

Forum:[Hack Forums](#)

Thread: [Save Tumblr jpg png gif Before 17th Dec \(Python\)](#)

FriendZone - December 13, 2018, 3:33 pm

Save this as HexeTumblerDownloader.py

And from terminal run:

Code python3 HexeTumblerDownloader.py It will save files in the directory where you have the script.

Login to get cookies session

Give it a link like: <https://ILoveCats.tumblr.com>

And a max number pages to scan (without a limit there are tumblr pages with 400+ pages lol tested lmao i had to stop it)

To use it from windows:

<https://www.python.org/downloads/>

download python from the link, and install it in the C:\ folder

inside Python37-32 folder (just installed) there will be a Scripts folder, inside there will be a pip.py, copy it and paste outside the folder and inside Python37-32

you should have python.exe and pip.py inside the same directory now

open cmd from start search bar

and move ourself in the Python37-32 directory

Code cd C:\Python37-32 time to install required modules

Code python.exe pip.py install requests

python.exe pip.py install beautifulsoup4

fine now copy the following script as HexeTumblerDownloader.py inside Python37-32

and run it

Code python.exe HexeTumblerDownloader.py

HOTFIX NOTES:

1-seems some pages dropped https, to move on http causing Page not found after the first one, added some code to analyze the problem and always find pics in https or http

2-as requested in PM added directory input to save each link in a different directory

```
Code #####  HEXECUTOR  #####
##### Hackforums.net #####
#####  Tumblr  #####
#####  jpg,png,gif  #####
#####  DOWNLOADER  #####
```

```
import os, requests, time, re
from urllib.request import urlopen
from bs4 import BeautifulSoup as bs
```

```
def HexeTumblrDownloader():
```

```
    username = input("Email/Username: ")
```

```
    password = input("Password: ")
```

```
    ua = 'Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/41.0.2228.0 Safari/537.36'
```

```
    s = requests.Session()
```

```
    s.headers = {'User-Agent': ua}
```

```
    # form_key
```

```
    print('Start...')
```

```
    r = s.get('https://www.tumblr.com/login')
```

```
    formKey = r.text.split('<input type="hidden" name="form_key" value="')[1]
```

```
    formKey = formKey.split('"')[0]
```

```
    data = {'user[email]': username, 'user[password]': password, 'user[tos]': 1,
            'context': 'login', 'version': 'STANDARD', 'follow': "",
            'http_referer': 'https://www.tumblr.com/login', 'form_key': formKey}
```

```
    # login
```

```
    print('Logging...')
```

```
    r = s.post('https://www.tumblr.com/login', data=data)
```

```
    if 'logged_in' not in s.cookies or s.cookies['logged_in'] is not '1':
```

```
        print('Error logging in.')
```

```
        return
```

```
    else:
```

```
        print('Successfully logged in.')
```

```

data = {}
email = ""
password = ""
i=0

#urls input cycle
while True:
    page=input("Please paste the URL or Ctrl+z to stop (IT MUST END WITH .com NOT WITH .com/): ")
    maxpage=input("Scanning max pages: ")
    directory=input("Please write the directory Path e.g. C:\\Users\\NAMEUSER\\Desktop\\ IT MUST END WITH \\ to save inside the last directory specified (if empty every pic will be saved into the python folder):")
    pagenumb=1
    while pagenumb <= int(maxpage):
        #the first page doesn't require page number specification (it can cause no contents)
        if pagenumb!=1:
            r = s.get(page+'/page/'+str(pagenumb))
        else:
            r = s.get(page)
        #save jpg png gif
        linkregex = re.compile('([a-z]+:V66(?:[^\V]+V)*[^\V]+\.(?:jpg|png|gif))')
        matched = linkregex.findall(r.text)
        #sometimes it drops https after the first page so let's use http if the output is empty
        if matched == [] :
            r = s.get(page.replace('https','http')+'/page/'+str(pagenumb))
            matched = linkregex.findall(r.text)

        for img in matched:
            print('----- Page '+str(pagenumb)+' Pic#'+str(i)+' -----')
            if img.endswith('jpg'):
                extension='.jpg'
            elif img.endswith('png'):
                extension='.png'
            else:
                extension='.gif'
            f = urlopen(img)
            with open(directory+'img'+str(i)+extension, 'wb') as w:
                w.write( f.read() )
            i=i+1
            print(img)
        pagenumb=pagenumb+1
    return

```

HexeTumblerDownloader()