方正当代报纸库高级分析工具

操作手册

目 录

—,		编写目	的	3
二、		阅读对	象	3
三、		操作说	明	3
	(—)	上传	数据	3
	(<u></u>)	数据	清洗	4
	1.	数据编辑		4
	2.	列裁减		5
	3.	批量替换		6
	4.	字段合并		6
	5.	数据筛选		7
	6.	异常值处理		8
	7.	字段类型转换		9
	(三)	三) 自然语言处理		9
	1.	词云分析		9
	2.	实体云		11
		2.1	人物云	11
		2.2	地区云	13
		2.3	机构云	15
	3.	关系图谱		17
		3.1	人物-人物的关系	17
		3.2	人物-机构的关系	20
		3.3	人物-地区的关系	22

一、编写目的

本文档编写目的,对方正当代报纸库高级分析工具进行操作说明描述,为使用该工具用户提供查阅指导。

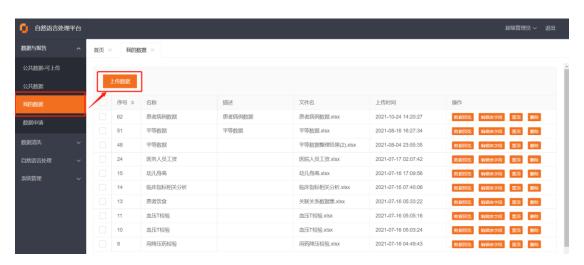
二、阅读对象

使用该工具所有用户

三、 操作说明

(一) 上传数据

- 1) 准备好要上传的数据表。
- 2) 点击【数据与报告】—【我的数据】菜单,点击上传数据按钮。



弹出添加数据窗口,选择(1)上传准备好的数据,点击(2)进行提交,数居上传完成。 在【我的数据】列表里进行查看。

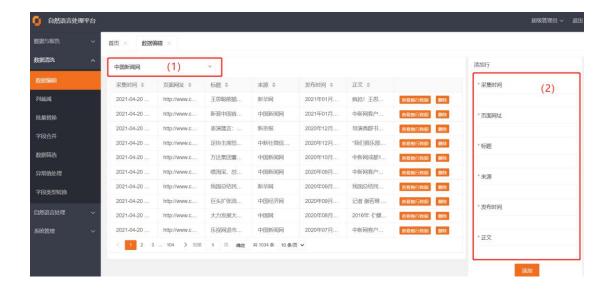


(二) 数据清洗

1. 数据编辑

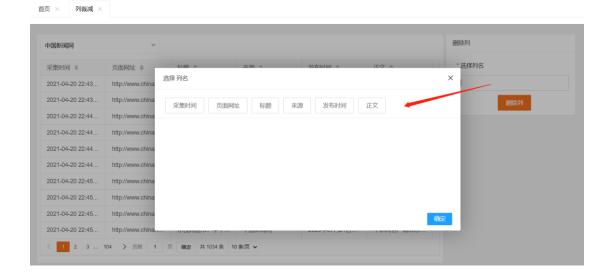
- 1) 数据编辑操作为数据集进行新增单条数据。
- 2) 点击菜单项【数据清洗】——【数据编辑】进入数据列表,对选中的数据集进行单条数据的填加操作。
- 3) 在数据编辑页面选择要编辑的数据集,点击(1)下拉项进行选择。

在列表右侧(2)区域进行数据的添加,带*号的为必填项。数据填写完后点击【添加】操作,数据填加成功。在列表最后一条进行查看。



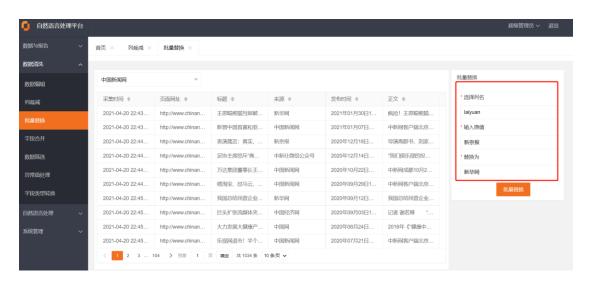
2. 列裁减

- 1) 列裁减为数据集进行删除某列操作。
- 2) 点击菜单项【数据清洗】——【列裁减】进入数据列表,对选中的数据集进行删除列操作。
- 在数据编辑页面选择要删除的列名,弹出层进行选择,点击【确定】数据中该列删除成功。



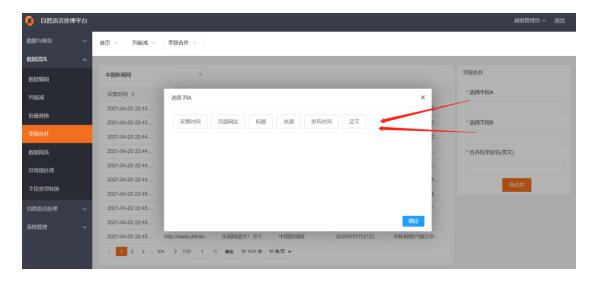
3. 批量替换

- 1) 批量替换操作为数据集中选择的某一列数据进行新数据替换旧数据的操作。
- 2) 点击菜单项【数据清洗】——【批量替换】进入数据列表,在右侧批量替换操作区进行 列项的选择,原值和替换值的输入。
- 3) 点击【批量替换】该列的数据替换成功。

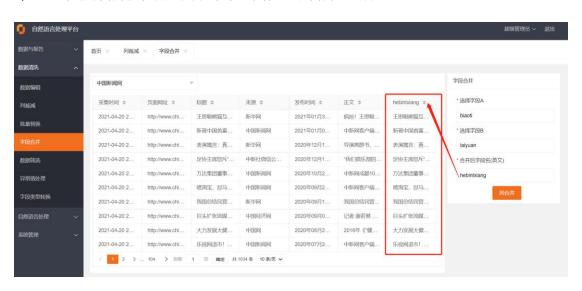


4. 字段合并

- 1) 字段合并操作为数据集中选择的两个字段值合并成一个字段值。
- 2) 点击菜单项【数据清洗】——【字段合并】进入数据列表,在右侧字段合并操作区进行要合并列项的选择,输入合并后字段的名称



3) 点击【列合并】在数据列表中最后面将显示合并后的数据。



5. 数据筛选

- 1) 数据筛选可以将某个字段数据进行符合条件的筛选,可以列表中直接展示筛选后的结果 , 也可以将筛选后的结果数据另存为我的数据中。
- 2) 点击菜单项【数据清洗】——【数据筛选】进入数据列表,在右侧选择要筛选的列,输入相应的条件,输入筛选值。
- 3) 点击【筛选】左侧列表数据显示符合筛选条件的数据。

4) 点击【数据另存为】将符合条件的数据保存到我的数据里。



6. 异常值处理

- 1) 异常值处理可对数据里异常的内容进行查询和删除操作。
- 点击菜单项【数据清洗】——【异常值处理】进入数据列表,选择数据集后,在右侧对 异常值进行条件选择。



- 3) 点击【查询】左侧数据显示查询后的数据。
- 4) 点击【删除异常值】将批量删除符合条件的数据。

7. 字段类型转换

- 1) 字段类型转换是对数据里字段类型的转换操作。字符串和数字类型进行转换。
- 2) 点击菜单项【数据清洗】——【字段类型转换】进入字段列表,在要转换的字段类型里进行点击,即可以字符串和数字类型进行转换。

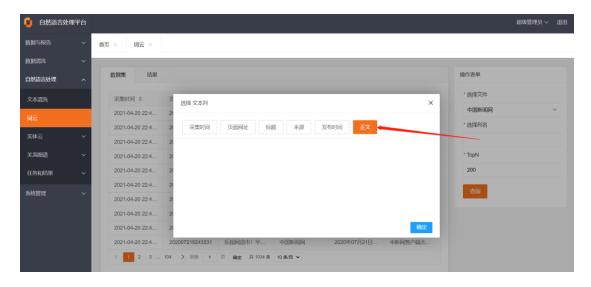


(三) 自然语言处理

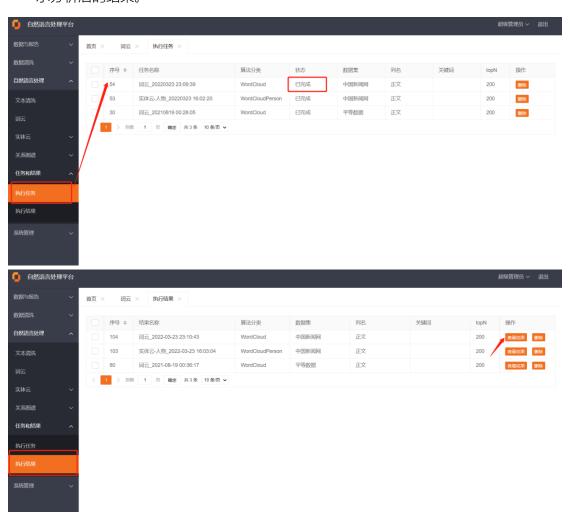
1. 词云分析

功能说明:词云是对选择字段里的内容进行词频统计,按照关键词出现的次数进行排序,出现越多的"关键词"字体显示越大。

1) 点击菜单项【自然语言处理】——【词云】进入数据集列表,在右侧需要选择分析的数据集,选择列名,点击【查询】,显示分析结果。



- 2) 在【任务和结果】——【执行任务】查看执行状态。
- 3) 状态为【已完成】到任务和结果菜单下,点击执行结果菜单 ,点击【查看结果】,将显示分析后的结果。



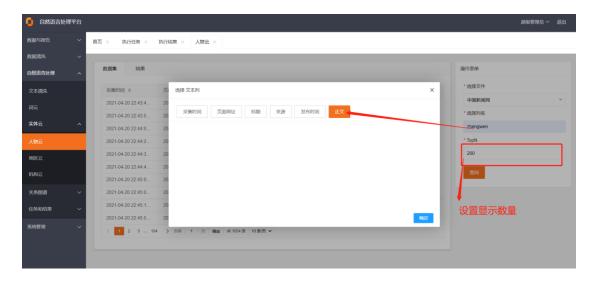


2. 实体云

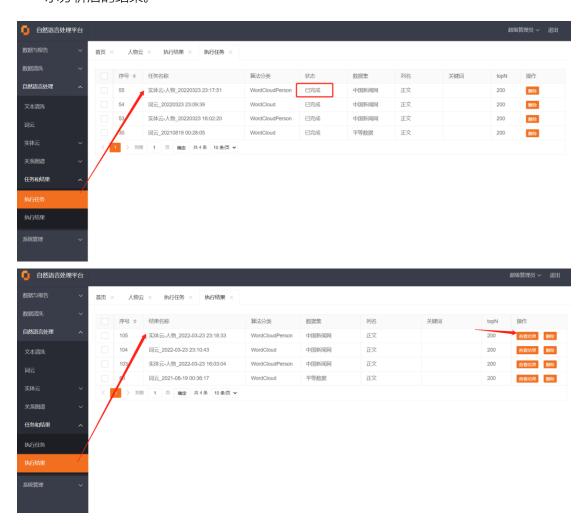
2.1 人物云

功能说明:人物云对选择字段里的内容进行【人物姓名】词频统计,根据设置的显示数量,按出现的词频数量由高到低进行展示,出现次数最多的人名字体显示越大。鼠标移入姓名时将显示出现的次数。所谓词频是对全部文章进行统计。

1) 点击菜单项【自然语言处理】—【实体云】—【人物云】进入数据集列表,在右侧需要选择分析的数据集,选择列名,设置要显示的数量。点击【查询】,显示分析结果。



- 2) 在【任务和结果】——【执行任务】查看执行状态。
- 3) 状态为【已完成】到任务和结果菜单下,点击执行结果菜单 ,点击【查看结果】,将显示分析后的结果。

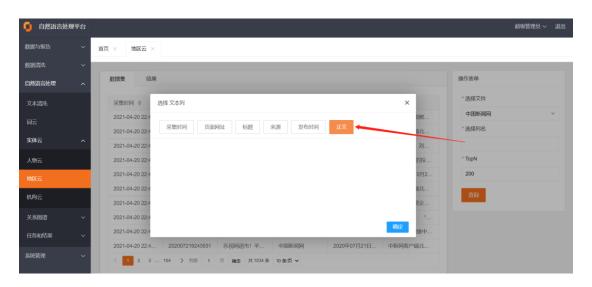




2.2 地区云

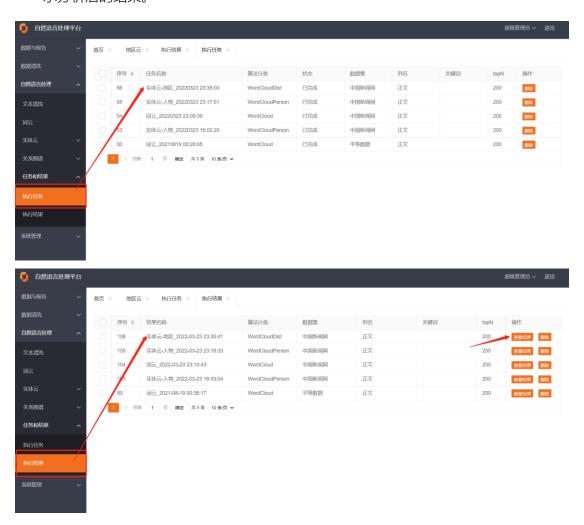
功能说明: 地区云对选择字段里的内容进行【地区名称】词频统计,根据设置的显示数量,按出现的词频数量由高到低进行展示,出现次数最多的地区名称字体显示越大。鼠标移入地区名称时将显示出现的次数。所谓词频是对全部文章进行统计。

1) 点击菜单项【自然语言处理】—【实体云】—【地区云】进入数据集列表,在右侧需要选择分析的数据集,选择列名,点击【查询】,显示分析结果。



2) 在【任务和结果】——【执行任务】查看执行状态。

3) 状态为【已完成】到任务和结果菜单下,点击执行结果菜单,点击【查看结果】,将显示分析后的结果。

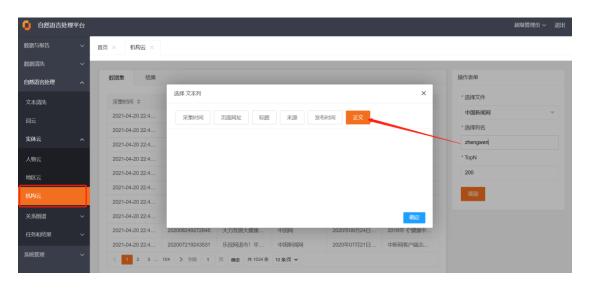




2.3 机构云

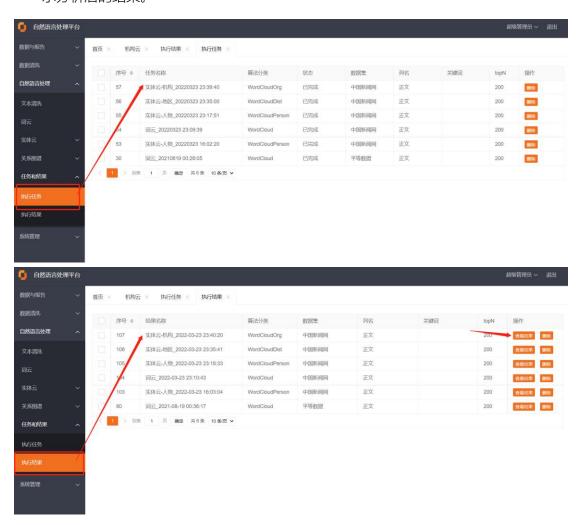
功能说明:机构云对选择字段里的内容进行【机构名称】词频统计,根据设置的显示数量,按出现的词频数量由高到低进行展示,出现次数最多的机构名称字体显示越大。鼠标移入机构名称时将显示出现的次数。所谓词频是对全部文章进行统计。

1) 点击菜单项【自然语言处理】—【实体云】—【机构云】进入数据集列表,在右侧需要选择分析的数据集,选择列名,点击【查询】,显示分析结果。



2) 在【任务和结果】——【执行任务】查看执行状态。

3) 状态为【已完成】到任务和结果菜单下,点击执行结果菜单,点击【查看结果】,将显示分析后的结果。





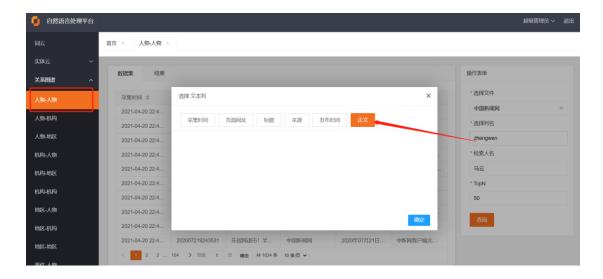
3. 关系图谱

3.1 人物-人物的关系

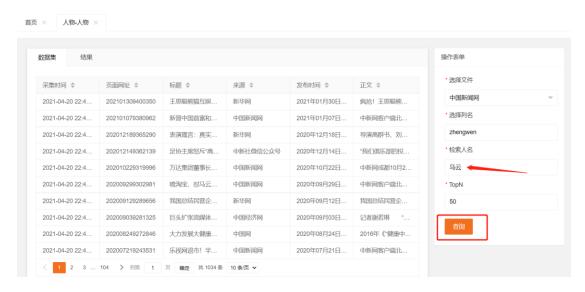
功能说明:根据检索人名分析出与之相关联人物关系图谱。根据设置的显示数量,将与检索人物同时出现的人物姓名按出现的词频数量由高到低进行展示,同时显示关联次数。

所谓关联是在同一篇文章中同时出现。

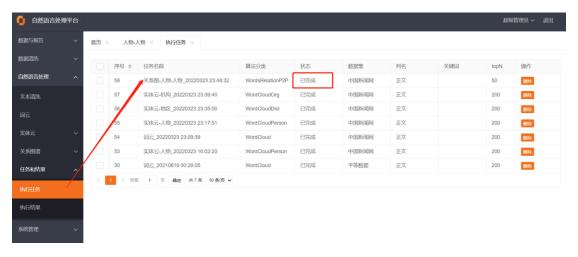
1) 点击菜单项【自然语言处理】—【关系图谱】—【人物-人物】进入数据集列表,在右侧需要选择分析的数据集,选择列名,点击【确定】。

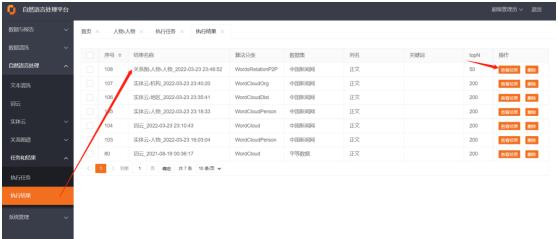


2) 输入要检索的人名,选择显示的数据数量,点击【查询】。



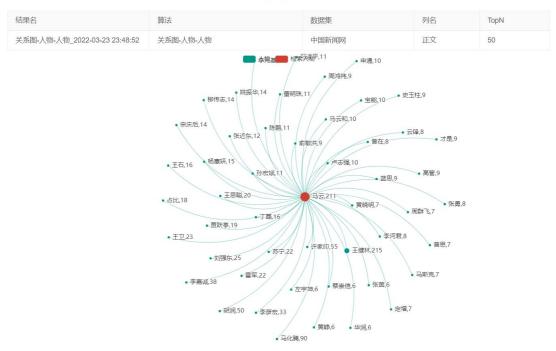
- 3) 在【任务和结果】——【执行任务】查看执行状态。
- 4) 状态为【已完成】到任务和结果菜单下,点击【执行结果】菜单 ,点击【查看结果】,将显示分析后的结果。





5) 查看结果

关系图谱结果

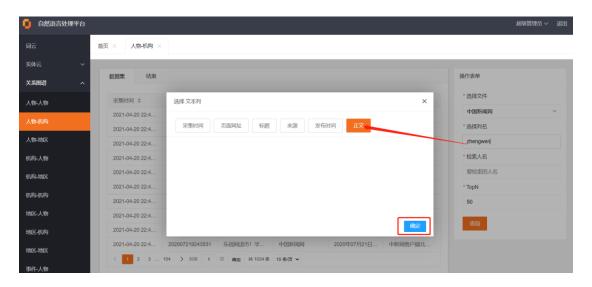


3.2 人物-机构的关系

功能说明:根据检索人名分析出与之相关联机构关系图谱。根据设置的显示数量,将与检索人物同时出现的机构名称按出现的词频数量由高到低进行展示,同时显示关联次数。

所谓关联是在同一篇文章中同时出现。

1) 点击菜单项【自然语言处理】—【关系图谱】—【人物-人物】进入数据集列表,在右侧需要选择分析的数据集,选择列名,点击【确定】。

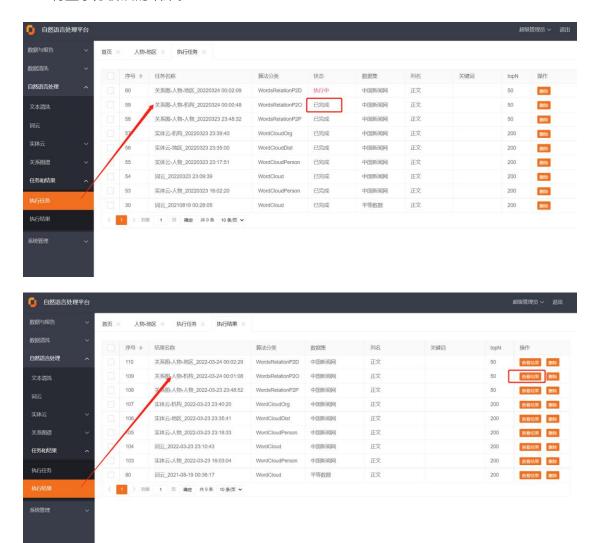


2) 输入要检索的人名,选择显示的数据数量,点击【查询】。



3) 在【任务和结果】——【执行任务】查看执行状态。

4) 状态为【已完成】到任务和结果菜单下,点击【执行结果】菜单 ,点击【查看结果】,将显示分析后的结果。



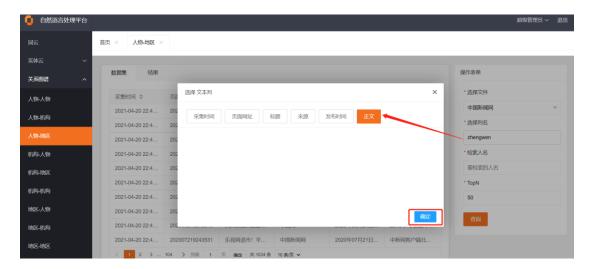


3.3 人物-地区的关系

功能说明:根据检索人名分析出与之相关联地区关系图谱。根据设置的显示数量,将与检索人物同时出现的地区名称按出现的词频数量由高到低进行展示,同时显示关联次数。

所谓关联是在同一篇文章中同时出现。

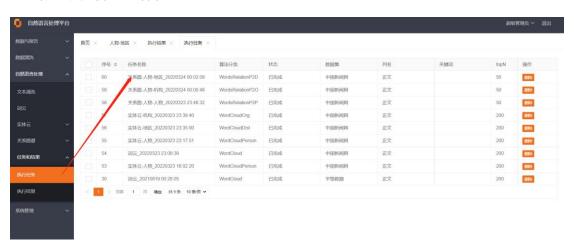
1) 点击菜单项【自然语言处理】—【关系图谱】—【人物-人物】进入数据集列表,在右侧需要选择分析的数据集,选择列名,点击【确定】。

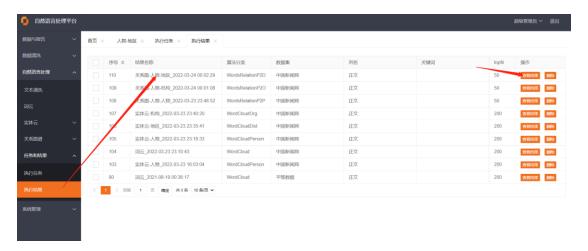


2) 输入要检索的人名,选择显示的数据数量,点击【查询】。



- 3) 在【任务和结果】——【执行任务】查看执行状态。
- 4) 状态为【已完成】到任务和结果菜单下,点击【执行结果】菜单 ,点击【查看结果】,将显示分析后的结果。





5) 查看结果

关系图谱结果

