



Accessing Twitter API

Jisu Kim¹

¹Max Planck Institute for Demographic Research

Outline

- ► Twitter API
- ► Academic Research track
- ► Twitter data format
- Collecting data

Twitter API

What is Twitter API (Application programming interface)?

- \rightarrow API is a software that allows two applications to talk to each other. E.g., *Twitter API*, Twitter Ads API...
 - → Standard product track and Academic Research track

Standard product track

- ► Currently, there is a limit up to 500,000 tweets per month
- ▶ Query limit up to 512 characters

Academic Research track²

- Access to full-archive search and tweet counts
- ► Monthly-tweet cap usage: "Certain endpoints (like filtered stream and recent search) have a limit on how many Tweets they can pull per month." == 10 million tweets¹
- ▶ Query limit up to 1024 characters

¹https://developer.twitter.com/en/docs/twitter-api/rate-limits

²https://blog.twitter.com/developer/en_us/topics/tools/2020/

API V1 vs. V2

Version 2 of Twitter API available through Academic research Track!³ What's new?

- more detailed data objects: e.g., conversation ID, Spaces, Media ...
- new parameters to request objects and fields: e.g., "has:links OR is:retweet" 4
- ▶ advanced metrics: e.g., of impressions, retweets, likes, replies...
- Tweet reply threads
- Contextual tweet annotations

³https://developer.twitter.com/en/docs/twitter-api

 $^{^4} https://developer.twitter.com/en/docs/twitter-api/tweets/search/integrate/build-api/tweets/search/search/search/search/search/search/sear$

Contextual tweet annotations

Entity, for example:

- Person Barack Obama, Daniel, or George W. Bush
- ▶ Place Detroit, Cali, or "San Francisco, California"
- Product Mountain Dew, Mozilla Firefox
- Organization Chicago White Sox, IBM...

Based on what is mentioned in the Tweet text.

Context, for example:

- ► TV shows
- Sport
- Politicians
- Brand
- ► Movie ...

Based on the analysis of a Tweet's text.



Keys and tokens

https://developer.twitter.com/en/apply-for-access

- ► API key and secret key: identify the calling project -the application or site⁵
- ▶ Bearer token (Authentication): User authentication and authorization

Consumer Keys ①	
API Key and Secret	Regenerate
Authentication Tokens ①	
Bearer Token Generated October 15, 2021	Regenerate Revoke

Twitter data format

- ► Tweet object
- User object
- Entity object

Tweet object

```
"id": "1050118621198921728"
"text": "#paa2022 is happening soon! This also means that our
Workshop is approaching! If you haven't registered yet, there is still time!
For those already registered, @EbruSanliturk and I cannot wait to meet
you all in person! - :-) @MPIDRnews"
"Context annotations": 'name': 'Interests and Hobbies Category',
'description': 'A grouping of interests and hobbies entities, like Novelty
Food or Destinations', 'entity': 'id': '852291840472629248', 'name':
'Online education'. 'description': 'Online education'
"created_ at": "2022-03-28T15:08:xx.000z"
"conversation id": "1435336531519197188"
"lang": "en", "geo": e.g., "coordinates": [-73.999xx, 40.7416xxx]
"place_id": "01a9a39529b27f36"
```

Entity object

```
"hashtags": ["tag":
"paa2022"],
"mentions": "tag":
"@EbruSanliturk,
@MPIDRnews",
"url": None,
"expanded_url": None,
"description": "paa2022 is
happening soon..."
```



Collecting tweet object

- ► Streaming keywords/hashtags: e.g. #migrants, #refugees
- Streaming for specific geo-locations: place, country, point radius, bounding box
- Collecting tweets for a specific user: by user ID or user name
- Searching for historic tweets for a keyword

Query:

- ► from: twitter user
- ► −is : excluding retweets
- place_country: country
- has: e.g. media, geo, images
- conversation_id: returns all tweets in the conversation thread with conversation id. xxx
- lang: en, kr, ch, etc.



User object

- **"id"**: "2244994945",
- ► "name": "Twitter Dev",
- **"username"**: "TwitterDev",
- ► "location": "127.0.0.1"
- "verified": true,
- "protected": false,



- ► "description": "The voice of the #TwitterDev team and your official source for updates, news, and events, related to the #TwitterAPI."
- "url": https://t.co/3ZX3TNiZCY,
- "profile_image_url": https://pbs.twimg.com/profile_images/
 1267175364003901441/tBZNFAgA_normal.jpg,
- "created_at": "2013-12-14T04:35:55.000Z"



Rate limit

Rate limits⁶ are the number of requests (or API calls) that are allowed. App limits apply only to a single App, while user limits apply to a combination of all Apps.

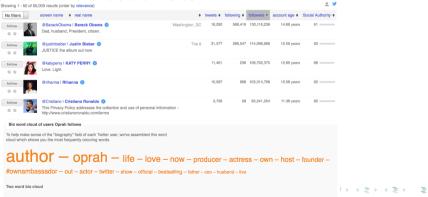
- ► Recent search (tweets): 450
- ► Full-archive search: 300 (Full-archive also has a 1 request / 1 second limit)
- ► User lookup: 900
- ► Get list followers: 180

E.g., If an endpoint has a rate limit of 900 requests/15-minutes, then up to 900 requests over any 15-minute interval is allowed.

⁶https://developer.twitter.com/en/docs/twitter-api/rate=limits < > > > > >

Collecting user object

- Searching by user: user ID, user name, https://followerwonk.com/bio
- Social network: followers and friends



R and Python packages for Twitter API

Python:

Twerc, Tweepy, python-twitter-v2, search-tweets-python, Python client **R**:

rTweet, academictwitteR

archive of Twitter data freely available:

https://catalog.docnow.io

So what can we do with these data for the studies of migration?

To be continued...