

03_saving_tweets

September 19, 2016

1 Guardando Tweets en un fichero

```
In [ ]: import tweepy
import json

consumer_key = ""
consumer_secret = ""
access_token = ""
access_token_secret = ""

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

API = tweepy.API(auth)

In [ ]: # Hacemos la consulta

query = "valencia"
resultado = API.search(query, count=10)
```

1.1 Guardamos los resultados en formato JSON

```
In [ ]: # guardando los tweets en formato JSON

def save_json(filename, j):
    with open(filename, 'w') as fh:
        json.dump(j, fh, ensure_ascii=False)

def load_json(filename):
    with open(filename) as fh:
        return json.load(fh)

# puedes guardarlos como una lista
save_json("elnombredelfichero", [t._json for t in resultado])

# o como un diccionario
save_json("elnombredelfichero", {t.id: t._json for t in resultado})
```

1.2 Guardamos los resultados en texto plano

```
In [ ]: # guardando los tweets en texto plano

import re

er = re.compile('\s')

def save_tweets(filename, tweets):
    with open(filename, 'w') as fh:
        for tweet in resultado:
            tid = str(tweet.id)
            user = tweet.user.screen_name
            #text = ' '.join(tweet.text.split()) # eliminamos separadores '
            text = er.sub(' ', tweet.text) # eliminamos separadores "raros"
            line = '\t'.join((tid, user, text))
            fh.write(line + '\n')

def load_tweets(filename):
    tweets = []
    for line in open(filename):
        print (" "+line+" ")
        print (line.split('\t'))
        tid, user, text = line.split('\t')
        tweets.append({'id':tid, 'user':user, 'text':text})
    return tweets

save_tweets("elnombredelfichero", resultado)

#load_tweets("elnombredelfichero")
```

2 Minando Twitter

2.0.1 Lluís F. Hurtado (lhurtado at dsic.upv.es)

```
In [ ]:
```