

Computer Science 290 Team 3 Final Project Requirements Document

Design Specification

For the final project, the TwitterAnalytics system implements five analysis tools on Twitter tweets by a user. The system must be able to perform these analyses on a .zip archive provided by the user, and also on their live timeline. It should store all of these tweets in a SQLite 3 database.. The tool will be required to authenticate with the Twitter REST API version 1.1 to connect a user with their most recent tweets. In addition, the system must implement a clearing functionality to reset the internal database, and must also allow the use of multiple zip file inputs. Additionally the system must implement a command line interface and parse these inputs using the JCommander software tool.

System Behavior

Our system will implement a graphical user interface (GUI) and a command line interface for our users to interact with. The user interfaces will expose all system functionality as specified above, giving the user a large amount of control over their data and analytics.

1. Command Line Interface

Usage:

```
java -jar burdwatch.jar [OPTION...] [ANALYZER]
```

Performs analysis on the available tweet data using the specified analyzer

date	Identifies the date on which the most tweets occurred
pattern PATTERN	Prints all tweets that contain the given pattern
ratio	Prints the user most retweeted. Requires an Internet connection to retrieve the username, otherwise prints the ID number.
device	Prints the ratio of retweets to non-retweet tweets
most-retweeted	Identifies and prints the device most used to post tweet

OPTIONS:

-auth FILE	Connect to the Internet by performing Twitter authentication, and write token result to the specified file for later
-oa FILE	Connect to the Internet using the OAuth file created previously by the -auth option
-zip FILE	Adds the contents of the specified Twitter archive file to the internal database.
-verbose -v	Verbose mode, which prints out technical errors messages
-clear -c	Clears the database first before doing anything else

2. Analytics

The following twitter analytics will be supported by the system:

- All tweets that contain a specified character pattern
- The localized string representing the day of the week on which this user tweeted the most, formatted to the user's locale
- The user most commonly retweeted by this user
- The ratio of retweets to non-retweets
- The device most used by this user to send tweets