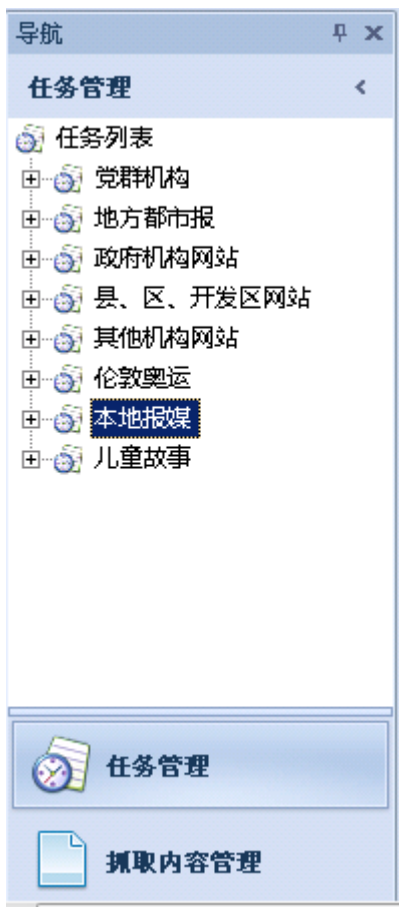


客户端使用说明

一、采集

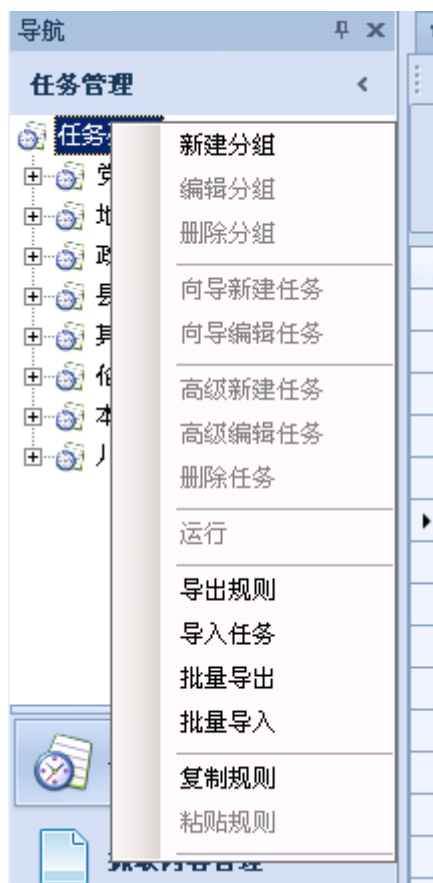
(一) 配置规则



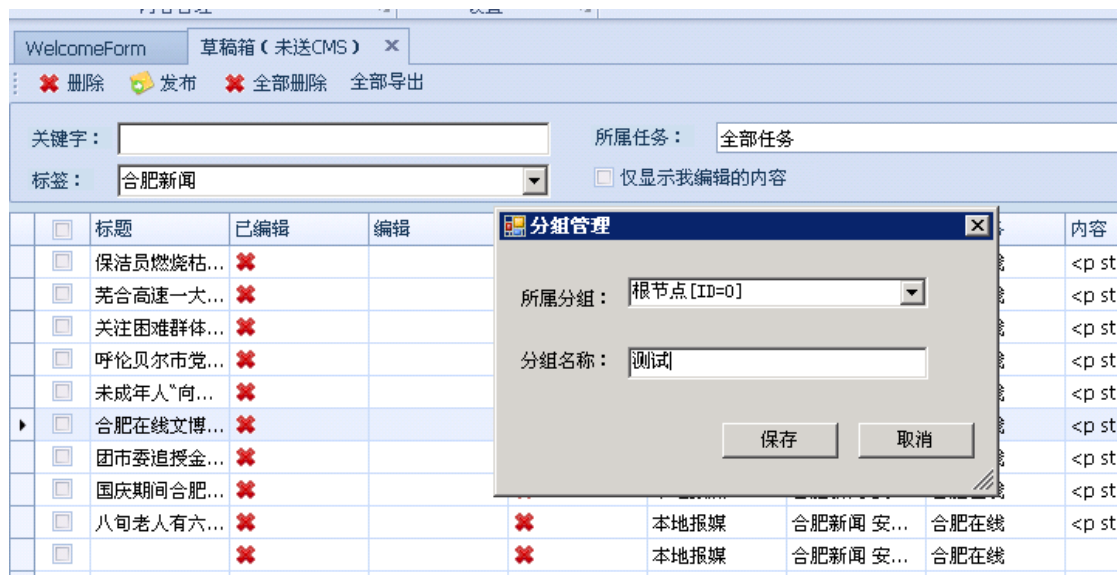
要开始采集，首先需要配置采集规则，无规矩不成方圆，请按下面的步骤来配置规则

1. 分组管理规则

规则可能有数十个，你可以建立一系列的分组来管理这些规则



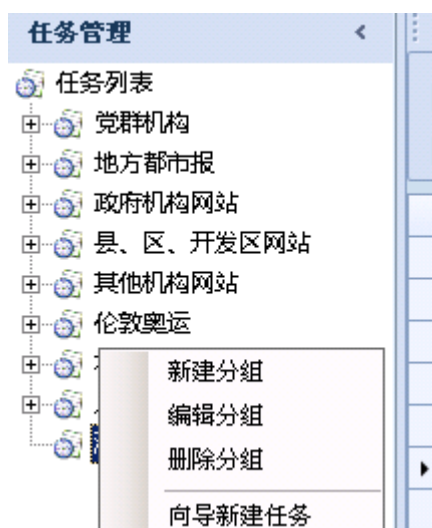
鼠标选中任务列表，右键菜单选择新建分组就可以新建一个任务分组



输入分组名称，点击确定，这个时候任务列表会多了一个测试分组



如果需要更改分组名称，选中分组，右键菜单选择编辑即可



2. 新建规则

1) 填写任务基本信息

先选中一个分组，邮件菜单选择“向导新建任务”



点击后会打开任务规则向导

按提示填写任务名称

任务标签：相当于内容分类，你可以为一个站点添加多个标签，用空格分开即可；标签的好处在于你可以按标签来区分和组织不同站点的内容

任务定时执行：你可以开启定时执行功能，这样每隔指定的分钟数，程序会自动执行这个任务

填写完毕，单击下一步

2) 设定起始页（列表页）



首先，在起始地址输入要采集网站的列表页面，然后单击“前往”按钮

列表页面：列表页面指新闻的列表页面，而不是栏目页面，例如：

新浪新闻——国内新闻：<http://roll.news.sina.com.cn/news/gnxw/gdxw1/index.shtml>



而不是 <http://news.sina.com.cn/china/>



香港学生参加“九一八”事变纪念



马拍隆奥选手坐轮椅从 日首相坦言中国海监船 巡航令其十分紧张

东海舰队举行海上演练 发射导弹40余枚 兼任日本驻华大使西宫伸一因病去世

国内要闻

我国公布钓鱼岛及部分附属岛屿坐标

【联合国将妥为公布钓鱼岛坐标 专题 滚动】

全国铁路动车网上售票尚未实现选票功能

【全国动车座席号统一 官网因系统升级出故障】

- 台湾要求菲律宾遣返近300名被捕网络诈骗嫌犯
- 广州悬赏征集游行中打砸人员身份线索
- 西安投资600亿改善城市水系生态规划遭质疑
- 反日活动细节：市民举牌提示日系车绕行
- 国台办主任：民族大义面前两岸应一致对外
- 各地政务微博就涉日游行发声 青岛西安保持沉默
- 湄公河惨案主犯糯康等人将于20日受审
- 美国防长称将扮演中日关系调停者角色
- 广州中山官员1年遭3次投诉人事任命或被打回
- 各大银行被指漠视国务院指令拖延降刷卡手续费
- 世界最大跨度公铁两用桥主跨钢梁在湖北合龙
- 宁夏海原县城幼儿园上学前班被要求买赞助票

点击排行

今天 昨天 一周

- 1 广州女子煽动破坏地铁设施
- 2 日本首相：抗议焚烧日本国旗
- 3 专家：美军或4种方式支援日本
- 4 野田：日将拘捕中国登岛渔民
- 5 中国完成捍卫主权法律手续
- 6 美国防长称将当中日调停人
- 7 据称多数自助捐赠者动机不纯
- 8 河北白头百年火柴厂破产背后
- 9 北京部分日式餐厅停止营业
- 10 西安禁止部分区域内游行集会

热门评论

- 陈小二：行动表达不应失去底线
- 汪昌莲：“冒雨工作照”别只当笑话看
- 吴江：“国门成景区”旅游无边界？
- 王石川：节假日免费“损失”200亿的秘密
- 齐鲁晚报：爱国 420547098@qq.com
- 沙元森：爱国就
- 干栋雷：我们是

更新：Interspeed 陈伟伟 chen@isf.net

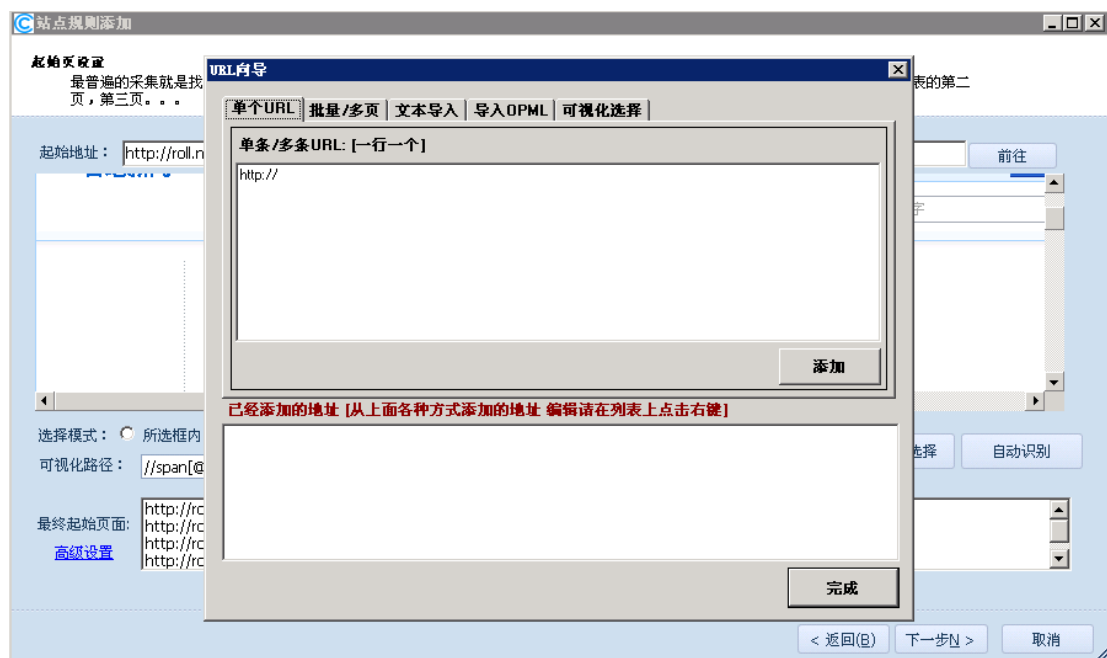
列表页面打开后，如果该列表页有分页，你可以单击“自动识别”按钮，软件会自动识别页面的分页链接，并加入起始页面库。如果不能自动识别，则需要手动选取了，点击“点此开始选择”按钮，将鼠标移动到分页链接所在的地方，单击两个不同的分页即可。（如下面的两张图片所示，分别点击 2，3 即可）





完成操作后最终起始页面自动填充了识别的结果

高级设置：厌烦了选来选取的麻烦？那么使用高级设置吧！单击“最终起始页面”下面的“高级设置”链接，打开高级设置窗口：



在单个 URL 里，你可以一行一个将 URL 地址拷贝进去，然后单击添加，最后单击完成。

在批量多页里，你可以通过通配符完成网站的批量添加，例如，我们知道新浪新闻的列表分页是这样的：

第二页：

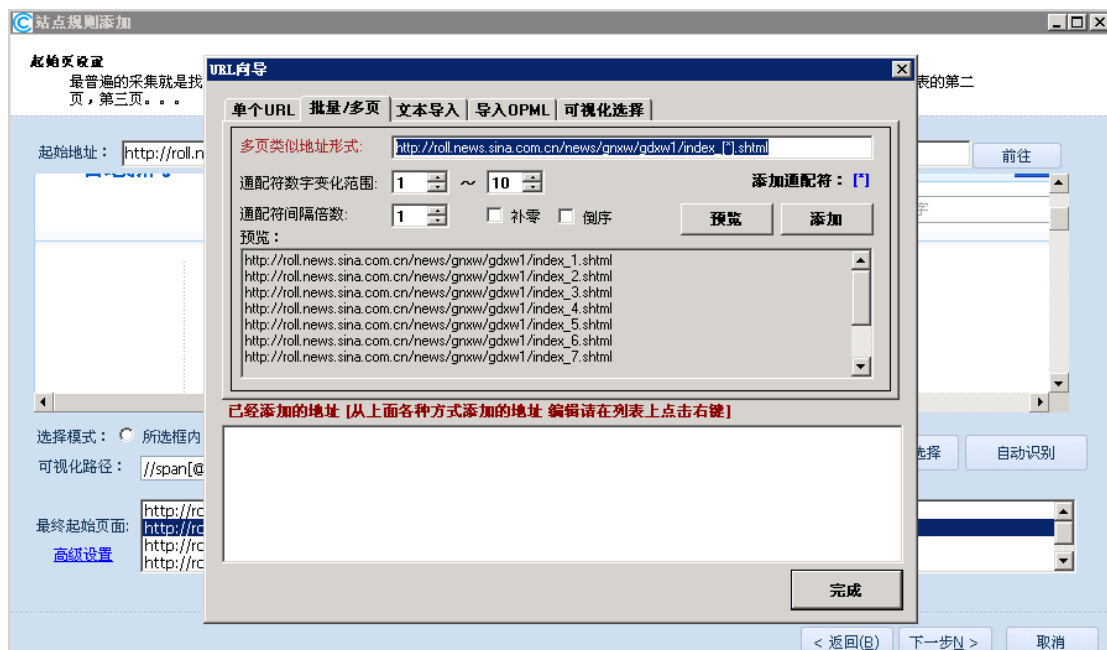
http://roll.news.sina.com.cn/news/gnxw/gdxw1/index_2.shtml

第三页

http://roll.news.sina.com.cn/news/gnxw/gdxw1/index_3.shtml

那么，我们可以这样设定规则：将数字的部分用通配符[*]代替：

http://roll.news.sina.com.cn/news/gnxw/gdxw1/index_*.shtml

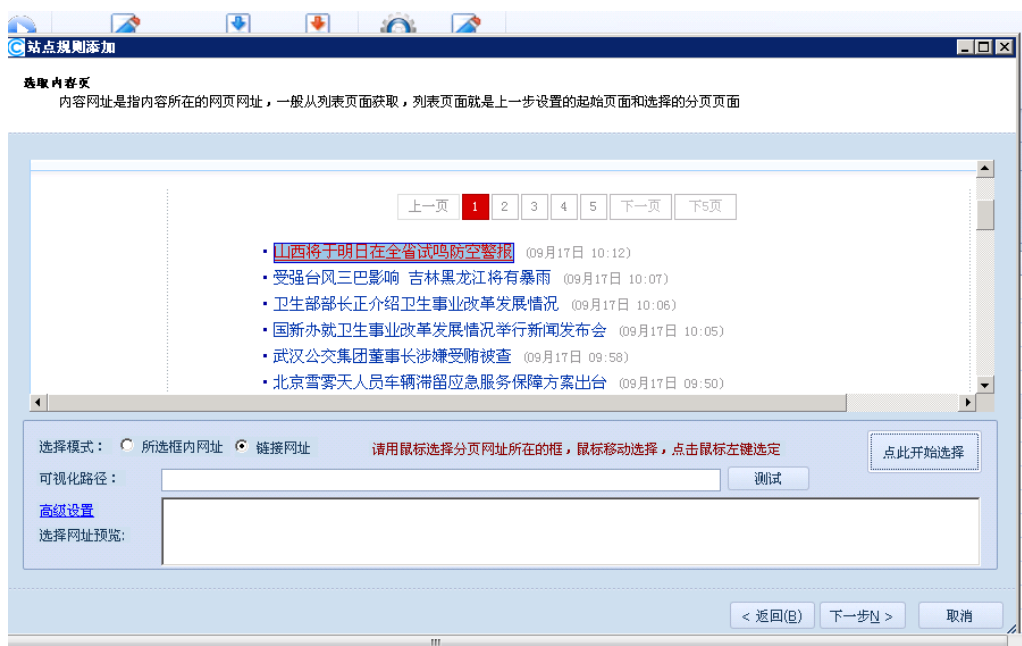


然后设定通配符变化范围，比如 1 到 10，单击预览即可查看效果，如果符合邀请，单击添加，完成。

设定完成后单击下一步。

3) 选定内容页 URL 选择规则（列表页）

这里的规则决定怎么从列表页面找出内容页的 url，你可以方便的点几下鼠标即可完成操作。



同样的，先点击“点此开始选择”，然后点击浏览器里第一条新闻，然后点击第二条新闻：



操作完成后，在预览框里会有选取的 url 结果，你可以查看这里的结果，如果不正确，可以重新操作。

高级设置：

可视化选择对比较新的网站一般可以正确识别，但是对于比较老的 table 嵌套页面，往往心不从心，这个时候，有个万能的规则：

- (1) 在可视化路径里输入 “//a”，点击测试，会出来数百条结果
- (2) 仔细浏览这里结果，可以发现，内容页以 “.shtml” 结束，并且不包含 “index”，所以，

我们点击高级设置：



输入相关配置，点击确定即可。

完成设置后单击下一步，这里可以预览结果：



OK 后单击下一步。

4) 配置内容页采集规则

选好了内容页，现在进入最重要的一步，配置内容采集的规则。



你可以依次配置标题、时间，来源，摘要，内容

配置方式非常简单，要切换到要配置的内容项目，比如标题：

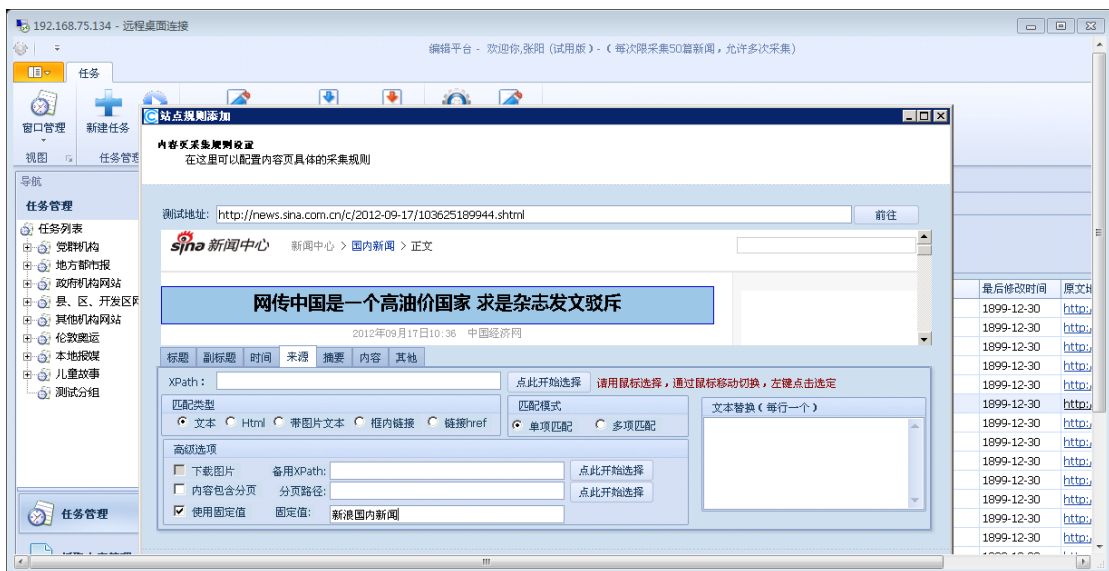
- (1) 点击“点此开始选择”
- (2) 在浏览器找到标题，单击即可



依次配置其他项目即可

高级设置：

- (1) 设定固定值 一般来说项目可以从网页选取，但是诸如来源等字段，你可以设定一个固定值



勾选上使用固定值，输入名称即可

- (2) 文本替换

文本替换（每行一个）

在文本替换框中输入要替换的文本即可，注意每行一个，不要留空白

设置完成，点击下一步即可预览结果

结果预览
在这里可以预览采集到的结果，如果不正确，请回到上一步修改

【标题】：求是驳斥中国是高油价国家说法:实际处较低水平

【副标题】：

【关键字】：

【来源】：新浪国内新闻

【时间】：2012-09-17 11:28

【内容】：<p style="text-indent: 28px;">中国经济网北京9月17日讯 “中国是一个高油价国家”的说法充斥着舆论和网络，最新出版的《求是》杂志对这个说法予以“驳斥”。《红旗文稿》杂志总编辑助理吴强在题为《关于垄断行业改革的几个问题》的文章中引用2012年5月美国彭博新闻社列出的一份数据，全球55个国家的油价排名中中国的油价排名第42位，是油价较低的国家。中国是人均能源极度匮乏的国家，维持较低的油价相当不容易。对于“油价低”的原因，文章认为，“很难想象，如果我国的三大石油公司都是私企，还能保持目前较低的油价吗？”</p><p style="text-indent: 28px;">2012年来发改委共有7次调整油价，累计汽油上涨560元/吨，柴油上涨590元/吨，油价总体呈上涨趋势。其中，油价上调有4次，多于下调次数，分别是：2012年2月08日汽油柴油每吨均上调300元；2012年3月20日汽油柴油每吨均上调600元；2012年8月10日汽柴油每吨分别升390元、370元；9月11日汽柴油每吨分别升550元、540元。</p><p style="text-indent: 28px;">此前也有专家表示，与发达国家相比，我国汽柴油零售价处于较低水平；从全球来看，在160多个国家中，中国汽柴油零售价处于中等水平。专家详细介绍了周边20个国家的汽油价格，中国不仅低于日韩，而且也低于印度、泰国、阿富汗等。此外，如果考虑俄罗斯、哈萨克斯坦等资源国，中国汽油零售价较周边国家属中等偏低水平。</p><p>【摘要】：中国经济网北京9月17日讯 “中国是一个高油价国家”的说法充斥着舆论和网络，最新出版的《求是》杂志对这个说法予以“驳斥”。《红旗文稿》杂志总编辑助理吴强在题为《关于垄断行业改革的几个问题》的文章中引用2012年5月美国彭博新闻社列出的一份数据，全球55个国家的油价排名中中国的油价排名第42位，是油价较低的国家。中国是人均能源极度匮乏的国家，维持较低的油价相当不容易。对于“油价低”的原因，文章认为，“很难想象，如果我国的三大石油公司都是私企，还能保持目前较低的油价吗？”</p><p>【其他】：

< 返回(B)

下一步N >

取消

如果结果正确，点击下一步。

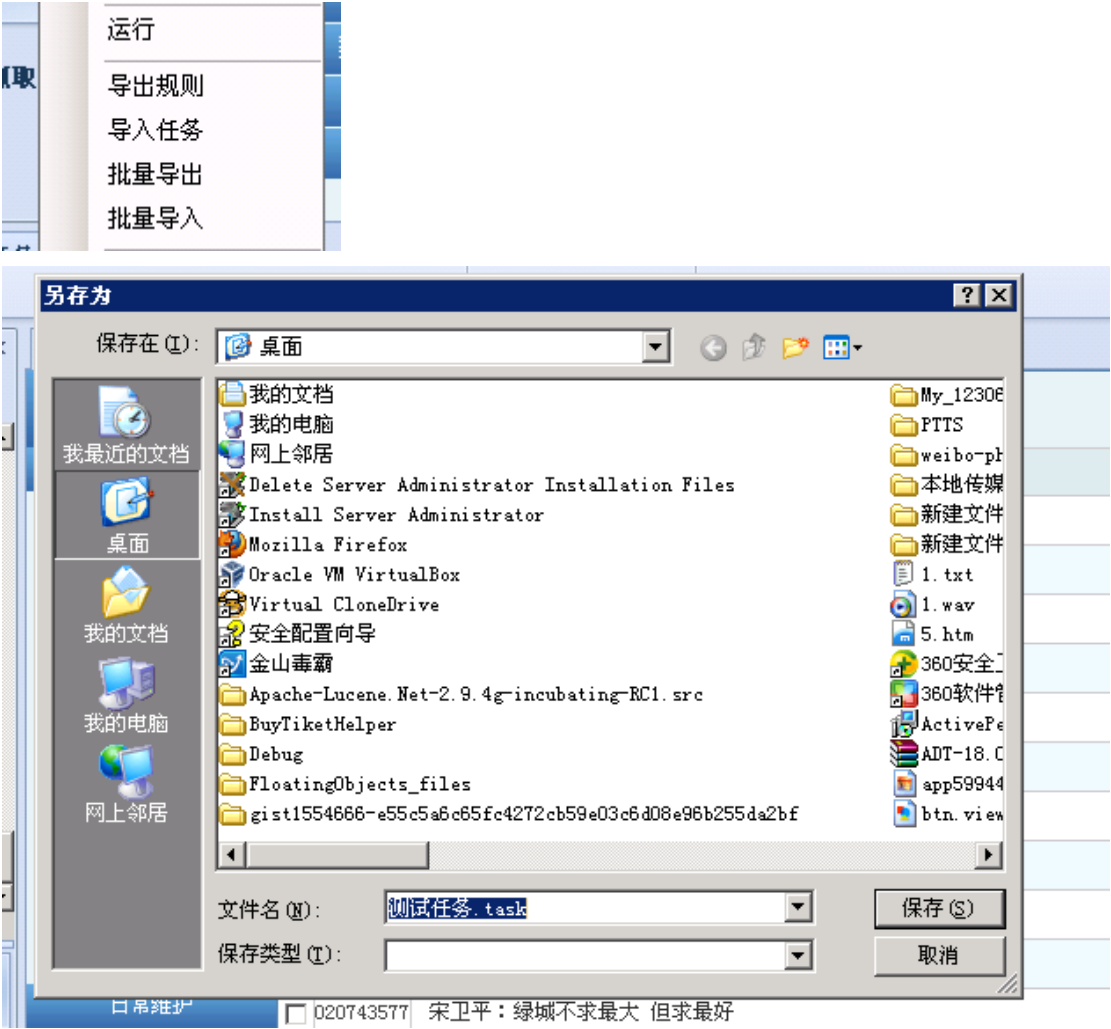


点击完成。

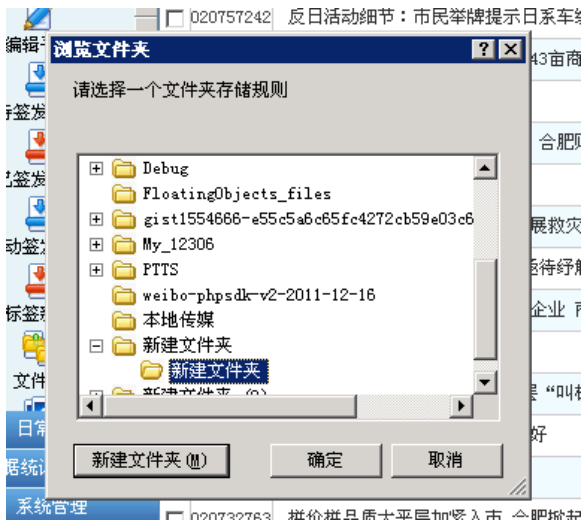
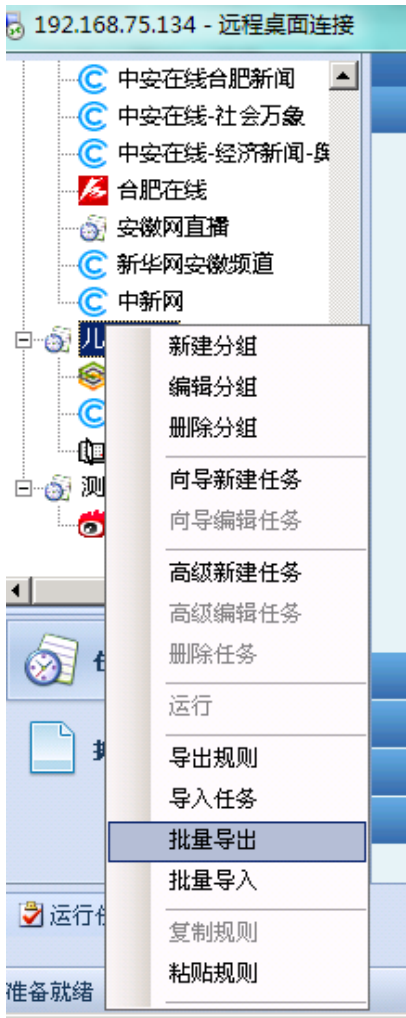


(二) 导入导出

1. 导入导出任务

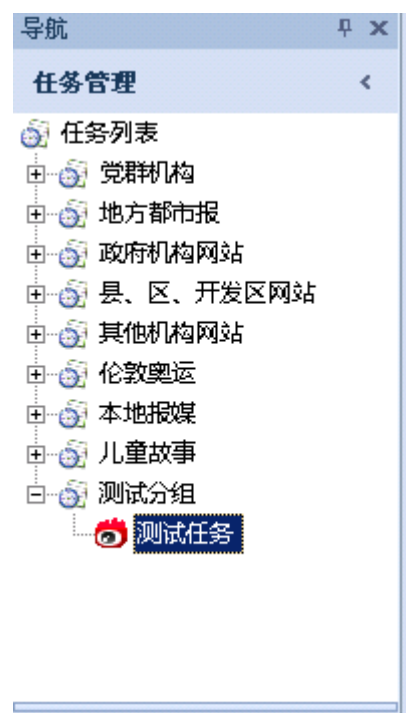


2. 批量导入导出任务（按分组）



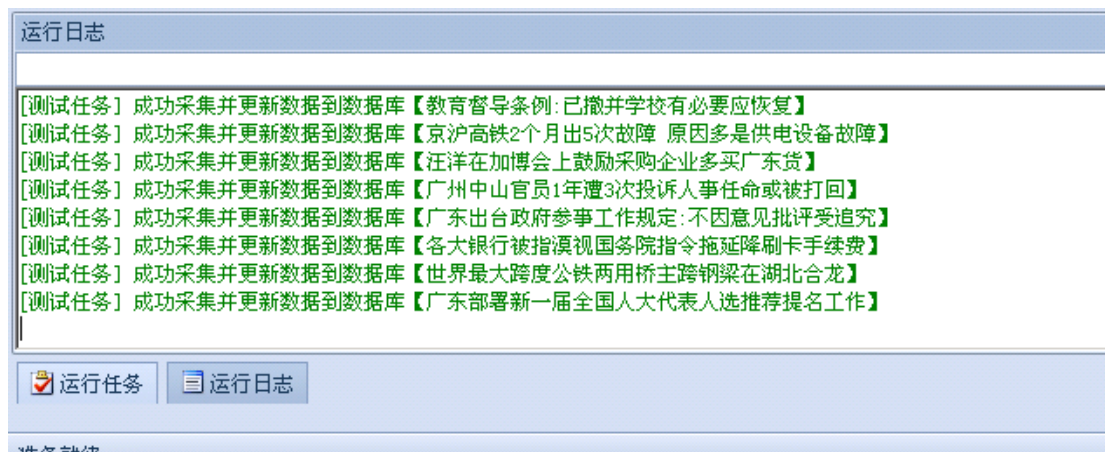
(三) 采集

1. 选择要采集的任务



2. 右键菜单，选择运行

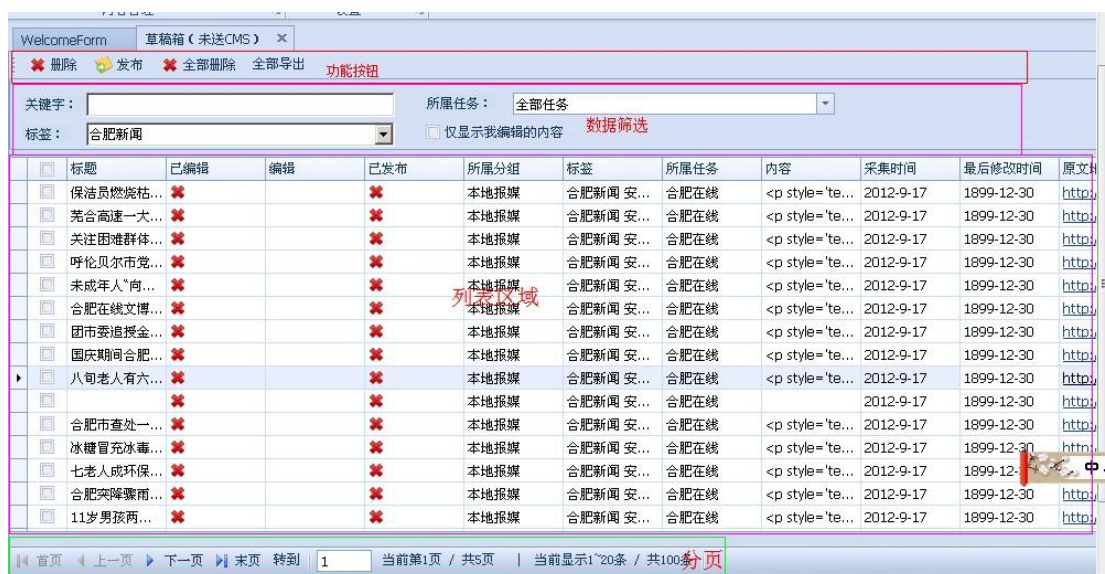




二、 编辑与发布

(一) 草稿箱

草稿箱是指已采集，但是没有发布的内容



从上到下分为功能操作，筛选，列表和分页四个部分

1. 筛选新闻

你可以从多个维度筛选新闻

1) 根据任务筛选

所属任务： 全部任务

☐ 仅显示我编辑

	任务名称	任务编号
	全部任务	0
所属分类	中新网	121
	一千零一夜	123
本地报	安徒生童话	124
本地报	十万个为什么	122
本地报	安徽网直播	119
本地报	合肥在线	118
本地报		
本地报		

2) 根据任务标签筛选

关键字：

标签： 合肥新闻

<input type="checkbox"/>	全部
<input type="checkbox"/>	合肥新闻
<input type="checkbox"/>	安徽新闻
<input type="checkbox"/>	国际国内新闻
<input type="checkbox"/>	儿童
<input type="checkbox"/>	童话故事
<input type="checkbox"/>	童话
<input type="checkbox"/>	呼伦贝尔市党...

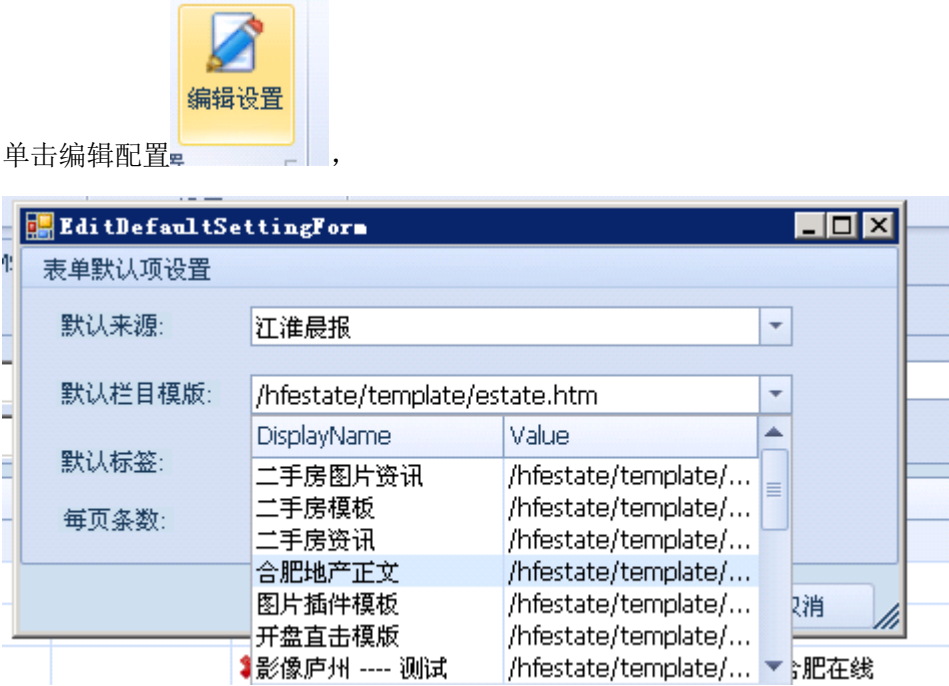
3) 根据关键字筛选

关键字：

4) 根据编辑筛选

☐ 仅显示我编辑的内容

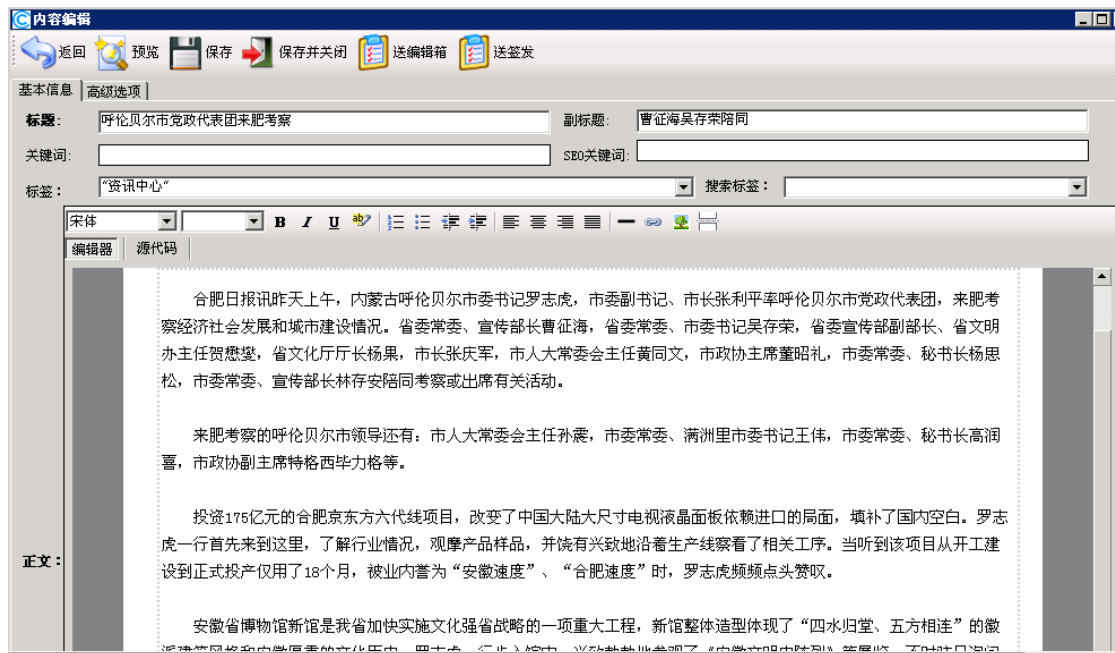
2. 编辑配置



配置默认标签、默认模版等

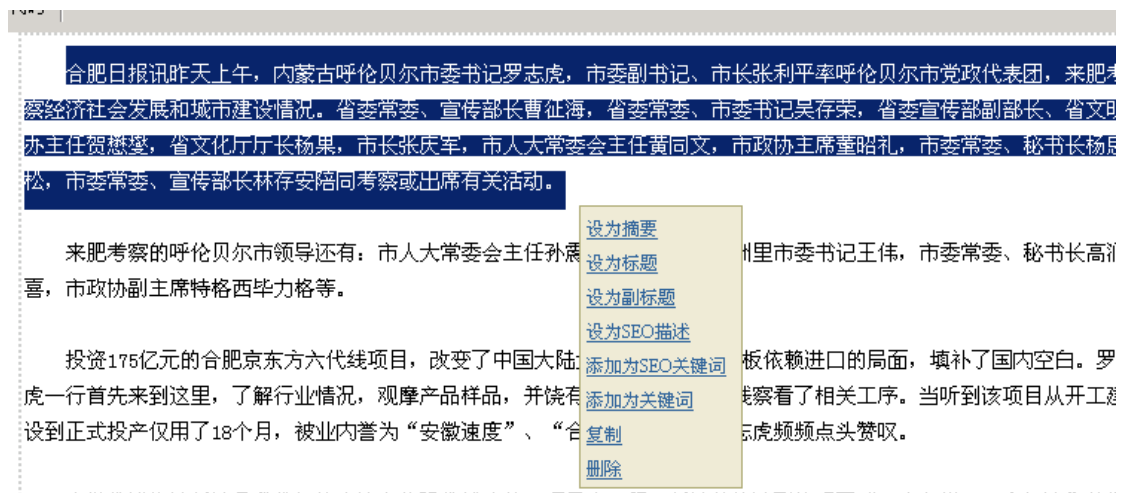
3. 编辑新闻

双击一篇新闻即可打开编辑界面，不过在编辑新闻之前，请做好配置



1) 快捷菜单

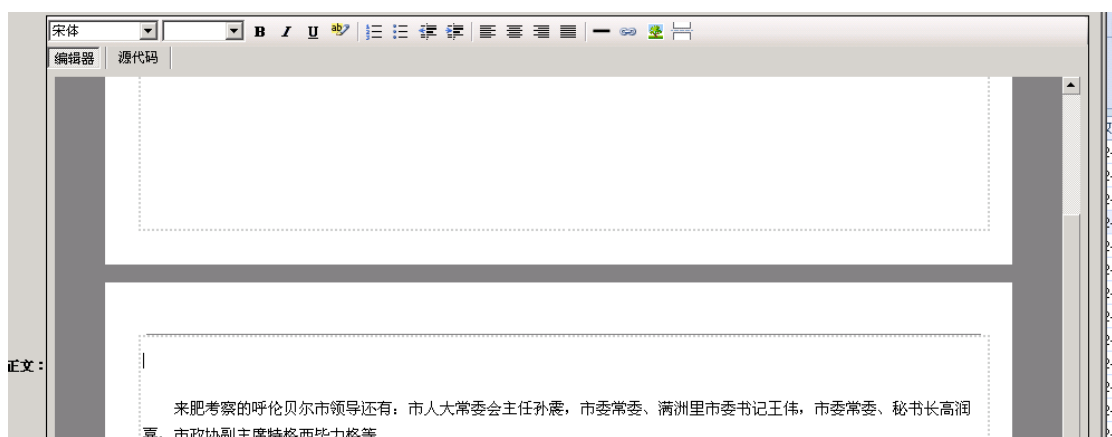
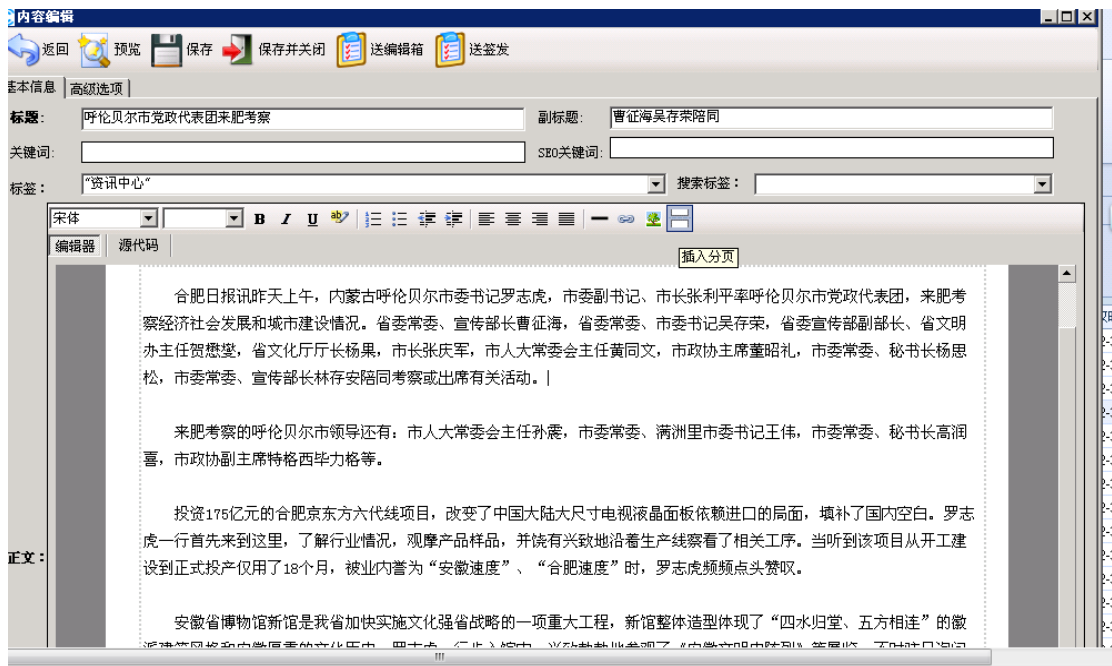
选中一段文字，会自动出现快捷菜单



可以设置摘要等选项，也可以复制删除。

2) 分页

选择一个插入分页的位置，点击分页按钮，即可插入分页



3) 分页标题

在分页的顶部直接输入即可



4) 插入图片





许多网友对这种“以爱为名”的暴力行为予以谴责。网友“田田”说：“爱国不是非要砸东西来体现的呀，更何况砸的还是made in china自己生产的made in china，都是中国工人生产出来的。”知名演员徐静蕾也提醒：“有多少家庭的内讧是由爱的名义开始的，有多少战争是以正义的名义发动的？这已经被历史证明了无数次。全天下认为自己有权利打别人，就意味意味着有人认为自己有权利打你。我们都有家人朋友至亲的人，搞不好就搞到自己身上。”

5) 送签发与送编辑箱

送签发是送往 CMS 的待签发箱，送编辑箱是送往 CMS 的编辑箱

送签发的结果：



送编辑箱的结果

