

Facepager Instruction

1. Download and install facepager on the releases page (<https://github.com/strohne/Facepager>). Choose the version corresponding to your operation system.

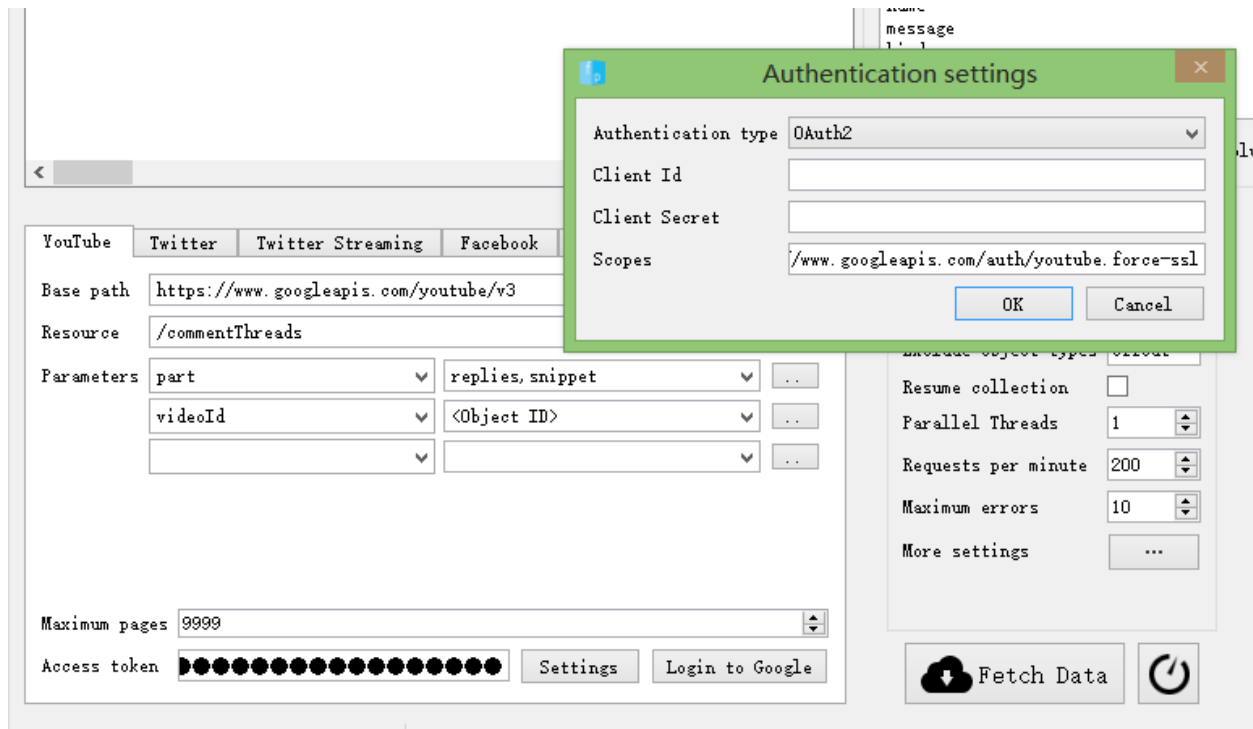
Installer

Installation packages for each version are available on the [releases page](#). Database files may be incompatible between versions.

- **Windows:** Download and execute the latest exe-installer from the [releases page](#). If Windows complains about an unknown publisher and refuses to launch the app click "More info" and start anyway.
- **Mac OS X:** Unzip the app.zip-file from the [releases page](#), drag & drop the app to your "Applications" folder. When a warning about the the source pops up you should allow the installation of non-app-store applications (go to "System Preferences>Security & Privacy" and check "Anywhere"). If you do not want to change your security settings permanently, you may open Facepager by CTRL+Click on the icon (each time on startup).
- **Linux:** There is no binary release, see [src/readme.md](#) for steps to run under linux.

If you want to run from source, see [src/readme.md](#).

2. Click "setting" to adjust Authentication setting. Make sure "Scopes" should be set with "<https://www.googleapis.com/auth/youtube.force-ssl>"



3. Open Facepager and connect to your Google account through "Login to Google". Allow Facepager to access your Google account.

4. Set the parameters as shown in the figure below.

YouTube Twitter Twitter Streaming Facebook Amazon Generic

Base path

Resource

Parameters

part	replies, snippet	...
videoId	<Object ID>	...
		...

Settings

Node level

Select all nodes ☐

Exclude object types

Resume collection ☒

Parallel Threads

Requests per minute

Maximum errors

More settings ...

5. Click “New Database” and create a database with the name of the new media, which you are going to collect comments from.

Open Database New Database Export Data Add Nodes Delete Nodes

Expand nodes Collapse nodes Find nodes

Save DB File

« comments » database

搜索"database"

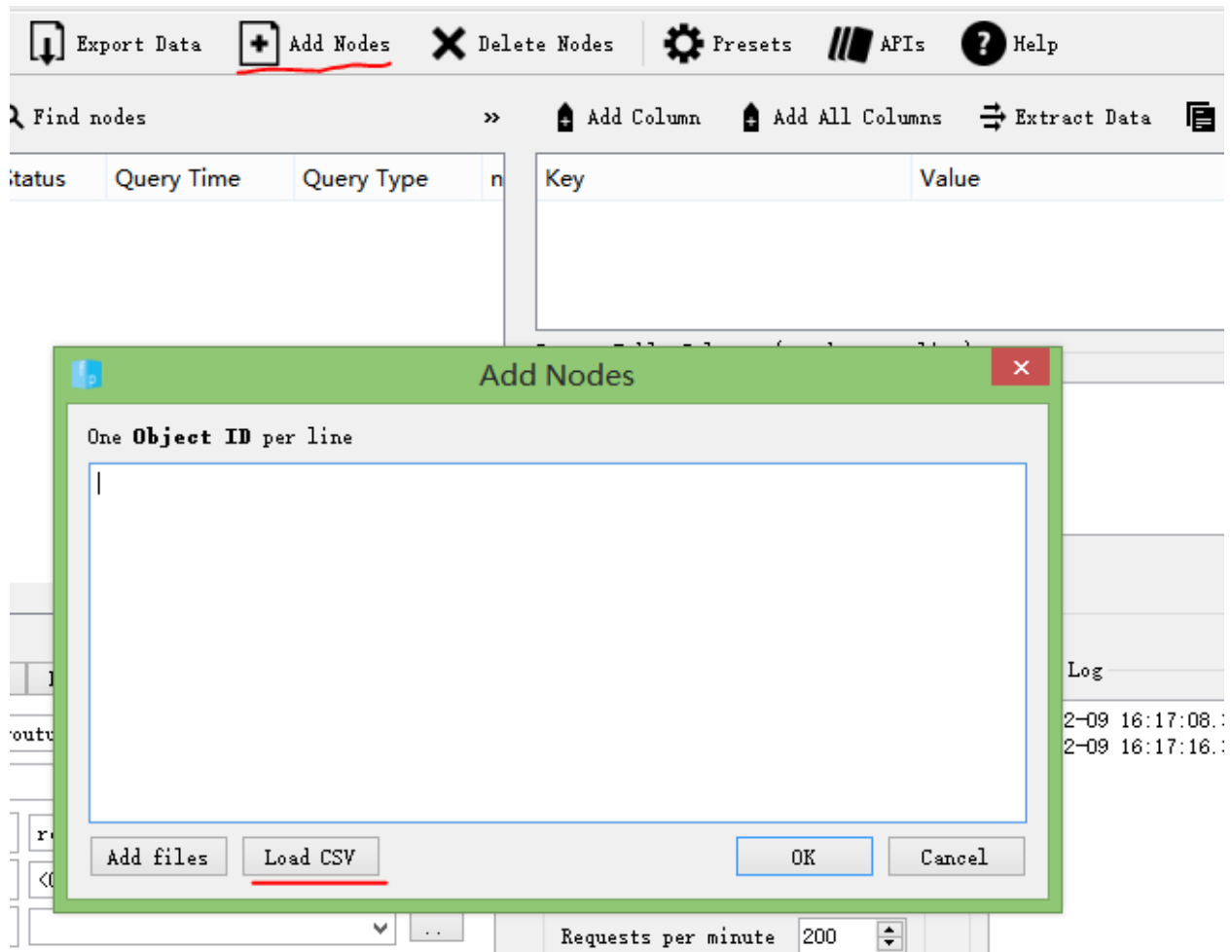
名称	修改日期	类型
abc	12/7/2020 12:17...	Data Base File
Breitbart	12/7/2020 2:38 ...	Data Base File
CBS	12/8/2020 2:18 ...	Data Base File
CNN	12/8/2020 1:52 ...	Data Base File
test	12/5/2020 4:23 ...	Data Base File

NewYorkTimes

DB files (*.db)

保存(S) 取消

- Click “Add Notes” and load the videoID list for one news media. For example, load “**abcNews.csv**” file, if you are going to collect comments for abc News.



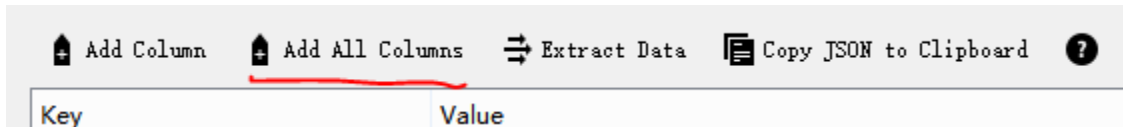
- Make sure all the ObjectIDs are selected using “Ctrl+A”.

Object ID	Object Type	Query Status	Query Time	Query Type
5gHehjiX1-8	seed			
929JS35DS...	seed			
khrHs98D...	seed			
aD03gPr_...	seed			
F1IurSdJhug	seed			
jfWtqrn6M...	seed			
lehEpNjGI...	seed			
D2axuhEG...	seed			
97zsLVrinEA	seed			
WwwWLn1...	seed			
0793ss7H6	seed			

- Now click “Fetch Data”.

Note: during collecting, if there is a window view prompted up said “**there are consecutive 10 errors**” (maybe more), please click “cancel” to stop this collection, since the API may get its limit.

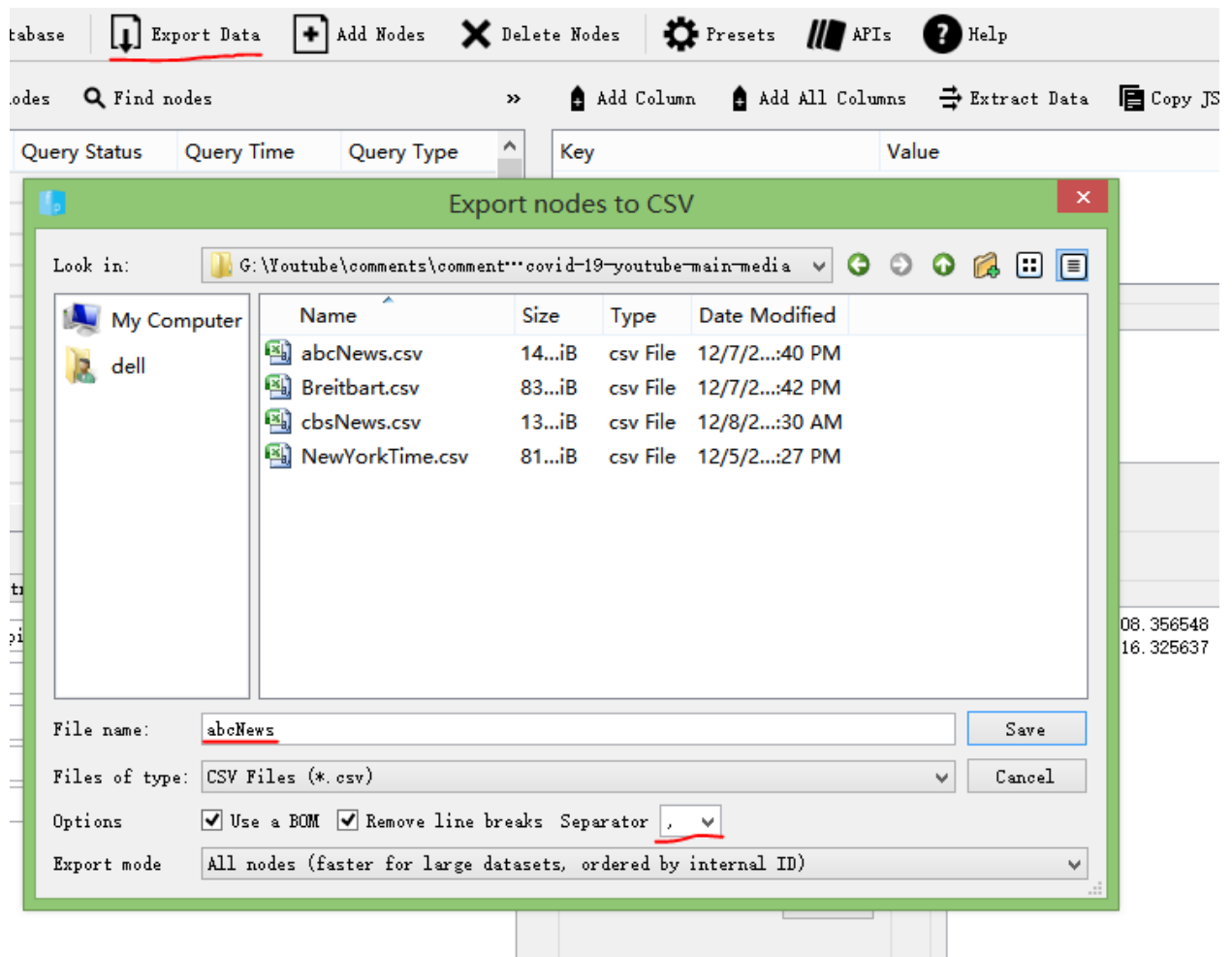
9. Click “Add All Columns” to collect all results.



10. Click “Export Data” to export nodes to csv. Name the file with the name of the news media. For example “abcNews”. Don’t forget to set **comma** as the separator.

Note:

- Please create a folder named “**covid-19-youtube-main-media**” and save the exported csv files from Facepager in this folder.
- In this step, make sure name the exported file with the **same name of its corresponding videoID csv file**.

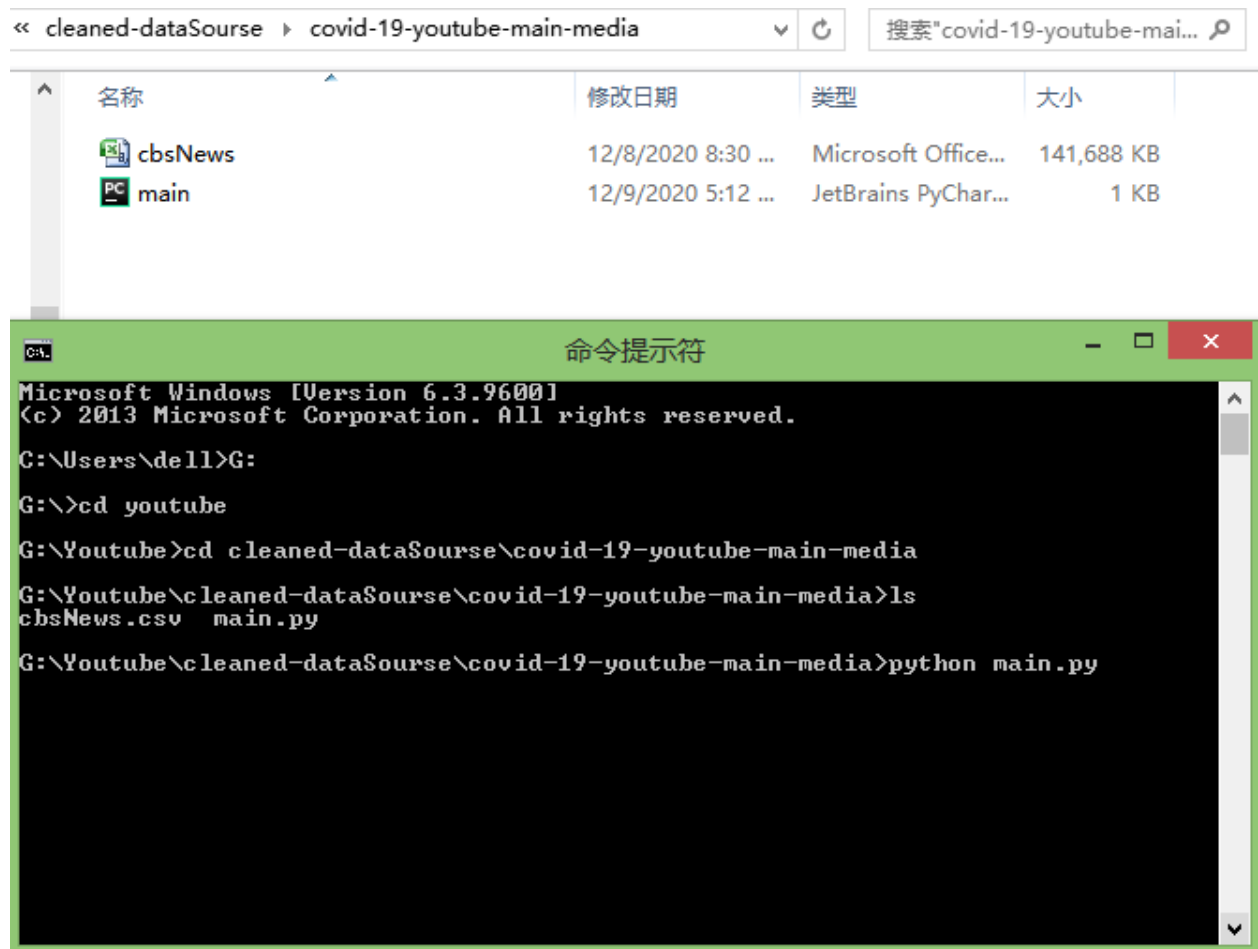


11. Now open terminal to clean the data using **main.py** file.

Note:

- Please put **main.py** into the same folder "**covid-19-youtube-main-media**".
- Make sure you have **pandas** module installed for python in your device. If you don't have this module, you can type "pip install pandas" in your terminal to install it first.




Retrieve to the folder using "cd" command and check files under the folder using "ls" as shown in the figure below.



Please make sure to give the complete and correct media.

```
G:\Youtube\cleaned-dataSource\covid-19-youtube-main-media>python main.py
Enter the News media name completely and correctly:
cbsNews
```

After running, you should get another cleaned file, for example, called "CL-cbsNews.csv"

	cbsNews	12/8/2020 8:30 ...	Microsoft Office...	141,688 KB
	CL-cbsNews	12/9/2020 5:35 ...	Microsoft Office...	26,586 KB
	main	12/9/2020 5:35 ...	JetBrains PyChar...	1 KB