

《信息检索》大作业 Part2 实验报告

施韶韵 计34 2013011358

实验步骤

前端输入

增加了几个输入框，可以输入作者、出版单位和起止年份以便查询

建立索引

在part1中已对所有域建立索引，此处无需更改

建立查询

基本的Query使用 `MultiFieldQueryParser` 对多个 `题名`、`英文篇名`、`摘要`、`英文摘要` 进行查询

其他输入条件与之用 `BooleanQuery` 合并，条件使用 `MUST`

代码属于 `IRSearcher.java`

引入Word2Vec

用作业3中已训练好的 `.bin` 文件，使用网络上的代码载入

Word2Vec java: [GitHub Ansj Word2VEC java](#)

临近词查询

对于基本的Query字符串，先进行分词

然后对每个词找到与其最近的三个词，在原Query字符串中进行替换得到新的Query字符串

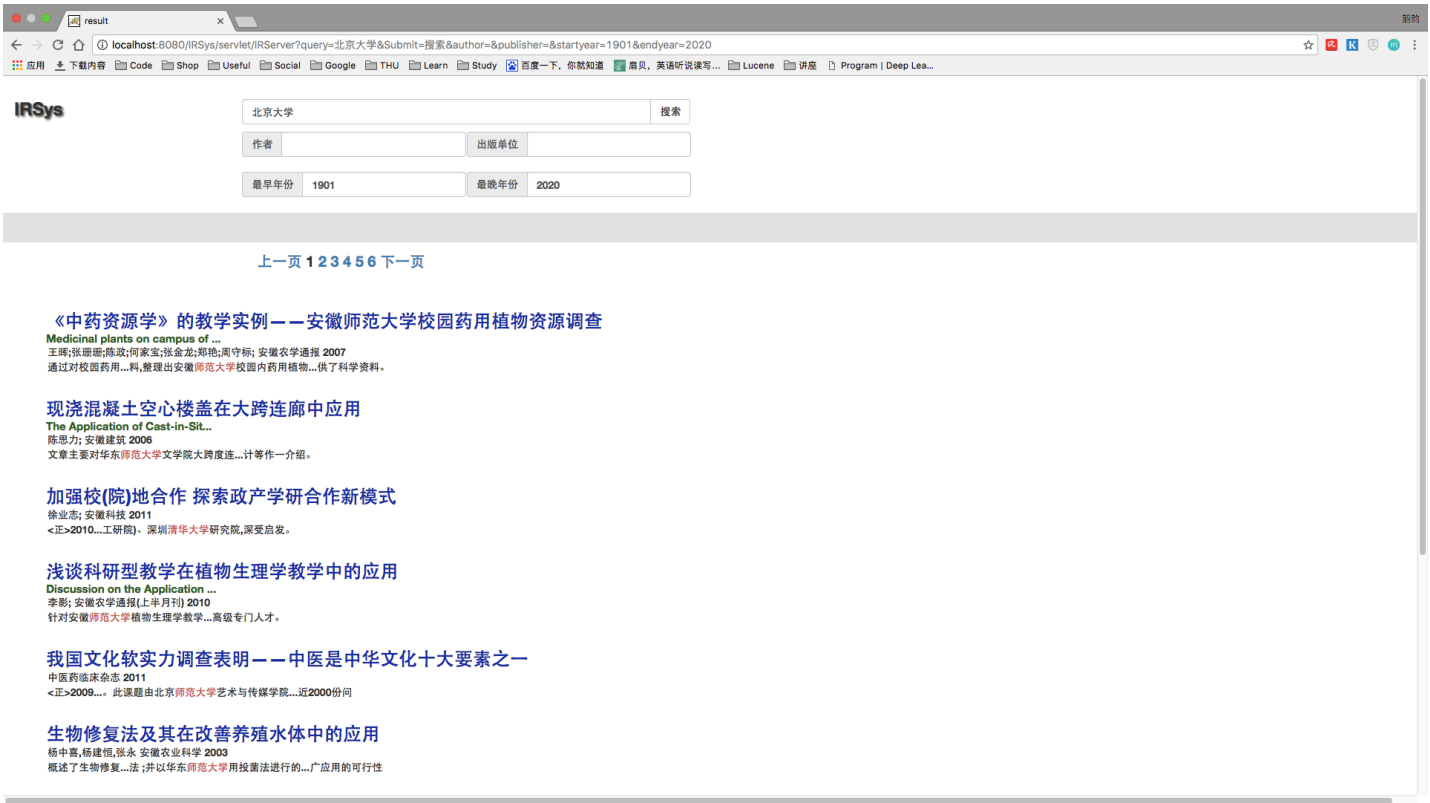
对每个Query字符串使用 `BooleanQuery` 结合，条件使用 `SHOULD`

基本的Query字符串boost为1，其他为0.5

再讲这个 `BooleanQuery` 与其他条件用 `BooleanQuery` 结合，条件使用 `MUST`

实例结果分析

北京大学



因为与北京大学最近的三个词是“复旦”、“师范大学”、“清华大学”。虽然原本“北京大学”搜索不到结果（标题和摘要中确实没有北京大学）。但是结果可以搜索到这些词。

北京大学 2010-2020



可见结果只剩2010年以后的。

北京大学 2010-2020 安徽农学通报

IRSys

北京大学

搜索

作者

出版单位

安徽农学通报

最早年份

2010

最晚年份

2020

上一页

1

2

3

4

5

6

下一页

浅谈科研型教学在植物生理学教学中的应用

Discussion on the Application ...

李影; 安徽农学通报(上半月刊) 2010

针对安徽师范大学植物生理学教学...高级专门人才。

适应学科发展需要 改革园林生态学教学

Reforming the Teaching of Land...

卫新菊; 安徽农学通报(上半月刊) 2010

园林生态学应用...进行改革。山西师范大学生命学院的园林...综合评定模式。

《植物生理学实验》教学中的思考

高轩; 于红梅; 安徽农学通报(上半月刊) 2012

植物生理学实验课是安徽师范大学园艺专业的基础...有了明显提高。

加强校(院)地合作 探索政产学研合作新模式

徐业忠; 安徽科技 2011

<正>2010...工研院)、深圳清华大学研究院,深受启发。

低碳大学城推进策略思考

Advancing Strategies Thinking ...

袁欣; 安徽建筑 2011

低碳城市是当前...区和同济大学、复旦大学实地调研...供有益的借鉴。

上一页

1

2

3

4

5

6

下一页

可见靠前结果只剩安徽农学通报上发表的。
之后结果是类似出版单位的。

北京大学 2010-2020 安徽农学通报 卫新菊

result

localhost:8080/IRSys/servlet/IRServer?query=北京大学&Submit=搜索&author=卫新菊&publisher=安徽农学通报&startyear=2010&endyear=2020

应用 下载内容 Code Shop Useful Social Google THU Learn Study 百度一下, 你就知道 南京, 英语听说读写... Lucene 讲座 Program | Deep Lea...

IRSys

北京大学

搜索

作者

卫新菊

出版单位

安徽农学通报

最早年份

2010

最晚年份

2020

上一页 1 2 3 4 5 6 下一页

适应学科发展需要 改革园林生态学教学

Reforming the Teaching of Land...
卫新菊; 安徽农学通报(上半月刊) 2010
园林生态学应用...进行改革。山西师范大学生命学院的园林...综合评定模式。

上一页 1 2 3 4 5 6 下一页

可见结果只剩卫新菊发表的。