

伪装请求头-简单实用的fake-useragent库



邓旭东HIT

【公众号：大邓和他的Python】

35 人赞了该文章

详情可关注公众号：大邓带你玩Python

前几天意外找到一个简单实用的库- fake-useragent,可以伪装生成headers请求头中的User Agent值。

安装

```
pip3 install fake-useragent
```

各浏览器的user-agent值

```
from fake_useragent import UserAgent
ua = UserAgent()
#ie浏览器的user agent
print(ua.ie)
```

```
Mozilla/5.0 (Windows; U; MSIE 9.0; Windows NT 9.0; en-US)
```

```
#opera浏览器
print(ua.opera)
```

```
Opera/9.80 (X11; Linux i686; U; ru) Presto/2.8.131 Version/11.11
```

```
#chrome浏览器
print(ua.chrome)
```

```
Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.2 (KHTML, like Gecko) Chrome/22.0.1216.0
```

```
#firefox浏览器
print(ua.firefox)
```

```
Mozilla/5.0 (Windows NT 6.2; Win64; x64; rv:16.0.1) Gecko/20121011 Firefox/16.0.1
```

```
#safari浏览器
print(ua.safari)
```

```
Mozilla/5.0 (iPad; CPU OS 6_0 like Mac OS X) AppleWebKit/536.26 (KHTML, like Gecko) Ve
```

最实用的

但我认为写爬虫最实用的是可以随意变换headers，一定要有随机性。在这里我写了三个随机生成user agent，三次打印都不一样，随机性很强，十分方便。



```
print(ua.random)
print(ua.random)
print(ua.random)

Mozilla/5.0 (X11; Ubuntu; Linux i686; rv:15.0) Gecko/20100101 Firefox/15.0.1
Mozilla/5.0 (Windows NT 6.2; Win64; x64; rv:16.0.1) Gecko/20121011 Firefox/16.0.1
Opera/9.80 (X11; Linux i686; U; ru) Presto/2.8.131 Version/11.11
```

爬虫中具体使用方法

```
import requests
from fake_useragent import UserAgent
ua = UserAgent()
headers = {'User-Agent': ua.random}
url = '待爬网页的url'
resp = requests.get(url, headers=headers)
#省略具体爬虫的解析代码，大家可以回去试试
...
```

编辑于 2017-06-17

[Python](#) [Python 入门](#) [Python 库](#)

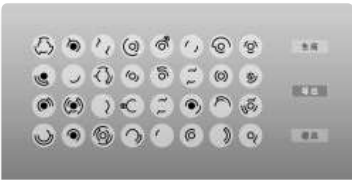
▲ 赞同 35 ▼ ● 10 条评论 ↗ 分享 ★ 收藏 ...

文章被以下专栏收录

 **爱上数据科学**
内容涵盖数据采集、文本分析和机器学习

进入专栏

推荐阅读



二厂开工——Fake Hymmnos

聂渲南 发表于造字二厂



二厂开工——Fake Cuneiform


聂渲南

Fal
从才
缠着
同时
钱，
非商
作伴
Chc

10 条评论

⇌ 切换为时间排序

写下你的评论...

 **Neo Ko** 1 年前

以上功能就算自己实现也不出20行，为了20行代码引入一个库？

👍 赞



那写起来省事不?

👍 2 💬 查看对话



知乎用户

1 年前

最实用? 说实话, 这个功能, 我在刚刚接触python的时候自己写了一个模块, 美滋滋的觉得自己挺厉害。每次只要import一下就可以随机UA。结果从来没见过反扒策略有针对UA的, 这个真的太好伪装了。而且市面上浏览器的热门UA本来就不多, 如果你来设计限制策略会怎么设置? 某一个UA单位时间内访问次数达到阈值, 就302到验证码? 这被人发现了, 分分钟搞得你所有热门UA都不能访问, 用户咋办? 通过UA识别访问对象太不准确了, 还是得用IP。后来那个模块我再也没用过。

👍 2



知乎用户

1 年前

这个可以有

👍 赞



Neo Ko 回复 邓旭东HIT (作者)

1 年前

反正也是几分钟的事情, 倒是这个功能就完全是我自己掌控了, 而你的这个库, 却是纯靠信仰。我并不觉得几行代码的事情应该引入一个新的依赖, 无论对部署或者维护都没有什么好处。

👍 1 💬 查看对话



无敌好尴尬

1 年前

很实用 赞

👍 赞



下一步

1 年前

```
user_agent = [
    "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_6_8; en-us) AppleWebKit/534.50 (KHTML,
    like Gecko) Version/5.1 Safari/534.50",
    "Mozilla/5.0 (Windows; U; Windows NT 6.1; en-us) AppleWebKit/534.50 (KHTML, like
    Gecko) Version/5.1 Safari/534.50",
    "Mozilla/5.0 (Windows NT 10.0; WOW64; rv:38.0) Gecko/20100101 Firefox/38.0",
    "Mozilla/5.0 (Windows NT 10.0; WOW64; Trident/7.0; .NET4.0C; .NET4.0E; .NET CLR
    2.0.50727; .NET CLR 3.0.30729; .NET CLR 3.5.30729; InfoPath.3; rv:11.0) like Gecko",
    "Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; Trident/5.0)",
    "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0)",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 6.0)",
    "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)",
    "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.6; rv:2.0.1) Gecko/20100101 Firefox/4.0.1",
    "Mozilla/5.0 (Windows NT 6.1; rv:2.0.1) Gecko/20100101 Firefox/4.0.1",
    "Opera/9.80 (Macintosh; Intel Mac OS X 10.6.8; U; en) Presto/2.8.131 Version/11.11",
    "Opera/9.80 (Windows NT 6.1; U; en) Presto/2.8.131 Version/11.11",
    "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_7_0) AppleWebKit/535.11 (KHTML, like Gecko)
    Chrome/17.0.963.56 Safari/535.11",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; Maxthon 2.0)",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; TencentTraveler 4.0)",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1)",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; The World)",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; Trident/4.0; SE 2.X MetaSr 1.0; SE 2.X
    MetaSr 1.0; .NET CLR 2.0.50727; SE 2.X MetaSr 1.0)",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; 360SE)",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; Avant Browser)",
    "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1)",
    "Mozilla/5.0 (iPhone; U; CPU iPhone OS 4_3_3 like Mac OS X; en-us) AppleWebKit/533.17.9
```



```
(KHTML, like Gecko) Version/5.0.2 Mobile/8J2 Safari/6533.18.5",
"Mozilla/5.0 (iPad; U; CPU OS 4_3_3 like Mac OS X; en-us) AppleWebKit/533.17.9 (KHTML,
like Gecko) Version/5.0.2 Mobile/8J2 Safari/6533.18.5",
"Mozilla/5.0 (Linux; U; Android 2.3.7; en-us; Nexus One Build/FRF91) AppleWebKit/533.1
(KHTML, like Gecko) Version/4.0 Mobile Safari/533.1",
"MQQBrowser/26 Mozilla/5.0 (Linux; U; Android 2.3.7; zh-cn; MB200 Build/GRJ22;
CyanogenMod-7) AppleWebKit/533.1 (KHTML, like Gecko) Version/4.0 Mobile
Safari/533.1",
"Opera/9.80 (Android 2.3.4; Linux; Opera Mobi/build-1107180945; U; en-GB)
Presto/2.8.149 Version/11.10",
"Mozilla/5.0 (Linux; U; Android 3.0; en-us; Xoom Build/HRI39) AppleWebKit/534.13
(KHTML, like Gecko) Version/4.0 Safari/534.13",
"Mozilla/5.0 (BlackBerry; U; BlackBerry 9800; en) AppleWebKit/534.1+ (KHTML, like Gecko)
Version/6.0.0.337 Mobile Safari/534.1+",
"Mozilla/5.0 (hp-tablet; Linux; hpwOS/3.0.0; U; en-US) AppleWebKit/534.6 (KHTML, like
Gecko) wOSBrowser/233.70 Safari/534.6 TouchPad/1.0",
"Mozilla/5.0 (SymbianOS/9.4; Series60/5.0 NokiaN97-1/20.0.019; Profile/MIDP-2.1
Configuration/CLDC-1.1) AppleWebKit/525 (KHTML, like Gecko) BrowserNG/7.1.18124",
"Mozilla/5.0 (compatible; MSIE 9.0; Windows Phone OS 7.5; Trident/5.0; IEMobile/9.0; HTC;
Titan)",
"UCWEB7.0.2.37/28/999",
"NOKIA5700/ UCWEB7.0.2.37/28/999",
"Openwave/ UCWEB7.0.2.37/28/999",
"Mozilla/4.0 (compatible; MSIE 6.0; ) Opera/UCWEB7.0.2.37/28/999",
]
headers = {'User-Agent': random.choice(user_agent)}
```

我以前也用你这个库，后来感觉太少太单一就自己收集了一些，以上，各种PC浏览器，手机浏览器，连诺基亚绝版手机的都有。

👍 2

 cakesweet

1 年前

确实不针对ua，你用curl默认ua也不会有问题

👍 赞

 ZKeeer

1 年前

获取代理IP才是正道，这个作用不是很大

👍 赞

 自由的陌生人 回复 邓旭东HIT (作者)

3 个月前

你好，我刚学爬虫遇到一个一直解决不了的问题，想咨询下你可以吗
[zhihu.com/question/2847...](https://www.zhihu.com/question/2847...)

👍 赞 💬 查看对话