

## دليل استعادة قاعدة بيانات التغريدات

في هذا الدليل سنشرح طريقتين لاستعادة كامل التغريدات :

### 1 . Hydrator

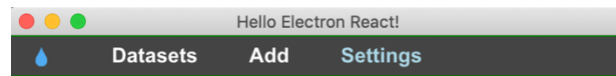
Hydrator هو تطبيق سطح مكتب يقوم باسترجاع كامل التغريدات عن طريق Tweet id

#### 1 - تثبيت تطبيق Hydrator

يمكنك تثبيت التطبيق بالاطلاع على [قائمة الإصدارات](#) المتاحة لنظام التشغيل OS X و Windows و Linux.

#### 2 - طريقة الاستخدام

- أولا : انسخ المستودع في جهازك
- ثانيا: الاتصال بواجهة برمجة تطبيقات تويتر (API) الخاصة بك باستخدام حساب تويتر الخاص بك



#### ▼ Settings

These are your settings that control who you connect to the Twitter API as. You will be directed over to Twitter to grant your Hydrator application permission to use your account to retrieve Twitter data.

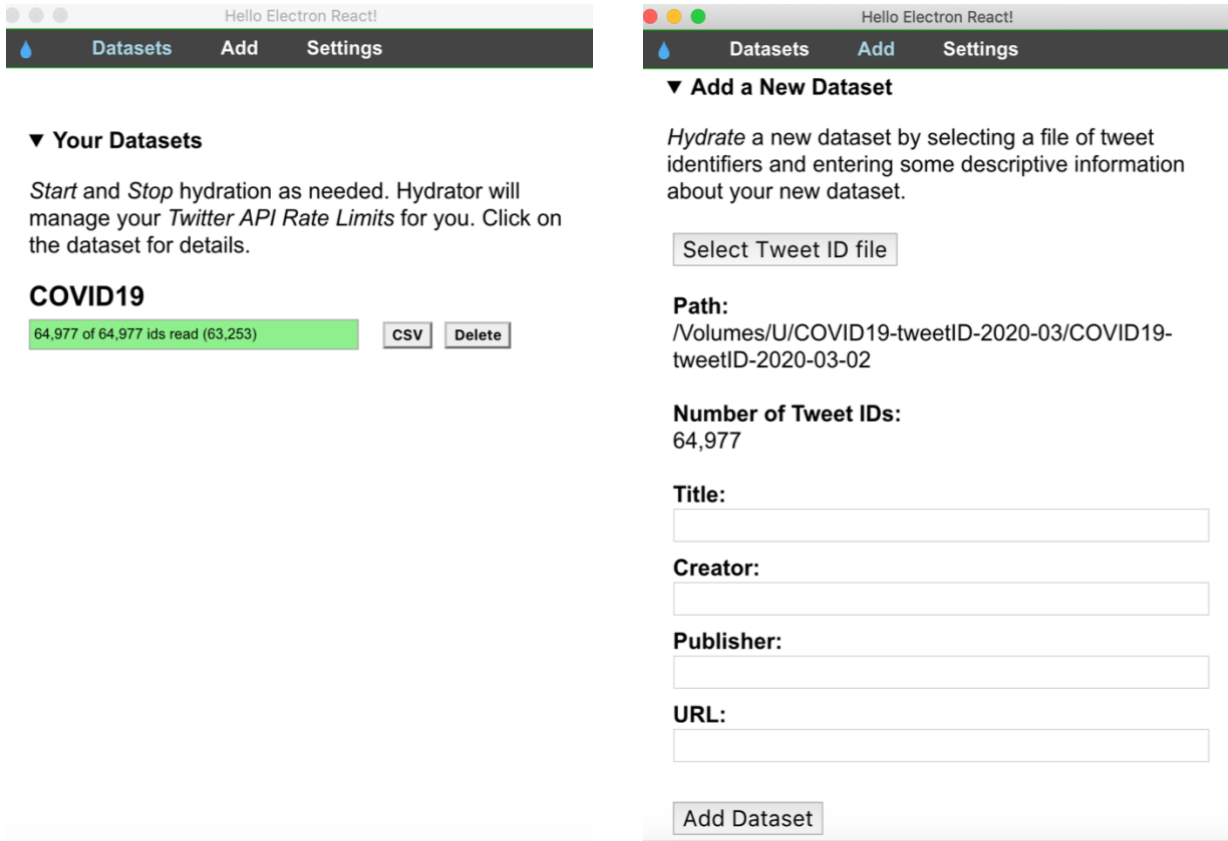
**Access Key:**

**Access Key Secret:**

---

رسم توضيحي 1 واجهة Hydrator

- ثالثًا: اختيار ملف الذي ترغب في استرجاعه بعد انتهاء التحميل يمكنك حفظه كملف CSV



رسم توضيحي 2 استرجاع مستودع التغريدات

## 2. TWARC

مكتبة بايثون لأرشفة بيانات تويتر تمكن من استرجاع كامل التغريدات باستخدام الـ TweetID

### نوت بوك لشرح طريقة استعادة التغريدات عن طريق الـ Tweet id

- هذا النوت بوك يساعد في استعادة التغريدات تلقائيًا باستخدام الـ TweetID من مستودع "COVID-19- Arabic-Tweets-Dataset on Github" باستخدام TWARC
- لاستعادة التغريدات يجب عليك الحصول على حساب مطور تويتر

### 1- تثبيت مكتبة Twarc

```
%pip install twarc
```

```
Form google.colab import drive
From twarc import Twarc
Import pandas as pd
Import json
Import csv
```

## Mount Drive -2

لنسخ مستودع تغريدات كوفيد ١٩ باللغة العربية من القت هاب على قوقل درايف . قم بتنفيذ الكود التالي واتبع الرابط الذي سيعرض لك كمخرج

```
Drive.mount("/content/drive")
```

هذا الكود ينسخ مستودع تغريدات كوفيد ١٩ باللغة العربية من القت هاب على قوقل درايف

```
!git clone https://github.com/SarahAlqurashi/COVID-19-Arabic-Tweets-Dataset
```

## Twitter API Keys -3

بمجرد إنشاء تطبيقك كمطور تويتر ، اكتب المعلومات التالية:

```
# Variables that contains the user credentials to access Twitter API
Access_token = "" #@param {type:"string"}
Access_token_secret = "" #@param {type:"string"}
Consumer_key = "" #@param {type:"string"}
Consumer_secret = "" #@param {type:"string"}
```

```
T = Twarc ( Consumer_key , Consumer_secret , Access_token , Access_token_secret )
```

## Hydrate TweetID -4

قم بإعداد ملفات الإخراج والإدخال

```
InputFile = "" #@param {type:"string"}
OutputFile = "" #@param {type:"string"}
```

يعمل الكود التالي على استرجاع التغريدات كامله من خلال الـ id الخاص بالتغريدة  
ملاحظة : وقت تنفيذ الكود يعتمد على عدد التغريدات المسترجعة لإن واجهة برمجة التطبيقات الخاص  
بتويتر ( API ) يفرض حدود لاسترجاع التغريدات فإن الكود سيظل في وضع السكون لبضع دقائق ويستمر  
تلقائيًا.

```
Ids = []
HydratedTweets = []
With open (InputFile,"r") as IdsFile : # Open the input file
    Ids= IdsFile.read().split() #read and store the ids
Count=len(HydratedTweets) #set up counter
EndIndex=len(ids)
Print("Total Ids : " + str(len(ids)))
For tweet in T.hydrate(iter(ids)):
    HydratedTweets.append(tweet) #append tweet json into a list
    Count += 1
    If (count == end_index): # stop condition
        Break
Print(str(len(HydratedTweets)) + "Tweet Saved ")
With open ('JsonOutputFile.txt','w') as outfile : #open json file
    Json.dump(HydratedTweets, outfile) #write the data into json file
```

هذا الكود يقوم بحفظ البيانات في ملف CSV

```
FileInput = "JsonOutputFile.txt"
InputFile = open () #open json file
OutputFile =open (OutputFile , 'w' ) #load csv file
Data = json.load() #load json content
InputFile.close() #close the input file
Output= csv.writer(OutputFile) #create a csv write
Output.writerow(data[0].keys()) # header row
for row in data :
    Output. Writerow(row.values()) #values row
```