

Summary

The TwitterVane tool will allow you capture URL's related to a specific topic of interest in a collection. It extracts and analyses URLs embedded in the tweet to allow reporting on top URLs and domains for a given collection.

This is a [prototype](#) application and not yet fully developed. You are free to download the code for your own use but at your own risk.

Application Configuration

To access the application configuration page click on 'Configuration' link on the right side menu bar.

Configuration Options

Twitter API

The Twitter API OAuth fields are required.

You must supply your own API credentials. Create a developer account with twitter and creating a new application and use the supplied credentials.

Bit.ly API

This is optional. You can get your own Bit.ly credentials by creating an account at the bit.ly web site.

If you don't supply bit.ly credentials the application will still attempt to get the expanded URL from the bit.ly shortened URLs.

Proxy Settings

Optional. You can add proxy settings if needed to contact the Twitter API.

Collections

Create Collection

To access the collections page click on the 'Collection's link in the top menu.

Under the heading 'Add New Collection' fill in the Title, Start Date and End Date for the Collection and press add.

The date represent the period during which the application will collect tweets related to search terms for that collection.

Edit Collection

To edit a collection click on the collection title on the collections page.

Edit fields as desired and click save.

Add Term(s) To Collection

To add search terms to a collection click on the ‘Collections’ link on the top menu. On the collections page click on the title of the collection you want to add search terms to. This will take you to the collection edit page.

Add a term in the term field and click on add.

You can repeat this as many times as needed to add different search terms.

Twitter Stream Management

To access the twitter stream management page click on the ‘Twitter Stream’ link in the admin menu on the right.

This page will give you information on the status of the twitter stream and information of required application configuration for the twitter stream. It will also allow you to start and stop collecting tweets.

Under the heading of ‘Twitter Stream Required’ you will find any missing configuration required for starting collection in red.

In order to be able to collect tweets you will need to fill in the Twitter API OAuth credentials on the configuration page and also have at least one search term added to a collection.

When all requirements are met you can click on the ‘Start’ and ‘Stop’ links to start and stop twitter collection.

Twitter Analysis

After tweets have been collected they must be further processed to expand shortened URLs, match the URL to a collection and to extract the domain.

To access the twitter analysis page click on the 'Analysis Jobs' link on the right side menu.

The twitter analysis page will give you information on the number of tweets that have been collected, the number of URL's that have been analysed and the number of tweets waiting to be analysed.

To process remaining tweets enter the number of tweets to be processed and click on the submit button. *This can take a long time to process depending on the number you choose.

The number of analysed URLs can be higher than the total number of tweets. This is because a tweet can have multiple URLs.

It should be noted here again that this is a prototype application and with further development the tweets analysis should become an automated process.

Tasks

Purge processed tweets – This task will clear all tweets that have been processed.

Purge failed analysis – This task will clear all URLs that failed during the analysis process.

Purge all analysis - This will clear all analysis and reports. This is not reversible so make sure you want to do this, particularly if you have cleared processed tweets.

Purge all tweets - This will clear all tweets including unprocessed. Again, not reversible.

Reporting

The reporting page allows you to report on analysed URLs. There are four different types of reports. The results page of all reports gives you the first 100 results but you can use paging to see more results.

Top URL: This report gives you a listing of all URLs ordered by the number of times they appear.

TOP Domain: This report gives you a listing of top domains ordered by the number of times they appear.

TOP URL by Popularity: This report gives you a listing of top URLs based on the highest re-tweet number given per a single URL.

Failed: This report lists any URLs that failed during the analysis phase. A likely reason for this is that a shortened URL was unable to be expanded.

XML Response for TOP URL: On the reporting page there is 'XML Request' link at the bottom. You can use this link instead of the 'Generate Report' button and it will give you an XML response rather than HTML. This can be used as a method of automatically collecting report data from another application or script.

Application Administration

This section is for installation of the application.

Requirements

Tomcat 6 or higher

Installation

Deploy war file in tomcat.

Other

The application uses an odb database stored within the WEB-INF directory of your installation. Should you need to re-deploy your application make sure you have a backup of the db directory if you want to restore it.