FireAnt (Windows)

Build 1.1.3 (Released March 10, 2017)

Laurence Anthony, Ph.D.
Center for English Language Education in Science and Engineering, School of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan.
Help file version: 001.

Introduction

FireAnt is a freeware "[F]ilter, [I]dentify, [R]eport & [E]xport [An]alysis [T]oolkit" for processing small, large, and very large tabular and hierarchical data sets, such as those generated by the Twitter API for use in corpus, time series, and network graph analyses. Data sets in JSON, BSON, CSV, TSV (TXT), and Excel (XLSX) formats that are **UTF-8 encoded** can be loaded directly into FireAnt for processing. FireAnt can filter the data sets in various ways, such as by date, id, frequency, and column content, and export the results as JSON, CSV, TSV (TXT), or Excel (XLSX) files. For a detailed explanation of the program's features, see the explanation below.

FireAnt runs on any computer running Microsoft Windows (tested on Win 98/Me/2000/NT, XP, Vista, Win 7, Win 8) and Macintosh OS X computers (tested up to OS X 10.9 Mavericks). *FireAnt* is developed in Python and Qt using the *PyInstaller* compiler to generate executables for the different operating systems.

FireAnt is developed in collaboration with Claire Hardaker of Lancaster University, UK.

Getting Started (No installation necessary)

Windows

On Windows systems, simply double click the *FireAnt* icon to launch the program.

Macintosh OS X

On Macintosh systems, simply double click the *FireAnt* zip file. The zip file will unzip the *FireAnt* application. Then, you can drag the *FireAnt* application to your application folder, your desktop, or anywhere else you like. Throw away the zip file when you are finished.

Loading Data Sets

File can be loaded in any of the following ways:

1) Click on the File->Open Data File(s) menu option and select the files you want.
2) Click on the "Input Data File(s)" Open button in the main display and select the files you want.
3) Drag and drop files directly onto the *FireAnt* application.
   Note: If you click on the File->Close Data File(s) menu option or the "Input Data Files" Close button in the main display, the files will removed from the list.

Importing an Existing Database

A previously exported database can be loaded in any of the following ways:

1) Click on the File->Import Database menu option and select the database that you want to import.
2) Click on the "Input Data File(s)" Open button in the main display and select the database that you want to import.
3) Drag and drop files directly onto the *FireAnt* application.
   Note: If you click on the File->Close Data File(s) menu option or the "Input Data Files" Close button in the main display, the database will removed from the list.

Exporting a Loaded Dataset

Once a dataset (or database) is loaded into *FireAnt*, the database can be exported for future use by clicking on the File->Export Database menu option and selecting an appropriate name and location for the exported database.

Exporting Raw Filtered Data

Filtered and processed results can be saved as raw data via the File->Exporting Raw Filtered Data dialog box (see the figure on the right). To export your data, follow the steps below:
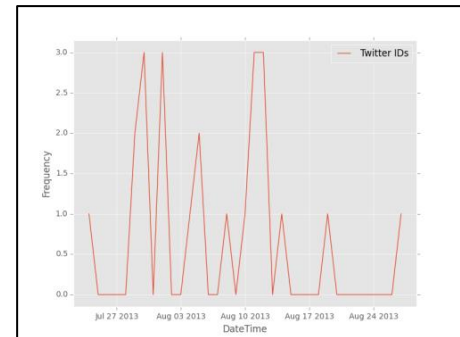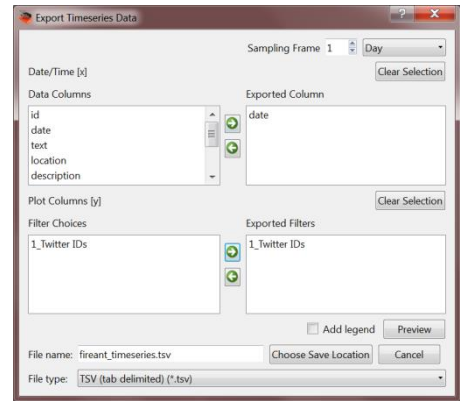
1. Choose which formats to use for the exported data (e.g. CSV, JSON) by checking one or more options of the Export options.
2. Choose which columns of the dataset to export by clicking on one or more of the column names in the left list and then clicking the right arrow button. Any selected column names in the right list can be removed by clicking on the left arrow.
3. Choose a file path for the selected output data. All data will be saved with an appropriate file extension (e.g. .txt, .csv, .json).
4. Click the "Save" button.
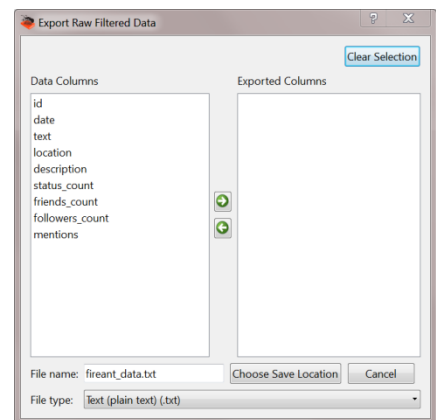
Exporting Timeseries Data

Filtered and processed results can be saved as time series data via the File->Export TimeSeries Data dialog box (see the figure on the right). To export your data, follow the steps below:
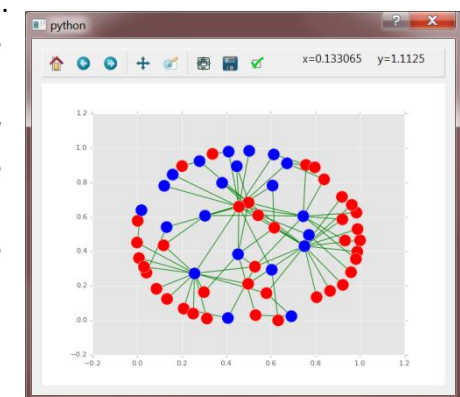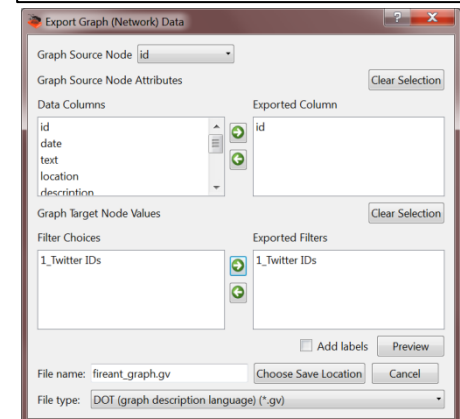
1. Choose which formats to use for the exported data (e.g. TSV, CSV) by checking one or more options of the Export options.
2. Choose the sampling frame (e.g. 1 day, 1 month)
3. Choose which data/time column to use for the time series x axis by clicking on one of the column names in the left list and then clicking the right arrow button.
4. Choose which filter(s) to use for the y axis of the plot.
5. Choose whether or not to add a legend to the plot using the checkbox option.
6. Preview the plot by clicking on the Preview button. The preview window can also be used to save a plot.
7. Choose a file path for the selected output data. All data will be saved with an appropriate file extension (e.g., .csv).
5. Click the "Save" button.

Exporting Graph (Network) Data

Filtered and processed results can be saved as network graph data via the File-> Exporting Graph (Network) Data dialog box (see the figure on the right). To export your data, follow the steps below:

1. Choose which formats to use for the exported data (e.g. TSV, CSV) by checking one or more options of the Export options.
2. Choose the source node column for the graph by selecting one of the columns in the combobox.
3. Choose the source node attributes to include in the graph by clicking on one of the column names in the left list and then clicking the right arrow button.
4. Choose which filter(s) to use for the target node values of the graph.
5. Choose whether or not to add labels to the graph using the checkbox option.
6. Preview the graph by clicking on the Preview button. The preview window can also be used to save a graph (see the figure on the right).
7. Choose a file path for the selected output data. All data will be saved with a DOT format .gv file extension.
8. Click the "Save" button

Exporting Geoposition Data (Maps)

Filtered and processed results can be saved as geoposition (map) data via the File->Export Geopositions Data dialog box (see the figure on the right). To export your data, follow the steps below:

1. Choose which formats to use for the exported data (e.g. TSV, CSV) by checking one or more options of the Export options.

2. Choose the geoposition latitude and longi tude columns as well as a column containing map lables by selecting the appropriate column names using the comboboxes.

3. Choose which filter(s) to use for color-coding the map points.

4. Preview the map by clicking on the Preview button. The preview window can also be used to save the map (see the figure on the right).

5. Choose a file path for the selected output data. All data will be saved with an appropriate file extension (e.g., .csv).

6. Click the "Save" button.
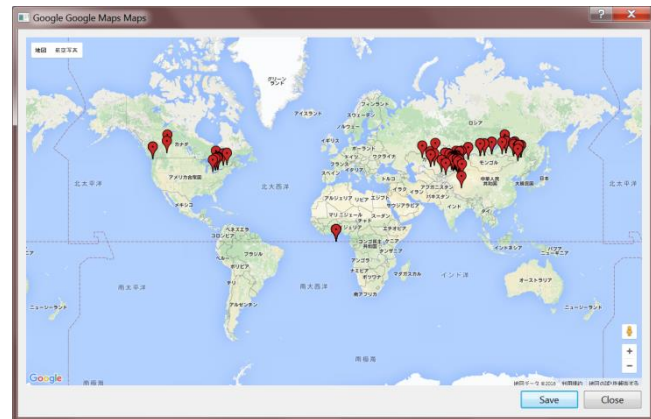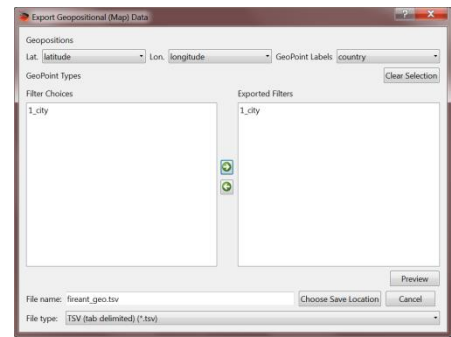
Importing Twitter Data

Data can be collected from Twitter directly through *FireAnt* via the File->Collect Twitter Data dialog box (see the figure on the right). To export your data, follow the steps below:

1. Click the "Authorize" button to load the Twitter authorization web page, where you can sign in and authorize *FireAnt* to use your account to collect Twitter data. Note that throughout the authorization process, *FireAnt* never sees or needs to store your User ID or Password information. **Your privacy details are completely safe**.

2. After authorizing *FireAnt*, the Twitter webpage will show you a PIN number. Copy and paste the PIN code into the textbox below the page and click "Generate Key File" to complete the authorization process. A key file with your authorization details are stored in a file in the same location where *FireAnt* was launched. Please keep this file together with *FireAnt* so that you do not need to go through the authorization process again.

3. Choose a file to store your Twitter data. This file will be created at the start of the collection process and updated whenever new data arrives. Note that if you choose a previous file, all data in the file will be erased.

4. Set the parameters for your Twitter collection. Here you can either collect a sample/filtered collection or the recent tweet history of specified users

   Sample/Filtered Collection

   a. Start Time: Start immediately or set a date/time in the future.
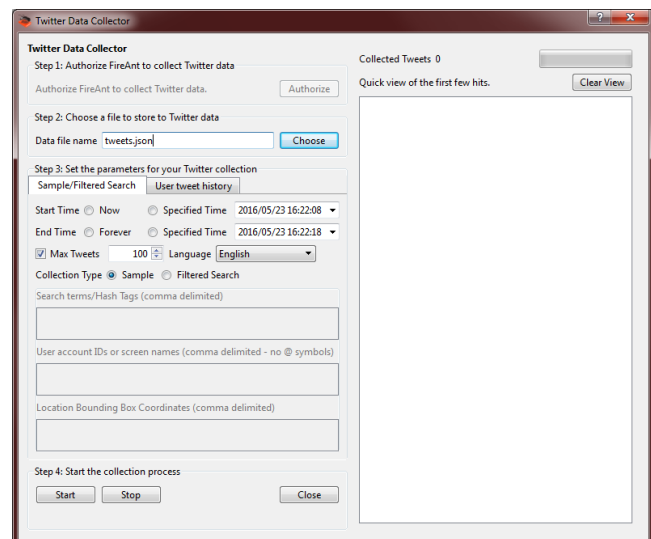
b. End Time: Stop the collection yourself, or set a date/time in the future.

c. Max Tweets: Set the maximum number of tweets to collect before the end time specified.

d. Language: Set the locale where the Tweets will be collected.

e. Search filters: Set the search terms, hash tags, account IDs/screen names, and locations boundaries for your   search criteria. All filter values should be separated by commas. For more details,   see here (https://dev.twitter.com/streaming/overview/request-parameters).

User Tweet History

a. Max Tweets: Set the maximum number of tweets to collect from the specified user(s). Note that Twitter has a limit of the most recent 3000+ tweets from a user. So, if you specify a maximum number above this limit, only the tweet limit will be returned.

b. User account IDs or screen names: List the user account IDs and/or screen names of the users you want to collect the tweet history for. The values should be separated by commas.

c. Include retweets: Include retweets in results set

5. Click the "Start" button, and wait for the collection to finish. The number of hits and the stream of tweets is shown in the status box on the right of the main window. The results are stored in the JSON format and can be loaded directly into *FireAnt* for further analysis.

Viewing Input Data

Once a dataset (or database) is loaded into *FireAnt*, the columns of the dataset can be seen in the top right table of FireAnt. The number of rows per page can be set using the spinbox, and clicking the "Show" button will update the page if the page size is changed. To see the previous and next pages, click one of the navigation buttons. Any column in a table page can be sorted by clicking on the column header.

Filtering Input Data

To filter the dataset on one or more columns, follow the steps below:

1. Click on either the "Add Filter" or "Add SubFilter" button in the left of the main screen of *FireAnt* to introduce a new filter condition. Each time you press the "Add Filter" button, you will introduce a new **main** filter condition. If you press the "Add SubFilter" button after adding a main filter, you will introduce a new **sub-filter** condition below the main filter. Similarly, if you press the "Add SubFilter" button after adding a sub-filter, you will introduce a new sub-filter condition below that, leading to ever more specific filter conditions. The diagram below illustrates the hierarchical relationship between filters and sub-filters:

```
FILTER
   SUB-FILTER
           SUB-FILTER
                   …
FILTER
   SUB-FILTER
           SUB-FILTER
```
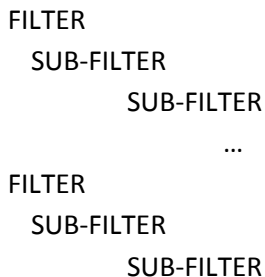
2. Once you have chosen to add a filter or sub-filter, *FireAnt* will display a new filter (or sub-filter) listing all possible columns in the dataset. Click on any column name in the filter to reveal a set of custom options that you can use to select data in that column. Depending on the type of data contained in your dataset column, *FireAnt* will pick the most suitable filter options for you. For example, if you dataset column

contains data/time information, *FireAnt* will present options to pick the starting and ending year, month, day, and time for your filter. Similarly, if your dataset column contains text information (e.g. tweets), *FireAnt* will present options to find words or phrases in that column. The following custom filters are currently available (see below for examples of each filter):

- Default filter: A text based filter with word/phrase search options (including regular expression searching)
- Date filter: A data/time filter with start date/time and end date/time options
  - ➢ column headers need to end with the string: "created_at" "createdon"
- Count filter: A filter designed to choose data based on min. and max values
  - ➢ column headers need to end with the string: "count"
- Group filter: A filter that groups names/values in a column and ranks these names/values their frequency of occurrence. Particular groups can be searched for and selected before filtering.
  - ➢ column rows containing lists of names/values will be expanded before the grouping in carried out
  - ➢ column headers need to end with the string: "id", " "name", "location", "hashtags" "mentions"

3. Click the "Apply" button to filter the results. The filtered results are shown in the bottom right table. The number of rows per page can be set using the spinbox, and clicking the "Show" button will update the page if the page size is changed. To see the previous and next pages, click one of the navigation buttons. Any column in a table page can be sorted by clicking on the column header.
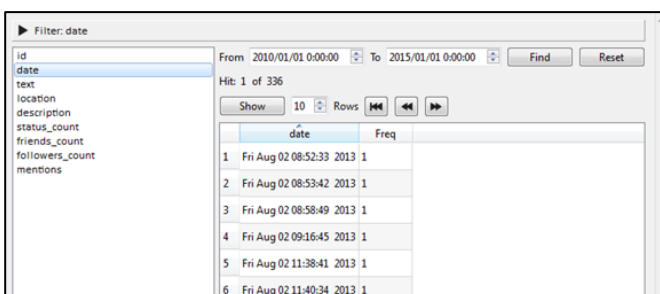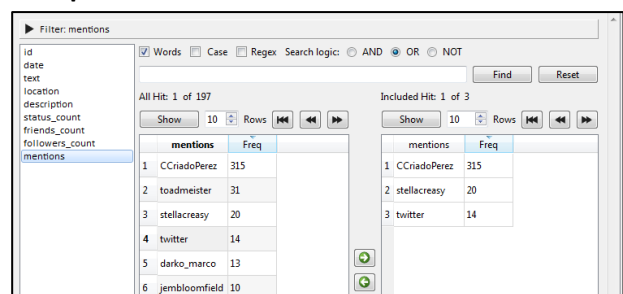
Default Filter
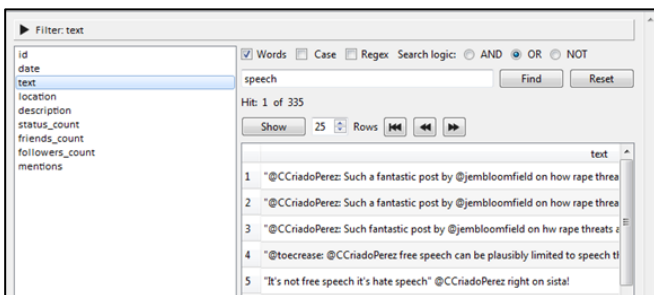
Count Filter
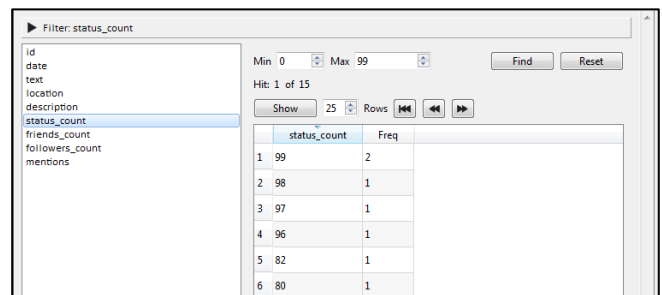
Date Filter

Group Filter

Shortcuts

Information in all tables and filters can be selected, copied, and pasted as is standard on the operating system:

Windows:	CTRL-A ➪ Select All	CTRL-C ➪ Copy	CTRL-V ➪ Paste
Macintosh:	CMD-A ➪ Select All	CMD -C ➪ Copy	CMD -V ➪ Paste

NOTES

Comments/Suggestions/Bug Fixes

All new editions and bug fixes are listed in the revision history below. However, if you find a bug in the program, or have any suggestions for improving the program, please let me know and I will try to address the issues in a future version. This software is available as 'freeware' according to the license below. It is important for my funding to hear about any successes that people have with the software. Therefore, if you find the software useful, please send me an e-mail briefly describing how it is being used.

ACKNOWLEDGMENT

CITING/REFERENCING *FireAnt*

Use the following method to cite/reference *FireAnt* according to the APA style guide:

Anthony, L. and Hardaker, C. (YEAR OF RELEASE). *FireAnt* (Version VERSION NUMBER) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/

For example if you download *FireAnt 1.0*, which was released in 2016, you would cite/reference it as follows:
Anthony, L. and Hardaker, C. (2016). *FireAnt* (Version 1.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from http:// www.laurenceanthony.net/

Note that the APA instructions are not entirely clear about citing, and it is debatable whether or not the "Available from ..." statement is needed. See here for more details:
http://owl.english.purdue.edu/owl/resource/560/10/

LICENSE for FireAnt

FireAnt 1.0 and any minor updates issued by AntLab Solutions (collectively 'the Software')

TERMS GOVERNING THE USE OF THE SOFTWARE
The Software is protected by copyright and must not be used, displayed, modified, adapted, distributed, transmitted, transferred or published or otherwise reproduced in any form by any means other than strictly in accordance with the terms set out below. By installing the Software, you agree to be bound by the terms of the license. This FireAnt License ("License") is made between AntLab Solutions, Tokyo, Japan as licensor, and you, as licensee, as of the date of your use of the Software. The Software is in use on a computer when it is loaded into the RAM or installed into the permanent memory of that computer, e.g., a hard disk or other storage device.

1. License Material
These terms govern your use of the Software but not including subsequent versions (e.g. FireAnt 2.0').

2. License Grant
AntLab Solutions grants to you a personal non-exclusive non-transferable license ('the License') to use the Software in the following specific contexts.

a)   Non-Commercial (Freeware) Use:
You may use the software for non-profit purposes on more than one computer or on a network so long as you are the sole user of the Software. (A "network" is any combination of two or more computers that are electronically linked and capable of sharing the use of a single software program.) You are not permitted to sell,

lease, distribute, transfer, sublicense, or otherwise dispose of the Software, in whole or in part, for any form of actual or potential commercial gain or consideration.

b)   Commercial Evaluation (Trial) Use:
You may evaluate (trial) the software for commercial purposes for a period of no more than fourteen (14) days from the date of download on more than one computer or on a network so long as you are the sole user of the Software.

c)   Commercial Use
When you pay the commercial license fee established by AntLab Solutions, you may use the software for non-profit or commercial purposes on more than one computer or on a network so long as you are the sole user of the Software. (A "network" is any combination of two or more computers that are electronically linked and capable of sharing the use of a single software program.) You will obtain a separate license for each additional user of the Software (whether or not such users are connected on a network). You are not permitted to sell, lease, distribute, transfer, sublicense, or otherwise dispose of the Software, in whole or in part, for any form of actual or potential commercial gain or consideration.

3. Termination
You may terminate this License at any time by uninstalling the Software and deleting it. The License will also terminate if you breach any of the terms of the License.

4. Proprietary Rights
The Software is licensed, not sold, to you. AntLab Solutions reserves all rights not expressly granted to you. Ownership of the Software and its associated proprietary rights, including but not limited to patent and patent applications, are retained by AntLab Solutions. The Software is protected by the copyright laws of Japan and by international treaties. Therefore, you must comply with such laws and treaties in your use of the Software. You agree not to remove any of AntLab Solutions' copyright, trademarks, and other proprietary notices from the Software.

5. Distribution
Except as may be expressly allowed in Section 2, or as otherwise agreed to in a written agreement signed by both you and AntLab Solutions, you will not distribute the Software, either in whole or in part, in any form or medium.

6. Transfer and Use Restrictions
You may not sell, license, sub-license, lend, lease, rent, share, assign, transmit, telecommunicate, export, distribute or otherwise transfer the Software to others, except as expressly permitted in this License Agreement or in another agreement with AntLab Solutions. You may not modify, reverse engineer, decompile, decrypt, extract, or otherwise disassemble the Software.

7. Warranties
ANTLAB SOLUTIONS MAKES NO WARRANTIES WHATSOEVER REGARDING THE SOFTWARE AND IN PARTICULAR, DOES NOT WARRANT THAT THE SOFTWARE WILL FUNCTION IN ACCORDANCE WITH THE ACCOMPANYING DOCUMENTATION IN EVERY COMBINATION OF HARDWARE PLATFORM OR SOFTWARE ENVIRONMENT OR CONFIGURATION, OR BE COMPATIBLE WITH EVERY COMPUTER SYSTEM. IF THE SOFTWARE IS DEFECTIVE FOR ANY REASON, YOU WILL ASSUME THE ENTIRE COST OF ALL NECESSARY REPAIRS OR REPLACEMENTS.

8. Disclaimer
ANTLAB SOLUTIONS DOES NOT WARRANT THAT THE SOFTWARE OR SERVICE IS FREE FROM BUGS, DEFECTS, ERRORS OR OMISSIONS. THE SOFTWARE OR SERVICE IS PROVIDED ON AN "AS IS" BASIS AND ANTLAB

SOLUTIONS MAKES NO OTHER WARRANTIES OR CONDITIONS, EXPRESS OR IMPLIED, WITH RESPECT TO THE SOFTWARE INCLUDING WITHOUT LIMITATION THE IMPLIED WARRANTIES OR CONDITIONS OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

9. Limitation of Liability
ANTLAB SOLUTIONS WILL HAVE NO LIABILITY OR OBLIGATION FOR ANY DAMAGES OR REMEDIES, INCLUDING, WITHOUT LIMITATION, THE COST OF SUBSTITUTE GOODS, LOST DATA, LOST PROFITS, LOST REVENUES OR ANY OTHER DIRECT, INDIRECT, INCIDENTAL, SPECIAL, GENERAL, PUNITIVE OR CONSEQUENTIAL DAMAGES, ARISING OUT OF THIS LICENSE OR THE USE OR INABILITY TO USE THE SOFTWARE OR SERVICE. IN NO EVENT WILL ANTLAB SOLUTIONS'S TOTAL AGGREGATE LIABILITY (WHETHER IN CONTRACT (INCLUDING FUNDAMENTAL BREACH), WARRANTY, TORT (INCLUDING NEGLIGENCE), PRODUCT LIABILITY, INTELLECTUAL PROPERTY INFRINGEMENT OR OTHER LEGAL THEORY) WITH REGARD TO THE SOFTWARE AND/OR THIS LICENSE EXCEED THE LICENSE FEE PAID BY YOU TO ANTLAB SOLUTIONS. FURTHER, ANTLAB SOLUTIONS WILL NOT BE LIABLE FOR ANY DELAY OR FAILURE TO PERFORM ITS OBLIGATIONS UNDER THIS LICENSE AS A RESULT OF ANY CAUSES OR CONDITIONS BEYOND ANTLAB SOLUTIONS' REASONABLE CONTROL

10. Jurisdiction
These terms will be governed by Japanese law and the Japanese courts shall have jurisdiction.

KNOWN ISSUES
JSON files need to be formatted with singular JSON objects on each line.

REVISION HISTORY
1.1.3
Bug fixes:
1. When using the "User Tweet History" function, unchecking the "Max Tweets" option caused no Tweets to be collected. Also, after all Tweets were collected, the program would wait until the Stop button was pressed. This is now fixed.
2. Due to changes in the Twitter API, tweets that were longer than 140 characters were returned with truncated text. This has now been fixed. Note that the new 'extended' text column is called 'full_text'.

1.1.2
Bug fixes:
1. In previous versions of the "User Tweet History" function in Twitter collector, the value of the "Max Tweets" option from the "Sample/Filtered Search" setting was being incorrectly used to pause the collection process when a maximum value was set. This is now fixed.

1.1.1
Bug fixes:
1. Only one radio button was allowed to be selected. This is now fixed.
2. Version 1.1.0 introduced a new bug that caused the Twitter collector to not collect tweets when no user handle was specified. This is now fixed.
3. Version 1.1.0 introduced a new bug that caused the max tweets on/off checkbox in the Twitter collector to not have no effect. This is now fixed.

1.1.0
New features
1. The Twitter collection tool now accepts both user IDs and user screen names in the filter search.

2. The Twitter collection tool now has a "User tweet history" option to allow collecting of the most recent 3000+ tweets of a target user.
3. The Twitter collection tool interface has now been redesigned to be more compact enabling it to fit on low resolution screens.

Bug fixes:
1. The network graph export function has been adjusted so that sub-filter results are correctly reflected in the output graph. In previous versions, only the main filter results were reflected in the graphs.
2. The export function now conflates all target nodes to a single filter name. This leads to clearer and easier to interpret results.
3. On Mac OSX, the Key File (key_file.json) generated by the Twitter collection tool was saved in a sub folder of the main application meaning that the main program itself could not always find it. The location has now been moved to match that on Windows systems.

1.0.2
Bug fixes:
1. All export menu options have been updated to output correctly formatted data that can be loaded back into *FireAnt* for further analysis. The interfaces for the output options have also been simplified to allow a better user experience. In particular, the Graph output is now a single DOT format .gv file that can be loaded into most popular graph visualization tools, such as Gephi.

1.0.1
Bug fixes:
1. The Twitter steam did not stop when a specified end time was reached.
2. A very fast Twitter stream slowed the main interface as it was updating the "most recent tweets" viewing. To improve the response and to not have memory problems when millions of tweets are streamed, the viewer now only shows the first few tweets. This is currently set at a maximum of 25. This screenshot of the Twitter collector viewer has been updated in this help page.

1.0
This is the first release of the program