

TwitterWars

Installation

Environment: Mac OS X (v10.7.5)

Development: NetBeans IDE 7.2

MAMP (Mac Apache MySQL PHP) Version 2.1.1

PHP

Version: 5.4.4

Apache Port: 8888

MySQL

Socket: /Applications/MAMP/tmp/mysql/mysql.sock

Edit /etc/php.ini to set this socket to mysql.default_socket

MySQL Port: 8889

Apache document root: /Applications/MAMP/htdocs

Configuration

Unzip TwitterWars.zip into Apache's document root folder.

Then open the Terminal to run some scripts, and finally open a browser (I used Mozilla Firefox 15.0.1) to view some results.

Database

Open phpMyAdmin() in browser and import twitter_wars.sql. Then run this to create a database with the tables london_tweets and exeter_tweets. They contain Twitter data for the last week.

Create a user and grant him full privileges to twitter_wars database. In my case the SQL connection is established through the following info in config.php:

```
define('DB_HOST', 'localhost');  
define('DB_USER', 'andrei');  
define('DB_PASS', 'wdtTxjd8wczPewfS');  
define('DB_NAME', 'twitter_wars');
```

Running

Fetching Twitter Data

To connect to and consume the Twitter stream via the Streaming API:

```
$ cd /Applications/MAMP/htdocs/TwitterWars  
$ php twitter_fetcher.php
```

This will add real-time statuses from London and Exeter areas into the database.

To disable printing the statuses to the console pass “h” as the 1st argument:

```
$ php twitter_fetcher.php h
```

Retrieving older statuses

To get twitter statuses for the last week:

```
$ php twitter_search.php [place]
```

where [place] is (optionally) either “London” or “Exeter”. If not specified then the statuses from both places are retrieved. Example:

```
$ php twitter_search.php London
```

This will retrieve 1000 statuses per day (that are not already in database) for the last 7 days. You can change this by altering *DAYS_COUNT* and *TWEETS_PER_DAY* constant members of *TwitterSearch* class.

Spellchecking

This work is done by *dbspell.php* for statuses already fetched and added to the database. From Terminal execute:

```
$ php dbspell.php [since] [until] [place]
```

[since] Analyze statuses after this “Y-m-d” date.

[until] Analyze statuses no older than this “Y-m-d” date.

[place] London or Exeter. or all statuses in the given interval, from Exeter.

Example:

```
$ php dbspell.php 2012-09-16 2012-09-20 Exeter
```

This will check the spelling for all statuses in the given interval, from Exeter.

Note: Only statuses that have not been analyzed yet will be considered.

Compare Spelling

In order to compare the spelling quality of statuses from London and Exeter, I generated some statistics in a graph representation, over several days. To obtain this, enter the following URL in your browser:

<http://localhost:8888/TwitterWars/index.php>

This will output the spelling correctness for the tweets over the last week, for those that are spellchecked.

You can also specify the number of days for which to generate the report, before a given date. Example:

<http://localhost:8888/TwitterWars/index.php?days=5&until=2012-09-19>

will output this:

