对H组的软件问题报告

# 问题报告表单

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 项目名称 | H组 EasySpider | | | | | |
| 系统交付日期 | 2020-05-18 | 测试日期 | 2020-05-19 | | | |
| 测试工具 | Chrome | 编制人 | 李坤浩 | | | |
| 测试环境 | Windows10 | 测试人员 | 李坤浩，张竹君，张文斌 | | | |
| 序号 | 测试步骤 | | 预期结果 | 实际结果 | 报告人 | 备注 |
| 1 | 注册账号时，邮箱字段输入ddd@ddd | | 邮箱不合法，注册失败 | 注册成功 | 张文斌 |  |
| 2 | 1、新建豆瓣电影/音乐/读书任务，填写必要字段；  2、启动任务；  3、点击我的任务；  4、任务执行完毕后查看采集数据页面，并滚动已采集数据预览页面中表格的滚动条 | | 表格的项与内容对应 | 表格的表头与内容对不齐 | 张竹君 | 详见说明2.1 |
| 3 | 1、新建豆瓣读书任务，填写必要字段，其中关键字填写“爵士”；  2、启动任务；  3、点击我的任务；  4、任务执行完毕后查看采集数据页面  5、查看表格的第1条数据“爵士吉他金曲” | | 数据的“出版社”字段值应当为空 | 数据的“出版社”字段值为“出版年:” | 李坤浩 | 详见说明2.2 |
| 4 | 1、新建豆瓣电影任务，填写必要字段，其中关键字填写“哈哈哈哈哈哈哈”；  2、启动任务  3、点击我的任务；  4、任务执行完毕后查看采集数据页面；  5、查看表格的第13条数据“哈哈大冒险 第二季” | | 数据的“又名”字段应当为空 | 数据的“又名”字段为“2” | 李坤浩 | 详见说明2.3 |
| 5 | 1、新建豆瓣电影任务，关键词填写“蝙蝠”，翻页次数填写“25”；  2、启动任务；  3、点击我的任务；  4、任务执行完毕后查看采集数据页面；  5、点击导出数据，选择excel格式  6、将导出数据与表格数据进行对比 | | 导出数据与表格数据相同 | 导出数据有198条，而表格里只有15条数据 | 李坤浩 | 详见说明2.4 |

# 2 说明

## 2.1 问题2

第一次进入到“已采集数据预览”界面，就会出现表格头和其对应值对不齐的情况，影响系统的可读性。

页面的部分截图如下：





## 2.2 问题3

豆瓣读书的爬虫问题。

对于采集数据中某一条数据的“出版社”字段值有误，问题截图如下：



通过查看该书的豆瓣页面<https://book.douban.com/subject/10755581/>，可以看到该条目无“出版社”字段，但是有“出版年”字段。



## 2.3 问题4

豆瓣电影的爬虫问题。

对于采集数据中某一条数据的“又名”字段有误，问题截图如下：



通过查看该影视的豆瓣页面[https://movie.douban.com/subject/35004179 /](https://movie.douban.com/subject/35004179%20/)，可以看到该条目为电视剧，与电影的采集数据字段不一样。“又名”字段的值被采集为“2”，猜测爬虫系统是将该页面的“季数”的值当做了“又名”。

备注：豆瓣电影包括电影、电视剧和综艺，这三种类型影视的字段有不同，建议做一下区分。



## 2.4 问题5

页面上表格呈现的数据只有15条：



导出数据后，excel表中共有198行（去除第一行表头，有197条数据）

