# 高校舆情分析系统开发任务书

## 任务概述

面向微博等信息平台实时产生的网络舆情数据，利用网络爬虫模拟完成微博高级搜索引擎的数据检索和采集，对爬取到的博文文本清洗之后进行文本情绪量化处理，并对分析结果进行合理的可视化展示。（允许学生采取自己的技术路线，实现高校舆情分析、发现功能）

## 二、任务要求

1. 设计、实现高效的网络爬虫，完成微博高级搜索引擎的数据检索和采集；
2. 对爬取到的博文文本进行必要的清洗；
3. 对清洗后的博文文本进行情绪量化分析、处理；
4. 对高校舆情分析的结果进行直观、合理的可视化展示。

### 实现要求

1. 开发语言、爬虫框架、数据库等不限，学生自选；
2. 可以面向特定的高校进行舆情发现和分析，也可以针对高校异地校区设置、线上教学、新冠疫情等特定的热点问题进行深入的分析；
3. 在分析视角、理论方法、实现技术等方面有创新的舆情分析系统，成绩可以评定为优秀。

## 课程设计报告提交说明

1．使用统一的报告模板，见附件1；

2．报告应阐述系统架构、系统功能、主要功能模块的设计和实现方法，展示重要的运行结果和系统界面；

3．字数不少于6000字（包括图表，大约6页）；

4．正文采用宋体小四号字体，单倍行距，报告用Word 排版，A4纸打印，左侧装订，打印1份；

5．9月30日前以班级为单位提交到研究院北424董开坤老师，同时提交课程报告电子版，电子版文件以学号-姓名命名。

6. 课程报告封面不要写报告具体内容。