

怎么用爬虫获得上市公司数据库，网络爬虫能获取企业后台网站数据么-股识吧

一、网络爬虫能获取企业后台网站数据么

1.用浏览器打开 *://weibo.cn/pub/ 这是新浪微博面对手机用户的一个入口.建议不要使用自己的账号来做爬虫.之所以选择weibo.cn来爬去，是因为这个手机版的限制较少，但是weibo.cn的账号密码都是以明文传输，所以为了防止盗号，建议新申请账号进行爬取.2.登陆进去后，进入个人微博界面(不是*://weibo.cn/pub/)，使用firebug查看网页的cookie，找到cookie中的gsid_CTandWM的属性记录gsid_CTandWM的值，以后爬虫登陆新浪，全靠这个值.3.代码：代码是基于WebCollector的：WebCollector官网假设我们要爬取周鸿祎的微博信息.周鸿祎的微博主页：*://weibo.cn/u/1708942053(不登陆只能看第一页)我们爬取前10页.

```
public static void main(String[] args) throws IOException{
    CollectionGenerator generator=new CollectionGenerator();
    for(int i=1 ;
    i<=10 ;
    i++){ generator.addUrl("*://weibo.cn/u/1708942053?page="+i);
    }
    Fetcher fetcher=new Fetcher();
```

二、爬虫抓数据时，详情页的图片有必要单独存下来吗

你是指搜索引擎的爬虫？爬虫不会抓取图片，只会读取ALT信息

三、网络爬虫能获取企业后台网站数据么

一般是不能的，但是代码里没加密会被收录，我经历过

四、源码中没有出现的数据如何爬虫

采集表格内容，包括列表形式的商品评论信息、正文中的表格等，凡是html代码采用表单形式的表格，都可以不写代码，通过可视化的方式完成采集。

首先，我们使用的工具是前嗅大数据的ForeSpider数据采集系统。

在前嗅大数据官网下载免费版，免费版就可以满足我们抓取图片的需求。

下载安装后，启动软件。

根据如下步骤，即可抓取到图片或图片的地址了。

在ForeSpider中，采集表格的功能称之为“多值”。

多值用于存储表格的数据，将表格的不同列对应存入不同字段，表格的不同行分别存储为数据表的多条记录。

本文以某表格为例。

【需要多值存储的表格】（一）创建表单

根据表格内容，创建一个存储表格数据的表单。

在选项卡“表单”中，创建一个表单。

【多值的表结构】1.主键 采集表格时，表格的一行作为一条数据。

由于整个表格属于同一个网页文档，而文档主键只有一个，因此不能像采集其他内容一样，取值类型选择“文档主键”。

表格的主键的变量类型，根据表格的行数长度，选择“Integer”或者“Long”。

取值类型选择“空”。

字段属性选择“主键字段”和“自动字段”（选择主键字段后，软件会自动选择“键值唯一”和“索引字段”。

）

五、金蝶账套启用，房地产企业的开发成本数据怎么录入

帐套已经启用，只能通过编制记帐凭证进行处理。

参考文档

[下载：怎么用爬虫获得上市公司数据库.pdf](#)

[《股票一直压盘是什么意思》](#)

[《涨停板怎么表示》](#)

[《股票app里的胜率指什么》](#)

[《股票老兵专用指标什么意思》](#)

[《重仓煤炭基金有哪些2023》](#)

[下载：怎么用爬虫获得上市公司数据库.doc](#)

[更多关于《怎么用爬虫获得上市公司数据库》的文档...](#)

声明：

本文来自网络，不代表

【股识吧】立场，转载请注明出处：

<https://www.gupiaozhishiba.com/book/26048822.html>