Twitter and YouTube Importers

Jeff Reminga
The CASOS Center
COS Program, School of Computer Science, Carnegie Mellon
Summer Institute 2020

Overview

- How to use the Data Import Wizard to import Twitter data
- Twitter data is packaged by a number of companies, and ORA has importers to parse a variety of sub-formats
- Open the Data Import Wizard and type “Twitter” in the search bar
- YouTube is another data format that is easily imported into ORA
Data Import Wizard

- Navigate to any sample twitter data provided by CASOS for the Summer Institute
- This data is in the Twitter JSON format, which is the official, standard format from Twitter
- Here the flood dataset is selected
- Note that multiple files, or an entire directory of files could be selected to import
- There are many import options
Twitter Importer – General Options

- The first tab, General Options, contains a scrollable list of options for importing.
- The red box contains the options for parsing the tweet message words using a thesaurus and/or delete list.
- The blue box contains options for Aggregating (i.e. Chunking or combining) tweets by date of posting.

Twitter Importer – Derived Networks

- The second tab “Derived Networks” lists the networks that ORA will create from the basic networks that it creates during import.
- Derived networks never need to be created in order to run the Twitter Report.
- Derived networks can be (re)computed at a later time using the main menu item:

Data Management \ Meta-Network Derived Data
Twitter Importer – Custom Attributes

- The third tab “Custom Attributes” allows you to create tweet and agent node attributes from the underlying Twitter JSON data.
- This is an advanced option.

Twitter Importer – Finish

- Click Finish
- The file will be imported and then a dialog will report the number of tweets parsed and the date range.
- Click OK
- Note that the Data Import Wizard closes, and a new meta-network is in the main ORA interface.
Twitter Importer – Again!

- Let’s import again, but this time Aggregating (chunking) the tweets into 15 minute time periods.
- Open the Data Import Wizard
- Note how it maintains the settings from the last import
- This time click “Aggregate” and select 15 Minute(s)

Twitter Importer – Aggregation

- Click Finish
- The ending dialog now describes that there were retweets placed into the time periods of their creation (there is an option to not do this)
- Click OK
Twitter Importer – Aggregation

- The main ORA interface now contains a dynamic meta-network
- It contains one meta-network per 15 minute time period
- It contains the meta-networks for the time periods in which the retweets were created

YouTube Importer

- Use the YouTube importer to import in a few clicks YouTube data
- This data contains video, channel, comment, and commenter information
- Open the Data Import Wizard and type YouTube in the search bar
YouTube Importer

- Click Next and select to create a new meta-network
- Then click Next

- Browse to the YouTube sample data
- Click Create Derived Networks
- Click Finish
- A new, single meta-network will have been added to the main ORA interface
YouTube Report

- Return to the main ORA interface and select the newly created meta-network
- Click the Generate Reports button
- This dialog appears
- Type YouTube in the search bar and click next

YouTube Report

- The report has no options and so the same Save panel appears as for the other reports
- Click Finish
- The HTML YouTube report will appear with simple analyses on the YouTube activity