企查查和数据中心舆情对比

## 舆情主要需求：

舆情信息作为新华信用平台企业信息的重要部分之一，需要相对较为精准的定位到企业主题。

## 对比分析

目前系统使用企查查的舆情接口，实时从企查查获取企业新闻信息。数据公司也有舆情接口。我们从接口内容、相关性、敏感信息过滤、消息总量分别对比两家公司的舆情接口。

#### 接口内容

如下表格，数据公司的舆情信息维度更加丰富，而且具有词条舆情的正负面的分析。

|  |  |
| --- | --- |
| 企查查舆情 | 数据公司舆情 |
| **标题** | **标题** |
| **图片url** | **图片url** |
| **新闻标题** | **新闻标题** |
| **新闻来源** | **新闻来源** |
| **新闻发布时间** | **新闻发布时间** |
| **新闻描述** | **新闻描述** |
| **新闻url** | **新闻url** |
|  | **新闻正负面分析（1：正面、-1负面、0中性）** |
|  | **境内外类型（是否是境外）** |
|  | **数据标红关键词，包含搜索词、地区词、事件词** |
|  | **发表人呢称** |
|  | **发表楼层，0表示主贴** |
|  | **回复数（新闻回复数量）** |
|  | **事件标签** |

#### 相关性

数据公司的舆情接口可以设置精准搜索，和非精准搜索。测试发现精准搜索的结果里面都是包含完整企业名称的，准确定位到企业主体。

企查查提供的舆情接口也可以定位到企业主体。而且每条新闻和企业的关联性更强。

#### 敏感信息过滤

均无过滤机制

#### 消息总量

数据公司的舆情接口数量更大，比如搜索中国经济信息社，数据公司返回71条舆情结果，企查查返回7条舆情结果。

## 存在的问题

1. 数据公司提供的舆情，还需要更精确到企业才可以使用。目前返回的新闻条数过多，大多不是在讲这个企业本身，相关性不够强。
2. 增加敏感词汇过滤机制，建立敏感词库，展示之前过滤一遍。这个难度比较大，自建敏感词库需要大量的人工干预和长时间的积累。建议和市场上相关领域较为成熟的公司合作。
3. 精准匹配剔除了大部分非相关公司，但是也有部分相关的公司被去除。建议形成企业简称的库，可以更精准的获得企业新闻。