



FireAnt (Windows)

Build 2.2.1

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.

November 14, 2022

The screenshot shows the FireAnt application window. The main interface is divided into several sections:

- Input Data File(s):** A text box containing 'dt_timeline_tweets.json' with 'Open' and 'Close' buttons.
- Filter Tools:** Buttons for 'Add Filter', 'Add SubFilter', 'Remove', and two green arrows.
- Filter:** A section titled 'Filter: FilterName [text]'. It includes checkboxes for 'Words', 'Case', and 'Regex', and radio buttons for search logic: 'AND', 'OR' (selected), and 'NOT'. A text box contains 'President', with 'Find' and 'Reset' buttons. Below this, it shows 'Hits (aggregated): 1 of 109' and a list of 10 hits related to President Obama.
- Input Table:** A table with columns 'created_at' and text. It shows 'Row: 1 of 3205' and a 'Show' button with '25' rows selected. The table displays 8 rows of tweet data.
- Filtered Table:** A table showing 'Filtered Table: 109 / 3205 (3 %)' and 'Hit: 1 of 109'. It includes checkboxes for 'Handles', 'Hash tags', and 'URLs'. It shows a 'Show' button with '10' rows selected and a 'Randomize' checkbox checked. The table displays 8 rows of filtered tweet data.
- Progress:** A green progress bar and buttons for 'Apply' and 'Stop'.
- Filters applied:** A status bar at the bottom left.

Introduction

FireAnt is a freeware "[F]ilter, [I]dentify, [R]eport & [E]xport [A]nalysis [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 JSONL (.json), BSON, CSV, TSV (TXT), and Excel (.xlsx, .xls) 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 JSONL (.json), CSV, TSV (TXT), or Excel (.xlsx, xls) files. For a detailed explanation of the program's features, see the explanation below.

FireAnt runs on any computer running Microsoft Windows (built on Win 10), MacOS (built on Mac Catalina), and Linux (built on Linux Mint). It 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

Windows - Installer

Double click the *FireAnt.exe* file and follow the instructions to install the application into your Programs folder. You can delete the .exe file when you are finished. You can start the application via the Start Menu.

Windows - Portable

Unzip the *FireAnt.zip* file into a folder of your choice. In the *FireAnt* folder, double click the *FireAnt.exe* file to launch the program.

Macintosh OS X

Double click the *FireAnt.dmg* file to create a FireAnt disk image on your desktop. Open the disk image and drag and drop the FireAnt app onto the Applications folder (or into another location if you desire). You can then launch the app by double clicking on the icon in the Applications folder or the Launchpad.

Linux

Decompress the *FireAnt.tar.gz* file into a folder of your choice. In the *FireAnt* folder, double click the *FireAnt.sh* file to launch the software. On the command line, type `./FireAnt.sh` to launch the software.

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 be 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.

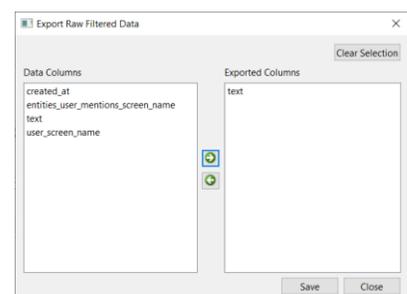
Note: If you click on the File->Close Data File(s) menu option or the "Input Data Files" Close button in the main display, the database will be removed from the list.

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:

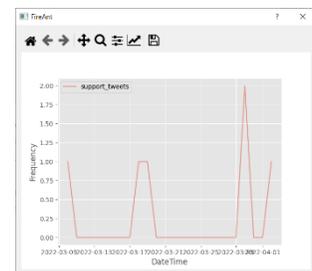
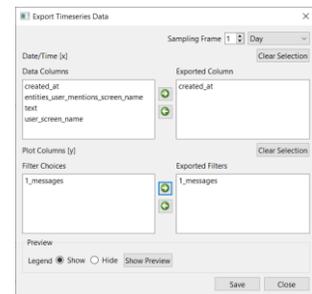


1. Choose which formats to use for the exported data (e.g. CSV, JSON) by checking one or more options of the Export options.
2. Choose which columns of the dataset to export by clicking on one or more of the column names in the left list and then clicking the right arrow button. Any selected column names 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:

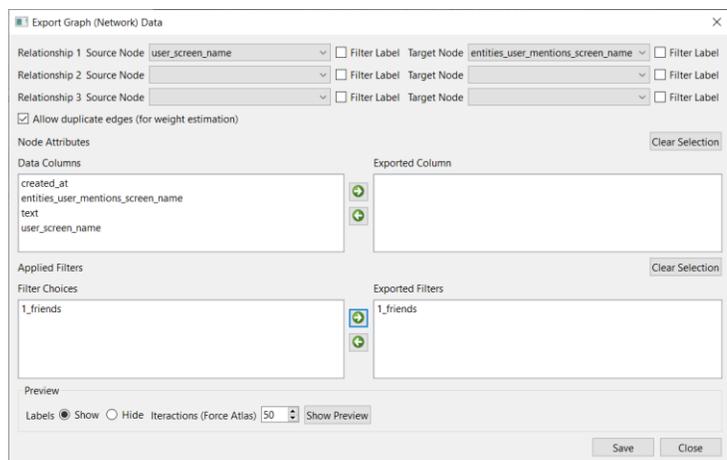
1. Choose which formats to use for the exported data (e.g., TSV, CSV) by checking one or more options of the Export options.
2. Choose the sampling frame (e.g., 1 day, 1 month)
3. Choose which data/time column to use for the time series x axis by clicking on one of the column names in the left list and then clicking the right arrow button.
4. Choose which filter(s) to use for the y axis of the plot.
5. Choose whether 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.



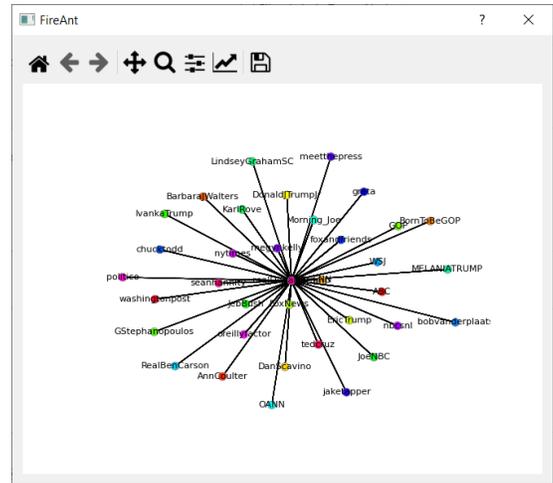
Exporting Graph (Network) Data

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 the relationships you want to visualize. You can draw up to three relationships. For each one, choose the parameters to serve as the source node and target nodes and decide if the data labels should be replaced by the filter names.
2. Choose to allow duplicate edges (which allows the duplicate count to be used as a weight measure)
3. Choose the node attributes to include in the graph.
4. Choose which filters to apply to the visualization



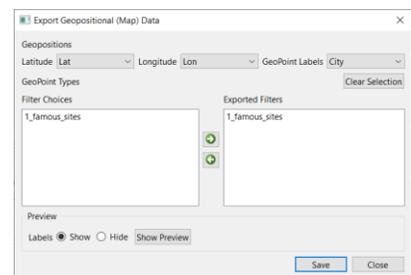
- Preview the graph by first choosing whether to add labels and then deciding how many iterations of the plotting algorithm to run before plotting the visualization. Finally, click the "Show" button. Note that more iterations will take longer to process but will result in a more stable/consistent preview. The preview window can also be used to save the previewed graph (see the figure on the right).
- Click "Save" to save the visualization data. All data will be saved with a DOT format .gv file extension. The saved data can be loaded into a visualization tool such as Gephi to produce a production ready graph.



Exporting Geo-positional Data (Maps)

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

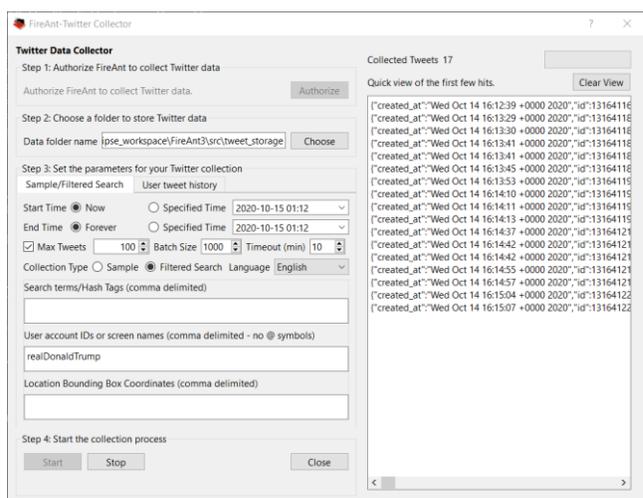
- Choose which formats to use for the exported data (e.g., TSV, CSV) by checking one or more options of the Export options.
- Choose the geo-position latitude and longitude columns as well as a column containing map labels by selecting the appropriate column names using the comboboxes.
- Choose which filter(s) to use for color-coding the map points.
- 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).
- Choose a file path for the selected output data. All data will be saved with an appropriate file extension (e.g., .csv).
- 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 collect your data, follow the steps below:

- 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.**



- 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 is 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 folder to store your Twitter data. This folder will be created at the start of the collection process and updated whenever new data arrives. Note that if you choose a folder containing data, you will be asked if you want to delete the data before you continue.
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

Start Time:	Start immediately or set a date/time in the future.
End Time:	Stop the collection yourself or set a date/time in the future.
Max Tweets:	Set the maximum number of tweets to collect before the end time specified.
Batch Size:	Set the number of tweets to be collected in a single tweet batch. After the number is reached, FireAnt will disconnect from Twitter, reconnect, and start collecting the more tweets in a new batch. This batch process will prevent data loss in the event of a system error during the collection process.
Timeout (min)	Set the number of min after which the tweet collector will restart if no tweets have been received. This setting is designed to overcome limitations with the Twitter API that can cause the feed to inadvertently stop returning tweets on occasion.
Language:	Set the locale where the Tweets will be collected.
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

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.
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.
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.

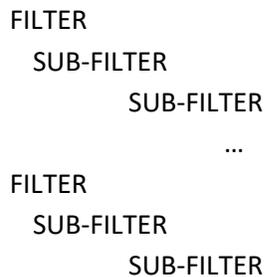
Viewing Input Data

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:

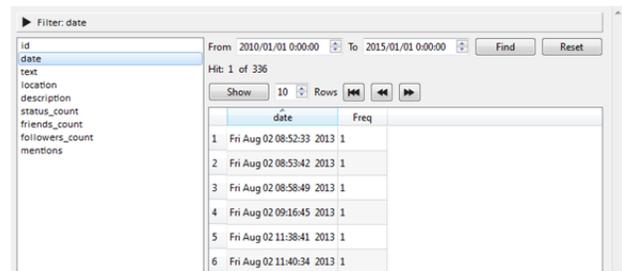
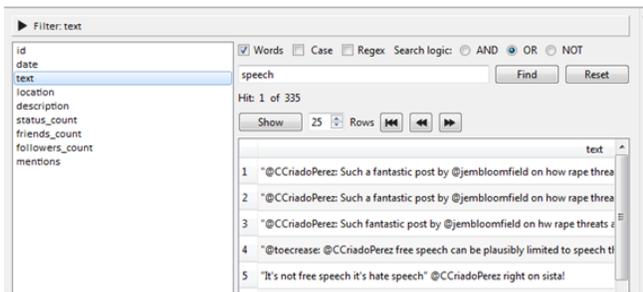
1. Click on either the "Add Filter" or "Add SubFilter" button in the left of the main screen of *FireAnt* to introduce a new filter condition. Each time you press the "Add Filter" button, you will introduce a new **main** filter condition. If you press the "Add SubFilter" button after adding a main filter, 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:



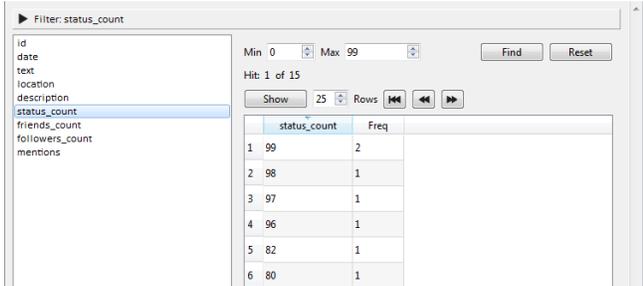
2. Once you have chosen to add a filter or sub-filter, *FireAnt* will display a new filter (or sub-filter) listing all possible columns in the dataset. Click on any column name in the filter to reveal a set of custom options that you can use to select data in that column. Depending on the type of data contained in your dataset column, *FireAnt* will pick the most suitable filter options for you. For example, if your 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. Groups can be searched for and selected before filtering.
 - column rows containing lists of names/values will be expanded before the grouping is 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

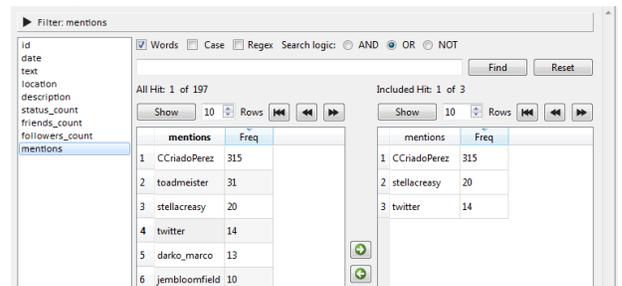
Date Filter



Count 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 |
| Linux: | CTRL-A ⇨ Select All | CTRL-C ⇨ Copy | CTRL-V ⇨ Paste |

NOTES

- If you have any suggestions for improving the software or notice any bugs, please post them in the FireAnt Discussion Group (<https://groups.google.com/g/fireant>). Indeed, many of the improvements and updates made to the software have been due to the comments of users around the world, for which I am very grateful. The FireAnt Discussion Group is also a good place to discuss how you are using the software and any challenges that your face.
- If you find the software useful in your research, teaching, or learning, you may consider making a small donation to support the future development of this tool. A link to the donation page can be found here: <https://www.laurenceanthony.net/software/fireant/>
- You may also be interested in becoming an AntLab patron. Depending on the level of support, this option will give you priority support with direct access to the developer (Laurence Anthony), and various other benefits. A link to the donation page can be found here
- <https://www.patreon.com/antlab>

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 2.0*, which was released in 2020, you would cite/reference it as follows:
 Anthony, L. and Hardaker, C. (2020). *FireAnt* (Version 2.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from [http:// www.laurenceanthony.net/](http://www.laurenceanthony.net/)

Note that the APA instructions are not entirely clear about citing, and it is debatable whether the "Available from ..." statement is needed. See here for more details:

<http://owl.english.purdue.edu/owl/resource/560/10/>

ACKNOWLEDGMENT

I would like to say thank you to the users of *FireAnt* who have taken the trouble to post feedback on the software and make suggestions for improvements and/or changes. A very special thank you goes to all those who have very generously supported the project either through single donations via [PayPal](#) or becoming a [Patreon supporter](#).

The development of this software was supported by the ESRC Centre for Corpus Approaches to Social Science, ESRC grant reference ES/L008874/1.

KNOWN ISSUES

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