

Grebe Social Data Aggregator

RESTful API Version 1.1.0 Documentation

Description:

Get tweets for a given date range and province with options to filter by keywords & fields

Endpoint (authentication required):

<http://199.116.235.207/grebe/api/?start=&end=&province=>

Supported Methods:

GET

Required URL Parameters:

[start=dd-mm-yyyy](#)

[end=dd-mm-yyyy](#)

The maximum date range is 30 days

[province=AB](#)

The geo-coordinate data is used to filter by province if needed. Official two-letter abbreviations are supported¹, and cascading queries are not supported, i.e. only one province can be specified for filtering.

Optional URL Parameters:

[keywords=string1+string2|string3|string4](#)

For keywords' search, both OR and AND operations are supported. For OR operations, i.e. string1 OR string2 OR string3, use the | (pipeline) separator, e.g. string1|string2|string3. For AND operations, i.e. string1 AND string2 AND string3, use the encoded space operator (plus) +, e.g. string1+string2+string3. Combined operations are also supported, e.g. string1+string2|string3|string4 implies string1 AND string2 OR string3 OR string4

¹ Two-Letter Abbreviations for Canadian Provinces <https://www.ncbi.nlm.nih.gov/books/NBK7254/>

Optional URL Parameters (continued):

[fields=field1,field2,field3](#)

The fields returned can be specified, and the field names need to be separated by commas. A subset of the Twitter Field Guide² is mapped to intuitively named Grebe database fields: id, tweet, longitude, latitude, created_at, collected_at, lang, place_name, country_code, user_ppid

The following five fields are returned by default when no fields are specified: tweet, longitude, latitude, created_at, place_name

[download=1](#)

The output from the API can either be displayed in the browser/client or be available as a downloadable file. Setting this option to 1 will create a downloadable file, while setting to any other value, or omitting this option will return the data in the client.

Returned Data Format:

All data is returned in JSON format³. The returned data is downloadable as plaintext. An example request, and returned JSON are given below (formatted and simplified for clarity).

http://199.116.235.207/grebe/api/?start=14-09-2016&end=13-10-2016&province=AB&keywords=edm+yeg|oilers&fields=tweet,user_ppid

```
[
  {
    "tweet": "Rushed it Oilers",
    "user_ppid":
    "b9d83569d7f8f5650bb5fcf1797f7e7963217c485d110a0d4f09a3873fcc64c6
    b9ec2c7b2e9635f5f82bd201c3254e169c17cb7e3e59070770c02eadaaa0487e"
  },
  {
    "tweet": "Late Night at the Edmonton yeg @ Alberta",
    "user_ppid":
    "d1768f81eedc66e59c74174216eb41f85c3b94465231716ede05ac880c1bdbdc
    f2b082b686f902b60df027fc1fd06172410f9df7cf770e80cb2b54eaddf0361f"
  }
]
```

² Twitter Field Guide <https://dev.twitter.com/overview/api/tweets>

³ Validated JSON RFC-4627, RFC-7159, and ECMA-404 <https://jsonformatter.curiousconcept.com>

API Change Log:

Version	Date	Changes
1.1.0	2018-10-01	<ul style="list-style-type: none">• User ID, screen name, and other details replaced with Privacy Preserving ID (PPID) hashed representation• Aggregator replaced with new status, search, and stream scrapers for getting tweets from public organizational pages, searching for hashtags, and getting live streaming tweets• Switched from Tweepy to TwitterAPI Python library• Twitter app OAuth 2.0 login for more balanced rate limits• Database schema updated with new indexes and data types
1.0.4	2017-09-03	<ul style="list-style-type: none">• Bug fix: new CDN for OpenLayers API to show map demo• Bug fix: strict filtering for Canada-only coordinates on map• Bug fix: Non-relevant keywords removed when filtering by keyword on graph demo• New feature: visual tracker/marker for currently selected date on the time map demo
1.0.3	2017-08-29	<ul style="list-style-type: none">• Move from MongoDB to MariaDB (rationale of more efficient disk usage with comparable performance)• Subfield/dot notation support removed, subset of Twitter Field Guide is mapped to MariaDB fields• Some fields renamed, e.g. <code>text</code> is now <code>tweet</code>, etc.• Added optional output download parameter• Documentation & REST API updated to support new fields
1.0.2	2017-06-18	<ul style="list-style-type: none">• Endpoint change due to move from Digital Ocean to Cybera<ul style="list-style-type: none">◦ New value: <code>http://199.116.235.207/grebe/api/</code>• Minor configuration fixes and documentation updates
1.0.1	2017-06-10	<ul style="list-style-type: none">• Changed endpoint, port number no longer needed<ul style="list-style-type: none">◦ Old value: <code>http://159.203.60.220:5000/grebe/api/</code>◦ New value: <code>http://159.203.60.220/grebe/api/</code>• Province field is now required• Removed paging limitations• Returned data is now downloaded to text file• Resolved field dot notation compatibility issues in returned data
RC 1.0.0	N/A	First stable production version release