

Что делать, когда тебя вычисляют по IP

Дмитрий Сергеев
ZeptoLab



zeptolab®

Зачем это нужно?

Зачем это нужно?

- Вы решили собрать себе немножко данных

Зачем это нужно?

- Вы решили собрать себе немножко данных
- Сервер не в восторге от ковровой бомбардировки автоматическими запросами

Зачем это нужно?

- Вы решили собрать себе немножко данных
- Сервер не в восторге от ковровой бомбардировки автоматическими запросами
- Error 403, 404, 503...

Зачем это нужно?

The internet will make those bad words go away



Essential

Googling the
Error Message

O RLY?

The Practical Developer
@ThePracticalDev

Предыстория вопроса

- Небольшой исследовательский проект

Предыстория вопроса

- Небольшой исследовательский проект
- Предсказание популярности мема по его описанию

Предыстория вопроса

- Небольшой исследовательский проект
- Предсказание популярности мема по его описанию
- Зачем - «for research purposes»

Предыстория вопроса

- Небольшой исследовательский проект
- Предсказание популярности мема по его описанию
- Зачем - «for research purposes»
- Надо собирать датасет



Applied Memology

CHEEZ
burger

I Can Has | Pop Culture | FAIL Blog | Memebase | Geek Universe | Loquillo | Know Your Meme

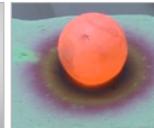
About Rules Chat Random Activity

Welcome! [Login](#) or [signup now!](#)

Search the database... Q

Advanced Search Protips

HOME MEMES IMAGES VIDEOS FORUMS BLOG EPISODES

WHO WOULD WIN IN A FIGHT?  EASY D    #SHEPERSISTED  B  PUNCH-A-NAZI 

JET FUEL STEEL BEAMS

Also Trending: Dear White People Richard Spencer's Interview The Emoji Movie Salt and Ice Challenge B Button Emoticon Antifa See You In Court He Will Not Divide Us...

All Entries

+ Add an Entry

Browsing all 15,266 meme entries.

Like us on Facebook!

Confirmed Submissions Deadpool All Entries Researching Most Popular

1 2 3 4 5 6 7 8 9 ... 381 382

Entry Categories

Events 1,031

Memes 11,693

People 1,039

Sites 364

Subcultures 1,136



Submission it wasn't because of any bed save report give gold reply
feathMrGoogle 8 points 1 day ago
you too have joined the legion
link embed save parent report give gold reply
Jhummie 8 points 1 day ago
Ah, a fellow brother
permalink embed save parent report give gold reply
LukeDrywaker 17 points 1 day ago
What the hell is this
permalink embed save parent edit delete reply

Submission SEE YOU IN COURT, THE SECURITY OF OUR NATION IS AT STAKE!
Official Twitter Account of the President of the United States
Twitter



Applied Memology



Doge

*Part of a series on Interior Monologue Captioning. [View Related Entries]
[View Related Sub-entries]*

Updated 9 months ago by Asdfghjkl.

Added 3 years ago by NovaXP.

281

11,731,983

54

1,633

895

[Share this Entry](#)

27,394

G+1

573

Like us on Facebook!

[Like](#) 834K

PROTIP: Press 'i' to view the image gallery, 'v' to view the video gallery, or 'r' to view a random entry.

About

Doge is a slang term for “dog” that is primarily associated with pictures of Shiba Inus (nicknamed “Shibe”) and internal monologue captions on [Tumblr](#). These photos may be photoshopped to change the dog’s face or captioned with [interior monologues](#) in [Comic Sans](#) font.

Origin

The use of the misspelled word “doge” to refer to a dog dates back to June 24th, 2005, when it was mentioned in an episode of [Homestar Runner](#)’s puppet show. In the episode titled “Biz Cas Fri 1”^[2], Homestar calls Strong Bad his “d-o-g-e” while trying to distract him from his work.

Meme

Status

Confirmed

Type:

Character

Year

2013

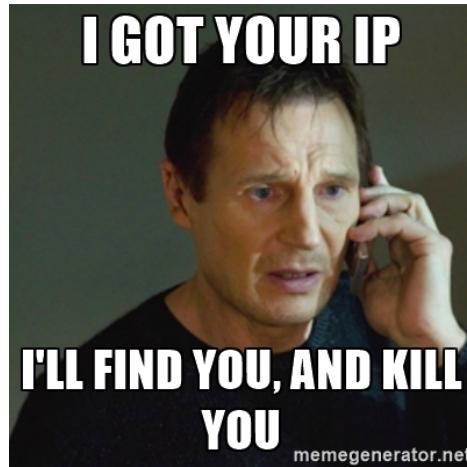
Origin

Tumblr

Tags

animal, dog, shiba inu, shibe, such doge, super shibe, japanese, super, tumblr, much, very, many, comic sans,

First encounter



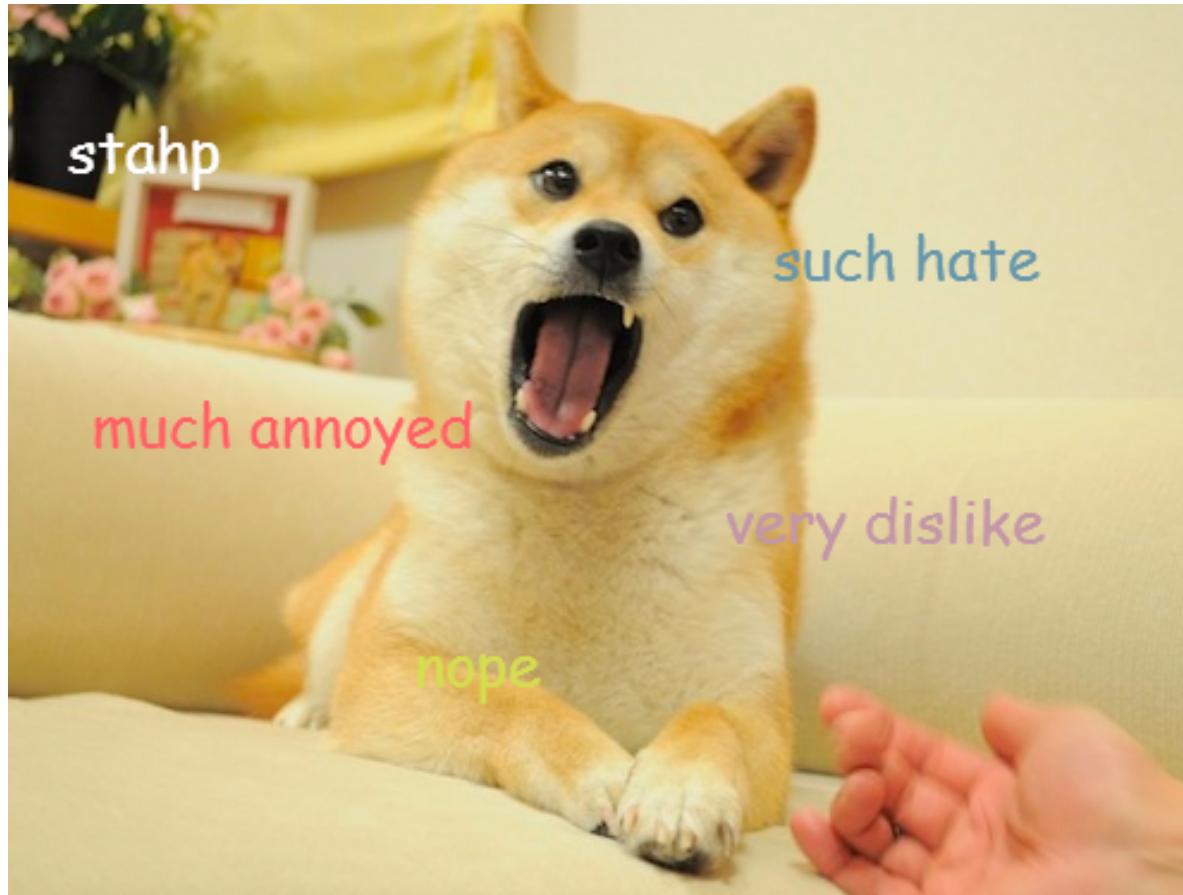
Your ip address **46.242.87.199 has been banned.**

If you think this is a mistake then please contact bans@knowyourmeme.com immediately and *make sure to include your ip address.*

This should be obvious but if you don't include your ip address then there's no way to unban you.

Don't include a screenshot of this page or any other page in the email; it's just a waste of time.

Как не разозлить сервер?



1. Быть терпеливым



1. Быть терпеливым



time.sleep(3)

1. Быть терпеливым

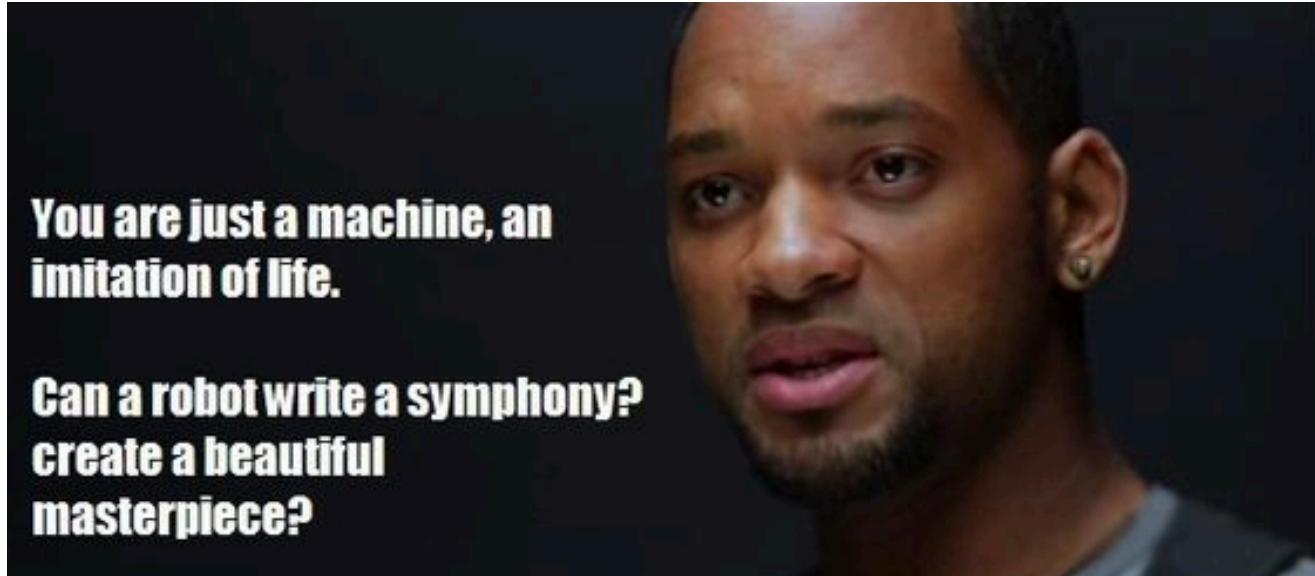


И пусть весь мир подождёт

1. Быть терпеливым

- Заполнение форм
- Переходы по страницам
- Клики на кнопки
- Параллельные запросы
- Запросы к API и т.д.

2. Быть похожим на человека



2. Быть похожим на человека

Request-headers здорового человека

Host	https://www.google.com/
Connection	keep-alive
Accept	text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8
User-Agent	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_5) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/39.0.2171.95 Safari/537.36
Referer	https://www.google.com/
Accept-Encoding	gzip, deflate, sdch
Accept-Language	en-US,en;q=0.8

2. Быть похожим на человека

Request-headers курильщика

Accept-Encoding	identity
User-Agent	Python-urllib/3.4

2. Быть похожим на человека

Генерация случайных юзер-агентов:

```
from user_agent import generate_user_agent  
  
requests.get('http://example.org', headers={'User-Agent': generate_user_agent()})
```

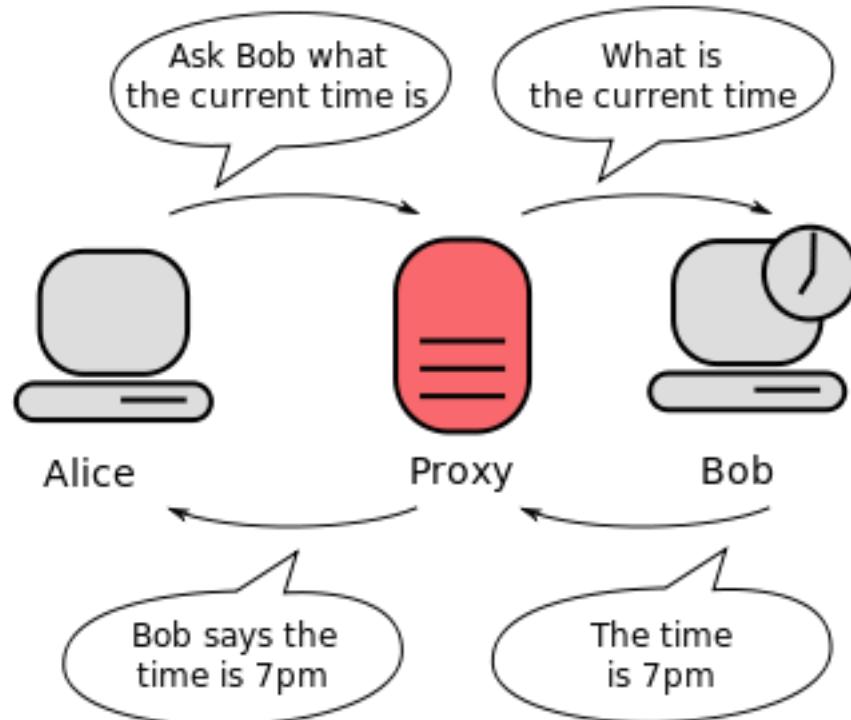
Mozilla/5.0 (Windows NT 10.0; rv:43.0) Gecko/20100101 Firefox/43.0

Mozilla/5.0 (Macintosh; Intel Mac OS X 10_11_1) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/39.0.2193.98 Safari/537.36

Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/
53.0.2798.6 Safari/537.36

https://pypi.python.org/pypi/user_agent

3. Общаться через посредников



3. Общаться через посредников

```
import requests

proxy = {'http' : 'http://102.32.3.1:8080',
          'https': 'https://102.32.3.1:4444'}
response = requests.get('http://google.com', proxies=proxy)
```

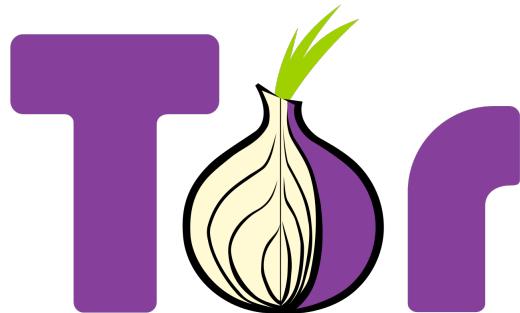
<https://www.myprivateproxy.net/>

<http://docs.python-requests.org/en/latest/user/advanced/#http-verbs>

4. Уходить глубже



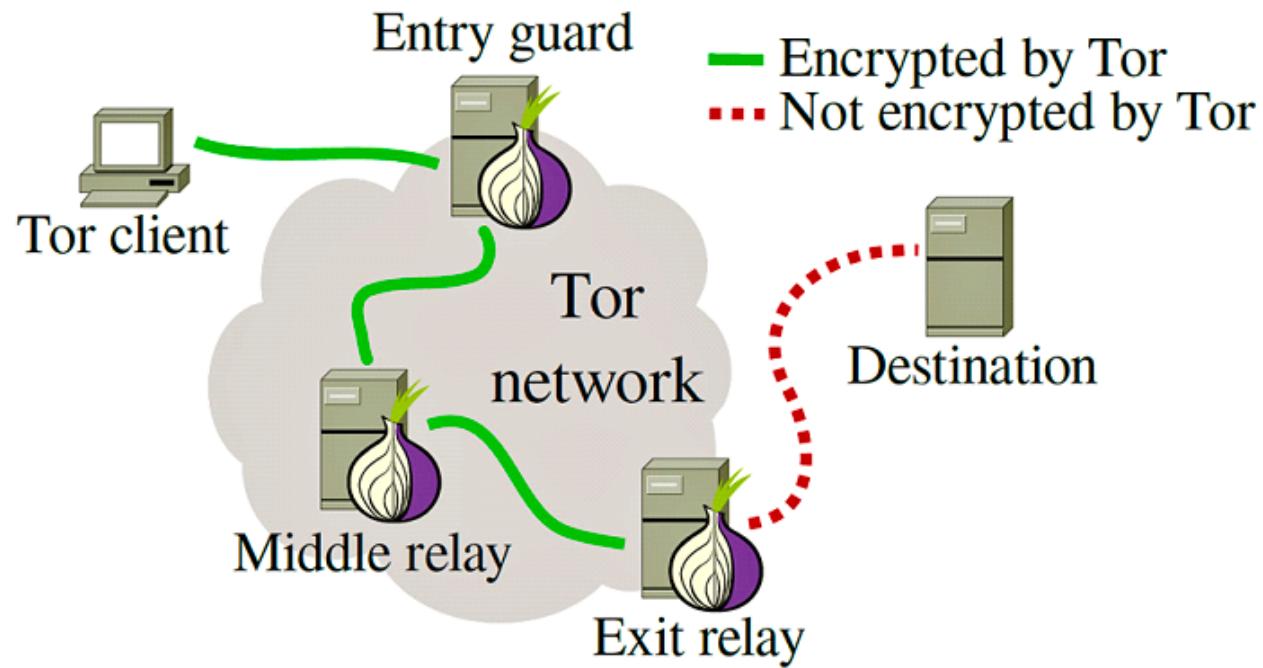
4. Уходить глубже



Tor circuit for this site

(google.com):

- This browser
- Germany (148.251.214.53)
- Switzerland (81.17.16.43)
- Switzerland (176.10.99.205)
- Internet



4. Уходить глубже

Задаем дефолтный порт для питона

TOR по умолчанию использует 9150

```
import socks
import socket
socks.set_default_proxy(socks.SOCKS5, "localhost", 9150)
socket.socket = socks.socksocket

print(urlopen('http://icanhazip.com').read())
```

```
>> b'185.100.85.220\n'
```

4. Уходить глубже

 **Your IP Address Information**

IP Address	185.100.85.220
Country Code	RO
Country Name	 Romania
Region Name	Bucuresti
City Name	Bucharest 

[Report Incorrect Location](#)

4. Уходить глубже

Также работает вместе с selenium

```
import os
from selenium.webdriver.firefox.firefox_binary import FirefoxBinary
from selenium import webdriver

# path to the firefox binary inside the Tor package
binary = '/Applications/TorBrowser.app/Contents/MacOS/firefox'
firefox_binary = FirefoxBinary(binary)

browser = webdriver.Firefox(firefox_binary=firefox_binary)
browser.get('https://check.torproject.org/')
```

5. Совместить всё?

5. Совместить всё?

- Начинать стоит с малого

5. Совместить всё?

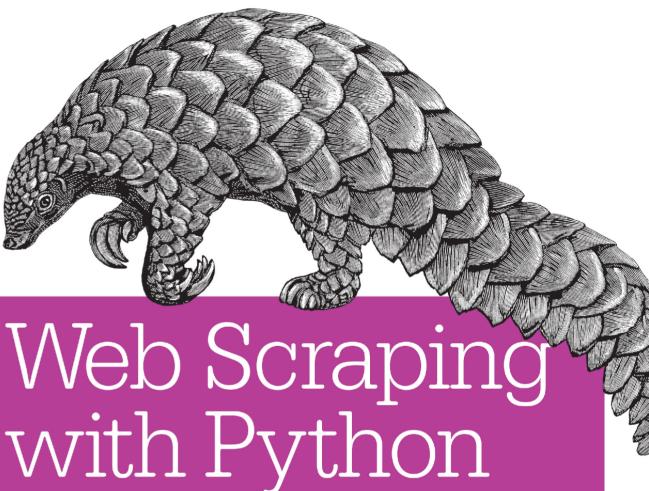
- Начинать стоит с малого
- Если продолжают баниТЬ – добавлять больше

5. Совместить всё?

- Начинать стоит с малого
- Если продолжают баниТЬ – добавлять больше
- ТеряЕМ в скорости

Чего почитать

O'REILLY®



Ryan Mitchell

GO FUN YOURSELF.

<https://github.com/DmitrySerg/memology>