

Potato Web Spider 应用开发文档 1.1.0



进入网站

工具和本文档均由[Potato Studio](#)制作

欢迎使用子雨网页爬虫制作工具

本文档将教会你怎么使用这个工具，并展示工具的代码以方便开发者开发。前往官网可以获得以编写好的爬虫，可供使用者们使用和开发

工具的介绍

子雨网页爬虫工具是一个Potato Studio主要使用python语言编写的面相于对计算机感兴趣的普通用户和入门程序员编写和测试简单地网站网页爬虫。此工具是一个工具集，可以使用特定的代码调用并执行，类似：

```
title=My first spider
url&get=http://www.potatost.xyz
file&new=result.txt
file&writevalue=textweb
```

新建一个文件，随意命名为test.posp将上面这部分代码写入(请使用utf-8编码)，保存文件，打开工具，输入go，再输入你创建文件的路径。比如:E:\test.posp，点击回车。结果如下：

```
请输入子雨爬虫配置文件XXX.posp位置
>>>test.posp
My first spider
正在抓取http://www.potatost.xyz.....
##抓取成功##
##网站抓取内容信息已写入##
```

网页内容已经抓取到了软件根目录的result.txt文件

posp语言介绍

Potato Spider 是 Potato Studio 开发的本工具中工具调用的表达语言，所有的爬虫工具调用都要使用 posp 语言表达。语言格式如下：

key=value(key = 键--工具代码; value = 值--参数)

`file&new=E:\\hello.posp` (调用file包的新工具，在E盘创建子雨网页爬虫配置文件hello.posp)

!!!需要注意：前一行与后一行之间不要有空行，也不要有其他字符

posp代码说明：

此处列出所有可以调用的工具和如何书写使用工具的代码

注：[]里的是数据类型，实际写代码的时候没有[]

一、文字展示

```
title=[text]-----显示一段[text]
```

```
title=hello(在控制台显示hello)
```

二、爬虫基本操作

1.爬虫获取网页

```
url&get=[url]-----抓取[url]
```

```
url&get=http://www.potatost.xyz(抓取http://www.potatost.xyz网站)
```

一、文件的操作

!!!需要注意：使用文件前必须创建文件，在上一文件使用结束之前不能创建新的文件

```
文件1创建：-----
```

```
文件1操作A
```

```
文件1操作B
```

```
文件2创建：-----
```

```
文件2操作A
```

```
...
```

1.创建文件

file&new=[filePlace+fileName]-----在[filePlace]位置创建文件[fileName]

比如要在E盘创建potsto.txt文件代码应该为:

file&new=E://potsto.txt

!!!温馨提示, 如果您在C盘中系统文件夹或根目录创建文件, 很可能因为没有管理员权限而失败。

2.给文件写入一段文字(请在此操作之前先创建文件)

file&writeword=[text]-----将一段[text]写入文件

file&writeword=potato(在文件中写入potato)

3.给文件写入一个数据(请在此操作之前先创建文件)

file&writevalue=[value]-----将特定的数据[value]写入文件

file&writevalue=textweb(在文件中写入抓取网站的主体信息)

(1)以下是[value]的数据和其意义表:

value	意义	备注/用法
time	时间: 年月日	file&writeValue=time
headers	访问头	file&writeValue=headers(先抓取)
textWeb	抓取网站的主体信息	file&writeValue=textWeb(先抓取)
result	计算分析结果	file&writeValue=result(待更新)

(2)数据来源说明表:

value	信息来源
time	计算机内置时间
headers	通过调用url&get或其他爬取工具获得
textWeb	通过调用url&get或其他爬取工具获得
result	通过一系列数据处理后的结果(待更新)

The End

感谢您使用 Potato Web Spider 工具, 我们会继续更新工具, 你可以通过 potatostuser@163.com 来联系我们, 再次感谢您的使用

^_^