**基于Lucene的分析与应用**

**配置管理总结报告**

小组成员：

刘宏宇

滕延林

顾泽鹏

杨帆

周晓懿

版本变更历史

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本 | 提交日期 | 主要编制人 | 审核人 | 版本说明 |
| 1.0 | 20160601 | 顾泽鹏 | 滕延林、杨帆、周晓懿、刘宏宇 | 配置管理报告v01 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

# 版本变更内容分析

## 需求规格说明书

**第一版变更：**

**变更日期：**2016-03-16至2016-03-20

**变更版本编号：**1

**变更请求：**对项目计划书的第一次修改

**变更范围：**明确项目计划书内容、修改排版问题

**评审意见：**

1、 数据的获取问题

接受

解决方案：通过爬虫爬取一定数量的非结构化数据（文本、图像或音频）构造测试数据集供测试使用，初定数据规模在1W左右

2、 上层程序的测试问题，

接受

解决方案：

衡量标准：检索时间和检索准确性（准确率、召回率）

具体实施：（1）部署原始lucene架构，从测试集中选取测试数据，在原始架构上使用测试数据作为query，将检索结果以及运行时间作为groundtruth。（2）通过运行改进后的程序，使用相同的测试数据作为query，将检索结果以及运行时间与groundtruth进行对比从而获得改进效果。

**变更的详细描述：**

* 明确编写该文档的目的
* 丰富了术语与缩略语
* 减少了部分遵循的标准
* 修改了项目分工中的用词
* 修改了交付产品的名称，使之更加标准化
* 添加了待定问题与解决方案
* 统一了文字排版，图表的标号

**因果分析：**

第一版的项目计划书的目的不够明确，使得小组成员对于后续工作不明确，所以加以修改；术语与缩略语较少，不能涵盖所有项目中所涉及的概念；许多的标准之间互相矛盾，所以减少了标准数量；将“阅读代码”改为“研读代码”；根据老师的PPT中内容作出相应修改；提出问题，使目标更加明确；统一排版。

**第二版变更**：

**变更日期：**2016-03-16至2016-03-20

**变更版本编号：**2

**变更请求：**对软件需求文档的第二次修改

**变更范围：**对软件需求文档中用例图修改，文档结构调整

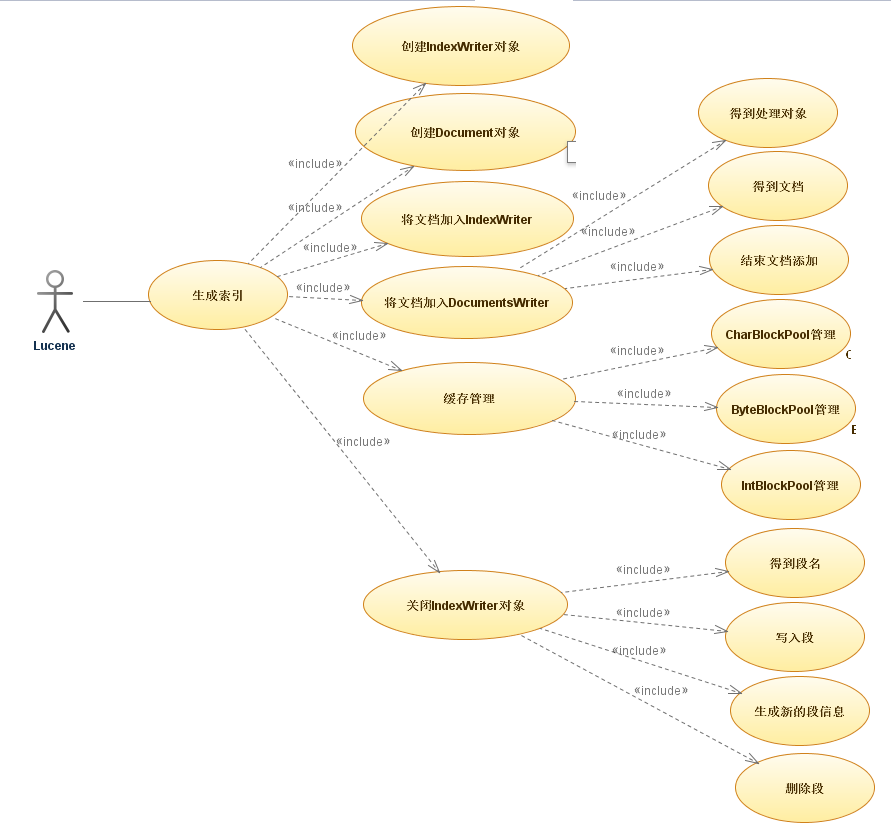
**评审意见**：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 评审对象 | 评审意见 | 处理建议 | 小组意见 |
| 1 | 《基于 Lucene 的分析与应用需求规格说明书》 | 上层应用的检索效率问题是如何考量的 | 我们已经在第一次的报告中回答了这些问题，我们会将具体的细节呈现在文档中 | 接受 |
| 2 | 《基于 Lucene 的分析与应用需求规格说明书》 | 文档中存在一些语言上的小错误 | 我们将改正错误并将改好的文档提交 | 接受 |
| 3 | 《基于 Lucene 的分析与应用需求规格说明书》 | 用例图中的Actor的概念没有理解正确 | 我们会重新修改并将改好的文档提交 | 接受 |

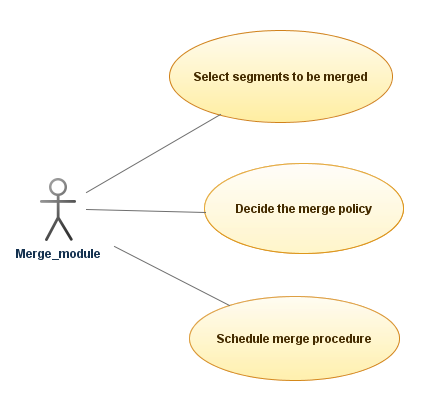
**变更的详细描述：**

* 对于功能需求的修改
* 将非功能需求单独作为一个章节
* 用例图修改

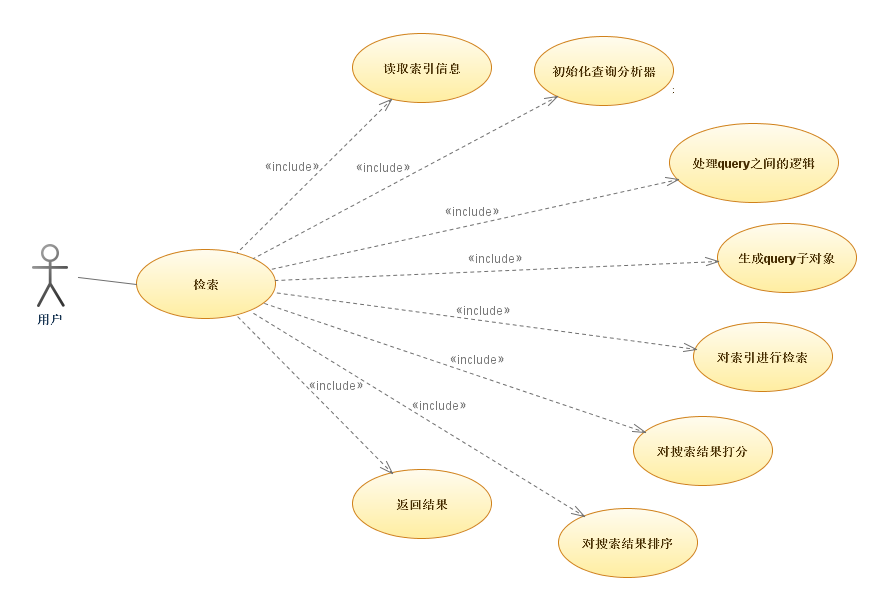
**生成索引用例图**



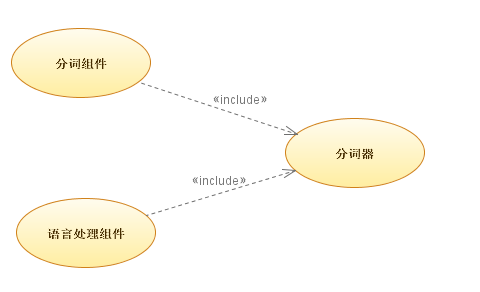
**段合并用例图**



**检索用例图**



**分词器用例图**



**因果分析：**

功能模块不明确，部分功能相互牵连；考虑到非功能需求为性能上要求，属于提高需求，并且不能以用例图的形式表示；对于部分用例图做出了修改，更加准确描述系统功能。

**第三版变更：**

**变更日期：**2016-03-20至2016-04-03

**变更版本编号：**3

**变更请求：**对软件需求文档的第三次修改

**变更范围：**修改软件需求文档的用例图，添加数据需求以及运行需求。

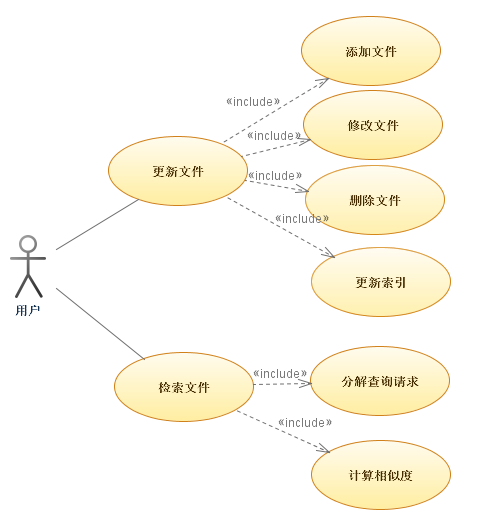
**评审意见**：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 评审对象 | 评审意见 | 处理建议 | 小组意见 |
| 1 | 《基于 Lucene 的分析与应用需求规格说明书V2.0》 | 需求的主要目的是作为设计开发的依据，因此目的部分需要调整。 | 我们将更正问题并提交。 | 接受 |
| 2 | 《基于 Lucene 的分析与应用需求规格说明书V2.0》 | 对于“将业务需求分解为功能性需求”： 需求和功能应是两个概念，要注意区分业务需求和功能需求的区别。“分解”的说法不确切，应是“转化”的关系。 | 我们将更正问题并提交。 | 接受 |
| 3 | 《基于 Lucene 的分析与应用需求规格说明书V2.0》 | 在用例图中，系统是否应该作为Actor？因为功能不是用例图应表达的事。 | 我们将对Lucene用例图中Actor的定位重新考量。 | 接受 |
| 4 | 《基于 Lucene 的分析与应用需求规格说明书V2.0》 | 用例图中的用例和RUCM的用例是有关系的。部分用例包含和模板里的包含关系对应不上，追踪性有问题。 | 我们将校验并更正用例图和RUCM中的用例，完善可追踪性。 | 接受 |
| 5 | 《基于 Lucene 的分析与应用需求规格说明书V2.0》 | 数据需求不应是测试需求，而应是开发所考虑的需求。（例如考虑文本特点、是否有图片等等） | 我们将更正问题并提交。 | 接受 |
| 6 | 《基于 Lucene 的分析与应用需求规格说明书V2.0》 | 解释自主构造一词。 | 我们将完善该部分。 | 接受 |

**变更的详细描述：**

* 大幅度调整用例图部分
* 添加数据需求
* 添加运行需求

**用例图**



**因果分析：**

原文档中的用例图有比较大的问题，例如“分词器”模块对于用户不可见，并不是用户的行为，不能以用例图来表示。另外在原用例图中出现了“系统”这样Actor，自己调用自己显然不符合逻辑，最终只保留一个用例图；该系统处理文件系统，对于数据有一定需求；该系统的运行环境的要求，包括软件环境与硬件环境。

**第四版变更：**

**变更日期：**2016-04-09至2016-04-09

**变更版本编号：**4

**变更请求：**对软件需求文档的第四次修改

**变更范围：**修改RUCM图、修改排版问题

**变更的详细描述：**

* 修改RUCM图，使之与用例图匹配
* 统一了文字排版，图表的标号

**第五版变更：**

**变更日期：**2016-04-03至2016-05-04

**变更版本编号：**5

**变更请求：**对软件需求文档的第五次修改

**变更范围：**根据项目实现计划更改对应的需求

**变更的详细描述：**

|  |  |  |
| --- | --- | --- |
| 变更原因 | 变更位置 | 变更内容概述 |
| 内因：  根据项目具体实现补充需求 | 第六章 | 增加了实现部分的说明：中文分词 |
| 内因：  根据项目具体实现补充需求 | 第六章 | 增加了实现部分的说明：图像特征选择 |
| 内因：  根据项目具体实现补充需求 | 第六章 | 增加了实现部分的说明；图像特征提取 |
| 内因：  根据项目具体实现补充需求 | 第六章 | 增加了实现部分的说明：数据预处理 |
| 内因：  根据项目具体实现补充需求 | 第六章 | 增加了实现部分的说明：距离计算公式的选择 |
| 内因：  根据项目具体实现补充需求 | 第六章 | 增加了实现部分的说明：网站搭建 |
| 内因：  根据项目具体实现补充需求 | 第六章 | 增加了实现部分的说明：Hash排序 |

**第六版变更：**

**变更日期：**2016-05-04至2016-05-09

**变更版本编号：**6

**变更请求：**对软件需求文档的第六次修改

**变更范围：**完善需求，补充基本需求及实现需求，并将已实现的部分落实到文档中

**变更的详细描述：**

|  |  |  |
| --- | --- | --- |
| 变更原因 | 变更位置 | 变更内容概述 |
| 外因（根据老师点评意见）：  需求应该范围两类：基本需求以及实现需求 | 第三章 | 在第3章添加对基本需求实现的说明 |
| 外因（根据老师点评意见）：  数据需求部分进行修改、添加场景需求说明 | 4.4、4.5 | 数据需求部分进行修改、添加场景需求说明 |
| 外因（根据老师点评意见）：  已经实现的内容要与需求文档中的需求结合起来 | 第六章 | 第6章修改为实现需求和实现方案，并添加部分实现需求方面的内容 |

## 测试规格说明书

**第一版：**

**变更日期：**2016-05-11

**变更版本编号：**1.0

**变更请求：**形成测试规格说明书初稿

**变更范围：**全文

**变更的详细描述：**

* 列出了基本的测试需求

**第一版：**

**变更日期：**2016-05-25

**变更版本编号：**1.1

**变更请求：**针对网评及课堂点评做出相应修改

**变更的详细描述：**

|  |  |  |
| --- | --- | --- |
| 变更原因 | 变更位置 | 变更内容概述 |
| 外因（根据网络互评意见）：  文档目录与文档实际排版不对应，导致查看时无法准确找到目录对应内容。 | 目录，各章节 | 更新目录，调整格式。 |
| 外因（根据网络互评意见）：  针对非功能需求模块的“高效性”需求用例对应的测试用例描述为“性能测试“ | 4.2 | 修改“性能测试” |
| 外因（根据网络互评意见）：  在系统部署测试测试用例中测试系统是否正确部署的步骤与部署的步骤混淆了，配置Web应用、导入Lucene包应该属于部署的步骤。 | 1.1.2测试用例 | 将部署的步骤放在了precondition中 |
| 外因（根据网络互评意见）：  测试需求文档中文件添加测试的描述与文本库的管理中的添加文件的描述、以及需求文档中RUCM模型中添加文件的描述不一致 | 1.2.2测试用例 | 标注基本需求模块与RUCM对应的章节号，用例名与需求文档统一 |
| 外因（根据网络互评意见）：  “Tester”描述为“测试员“，无实际意义。应该具体到组内成员中负责该测试用例的人员信息 | 各测试用例 | 测试员具体分配给个人 |
| 外因（根据网络互评意见）：  比较没有科学性与严谨性 | 2.3.1 测试策略描述 | 修改特征距离比较方式 |
| 外因（根据网络互评意见）：  建议与其他人选取的特征进行对比分析。 | 2.1.2节图像特征选择测试 | 添加与其他特征的比较 |
| 外因（根据网络互评意见）：  命名重复 | 2.3节 | 修改命名 |
| 外因（根据网络互评意见）：  Basic Flow(Test Sequence)中的Postcondition感觉不应该是“取图像特征”。 | 2.3.2测试用例 | 修改postcondition |
| 外因（根据网络互评意见）：  测试步骤都完全一致，为什么作为两个测试用例。 | 2.6图像检索测试,2 .4网站搭建测试 | 修改网站搭建测试用例 |
| 外因（根据网络互评意见）：  章节的名称发生了错误，出现两个第一章 | 目录 | 重新整理了编码 |

# 影响因素分析

## 外在因素

#### 评审意见：包括老师和同学在格式和内容上提出的意见

评审意见贯穿了项目进行的始终，老师和同学们的评审意见开阔了我们的视野，发现了许多我们小组内部讨论时没有想到的地方，帮助我们不断完善我们的项目需求、计划并确保我们在项目进行过程中没有偏离方向。

## 内在因素

#### 对系统的了解程度：

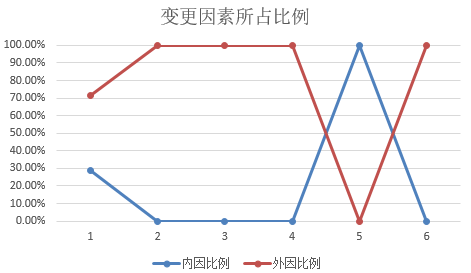
在项目刚开始的时候，对项目的熟悉程度较低，只是大致了解Lucene开源系统是做什么的，有哪些模块，根据这些模块倒推项目的需求分析。

随着对系统了解程度的加深，为了做出实质性的改进，我们对系统具体的使用场景以及数据样例的特点进行了分析，根据不同的使用场景以及数据样例的特点，有针对性地对需求进行了改进，进而变更了相应的需求规格说明书以及任务计划。

# 统计分析

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 版本变更记录 | 1 | 2 | 3 | 4 | 5 | 6 |
| 内 因 | 2 | 0 | 0 | 0 | 7 | 0 |
| 外 因 | 5 | 3 | 6 | 2 | 0 | 3 |
| 总 次 数 | 7 | 3 | 6 | 2 | 7 | 3 |
| 内 因 比 例 | 28.57% | 0.00% | 0.00% | 0.00% | 100.00% | 0.00% |
| 外 因 比 例 | 71.43% | 100.00% | 100.00% | 100.00% | 0.00% | 100.00% |

# 相关性分析



为了研究版本变更与影响因素之间的相关性，我们通过分析版本变更追踪链，发现：

1、第一至第四版本变更时，外因占据了主要地位，这是因为这一阶段主要是对需求规格说明书的格式、内容表达及阐述的准确性及合理性、用例图、RUCM等的格式及用法是否恰当进行的讨论及修改，因此是外因占了主导地位；

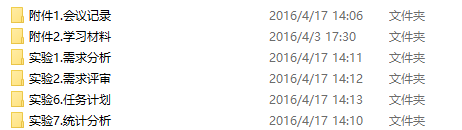
2、第五次版本变更时，由于涉及到了具体实现过程，项目的具体计划发生了改变，这一阶段内容主要是受到内因及项目组成员对系统的把握程度及其相应计划的影响，因此内因是主导地位；

3、第六次版本变更，是在项目实现工作已经完成的情况下，针对完成内容按照老师给出的指导意见进行规格说明书的增补，如增加“基本需求、实现需求”，将实现内容落实到需求文档中的需求，因此是外因占主导地位；

# 5.配置管理总结

为了便于对项目变更进行管控（版本控制），维护不同项目之间的版本关联，以及后期追溯软件开发过程进行总结分析，我们使用了GitHub作为基础配置管理工具。

1. 为了使项目文件具有合理的组织框架，我们通过讨论制定了文档分类组织结 构，依据各子实验进行组织分类。



1. 为了防止版本变更过程中出现命名混乱，我们通过开会讨论统一了本次软件工程综合实验文档的命名方法。
2. 为了方便追踪历次变更的变化，便于后期总结阶段进行分析及总结，在每次进行版本变更时，会在文档中附加版本变更说明表，详细记录变更的原因、内容及变更位置。

**版本变更说明表**

|  |  |  |
| --- | --- | --- |
| 变更原因 | 变更位置 | 变更内容概述 |
| 外因（根据老师点评意见）：  需求应该范围两类：基本需求以及实现需求 | 第三章 | 在第3章添加对基本需求实现的说明 |
| 外因（根据老师点评意见）：  数据需求部分进行修改、添加场景需求说明 | 4.4、4.5 | 数据需求部分进行修改、添加场景需求说明 |
| 外因（根据老师点评意见）：  已经实现的内容要与需求文档中的需求结合起来 | 第六章 | 第6章修改为实现需求和实现方案，并添加部分实现需求方面的内容 |

1. 为了确保项目文件完整性及安全性，我们采用了冗余存储的方式，除了将项目文件上传到Github上，还做好了备份工作，每次的项目文件小组成员分别存有备份。