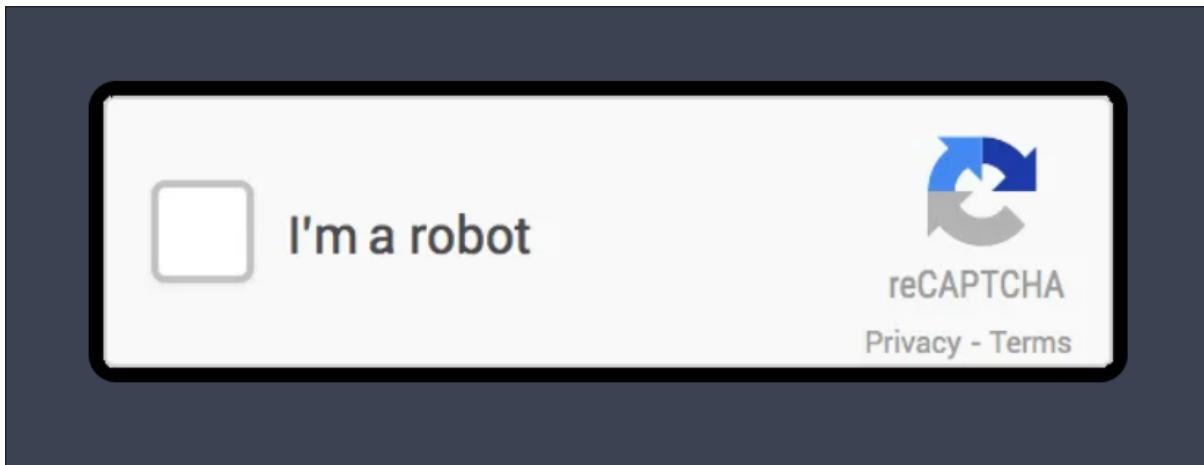


robot

#Web #HTTP #Header

/



Part #1



不是百度的robot我不认~

这里是对 User-Agent 请求头的修改，一般各大搜索引擎的蜘蛛都会带有特定的UA信息

[关于UA](#)

[百度的说法](#)

2. What is Baiduspider's user-agent?

Baidu uses different user-agents for different products:

Name of Products	User-agent
PC search	Baiduspider
Mobile search	Baiduspider
Image search	Baiduspider-image
Video search	Baiduspider-video
News search	Baiduspider-news
Baidu bookmark	Baiduspider-favo
Union baidu	Baiduspider-cpro
Business search	Baiduspider-ads
other search	Baiduspider

```
curl --user-agent Baiduspider http://<host>:<port>/verify
```

Part #2

```
<body>
  <div class="main">
    <a href="/verify"></a>
    <h1>不是百度来的还敢说自己是百度的robot(恼)</h1>
  </div>
</body>
 ~ 
```

为了声明自己从哪个网址到这里来，浏览器发送 Referer 头

Referer 的一个常见的作用是防盗链。比如一些图片网站会验证 Referer 头，如果不是自己的网站，就直接 403 —— ~~一种可行的反制措施是进行图片反代，推销下最近写的反代服务器~~

```
curl --user-agent Baiduspider --referer https://www.baidu.com http://<host>:<port>/verify
```

Part #3

```
<body>
  <div class="main">
    <a href="/verify"></a>
    <h1>你这曲奇多少钱一斤啊</h1>
  </div>
</body>
⌚ ~ |
```

这里是改 Cookie，参考Guidebook

要求是存在名为 price 的 Cookie (hint)，脑洞属实开大了>_<

```
curl --user-agent Baiduspider --referer https://www.baidu.com --cookie price=
http://<host>:<port>/verify
```

```
<body>
  <div class="main">
    <a href="/verify"></a>
    <h1>认证ok, 你要的东西: Spirit{ab4f7f67-edc0-46e1-8c8e-9d22cae19b03}</h1>
  </div>
</body>
⌚ ~ |
```