

memseto's Notebook

方知蓦然回首之时
那人却已不在灯火阑珊处

关于我
友情链接
文章聚合

Theme **Ringo** by **memset0**
Proudly powered by **Typecho**

下载 Oneindex 网站的整个文件夹

2019-02-07 | 开发

最近在别人的 Oneindex 下东西，懒得一个一个点于是就写了个脚本准备一次性下。

成品：github.com/memset0/oneindex-folder-spider

原理其实是很简单的，对于当前这个页面遍历每一个连接：如果是文件就下载，如果是文件夹就递归即可。下载出来的文件树会和 Oneindex 上的一样。

如果使用的话需要安装 Python3 & requests & wget 。

代码：

```
import requests, re, os, urllib

def mkdir(path):
    # print('mkdir', path)
    if not os.path.exists(path):
        os.makedirs(path)

def get(url, path):
    print('get', urllib.parse.unquote(url), urllib.parse.unquote(path))
```



方知蓦然回首之时
那人却已不在灯火阑珊处

方知蓦然回首之时

那人却已不在灯火阑珊处

关于我

友情链接

文章聚合

Theme **Ringo** by memseto

Proudly powered by [Typecho](#)

```

if url[-1] == '/':
    mkdir(urllib.parse.unquote(path))
    length = len(re.split(r'://[\S]*?/', url)[-1]) + 1
    content = requests.get(url).text
    children = re.findall(r'<li class="mdui-list-item mdui-ripple">\n
    folders = [children[i][52 + length:-2] for i in range(1, len(children)
    children = re.findall(r'<li class="mdui-list-item file mdui-ripple">\n
    files = [children[i][57 + length:-18] for i in range(0, len(children)
    for child in files + folders:
        get(url + child, path + child)
else:
    os.system('wget {url} -O {path}'.format(url=url, path=urllib.parse.un

__name__ == '__main__':
url = 'https://drive.bakaawt.com/Videos/%E6%B4%9B%E8%B0%B7%E6%98%A5%E5%A
path = os.getcwd() + '/down/'
get(url, path)

```

已有 2 条评论



z919126592

May 2nd, 2019 at 11:22 pm

貌似已经失效 使用Python3运行后马上停止 啥也没下载

回復



memset0

May 3rd, 2019 at 10:07 am

OneIndex 的前端有更新

略微修改即可。

回复

memseto's Notebook

方知蓦然回首之时
那人却已不在灯火阑珊处

用户名

邮箱

网址（选填）

可以在这里写评论哦 ~

提交评论

关于我
友情链接
文章聚合

Polya 定理学习笔记
上一篇 «

一句话题解 2019年1月
» 下一篇

Theme Ringo by memseto
Proudly powered by Typecho



memseto's Notebook

方知蓦然回首之时
那人却已不在灯火阑珊处

© 2017 - 2019 **memset0** 的博客.

浙ICP备19006255号-1

97700 visits · 24756 visitors · 74.48 W words

在这里输入关键字哦 ~ (回车搜索)

关于我

友情链接

文章聚合

Theme **Ringo** by **memset0**

Proudly powered by **Typecho**

