Social Media Data Analysis with Mozdeh

Contents

[1 Workshop 1: Downloading, searching and exporting tweets 1](#_Toc63157697)

[1.1 Installing Mozdeh and downloading tweets matching a query 1](#_Toc63157698)

[1.2 Searching the downloaded tweets 2](#_Toc63157699)

[1.3 Filtering the downloaded tweets 3](#_Toc63157700)

[1.4 Country filters and nonbinary gender detection 3](#_Toc63157701)

[1.5 Exporting the downloaded tweets for use outside Mozdeh 3](#_Toc63157702)

[2 Workshop 2: Word association mining applied to timeline tweets 4](#_Toc63157703)

[2.1 Downloading tweets from one or more users 4](#_Toc63157704)

[2.2 Searching the downloaded tweets 5](#_Toc63157705)

[2.3 Filtering the downloaded tweets 5](#_Toc63157706)

[2.4 Word association mining 5](#_Toc63157707)

[2.5 [optional] Drawing time series graphs 6](#_Toc63157708)

[2.6 Twitter exploration task 6](#_Toc63157709)

[3 Optional 1: Collecting high quality sets of tweets from queries [for reference] 7](#_Toc63157710)

[4 Optional 2: Advanced optional Twitter analysis methods 7](#_Toc63157711)

[4.1 Twitter networks 7](#_Toc63157712)

[4.2 Topic changes over time 8](#_Toc63157713)

[4.3 Content analysis or thematic analysis text saving steps 9](#_Toc63157714)

[5 Workshop 3a: Downloading YouTube comments on a video 10](#_Toc63157715)

[6 Workshop 3b: Downloading comments on all videos in a YouTube channel 10](#_Toc63157716)

[7 Appendix: Registering for a YouTube key 11](#_Toc63157717)

[8 Bibliography 15](#_Toc63157718)

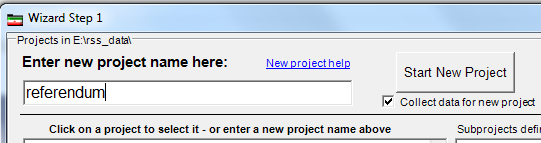
This booklet describes tasks to introduce the core features of Mozdeh for social media data analysis with Twitter and YouTube. If you hit problems, email Mike [m.thelwall@wlv.ac.uk](mailto:m.thelwall@wlv.ac.uk) or wave/shout in the online environment when you have a question. Please also share any interesting results or questions.

# Workshop 1: Downloading, searching and exporting tweets

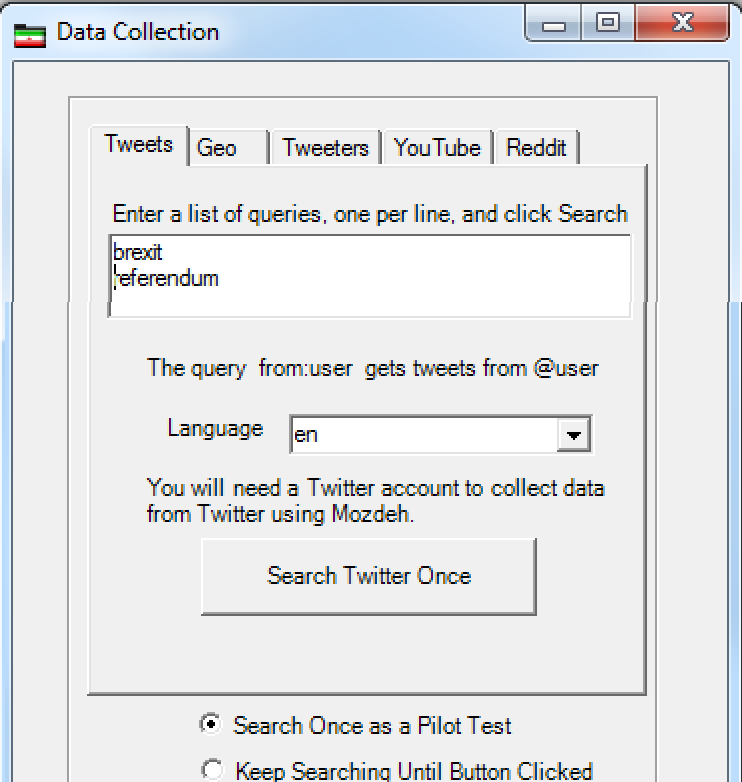
This section explains how to use Mozdeh to download tweets matching a query, then search and filter the downloaded tweets and export them to a simple file that can be loaded into Word or Excel for analysis outside Mozdeh.

## Installing Mozdeh and downloading tweets matching a query

1. Download *Mozdeh* from <http://mozdeh.wlv.ac.uk/> and save it to your computer or USB stick in a location, such as My Documents, where you can create folders. Mozdeh is free academic software from the University of Wolverhampton with no malware.
2. Start Mozdeh and follow the instructions about selecting a folder in which to store your data. Mozdeh will ask to create a folder called moz\_data on your computer. This folder will be initially empty but Mozdeh will populate it with subfolders, one for each Twitter project.
3. At Step 1, enter a name for a pilot test project (e.g., *referendum*) and click **Start New Project**.



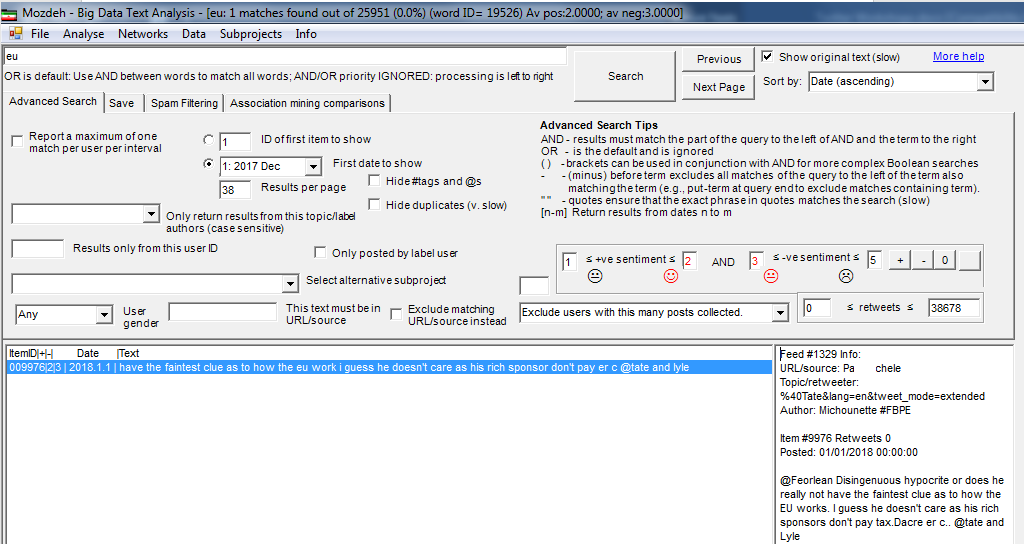
1. Choose your own set of queries (e.g., *“global warming”; or blackpink* and *bts)* in the Data Collection screen and select the matching language (e.g., *en* for English). Click the **Search Twitter Once** button. You will be taken to a web page asking you to logon to Twitter. This reveals a PIN to enter in Mozdeh for permission to download tweets on your behalf to your computer.



1. Mozdeh will now download tweets matching your queries to your computer. When finished, it will ask a series of questions – please click OK or give the suggested answer to these questions (the answers don’t matter for initial trials) and then you should get the main search screen.

## Searching the downloaded tweets

1. The initial screen for Mozdeh is its main search screen, which allows you to query the tweets that you have downloaded. Test this by entering a query (e.g., *EU*) and then clicking **Search** to see tweets that match that search. The query results are a bit strange – they reflect the words indexed by Mozdeh and the form in which they are indexed, which is not always the same as the original text. For instance, plural words are converted to singular versions. The original versions can be read by checking clicking a text in the search results list.



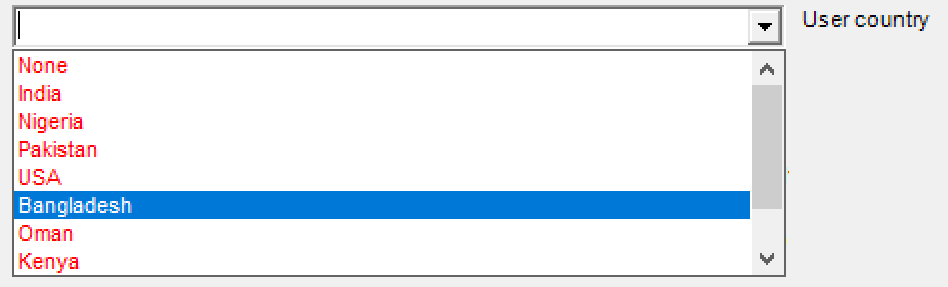
## Filtering the downloaded tweets

1. Mozdeh’s main search screen has a range of filters that can be applied to the downloaded tweets. Test this by clearing any query, selecting a **Female** filter (see User gender in the bottom left of the buttons) and clicking **Search** to see tweets authored by females. Click any result and check if the tweet author username is female. Change to a **Male** filter and click Search to see the new results. Clear the filter by selecting **Any gender** from the same dialog box. Notice that when a filter is set, it becomes red but when it is cleared, it returns to black. This makes set filters easier to see. Try the same with positive (+ button) and negative (- button) filters and with the minimum number of retweets increased from 0 to 10, clearing each one after use. In each case check a random tweet to see if it matches the query.



## Country filters and nonbinary gender detection

1. Projects sometimes need to focus on a country. Twitter reports country information for Tweeters, which Mozdeh can request. This requires another logon to Twitter. To detect user countries, select **Get countries of Twitter users** from the **Advanced** menu and follow the instructions. After this, try the Country: filter.

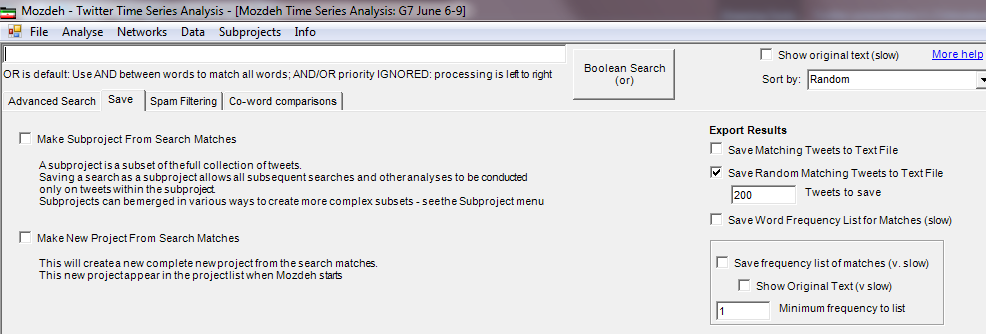


1. Projects sometimes need to identify nonbinary genders for Tweeters. Mozdeh normally guesses genders from first names but this does not work for nonbinary users. For this, Mozdeh can check for they/them pronouns in Twitter bios, tagging these tweeters as nonbinary. To activate this (after getting countries), select **Identify nonbinary, male and female genders from pronouns in bios** from the **Advanced** menu and follow the instructions. About 1 in 500 Tweeters declare they/them pronouns, so you would need a large sample of tweeters to identify many nonbinary tweeters. After this, try the nonbinary gender filter.

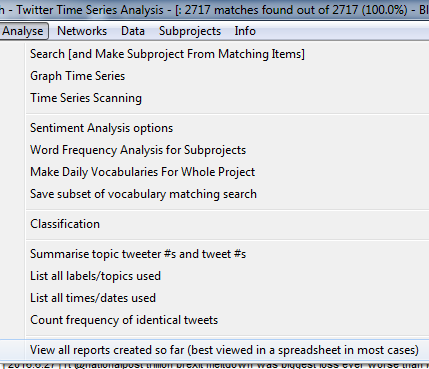


## Exporting the downloaded tweets for use outside Mozdeh

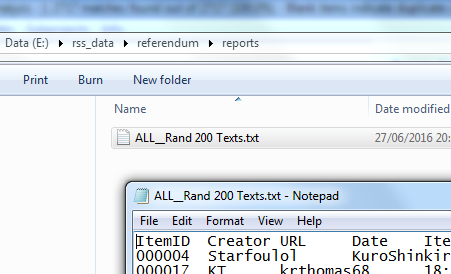
1. To export a random sample of tweets to a simple file for investigation outside Mozdeh, click the **Save** tab, and check **Save Random Matching Tweets to Text File (**or **Save Matching Tweets to Text File** to save all**)**. Now enter a blank search, select random from the **Sort by** drop-down box, and click the **Search** button if a Save As dialog box has not already appeared. Click OK when the Save As dialog box appears.



1. To find the file of random tweets for a content analysis, select **View all Reports** from the **Analyse** menu. A folder should appear containing a single file. This is the file for analysis outside Mozdeh.



1. Double click on the file name to read the random tweets. To copy the file of random tweets to Excel, open Excel and copy (Ctrl-A, Ctrl-C) the tweets from the text file and paste (Ctrl-V) them into Excel.



Now repeat the above for a set of your own tweets and explore them using various keyword searches. Your project will need a new name and there is no need to delete the exiting test project first.

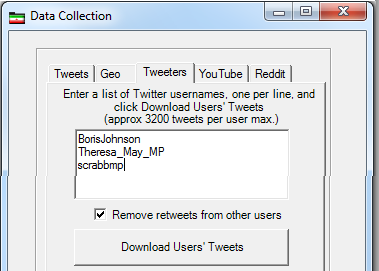
Each project is automatically and permanently saved in the moz\_data folder (a subfolder of the folder that you started Mozdeh from). The project name will appear in the project list box within Mozdeh whenever it is started. To open a previous project, start Mozdeh, select the project and click **Open Selected Project**.

# Workshop 2: Word association mining applied to timeline tweets

This part covers how to use Mozdeh to download posted by one or more specified tweeters (rather than by downloading tweets matching keywords, as above), and how to apply word association mining to the tweets.

## Downloading tweets from one or more users

1. Start Mozdeh, enter a name for a pilot test project (e.g., *Future PM*) and click **Start New Project**.
2. Click the Tweeters tab, enter up to three Twitter usernames (e.g., politicians *BorisJohnson, Kier\_Starmer* and *EdwardJDavey)* in the Data Collection screen and click the **Download Users’ Tweets** button. You will be taken to a web page asking you to logon to Twitter. This will give you a pin number to enter into Mozdeh that gives Mozdeh permission to search on your behalf. See the example in the PowerPoint slides or ask if you get stuck. You can enter as many usernames as you like in theory, but it can take a long time to download these tweets so please keep to a maximum of three users for practice.



1. Mozdeh will now collect tweets from the users. When it has finished, it will ask a series of questions – please click OK or give the suggested answer to these questions (the answers don’t matter at this stage – except that it is important to have **Day** rather than Hour as the time interval) and then you should get the main search screen.

## Searching the downloaded tweets

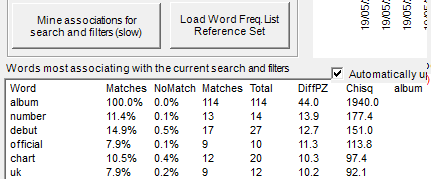
1. Mozdeh’s main search screen can be used to query the downloaded tweets. Test this by entering a query (e.g., *now*) and clicking **Search** to see tweets that match the query. Read a random tweet by clicking a text in the search results list. Check that there are tweets from all three users. Try any other searches that you think might give interesting results.

## Filtering the downloaded tweets

1. The gender filter will not be useful because there are only three tweeters in the data set, but the sentiment and retweet count filters work as before. Set the positive (+ button) filter and click search, checking a random result (is it negative?) and do the same for the negative (- button) filter, clearing it after use with the blank button. Try the retweet filter too, if you think it may be interesting.

## Word association mining

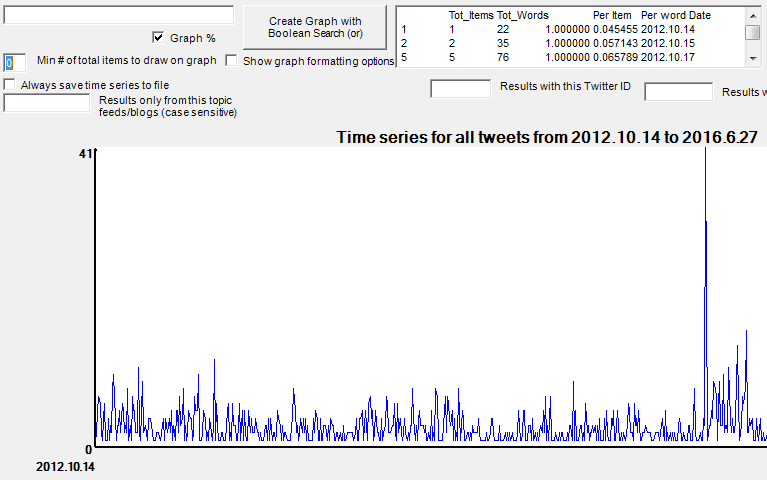
1. To find words that associate with tweets matching a query, enter the query in the search box (e.g., *album*), click **Mine associations for search and filters** button in the middle right of the screen and examine the text box underneath.



1. The *Words most associating* list shows terms that occur more often in tweets matching your search than in tweets not matching your search. It gives two association statistics, a differences in proportions test and a chi square test. Bigger values of both indicate a stronger association. For example, in the box above, other than album, the word that occurs most often in tweets containing the word album than in the remaining tweets is *number*, which occurs in 11.4% of tweets containing album in comparison to 0.1% of the non-matching tweets.
2. Copy the results of the above to Excel or another spreadsheet by opening Excel, right clicking in the box, selecting Copy, switching to Excel, and pasting the copied text (Ctrl-V). These results are now permanently saved in the spreadsheet for (hypothetical) later analysis.
3. Clear the query, and repeat this with other queries that you think might tell you something interesting about your tweets.
4. Clear the query, select the Female user gender and click Mine associations to see a list of words more used by females. Check the top 5 to see whether you think the results are plausible or due to spam/irrelevant texts. Try the same for words used by males.
5. Reset the gender filter to Any, select the Female user gender and click Mine associations to see a list of words more used by females. Check the top 5 to see whether you think the results are plausible or due to spam/irrelevant texts. Try the same for words used by males. The more tweets you gather, the more powerful and interesting the results will be.

## [optional] Drawing time series graphs

1. Mozdeh can draw time series graphs of the volume of tweets or the number of tweets matching a query. Select **Graph Time Series** from the **Analyse** menu to go to the screen that offers these.
2. To draw a graph of the volume of all tweets, click **Create Graph with Boolean Search**. This shows the number of tweets on each day. The left-hand side of the x axis shows the first date on which a Tweet was collected – probably about 2012.



1. To draw a time series graph of the proportion of all tweets matching a query (e.g., *vote*), enter it into the search box at the top left of the screen and then click **Create Graph with Boolean Search**. This will show the percentage of tweets collected on each day that match the query.



1. Return to the main search screen by selecting **Search** from the **Analyse** menu.

## Twitter exploration task

The above stages show how to download tweets matching a query or from a set of users, then search the downloaded tweets, filter them, association mine them for word lists, and draw time series graphs for them. Think of a new topic related to your research that you would like to explore on Twitter and investigate it using the above steps, as appropriate. Make sure to test any queries used so that they do not match irrelevant tweets. Here are some ideas if you can’t think of any.

* Vaccination
* Top musicians on Twitter https://en.wikipedia.org/wiki/List\_of\_most-followed\_Twitter\_accounts
* Selected UK politicians on Twitter. https://www.politics-social.com/list/party

Remember: No more than three tweeters when practicing. The optional section below contains ideas on generating suitable queries if this part is difficult for your topic.

# Optional 1: Collecting high quality sets of tweets from queries [for reference]

Collecting a useful, high quality set of tweets needs time to ensure that the queries used to generate the tweets are the best ones for the task. This is not needed for projects based on sets of usernames. Here are some tips for this critical task.

Start by constructing a set of queries to match tweets for a topic that you are interested in. For the queries, use the Twitter search interface <https://twitter.com/search-home>. This is a key step of the project that must be carried out carefully to ensure that all relevant queries are included, that queries matching too many irrelevant tweets are rejected, and that queries matching some irrelevant tweets are refined to eliminate most irrelevant tweets. For example, a study to investigate how the words kiss and hug are used in Twitter to express affection started with these two terms as queries. After pilot testing the queries, some terms were subtracted from them to remove lots of unwanted matches (below right). For example, subtracting ass removed many instances of the phrase "kiss my ass", removing 104.1 excluded references to the Kiss 104.1 radio station and subtracting bora removed lots of spam related to a TV meme at the time of checking. The final queries used were **kiss –ass –bora -104.1** and **hug –fanboy**. The recommended process to build a useful set of queries is:

1. Brainstorm a set of potentially relevant queries
2. Pilot the queries by testing them in <https://twitter.com/search-home> for relevant results.
3. Check each query for irrelevant content and remove if almost all content is irrelevant, or refine the query if there is some irrelevant content
4. Check for additional queries to add to the set
5. Repeat from 1. until the set of queries seems to be stable and satisfactory.

# Optional 2: Advanced optional Twitter analysis methods

## Twitter networks

Two types of networks can be created from the downloaded tweets in Mozdeh. These are not essential for most analyses but can be interesting to generate.

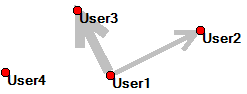
1) **Tweet network**. *This network illustrates which users tweet each other and how often.* If @user1 tweets mentioning @user2 then this creates an arrow from @user1 to @user2. If there are multiple tweets from @user1 mentioning @user2 then the width of the arrow is proportional to the number of tweets.

For example if @user1 - @user4 are in data set and there are two tweets from @user1:

* **@user1**: Hello @user3 how are you?
* **@user1**: Hello @user2 and @user3

Then there would be two arrows, one from @user1 to @user2 and a double thickness arrow from @user1 to @user3 but no arrow connecting @user4

@user3 <= @user1 -> @user 2 @user4



**2) Co-tweeted network**. *This network illustrates which users tweeted to within the same tweet and how often.* If any tweet mentions both @user1 and @user2 then this creates a line between @user1 to @user2. If there are multiple tweets from @user1 mentioning both @user1 and @user2 then the width of the line is proportional to the number of these tweets.

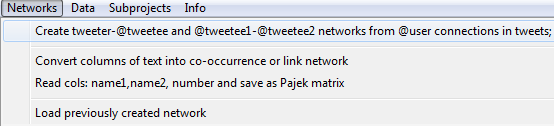
For example if @user1 - @user4 are in data set and there are three tweets:

* **@anyuser**: Hello @user3 how are you?
* **@anyuser**: Hello @user2 and @user3
* **@anyuser**: Hello @user3 and @user4

Then there would be two lines, one between @user2 and @user3 and one between @user3 and @user4.



Both types of networks can be created by clicking on the Network menu in Mozdeh and selecting the *Create a) tweeter-@tweetee and b) @tweetee1-@tweetee2 network from @user connections in tweets; calculate centrality scores; list tweeter-tweetee pairs* submenu item.



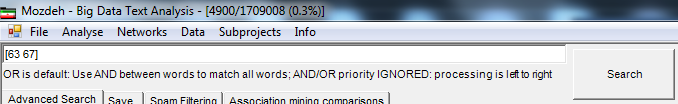
Here is an explanation of the questions that are asked after clicking this button.

1. "Select raw data file of tweets to use": Select a file in the folder that opens in the browser. This folder contains the tweets collected by Mozdeh in a simple plain text format. If there are multiple files, select the file containing “filtered” in the filename because it will have duplicate tweets removed.
2. "Which column contains the TWEETS (select Title col.)": Unless you are doing something fancy, accept the suggested answer of 4 (the 4th column of the file contains the tweets).
3. "Which column contains the tweetERs (select AuthorURL col.)": Unless you are doing something fancy, accept the suggested answer of 9. This column contains the username of the tweeter (or the original tweeter if it is a retweet).
4. A long question – answer 1 unless you are doing something fancy:
   * “For tweet from-to network, select users with”:
   * 1: Largest Minimum of {tweets FROM user to someone, tweets TO user from someone - @user}
   * 2: Most tweets sent + received
   * 3: Most tweets TO user (always used for co-network)
   * 4: Most tweets FROM user
   * 5 Enter your own list of tweeters to make network from
5. “Enter maximum number of nodes for network. More than 50 makes the network too cluttered, and large numbers (e.g., Over 20,000) may result in an out of memory error when creating the network. Unless you are doing something fancy, accept the suggested answer of 50. Bigger numbers make cluttered networks and may crash the computer. Although it would be nice to make a complete network with all nodes, this is usually impractical.
6. “Filename for main results”: Accept the suggestion.
7. “Use raw numbers in the networks rather than scaled numbers? Click YES ONLY if exporting the network to a spreadsheet or another graphics package” Click NO if you just want to draw the network. Click YES if you want to know the exact number of connections rather than needing the network diagram.

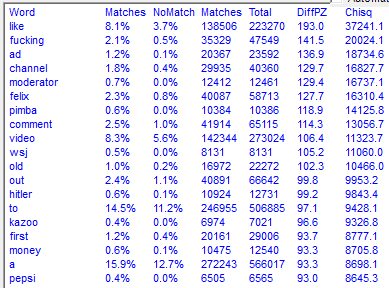
## Topic changes over time

This section includes techniques for discovering changes in topics between two time periods.

|  |  |
| --- | --- |
| To try this, you will need to have collected a Twitter project in Mozdeh with tweets from several years. This might be tweets from a set of users because Mozdeh can collect old tweets for them. To find out issues that attract more interest in one time period than another, choose a date range, identify the number of the start and end of the period in the First date to show dropdown box and enter the first and last numbers in square brackets in the search box. In the example below, to check for important words from 2017 the date range to enter as a search would be [63 67]. |  |



Finally, click **Mine associations for search and filters (slow)**. Words associating with the specified date range (i.e., 2017 in the above example) will be displayed in the bottom left association mining box.

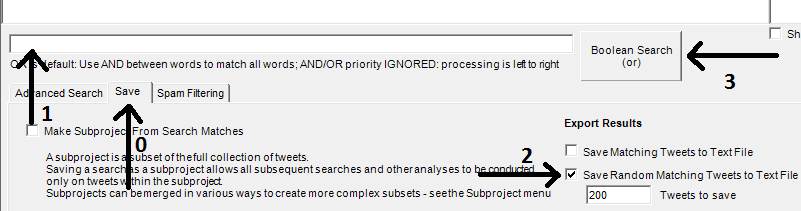


* *like* occurs in 8.1% of the 2017 comments compared to 3.7% of the earlier comments. This is due discussions in 2017 about why PewDiePie is popular. See the comparisons section below for information about how this was discovere3d.
* *ad* occurs in 1.2% of the 2017 comments compared to 0.1% of the rest. Browsing the results of the query *ad AND like [63 67]* suggests that this is due to discussions of some of PewDiePie’s sponsors withdrawing, and apparently increased comment spam.
* *Moderator* occurs in 0.7% of the negative comments compared to 0.0% of the rest. Browsing the results of the query *moderator AND like [63 67],* this is due to a PewDiePie’s call for comment moderators, with many applications taking the form of comments starting with “I would like to be a moderator because…”.

## Content analysis or thematic analysis text saving steps

The instructions below assume that you have tweets collected by Mozdeh to analyse.

1. **Content analysis or thematic analysis**. These need a random sample of tweets to examine. A random sample of texts can be saved by Mozdeh by clicking the Save tab (0 below), clearing the search text box (1), checking the Save Random Matching Tweets to Text File (2) and clicking the Boolean Search button (3). When asked for a location to save the results to, make sure that you remember and can find the location that you chose. Hint: If you can't find it, select View all Reports Created So Far from the Analyse menu.



1. The new text file just created can be loaded into a spreadsheet or word processor for the analysis steps. The easiest way to do this is to open a spreadsheet program, then open the text file, select and copy all the text in the text file and paste it into the spreadsheet. Now decide upon some categories for your tweets, write names and descriptions for the categories and categorise the random sample according to your descriptions. See the book above or content analysis books for more information on conducting a content analysis.

# Workshop 3a: Downloading YouTube comments on a video

This is like the Twitter procedure so only the download instructions are reported here. Start by getting a Developer key – see instructions at the end of this document.

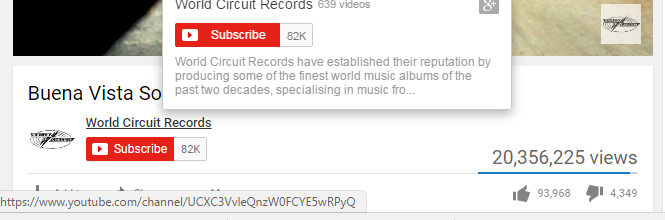
1. Follow the instructions here to get a YouTube Data API developer key. This is very tricky, sorry: http://lexiurl.wlv.ac.uk/searcher/YouTubeKeyRegister.html
2. Start Mozdeh, enter a name for a pilot test project (e.g., *YouTube test*) and click the **Start New Project** button.
3. Click the YouTube tab, *enter a video URL for a video with lots of comments* in the Data Collection screen (e.g., “Mary Poppins sings death metal” https://www.youtube.com/watch?v=iU2hy0L5lgg), select the Video IDs option button, enter your YouTube developer key (click the link to get one first) and click the **Get Comments…** button. [Note: you could also use a list of keywords, or channel names but these would probably take over an hour to download – perhaps several hours]



1. Mozdeh will now collect comments from the specified videos, or videos matching the searches. When it has finished, it will ask a series of questions – please click OK or give the suggested answer to these questions (the answers don’t matter at this stage – except that it is important to have **month** rather than hour as the time interval) and then you should get the main search screen.
2. Now try searching and filtering the comments, as for Twitter.

# Workshop 3b: Downloading comments on all videos in a YouTube channel

This takes several hours for the data collection so should be started when you can leave the computer running for this much time. It is easier to analyse the comments on all videos within a channel than comments for all videos matching a query because a channel will have no false matches but the query might. To search for a channel, you will need the channel ID. This is a strange string of characters and is not a username. To find a channel ID the easiest way is to go to a video from the channel that you want to investigate and click the owner link (World Circuit Records in the example below). You will then be taken to an URL ending in the channel ID. Copy the end of the URL after channel/. In the example below the channel URL is <https://www.youtube.com/channel/UCXC3VvleQnzW0FCYE5wRPyQ> and the channel ID is UCXC3VvleQnzW0FCYE5wRPyQ, which is the part that should be entered into Mozdeh.

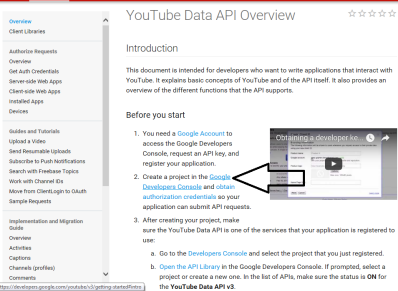


# Appendix: Registering for a YouTube key

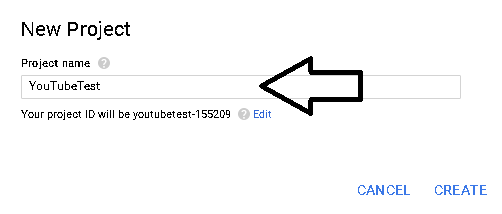
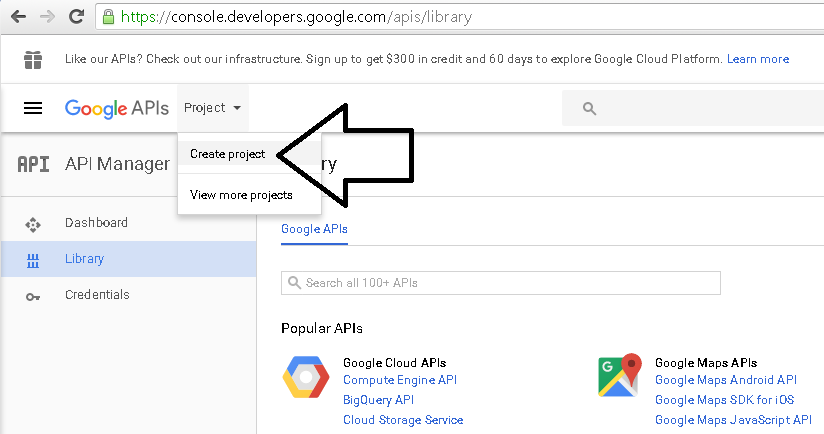
YouTube changes its instructions sometimes so you may need to experiment to work out how to get a YouTube key. This usually takes about half an hour. \*\*\***Make sure to add the YouTube Data API v3 after registering as a Google Developer and before getting a key**.\*\*\*

**Step 1**: Log on to Google with your Google email account (get one if you don't have one) and then open the [developer start page](https://developers.google.com/youtube/v3/getting-started) <https://developers.google.com/youtube/v3/getting-started>.

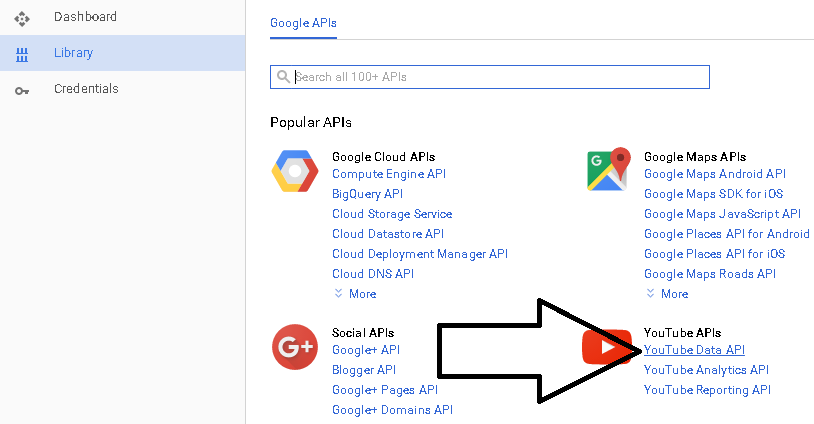
**Step 2**: Click on the Google Developers Console link in the middle of the Google Developer start page that you have just opened.



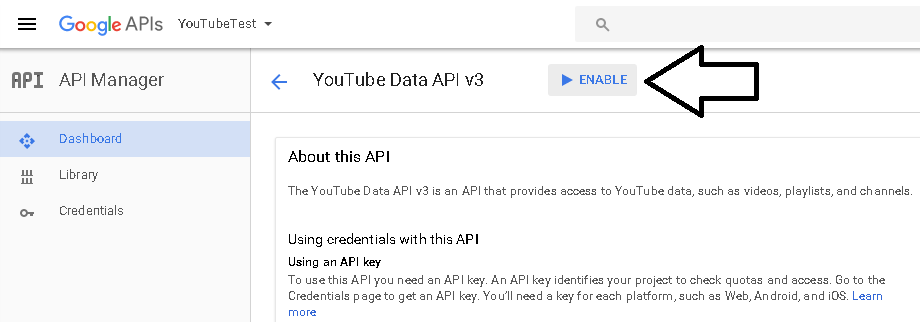
**Step 3**: Click the project button at the top left and then the Create Project link, giving the project your own name (e.g., YouTubeTest). Now wait 10 seconds for YouTube to process your request.



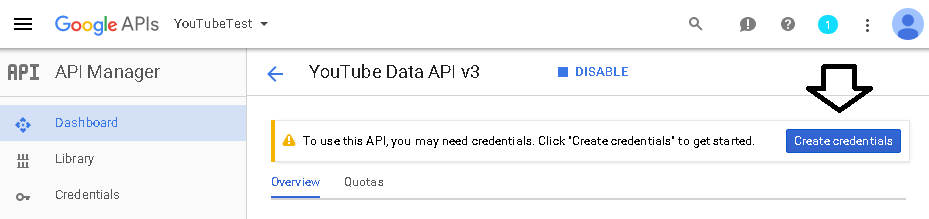
**Step 4**: If not selected already, click on the Library button on the left hand side to see all the APIs. Find the **YouTube APIs**section and click the YouTube Data API link. **THIS IS VERY IMPORTANT – YOU MUST DO THIS BEFORE CREATING A KEY OR THE KEY WILL NOT WORK**.



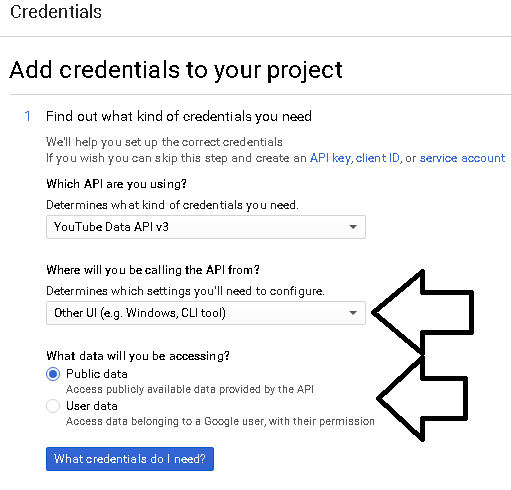
**Step 5**: Enable the YouTube Data API by clicking the **Enable** button so that it looks like below.



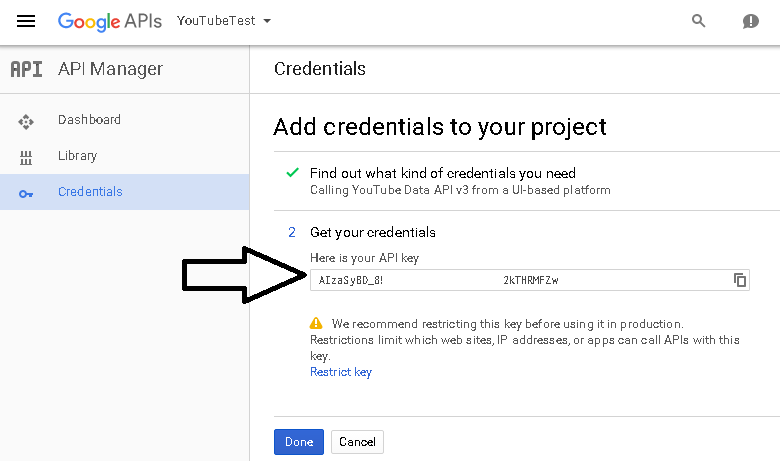
**Step 6**: Click the Credentials section at the top left, click the **Create credentials**button on the right.



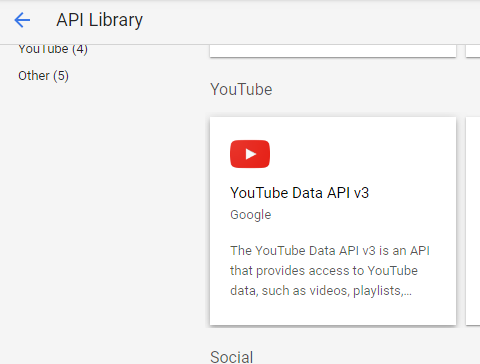
**Step 7**: Select **Other UI** from the drop-down list, and **Public data** option and click **What credentials do I need?**



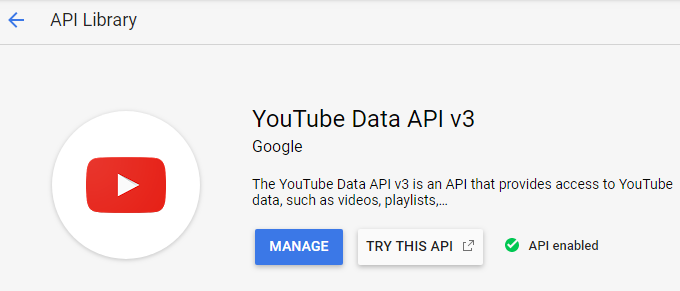
**Step 8**: You should now be shown an API key. Copy it (Ctrl-C) and email it to yourself so that you will not need to repeat these steps to get it again.



**Step 9**: Click on the Library link above to go into the library and scroll down to find the YouTube Data API v3.



**Step 10**: Make sure that the API is enabled (see below).



**Step 11**: You are now ready and can close this window. This key allows you to access YouTube videos forever. Paste (Ctrl-V) it into the yellow YouTube API key box in Webometric Analyst to run searches.

# Bibliography

*Books about social media analysis with Mozdeh.*

Thelwall, M. (2018). *Social web text analytics with Mozdeh*. University of Wolverhampton. <http://mozdeh.wlv.ac.uk/resources/SocialWebResearchWithMozdeh.pdf> [Free short book giving an overview of the social web text analytics capabilities of Mozdeh.]

Thelwall, M. (2021). *Word association thematic analysis: A social media text exploration strategy*. San Rafael, CA: Morgan & Claypool. [Focuses on the word association analysis component]

*Mozdeh was used to collect and analyse data for the following publications, which might give ideas for methods and applications.*

Holmberg, K., & Hellsten, I. (2014). Analyzing the climate change debate on Twitter: content and differences between genders. In Proceedings of the 2014 ACM conference on Web science (pp. 287-288). ACM.

Potts, G., & Radford, D. R. (2019). #Teeth&Tweets: the reach and reaction of an online social media oral health promotion campaign. British Dental Journal, 227(3), 217-222.

Thelwall, M. (2018). Can museums find male or female audiences online with YouTube? Aslib Journal of Information Management, 70(5), 481-497.

Thelwall, M. & Mas-Bleda, A. (2018). YouTube science channel video presenters and comments: Female friendly or vestiges of sexism? Aslib Journal of Information Management, 70(1), 28-46.

Ahmed, W. (2018). Using Social Media Data for Research: An Overview of Tools. Journal of Communication Technology, 1(1), 77-94.

Thelwall, M. & Cash, S. (to appear). Bullying discussions in UK female influencers’ YouTube comments. British *Journal of Guidance and Counselling*.

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. Journal of the American Society for Information Science and Technology, 62(2), 406-418.

Thelwall, M., & Cugelman, B. (2017). Monitoring Twitter strategies to discover resonating topics: The case of the UNDP. El profesional de la información.

Thelwall, M., Makita, M., Mas-Bleda, A., & Stuart, E. (2021). “My ADHD hellbrain”: A Twitter data science perspective on a behavioural disorder. Journal of Data and Information Science, 6(1). https://doi.org/10.2478/jdis-2021-0007

Thelwall, M. & Thelwall, S. (2020). Covid-19 tweeting in English: Gender differences. *El Profesional de la Información*, 29(3), e290301.

Thelwall, M. (2008). No place for news in social network web sites? Online Information Review, 32(6), 726-744.

Thelwall, M. (2018). Social media analytics for YouTube comments: Potential and limitations. International Journal of Social Research Methodology, 21(3), 303-316. doi:10.1080/13645579.2017.1381821