**踩坑记：终于怀着忐忑的心情完成了对 python 爬虫扩展库 pyspider 的安装**

在前面的系列文章中，笔者说了一些爬虫入门的相关知识点以及一些比较基本的爬虫方法和简单的框架使用等。下一个环节我么一起来看看 pyspider 框架。首先，这款框架的安装就十分令人头疼，花费了好几个小时本来要和女朋友聊天的时间，老王经过不断的踩坑终于安装成功了，所以，特别记录一下。

**1、pyspider介绍**

一个国人编写的强大的网络爬虫系统并带有强大的WebUI。

采用Python语言编写，分布式架构，支持多种数据库后端，强大的WebUI支持脚本编辑器，任务监视器，项目管理器以及结果查看器。

[docs.pyspider.org/en/latest/。](https://link.juejin.cn/?target=http%3A%2F%2Fdocs.pyspider.org%2Fen%2Flatest%2F%25E3%2580%2582)

**2、基础安装环境介绍**

操作系统：Windows10

python 环境：python3.6.8

操作人：老王的小跟班

**3、准备扩展库工具**

1*# 由于phantomjs官网下载太慢，可以到国内镜像站进行下载*

2

3http://npm.taobao.org/dist/phantomjs/

4

5phantomjs(版本=2.1.1，支持javascript操作)

6

7wheel(支持.whl文件安装)

8

9*# pycurl-7.43.0.4-cp36-cp36m-win\_amd64.whl下载地址*

10

11https://www.lfd.uci.edu/~gohlke/pythonlibs/*#pycurl*

12

13pycurl-7.43.0.4-cp36-cp36m-win\_amd64.whl(安装pycurl扩展库)

14

15wsgidav(版本=2.4.1)

16

17werkzeug(版本=0.15)

复制代码

**4、开始安装 pyspider**

1a.安装phantomjs

2

3*# 下载好phantomjs之后，解压之后配置好环境变量*

4

5b.安装wheel

6

7python -m pip install wheel

8

9c.安装下载好的pycurl扩展库

10

11python -m pip install c:/tools/pycurl-7.43.0.4-cp36-cp36m-win\_amd64.whl

12

13d.安装wsgidav依赖库

14

15*# 先卸载当前高版本*

16

17python -m pip uninstall wsgidav

18

19*# 安装指定低版本*

20

21python -m pip install wsgidav=2.4.1

22

23e.安装werkzeug

24

25*# 先卸载当前高版本*

26

27python -m pip uninstall werkzeug

28

29*# 安装指定低版本*

30

31python -m pip install werkzeug=0.15

32

33f.最后安装pyspider

34

35python -m pip install pyspider

复制代码

**5、启动问题**

1a.直接启动

2

3pyspider all

4

5b.启动中如果出现一直在 restarting的情况

6

7*# 处理办法：打开一个终端，输入命令启动。再打开一个新的终端启动，此时再关闭第一个终端，第二个终端会启动成功。*

8

9*# 为什么？我也不知道，这是在 Google 上找的解决办法。*

10

11c.第一次启动的时间比较长

12

13d.启动成功之后访问地址

14

15http://localhost:5000

16

17e.启动之后关闭终端，代表pyspider服务也会关闭