

## 第 3 周周记

周一	
完成内容	上网搜索法律案例文档的实例
内容描述	找到部分单独案例文档，以及汇法网
未解决问题	但没可以找到大量下载的案例

周二	
完成内容	阅读《基于加权 TextRank 的新闻关键事件主题句提取》
内容描述	词语位置加权的 Text Rank 方法。利用自然语言处理(NLP)技术对新闻文本进行处理，首先对新闻正文进行分句，并过滤新闻正文中的停用词。利用 ICTCLAS 中文处理模块进行分词、词性标注和命名实体识别。
未解决问题	

周三	
完成内容	复习 TCP/IP
内容描述	
未解决问题	

周四	
完成内容	复习 TCP/IP，考试
内容描述	
未解决问题	

周五	
完成内容	阅读论文《Online News Event Extraction for Global Crisis Surveillance》
内容描述	
未解决问题	

周末	
完成内容	阅读论文《Online News Event Extraction for Global Crisis Surveillance》
内容描述	
未解决问题	

工程汇总	
完成任务	搜索法律案例来源，阅读论文，尝试阅读英文文献。
任务描述	

代码量	
未解决问题	

论文汇总	
论文列表	<p>[1] 《基于加权 TextRank 的新闻关键事件主题句提取》蒲梅，周枫，周晶晶，严馨，周兰江</p> <p>[2] Online News Event Extraction for Global Crisis Surveillance, Jakub Piskorski, Hristo Tanev, Martin Atkinson, Eric van der Goot, and Vanni Zavarella</p>
论文摘要	<p>[1] 为了在大量的新闻中快速找到自己感兴趣的内容，提出在单文档中基于加权 Text Rank 算法提取主题句的方法，以得到新闻关键事件信息。通过计算新闻文本句子关键词的互信息值，对新闻报道进行事件句和非事件句的分类，过滤出非事件句。基于 Text Rank 算法的思想，构建一个事件句有向图，引入句子位置、句子相似度和关键词覆盖频率 3 个影响因子，以此计算句子之间的影响权重，利用 Text Rank 模型对图中的每个点计算权重，并选取排序最靠前的句子作为关键事件的主题句。实验结果表明，该方法的抽取效果优于基于词频-逆文档概率和新闻标题的主题句抽取方法。</p> <p>[2] This article presents a real-time and multilingual news event extraction system developed at the Joint Research Centre of the European Commission. It is capable of accurately and efficiently extracting violent and natural disaster events from online news. In particular, a linguistically relatively lightweight approach is deployed, in which clustered news are heavily exploited at all stages of processing. Furthermore, the technique applied for event extraction assumes the inverted-pyramid style of writing news articles, i.e., the most important parts of the story are placed in the beginning and the least important facts are left toward the end. The article focuses on the system's architecture, real-time news clustering, geo-locating and geocoding clusters, event extraction grammar development, adapting the system to the processing of new languages, cluster-level information fusion, visual event tracking, event extraction accuracy evaluation, and detecting event reporting boundaries in news article streams.</p>
未解决问题	

下周任务	
工作	寻找案例来源，阅读论文
论文	
其他	
汇总	

日期:2018/01/08-2018/01/20