USING THE TWITTER API

Queensland State Archives

Capturing content from Twitter

If you want to export content from Twitter, a simple way to do this is to use the Twitter API. You can access it through your usual browser. If you query or call the API, you get formatted results that you can copy and paste into documents that you can then save into your internal recordkeeping system.

If you want an archive of all your tweets, it might be easier to ask Twitter directly to provide you with this. You can make this request from your account settings page.
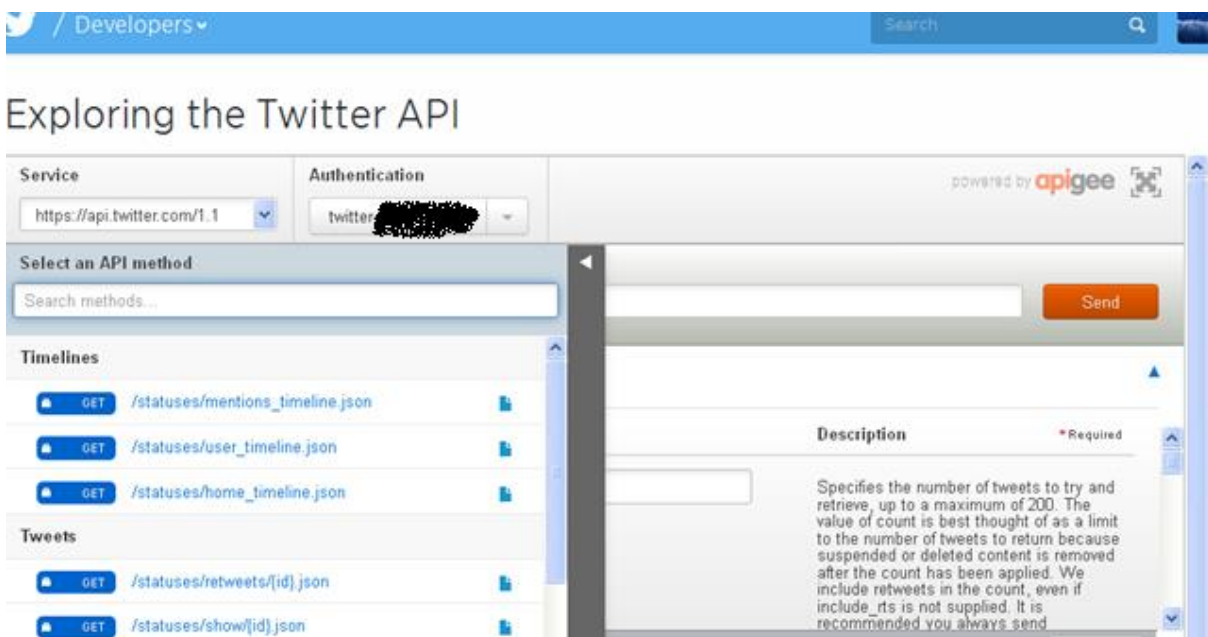
Using the API

To access the Twitter API, you have to first log in to Twitter. Then go to this page, which has the API console:
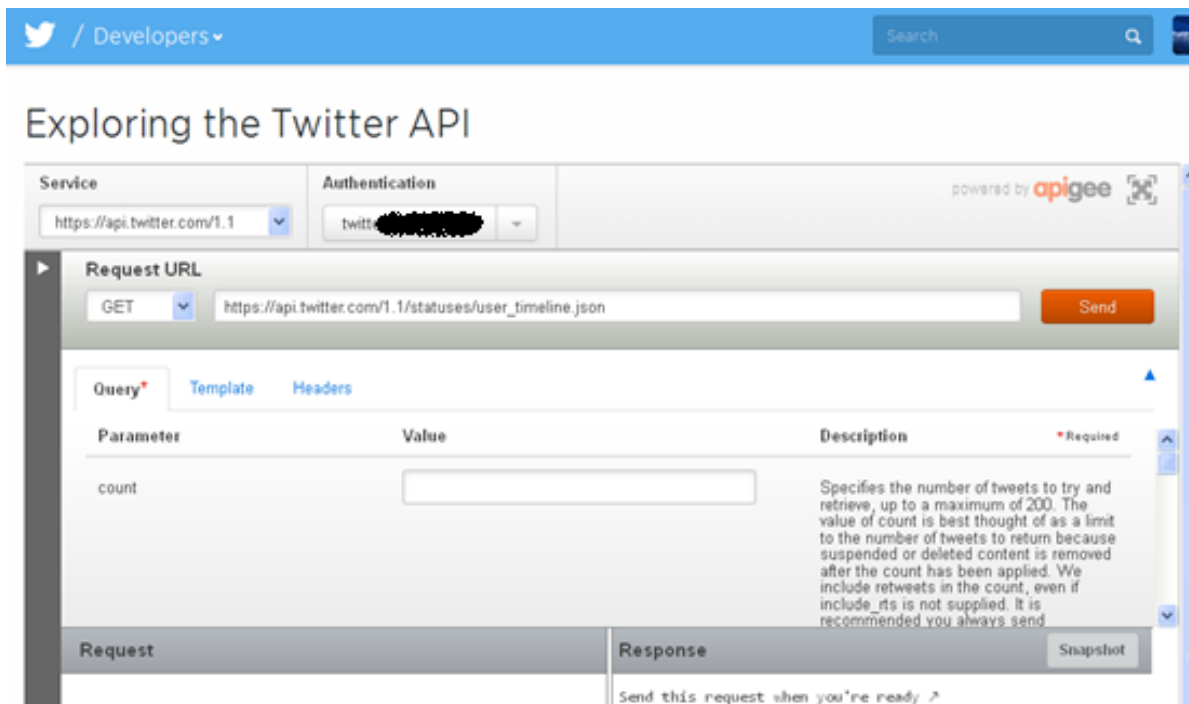
https://dev.twitter.com/rest/tools/console

Next, arrange for authorisation to query the API. Choose 'OAuth 1' under the 'Authentication' dropdown menu. The API will ask you to agree to authorise the Apigee Console to use your account. Agree and wait for the system to register your details and authorise access. Your username will appear in the 'Authentication' menu box when this has been done.

Now you can start querying the API. Click on the arrow in the grey bar on the left of the screen to see a list of queries you can make. The list looks like this:

To query the API, simply click on the query you want. This will create a template query and add it to the main console screen. To personalise the query, you'll have to modify it to specify, for example, a Twitter username or tweet ID. You can do this by editing it directly in the query bar, which is to the right of the 'Get' selection in the picture below:

When you've finished updating the query, click the 'Send' button on the right of the console and wait for it to run.

Useful API calls

Bulk query

One useful call you can make is **statuses/user_timeline.json**, which is in the 'Timelines' section of the overall list of calls. It returns a specified number of tweets sent by a user, starting with the most recent tweet. The full format for this call is:

https://api.twitter.com/1.1/statuses/user_timeline.json?count=[number]&screen_name=[user id]

To personalise this call, replace **[user id]** with your Twitter handle (eg QSArchives, leaving out the @) and **[number]** with the number of tweets you want to return, up to a maximum of 200 for each call. The results will be returned in .json format, which is a bit like .xml, and will look like this snippet:

```
"created_at": "Mon Jan 26 23:13:02 +0000 2015",

    "id": 559851535782256640,

    "id_str": "559851535782256640",

    "text": "Our #RecordkeepingSurvey launches 2 February 2015. See our
latest blog for more information http://t.co/aX5KJFkYHf
#KeepingRecordsTogether",
```

```
    "source": "<a href="http://twitter.com" rel="nofollow">Twitter Web
Client</a>"
```

Returning older tweets

As noted above, the maximum number of tweets you can return for each call is 200. If you've tweeted more than 200 times, you can return the tweets older than the last one in your original call by adding a **max_id** condition to the template call above:

https://api.twitter.com/1.1/statuses/user_timeline.json?count=[number]&screen_name=[user id]&max_id=[tweet id]

To find the ID of the last tweet in your original call, scroll down and look for information in this format:

"id": 559851535782256640

It will appear straight after the tweet's date and time stamp. Add the ID number you've found to the template query and make the call. If you've sent a large number of tweets, you'll have to repeat this process several times.

The total number of tweets you can access through the API is currently 3200. If there are more than 3200 tweets you want to capture, it's probably best to ask Twitter directly to provide you with an archive of them.

Returning newer tweets

If you've already exported some tweets and only want to capture more recent ones, you can use the **since_id** condition to tailor your query. The query looks like this:

https://api.twitter.com/1.1/statuses/user_timeline.json?count=[number]&screen_name=[user id]&since_id=[tweet id]

Add the ID number of the most recent tweet you've exported and make the call.

Returning tweets within a date range

If you want to return tweets within a specific date range, you can use a combination of the **max_id** and **since_id** conditions. The query will look like this

https://api.twitter.com/1.1/statuses/user_timeline.json?count=[number]&screen_name=[user id]&max_id=[tweet id]&since_id=[tweet id]

The **max_id** will be the ID of the most recent tweet you want to capture. The **since_id** will be the ID of the tweet one older than the oldest tweet you want to capture. This is because the **since_id** condition is exclusive while the **max_id** condition is inclusive.

Single query

If you want to capture a single tweet, you can use the **statuses/show/{id}.json** call. The full format for this call is:

https://api.twitter.com/1.1/statuses/show/[tweet id].json

The easiest way to find the ID of the tweet you want to capture, if you've not made any API calls up till now, is to find the tweet online, click on it to expand it and then click on 'details' link when that

appears. This will take you to a dedicated page for the tweet and the ID will be at the end of the page's URL, for example:

https://twitter.com/QSArchives/status/5606325000000

To make the call, replace **[tweet id]** in the template query with the number you've found and hit 'Send'.

Limiting the amount of information returned

Twitter API calls return a lot of information and metadata for each tweet. Much of it isn't important to capture and makes the results difficult to browse through.

There are a couple of conditions you can add to calls to limit the amount of information and metadata returned. These conditions are **trim_user** and **include_entities**. This bulk tweet call, for example, leaves out information about the Twitter user (eg name, location, description), hashtags and internal Twitter URLs:

https://api.twitter.com/1.1/statuses/user_timeline.json?count=[number]&screen_name=[user id]&trim_user=1&include_entities=false

Make sure that **trim_user** is set to '1' and **include_entities** is set to 'false'. You can use these conditions with the single tweet query as well.

Saving calls

If you want to save the content you've called through the Twitter API, you'll have to manually copy and paste the results into a separate document, like .txt. An easy way to select the content for copying is to click at the beginning of the results, and then again at the end while holding down CTRL+SHIFT.

Further information

More information on the Twitter API can be found here:

https://dev.twitter.com/rest/public