Gathering Data from Facebook

**There are many ways to extract data from facebook:**

1. Using python’s Scraping libraries like:
   1. Scarpy
   2. Beautifulsoup
2. Using Facebook’s graph api

(But it has many limitations, to have any king of information you need to have an Access token)

* **Facebook Graph API**

**Note:**

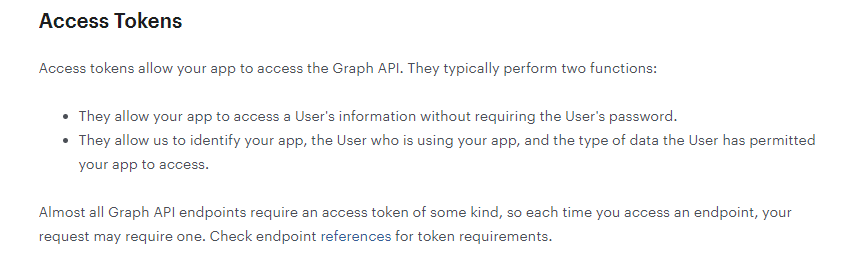
Public post search is no longer available from API version >=2.0.

Refer to the docs for more details:

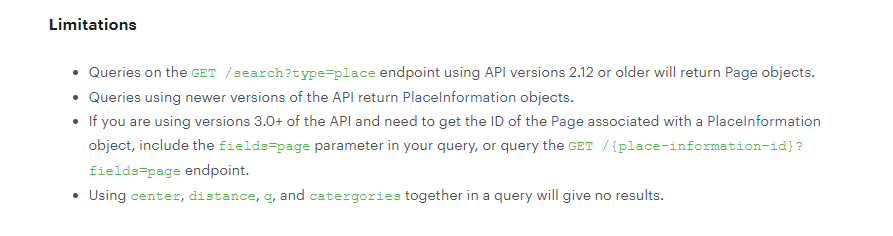
<https://developers.facebook.com/docs/graph-api/using-graph-api/v2.0#search>

*Basic Information:*

Be it any kind of information, without an access token it’s impossible to generate any information.



(As sir(Amit sir) told that I’ve to gather info on tenants, sellers and buyers from facebook pages and groups, so I thought to use fb graph api’s search querry so as to directly search for keywords such as rent, seller, buyers etc.)

 Search query:

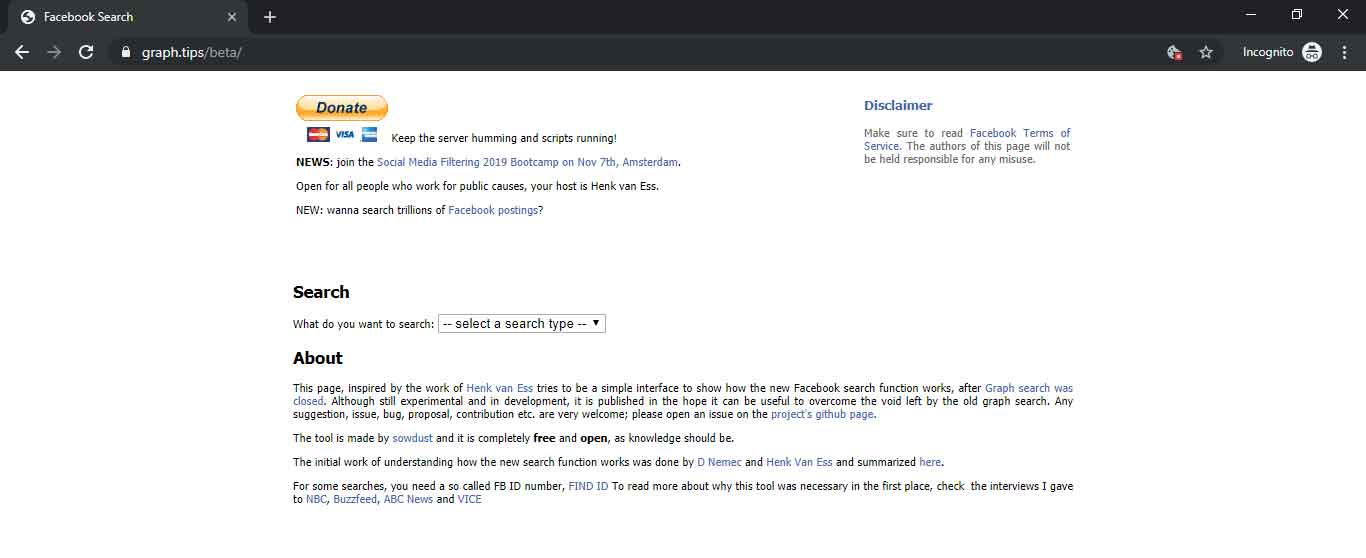
Facebook graph search has stopped working since June 7th, 2019. As a result, it has become near impossible to gather in-depth data about a fb user/page/group like before.

Although it’s still possible to search in Graph search like formats. The new process involves JSON and Base64 encoding. We shall understand the process in details as it helps even when our custom websites stops working.

**Custom Solution:**

Learning the above process proves handy when the custom websites stop functioning.

The website we recommend for custom graph search is the website - <https://graph.tips/beta/>



**The Behind the Scene Process Explained:**

**How to Find the ID Numbers:**  
Even with the new process, ‘entity\_id’, ‘page\_id’ and ‘group\_id’ are the most crucial parameters. You can find them using the following instructions:

* Open the Profile or page or group of interest in a web browser
* Right click somewhere on the page and choose “View page source”
* Search using Ctrl+F for the above parameters: ‘entity\_id’ (when looking at a profile), ‘page\_id’ (when looking at a page) and ‘group\_id’ (when looking at a group).  The number mentioned just behind ‘entity\_id’ etc is the number you need to copy.

Also similar to the previous Graph Search, choose ‘English US’ in the language settings for the best results.

**Let’s Get Started - The Base:**  
For every search, the base will always be this URL:  facebook.com/search/top/?q=people&epa=FILTERS&filters=

In this URL, we will make minor modification to search the different parameters. For example, replace “top” with:

* top/ - Searches top content
* posts/ - Searches public posts
* people/ - Searches for people
* photos/ - Searches for photos
* videos/ - Searches for videos
* pages/ - Searches for pages
* places/ - Searches for places

**An example:**  
facebook.com/search/top/?q=people&epa=FILTERS&filters=

When you want to search for more than one word, use as it represents a space. Similarly, you may use + between the words, for example, Ravi+Kumar.

**JSON and Base64:**  
We are now going to deal with the end part of our URL. An example:

facebook.com/search/posts/?q=cricket&epa=FILTERS&filters=

If we want to search the most recent posts on cricket  
‘search/posts/‘ - To search public posts  
‘?q=cricket‘ - As search query

To find the ‘Most recent’ posts, we need to translate a JSON string into Base64. The JSON string for ‘Most recent’ is:

{“rp\_chrono\_sort”:”{\”name\”:\”chronosort\”,\”args\”:\”\”}”}

To translate JSON into Base64, you may use the website CyberChef or any other alternate website.

**How to Use CyberChef Website:**

* **Step 1:**Choose “To Base64” from the Operations section. Drag this option to the Recipe section and drop.
* **Step 2:**Paste the JSON line in the blank Input field.
* **Step 3:**Copy the Base64 Output from the Output field.

An example: The result of JSON line for “Most Recent” shall look like: eyJycF9jaHJvbm9fc29ydCI6IntcIm5hbWVcIjpcImNocm9ub3NvcnRcIixcImFyZ3NcIjpcIlwifSJ9

Now the last step is to put the URL and the Base64 together: facebook.com/search/posts/?q=cricker&epa=FILTERS&filters=eyJycF9jaHJvbm9fc29ydCI6IntcIm5hbWVcIjpcImNocm9ub3NvcnRcIixcImFyZ3NcIjpcIlwifSJ9

Thus the result is the Most Recent posts on Cricket

**Possible Search Options:**  
Search/top/ - To search for most popular content

**Most recent popular content:**  
**JSON:**{“rp\_chrono\_sort”:”{\”name\”:\”chronosort\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9jaHJvbm9fc29ydCI6IntcIm5hbWVcIjpcImNocm9ub3NvcnRcIixcImFyZ3NcIjpcIlwifSJ9

**Most popular public content:**  
**JSON:**{“rp\_author”:”{\”name\”:\”merged\_public\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJtZXJnZWRfcHVibGljX3Bvc3RzXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Most popular content posted from your own profile:**  
**JSON:**{“rp\_author”:”{\”name\”:\”author\_me\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJhdXRob3JfbWVcIixcImFyZ3NcIjpcIlwifSJ9

**Most popular content viewed by your profile:**  
**JSON:**{“interacted\_posts”:”{\”name\”:\”interacted\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJpbnRlcmFjdGVkX3Bvc3RzIjoie1wibmFtZVwiOlwiaW50ZXJhY3RlZF9wb3N0c1wiLFwiYXJnc1wiOlwiXCJ9In0=

**Most popular content posted by your friends:**  
**JSON:**{“rp\_author”:”{\”name\”:\”author\_friends\_feed\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJhdXRob3JfZnJpZW5kc19mZWVkXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Most popular content from the pages you like or the groups you are a member of:**  
**JSON:**{“rp\_author”:”{\”name\”:\”my\_groups\_and\_pages\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJteV9ncm91cHNfYW5kX3BhZ2VzX3Bvc3RzXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Most popular content in the groups you are a member of:**  
**JSON:**{“rp\_group”:”{\”name\”:\”my\_groups\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9ncm91cCI6IntcIm5hbWVcIjpcIm15X2dyb3Vwc19wb3N0c1wiLFwiYXJnc1wiOlwiXCJ9In0=

**Most popular content from a specific page:**  
**JSON:**{“rp\_author”:”{\”name\”:\”author\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**Most popular content from a specific group:**  
**JSON:**{“rp\_group”:”{\”name\”:\”group\_posts\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**Search/posts/ - To search public posts:**

**Public posts:**  
**JSON:**{“rp\_author”:”{\”name\”:\”merged\_public\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJtZXJnZWRfcHVibGljX3Bvc3RzXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Posts from your friends:**  
**JSON:**{“rp\_author”:”{\”name\”:\”author\_friends\_feed\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJhdXRob3JfZnJpZW5kc19mZWVkXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Posts from the pages you like and the groups you are a member of:**  
**JSON:**{“rp\_author”:”{\”name\”:\”my\_groups\_and\_pages\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJteV9ncm91cHNfYW5kX3BhZ2VzX3Bvc3RzXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Posts in the groups you are a member of:**  
**JSON:**{“rp\_group”:”{\”name\”:\”my\_groups\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9ncm91cCI6IntcIm5hbWVcIjpcIm15X2dyb3Vwc19wb3N0c1wiLFwiYXJnc1wiOlwiXCJ9In0=

**Posts from a specific page:**  
**JSON:**{“rp\_author”:”{\”name\”:\”author\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**Posts from a specific group:**  
**JSON:**{“rp\_group”:”{\”name\”:\”group\_posts\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**Search/people/ - To search for people**

**People within a specific city:**  
**JSON:**{“city”:”{\”name\”:\”users\_location\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**People with a specific education:**  
**JSON:**{“school”:”{\”name\”:\”users\_school\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**People who work at a specific employer:**  
**JSON:**{“employer”:”{\”name\”:\”users\_employer\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**Find friends of your own profile:**  
**JSON:**{“friends”:”{\”name\”:\”users\_friends\”,\”args\”:\”\”}”}  
**Base64:**eyJmcmllbmRzIjoie1wibmFtZVwiOlwidXNlcnNfZnJpZW5kc1wiLFwiYXJnc1wiOlwiXCJ9In0=

**Find mutual friends of your own friends:**  
**JSON:**{“friends”:”{\”name\”:\”users\_friends\_of\_friends\”,\”args\”:\”\”}”}  
**Base64:**eyJmcmllbmRzIjoie1wibmFtZVwiOlwidXNlcnNfZnJpZW5kc19vZl9mcmllbmRzXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Find friends of a specific profile:**  
**JSON:**{“friends”:”{\”name\”:\”users\_friends\_of\_people\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**Search/photos/ - To search for photos**

**Public photos:**  
**JSON:**{“rp\_author”:”{\”name\”:\”merged\_public\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJtZXJnZWRfcHVibGljX3Bvc3RzXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Photos posted by your friends:**  
**JSON:**{“rp\_author”:”{\”name\”:\”author\_friends\_feed\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJhdXRob3JfZnJpZW5kc19mZWVkXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Photos posted by the pages you like and the groups you are a member of:**  
**JSON:**{“rp\_author”:”{\”name\”:\”my\_groups\_and\_pages\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9hdXRob3IiOiJ7XCJuYW1lXCI6XCJteV9ncm91cHNfYW5kX3BhZ2VzX3Bvc3RzXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Photos posted by the groups you are a member of:**  
**JSON:**{“rp\_group”:”{\”name\”:\”my\_groups\_posts\”,\”args\”:\”\”}”}  
**Base64:**eyJycF9ncm91cCI6IntcIm5hbWVcIjpcIm15X2dyb3Vwc19wb3N0c1wiLFwiYXJnc1wiOlwiXCJ9In0=

**Photos posted by a specific page:**  
**JSON:**{“rp\_author”:”{\”name\”:\”author\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**Photos posted by a specific group:**  
**JSON:**{“rp\_group”:”{\”name\”:\”group\_posts\”,\”args\”:\”PutIDHere\”}”}  
**Base64:**Depends on the Facebook ID of the page of interest.

**Search/videos/ - To search for videos**

**Search live videos:**  
**JSON:**{“videos\_source”:”{\”name\”:\”videos\_live\”,\”args\”:\”\”}”}  
**Base64:**eyJ2aWRlb3Nfc291cmNlIjoie1wibmFtZVwiOlwidmlkZW9zX2xpdmVcIixcImFyZ3NcIjpcIlwifSJ9

**Videos posted by your friends and the groups you are a member of:**  
**JSON:**{“videos\_source”:”{\”name\”:\”videos\_feed\”,\”args\”:\”\”}”}  
**Base64:**eyJ2aWRlb3Nfc291cmNlIjoie1wibmFtZVwiOlwidmlkZW9zX2ZlZWRcIixcImFyZ3NcIjpcIlwifSJ9

**Search/pages – To search for pages:**

**Search for ‘verified’ pages:**  
**JSON:**{“verified”:”{\”name\”:\”pages\_verified\”,\”args\”:\”\”}”}  
**Base64:**eyJ2ZXJpZmllZCI6IntcIm5hbWVcIjpcInBhZ2VzX3ZlcmlmaWVkXCIsXCJhcmdzXCI6XCJcIn0ifQ==

**Search for pages of local business or places:**  
**JSON:**{“category”:”{\”name\”:\”pages\_category\”,\”args\”:\”1006\”}”}  
**Base64:**eyJjYXRlZ29yeSI6IntcIm5hbWVcIjpcInBhZ2VzX2NhdGVnb3J5XCIsXCJhcmdzXCI6XCIxMDA2XCJ9In0=

**Search for pages of companies, organisations or institutions:**  
**JSON:**{“category”:”{\”name\”:\”pages\_category\”,\”args\”:\”1013\”}”}  
**Base64:**eyJjYXRlZ29yeSI6IntcIm5hbWVcIjpcInBhZ2VzX2NhdGVnb3J5XCIsXCJhcmdzXCI6XCIxMDEzXCJ9In0=

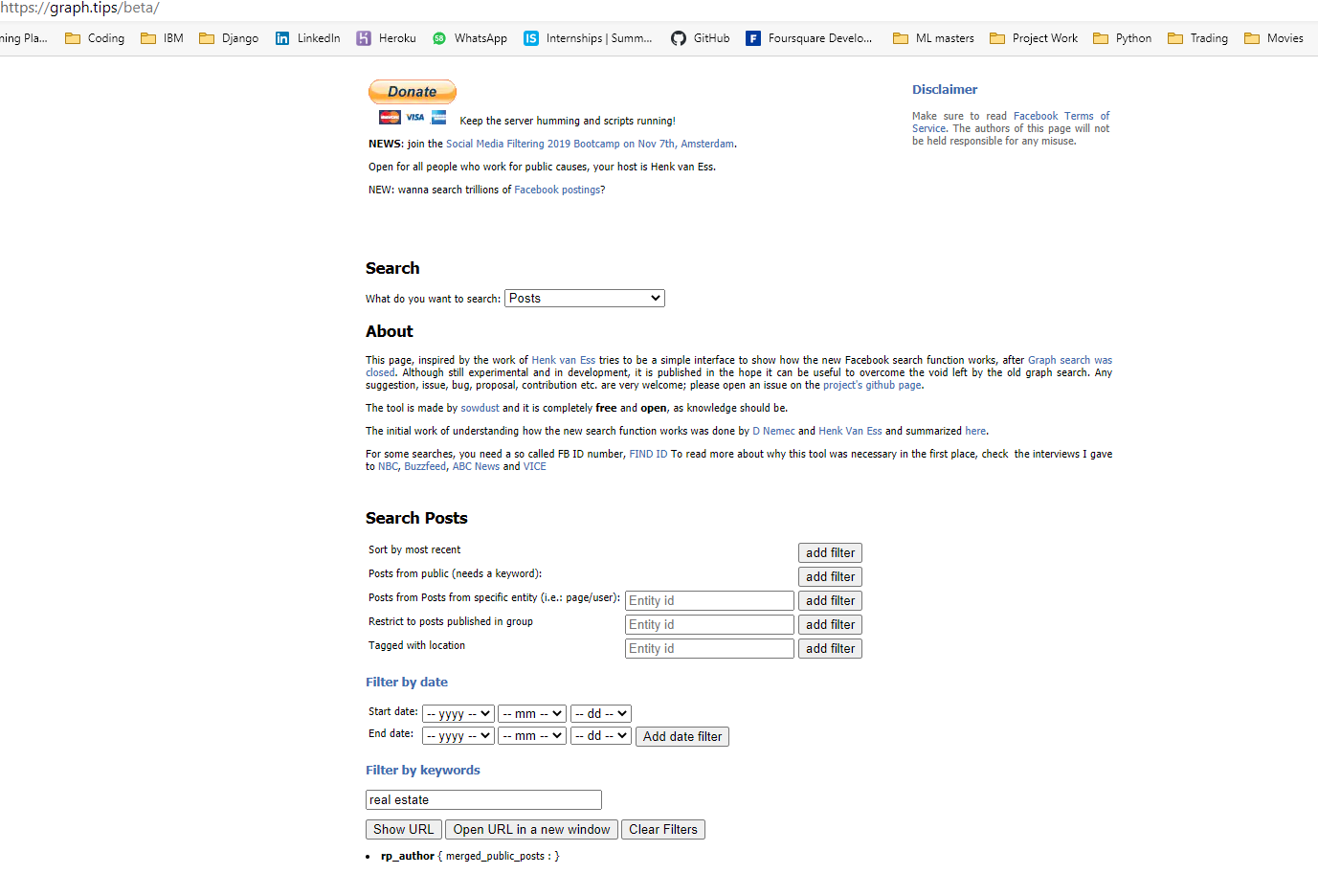
**Search for pages of a brand or a product:**  
**JSON:**{“category”:”{\”name\”:\”pages\_category\”,\”args\”:\”1009\”}”}  
**Base64:**eyJjYXRlZ29yeSI6IntcIm5hbWVcIjpcInBhZ2VzX2NhdGVnb3J5XCIsXCJhcmdzXCI6XCIxMDA5XCJ9In0=

**Search for pages of artists, bands or public figures:**  
**JSON:**{“category”:”{\”name\”:\”pages\_category\”,\”args\”:\”1007,180164648685982\”}”}  
**Base64:**eyJjYXRlZ29yeSI6IntcIm5hbWVcIjpcInBhZ2VzX2NhdGVnb3J5XCIsXCJhcmdzXCI6XCIxMDA3LDE4MDE2NDY0ODY4NTk4MlwifSJ9

**Search for pages in entertainment:**  
**JSON:**{“category”:”{\”name\”:\”pages\_category\”,\”args\”:\”1019\”}”}  
**Base64:**eyJjYXRlZ29yeSI6IntcIm5hbWVcIjpcInBhZ2VzX2NhdGVnb3J5XCIsXCJhcmdzXCI6XCIxMDE5XCJ9In0=

**Search for pages of a cause or a community:**  
**JSON:**{“category”:”{\”name\”:\”pages\_category\”,\”args\”:\”2612\”}”}  
**Base64:**eyJjYXRlZ29yeSI6IntcIm5hbWVcIjpcInBhZ2VzX2NhdGVnb3J5XCIsXCJhcmdzXCI6XCIyNjEyXCJ9In0=

For Examples:

We’ve to search for from pages about real estates. So we’ll fill in details like shown below:

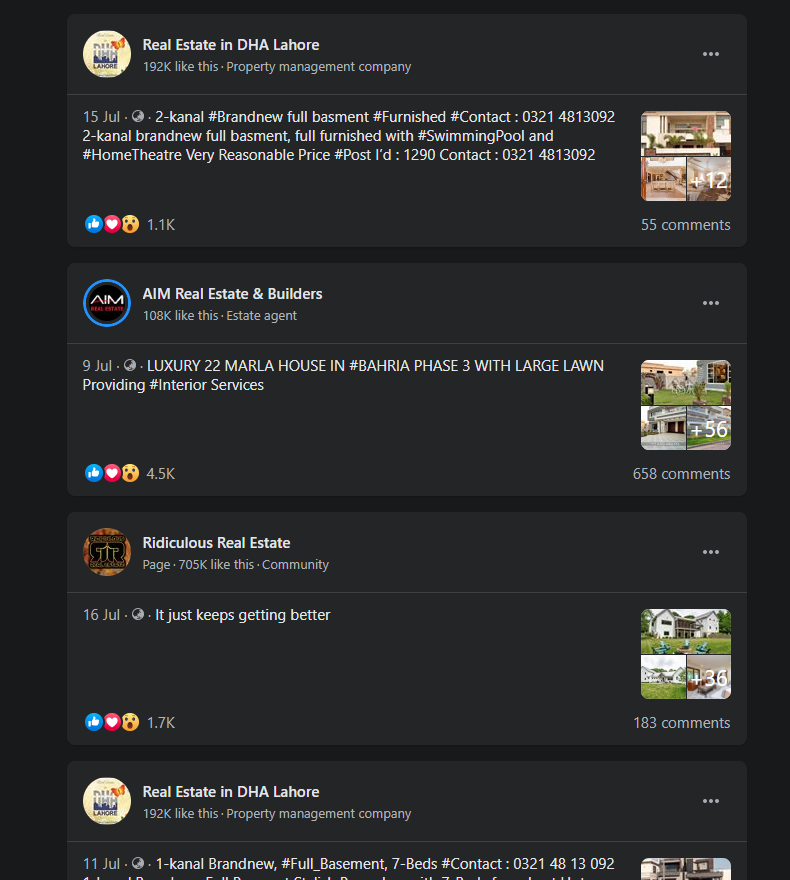
Select the entity from which you want to extract information from

Select the appropriate filters you need to generate an url.

Enter the keyword

Then it’ll generate an url like this:

And you’ll get something like this:



**Accessing Page info(with token):**

**Code:**



**Result:**

