任务名称： 爬虫与解析任务-爬取天眼查的企业全称信息

任务编号： TGNB\_PARSE\_13

1. 具体步骤：
2. 按股票代码查询：

<http://www.tianyancha.com/search?key=HK.[ddddd]&checkFrom=searchBox，>

其中[ddddd]是之前爬取得股票代码。



1. 按股票名称查询：

http://www.tianyancha.com/search?key=中国电子&checkFrom=searchBox



1. 按企业网址进行查询：

对之前爬取到的网址进行截取(去除开始的www及http:/.)，例如爬到到的中国长远的网址是www.fortunetele.com，则按如下网址查询：

<http://www.tianyancha.com/search?key=fortunetele.com&checkFrom=searchBox>



特别说明：

1) 以上三种方式均只能通过天眼查的方式来进行，用百度的site方式查询不到；

2) 以上三种方式对于每只股票都要进行，只需解析查询出来的公司全称即可，可能为空，也可能多个（多个取第一个结果）；

3) 天眼查有防刷机制，多访问几次可能会出现http://www.tianyancha.com/login?from=%2Fsearch这样的登录机制，请注意控制好频次。

二、访问要求： 主板+创业板，不到1900家企业，希望1天爬完。

三、整体目标：

主板+创业板，不到1900家企业，希望1天爬完。