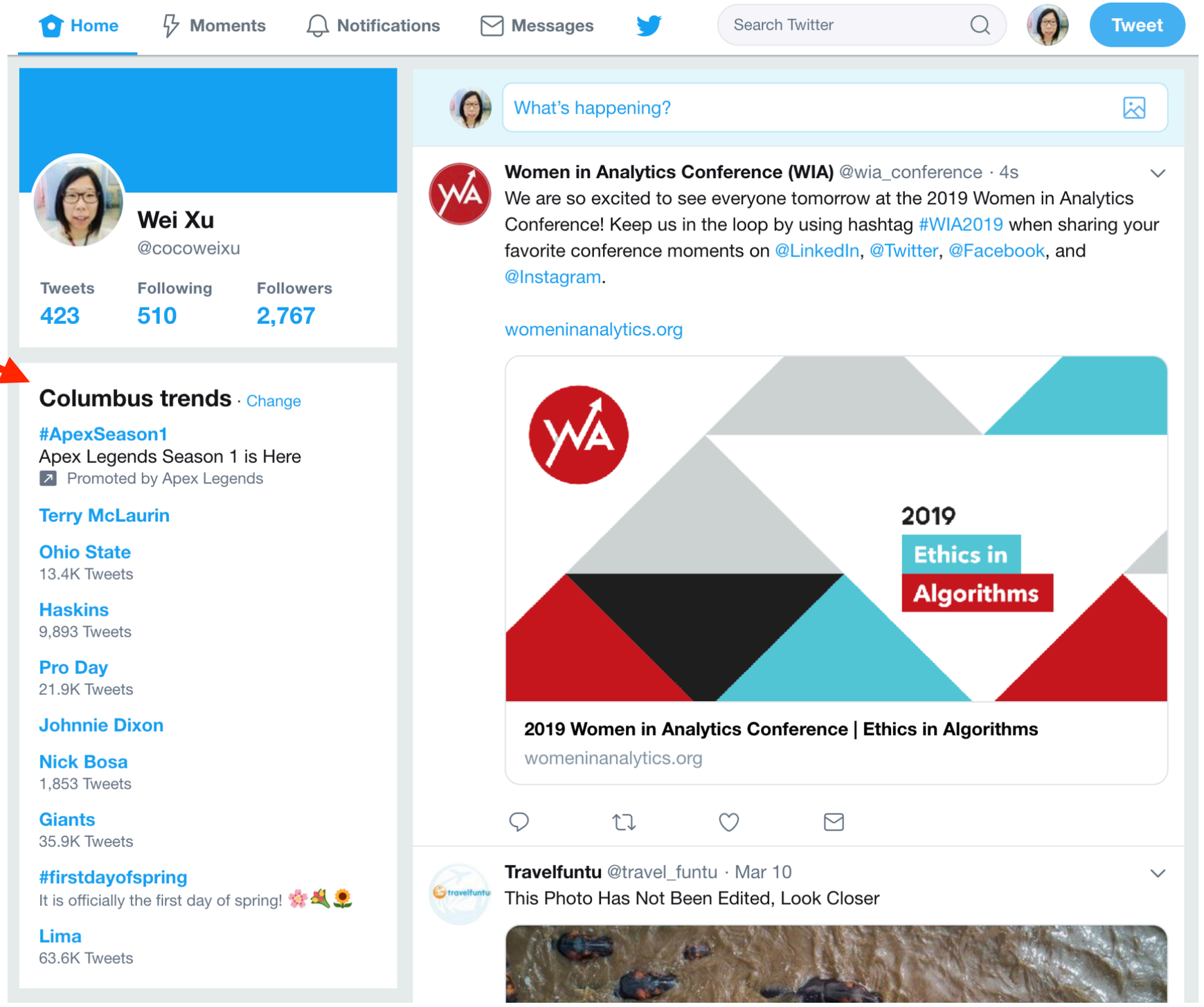
Search API

```
[ ]  # Search for latest tweets about "#nlproc"
     tweets = tweepy.Cursor(api.search, q='#nlproc')

     # Print out the latest 10 tweets that contain "#nlproc" hashtag
     for item in tweets.items(10):
       print(item._json)
```

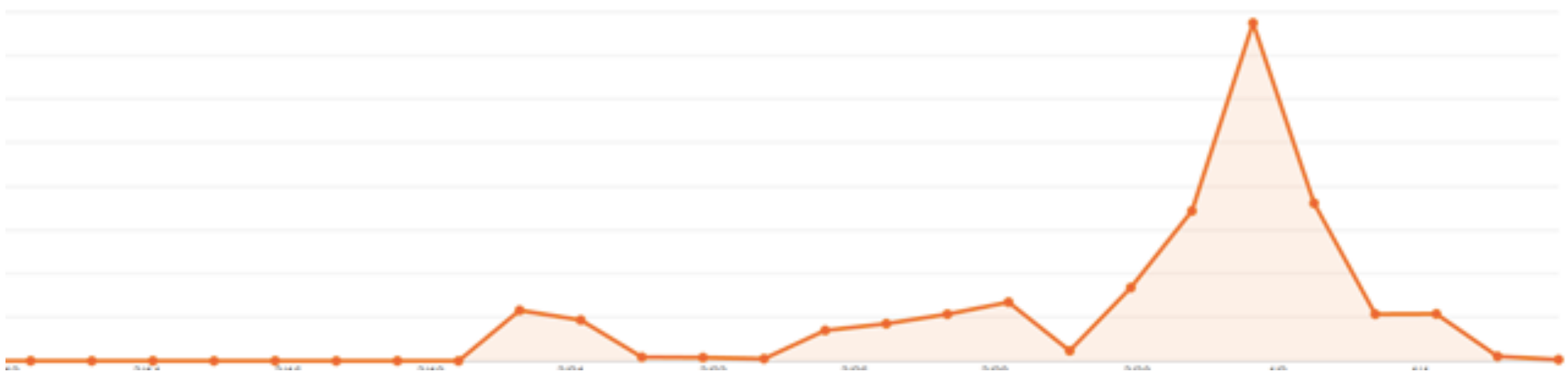

```
{'name': 'Dwayne Haskins', 'url': 'http://twitter.com/search?q=%22Dwayne+Haskins%22', 'promoted_content': None, 'query': '%22Dwayne+Haskins%22', 'twe
{'name': 'McCain', 'url': 'http://twitter.com/search?q=McCain', 'promoted_content': None, 'query': 'McCain', 'tweet_volume': 381900}
{'name': 'Lima', 'url': 'http://twitter.com/search?q=Lima', 'promoted_content': None, 'query': 'Lima', 'tweet_volume': 70085}
{'name': '#firstdayofspring', 'url': 'http://twitter.com/search?q=%23firstdayofspring', 'promoted_content': None, 'query': '%23firstdayofspring', 'tw
{'name': 'Daniel Caesar', 'url': 'http://twitter.com/search?q=%22Daniel+Caesar%22', 'promoted_content': None, 'query': '%22Daniel+Caesar%22', 'tweet_
{'name': '#InternationalDayOfHappiness', 'url': 'http://twitter.com/search?q=%23InternationalDayOfHappiness', 'promoted_content': None, 'query': '%2
{'name': 'AirPods', 'url': 'http://twitter.com/search?q=AirPods', 'promoted_content': None, 'query': 'AirPods', 'tweet_volume': 131265}
{'name': 'Pro Day', 'url': 'http://twitter.com/search?q=%22Pro+Day%22', 'promoted_content': None, 'query': '%22Pro+Day%22', 'tweet_volume': 19808}
{'name': '#SpringEquinox', 'url': 'http://twitter.com/search?q=%23SpringEquinox', 'promoted_content': None, 'query': '%23SpringEquinox', 'tweet_volu
{'name': 'Flume', 'url': 'http://twitter.com/search?q=Flume', 'promoted_content': None, 'query': 'Flume', 'tweet_volume': 37576}
{'name': '#HappinessInOneWord', 'url': 'http://twitter.com/search?q=%23HappinessInOneWord', 'promoted_content': None, 'query': '%23HappinessInOneWor
{'name': '#StrangerThings3', 'url': 'http://twitter.com/search?q=%23StrangerThings3', 'promoted_content': None, 'query': '%23StrangerThings3', 'twee
{'name': 'Eloy', 'url': 'http://twitter.com/search?q=Eloy', 'promoted_content': None, 'query': 'Eloy', 'tweet_volume': None}
{'name': 'Happy Spring', 'url': 'http://twitter.com/search?q=%22Happy+Spring%22', 'promoted_content': None, 'query': '%22Happy+Spring%22', 'tweet_vo
{'name': 'Bill & Ted 3', 'url': 'http://twitter.com/search?q=%22Bill+%26+Ted+3%22', 'promoted_content': None, 'query': '%22Bill+%26+Ted+3%22', 'twee
```

Trends


Home
Moments
Notifications
Messages
Search Twitter
Tweet
Wei Xu
@cocoweixu
Tweets
423
Following
510
Followers
2,767
Columbus trends · Change
#ApexSeason1
Apex Legends Season 1 is Here
Promoted by Apex Legends
Terry McLaurin
Ohio State
13.4K Tweets
Haskins
9,893 Tweets
Pro Day
21.9K Tweets
Johnnie Dixon
Nick Bosa
1,853 Tweets
Giants
35.9K Tweets
#firstdayofspring
It is officially the first day of spring!
Lima
63.6K Tweets
What's happening?
Women in Analytics Conference (WIA) @wia_conference · 4s
We are so excited to see everyone tomorrow at the 2019 Women in Analytics Conference! Keep us in the loop by using hashtag #WIA2019 when sharing your favorite conference moments on @LinkedIn, @Twitter, @Facebook, and @Instagram.
womeninanalytics.org
2019
Ethics in
Algorithms
2019 Women in Analytics Conference | Ethics in Algorithms
womeninanalytics.org
Travelfuntu @travel_funtu · Mar 10
This Photo Has Not Been Edited, Look Closer

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

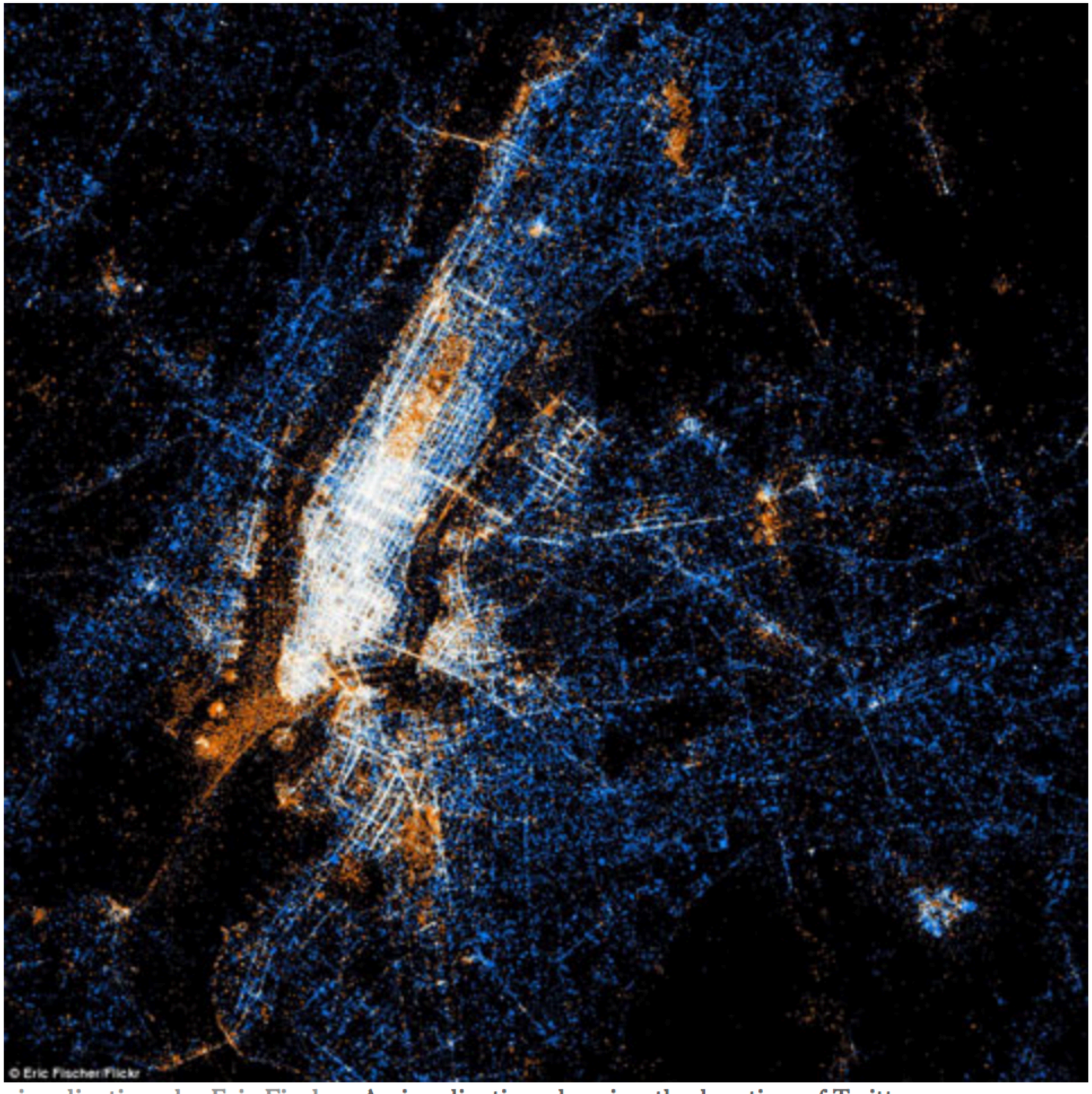
Where On Earth ID

```
# Where On Earth ID for Columbus, Ohio is 2383660.
COLUMBUS_WOE_ID = 2383660

columbus_trends = api.trends_place(COLUMBUS_WOE_ID)

trends = json.loads(json.dumps(columbus_trends, indent=1))

for trend in trends[0]["trends"]:
  print (trend)
```



visualizations by Eric Fischer A visualization showing the location of Twitter messages and Flickr photos in New York City.

Social Science

wonderfully delightfully beautifully fine well good nicely superbly

she says

he says

(also age & income)

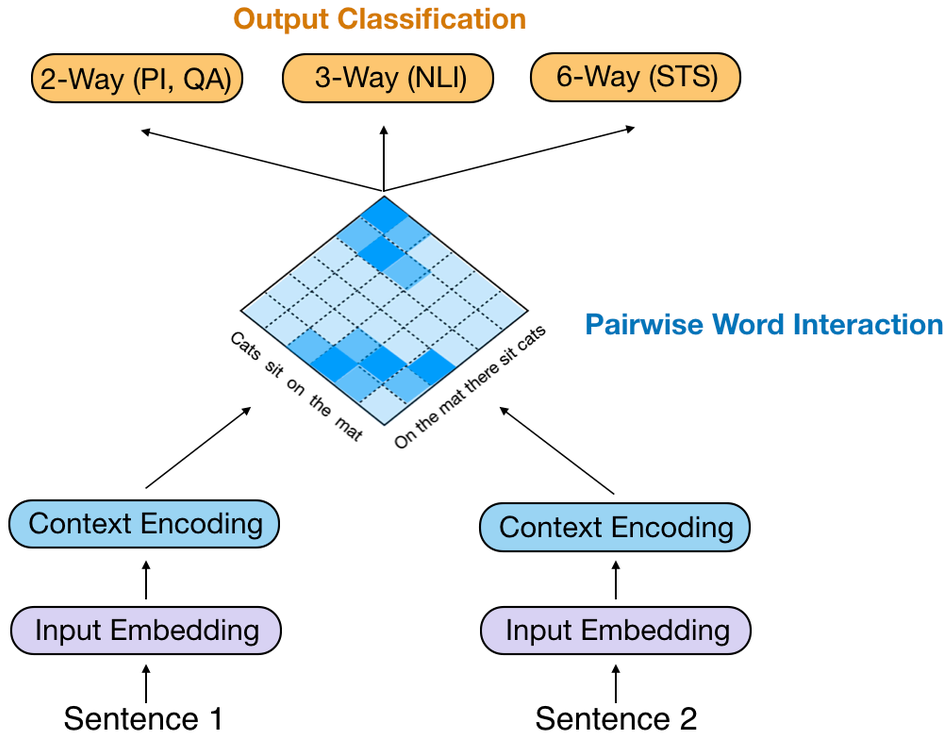
Daniel Preotiuc, Wei Xu, Lyle Ungar. “Discovering User Attribute Stylistic Differences via Paraphrasing” in AAAI (2016)

Detecting Similar Tweets

paraphrase ← *Ezekiel Ansah is wearing real3D glasses with the lenses punched out*

Ezekiel Ansah wearing 3D glasses wout the lens

non-paraphrase ← *I wore the 3D glasses wout lenses before Ezekiel Ansah*



Wei Xu, Alan Ritter, Chris Callison-Burch, Bill Dolan, and Yangfeng Ji. "Extracting Lexically Divergent Paraphrases from Twitter" in TACL (2014)

Wuwei Lan, Siyu Qiu, Hua He, and Wei Xu. “A Continuously Growing Dataset of Sentential Paraphrases” in EMNLP (2017)

Wuwei Lan and Wei Xu. “Neural Network Models for Paraphrase Identification, Semantic Textual Similarity, Natural Language Inference, and Question Answering” in COLING (2018) Best Paper Award