**知 识 库 管 理 系 统**

——**软件详细设计说明书**

# 目 录

[第1章 关键技术 1](#_Toc453419672)

[1.1 Lucene检索 1](#_Toc453419673)

[1.1.1 核心技术 1](#_Toc453419674)

[1.1.2 Lucene应用 1](#_Toc453419675)

[1.1.3 IKAnalyzer分词器 1](#_Toc453419676)

[1.1.4 Tika的使用 1](#_Toc453419677)

[1.1.5 具体流程 2](#_Toc453419678)

[1.2 Hadoop集群 2](#_Toc453419679)

[1.2.1 核心技术 2](#_Toc453419680)

[1.2.2 Hadoop集群运算 2](#_Toc453419681)

[1.2.3 具体流程 3](#_Toc453419682)

[1.3 AOP切面 4](#_Toc453419683)

[1.3.1 核心技术 4](#_Toc453419684)

[1.3.2 具体流程 4](#_Toc453419685)

[第2章 系统所用算法 5](#_Toc453419686)

[2.1 Mahout智能算法 5](#_Toc453419687)

[2.1.1 改进的PageRank算法--PeopleRank 5](#_Toc453419688)

[2.1.2 基于用户的协同过滤算法 5](#_Toc453419689)

[2.2 算法的应用 5](#_Toc453419690)

[第3章 系统数据结构设计 6](#_Toc453419691)

[3.1 数据字典 6](#_Toc453419692)

[3.2 数据结构详细列表 7](#_Toc453419693)

[3.2.1 sys\_user（用户信息表） 7](#_Toc453419694)

[3.2.2 sys\_role（角色信息表） 7](#_Toc453419695)

[3.2.3 acl\_permission（权限信息表） 7](#_Toc453419696)

[3.2.4 sys\_user\_role（用户角色关联表） 7](#_Toc453419697)

[3.2.5 acl\_role\_permission（角色权限关联表） 8](#_Toc453419698)

[3.2.6 theme\_parent\_type（主题类主表） 8](#_Toc453419699)

[3.2.7 theme\_type（主题类细表） 8](#_Toc453419700)

[3.2.8 file\_type（文件类型编号表） 8](#_Toc453419701)

[3.2.9 file\_info（文件信息表） 9](#_Toc453419702)

[3.2.10 theme\_relation（主题主细关联表） 9](#_Toc453419703)

[3.2.11 file\_theme\_relation（文件主题关联表） 9](#_Toc453419704)

[3.2.12 file\_relation（文件关联表） 10](#_Toc453419705)

[3.2.13 expert\_info（专家信息列表） 10](#_Toc453419706)

[3.2.14 user\_theme\_relation（用户主题关联表） 10](#_Toc453419707)

[3.2.15 expert\_theme\_relation（专家主题关联表） 10](#_Toc453419708)

[3.2.16 user\_focus（用户关注表） 11](#_Toc453419709)

[3.2.17 expert\_file\_relation（专家文件关联表） 11](#_Toc453419710)

[3.2.18 file\_general（用户常用表） 11](#_Toc453419711)

[3.2.19 user\_history（用户检索历史表） 11](#_Toc453419712)

[3.2.20 user\_recommend（用户推荐表） 12](#_Toc453419713)

[3.2.21 user\_evaluation（用户评价表） 12](#_Toc453419714)

[3.2.22 theme\_hot（热门主题表） 12](#_Toc453419715)

[3.2.23 people\_rank（主题专家排名表） 12](#_Toc453419716)

[第4章 系统模块详细设计 14](#_Toc453419717)

[4.1 主题模块 14](#_Toc453419718)

[4.1.1 主题知识浏览展示模块 14](#_Toc453419719)

[4.1.2 热门主题推荐 15](#_Toc453419720)

[4.2检索模块 16](#_Toc453419721)

[4.2.1 高级检索 16](#_Toc453419722)

[4.2.2 分类检索 18](#_Toc453419723)

[4.3 知识分类模块 19](#_Toc453419724)

[4.4 可视化展示模块 20](#_Toc453419725)

[4.4.1 主题专家网络 20](#_Toc453419726)

[4.4.2 主题知识图谱 20](#_Toc453419727)

[4.5 知识问答模块 22](#_Toc453419728)

[4.6 用户中心模块 23](#_Toc453419729)

[4.6.1 用户注册模块 23](#_Toc453419730)

[4.6.2 用户登录模块 24](#_Toc453419731)

[4.6.3 个人资料管理 26](#_Toc453419732)

[4.6.4 检索历史 27](#_Toc453419733)

[4.6.5 用户推荐 28](#_Toc453419734)

[4.6.6 用户上传 29](#_Toc453419735)

[4.6.7 用户关注和用户收藏 31](#_Toc453419736)

[4.6.8 知识关联模块 31](#_Toc453419737)

[4.7 后台管理模块 33](#_Toc453419738)

[第5章 系统开发阶段列表 34](#_Toc453419739)

[5.1 基本功能开发列表 34](#_Toc453419740)

[5.2 创新功能开发列表 35](#_Toc453419741)

# 第1章 关键技术

## 1.1 Lucene检索

### 1.1.1 核心技术

* 使用JSP的JSTL技术进行检索信息的展示
* 使用jquery infinite-scroll插件实现信息的无限滚动加载
* 使用mixitup和infinite-scroll插件实现主题知识的展示
* Jquery ajax技术检索信息展示

### 1.1.2 Lucene应用

本系统Lucene主要用于全局检索和分类检索中，包括对文件、图片、视频、音频的检索。通过对数据库中的文件、图片、视频、音频基本信息进行索引，将常可能查询到的知识存入索引中，方便知识的还原，并减少和数据库的交互。对不常用的信息和文本信息较多的字段进行分词索引，在查找时，通过对信息的索引进行检索，对命中条目提取具体的id值，再通过id值检索数据库获取相应的信息，对信息进行封装，并且高亮的展示给用户。通过设定定时任务，将定时对索引库进行更新，以便于“学习”新的“知识”，提高对知识检索的命中率和命中正确率，从而提升用户体验。

对于检索，本系统提供了两种方式，即基本检索和高级检索，基本检索是指，用户只需要输入关键字，系统会在标题和摘要里进行索引查找，命中即返回。高级检索是多字段的查找，用户可以输入标题关键字、摘要关键字、作者名字、时间范围等字段内容进行综合查找，还可以选择精确和模糊两种方式，高级检索能更快更好的检索出用户想要的知识信息。

### 1.1.3 IKAnalyzer分词器

Lucene检索对于英文内容可以很好地检索，但对于中文来说只会出现乱码，不能达到检索的目的，所以引进中文分词器IKAnalyzer，通过对文件、图片、视频、音频的关键字进行聚集，将其作为新词库添加到分词器词库中，由于关键字均存在于数据库中，只需定时的对词库进行重新初始化，便可动态的获取到变化的信息，实现实时词库的更新，由于关键字数量庞大，故不将所有的关键字全都存在数据库的同张表上，以提高加载词库的速度。在原有的基础上扩展了停用词库，使得分词更加准确。

### 1.1.4 Tika的使用

有了Lucene和IKAnalyzer，只可以对于数据库中有的字段内容进行索引，但本系统中涉及到的文件有许多种类型（如word，excel等），对于文本里面的内容，并不能有效的进行索引，故需要引入Tika，根据数据库中存入的文件的地址，使用Tika可以深入文件内容，对其里面的信息建立索引，当用户查询时，只有文章内容里出现了该关键字，那么这篇文章也会呈现给用户，避免了文章内容与关键字无关的现象，提高了查询的准确率。

### 1.1.5 具体流程

首先页面中获取到用户输入的关键字（时间、作者、标题、摘要等），获取页面中的搜索类型（文件、视频、音频、图片），搜索的精度（精确、模糊），将检索的请求交给搜索控制层，对检索任务进行分发。

到达检索服务层，通过lucene环境，获取单线层维护的IndexSearch索引检索接口，通过对检索信息的分页检索，检索索引库，获取根据相似度排好序的索引数据，并获取对应的Document，接着获取Document中存储的信息或获取相应的id，通过id查询数据库，结果集包装，返回给搜索页面。具体流程如图1-1所示：

## 1.2 Hadoop集群

### 1.2.1 核心技术

* Hadoop集群运算，通过执行Mahout中的智能算法获取专家网络中专家的权 值和用户相似邻居
* 使用mybatis与数据库以及Hadoop集群中的HDFS文件进行交互
* 通过ajax将后台数据进行封装
* 通过echarts展示给用户

### 1.2.2 Hadoop集群运算

Hadoop所擅长的并不是实时计算，系统将已经编写好的程序打成jar包在集群上运行。运行时所需的数据源和最后得出的结果存储在集群的HDFS中，在系统中，我们是通过后台来与HDFS进行通信，将Hadooop运行Mahout智能算法中的jar包所需的数据上传到HDFS中；另外，另有后台直接从HDFS中获取专家分数，并将其写入数据库，并且获得用户相似邻居。在集群中，我们定时执行Mahout智能算法中的jar包，后台也是定时的对数据库进行更新。



图 1-1 Lucene检索流程图

### 1.2.3 具体流程

首先通过后台与Hadooop集群进行交互。后台通过程序对数据库中的指定部分的数据进行处理后生成一个对应的.csv文件。然后将其上传到Hadoop集群的HDFS中。而后，在Hadoop集群进行设置，每隔固定的时间通过Hadoop执行已经编译好的Mahout智能算法中的jar包，而后将得到的结果存入HDFS中。后台程序也负责将Hadoop计算得到的结果从HDFS中取出，同时与数据库中的一部分数据进行整合后，存入到数据库中。

而后，前台程序通过与数据库交互。读取由Mahout智能算法中计算后得到的专家分数，专家网络中专家的关联情况，以及与用户相似的邻居用户。最后通过ajax与echarts来实现前端的展现。

## 1.3 AOP切面

### 1.3.1 核心技术

本部分采用了AOP（面向切面编程）技术，这是通过[预编译](http://baike.baidu.com/view/176610.htm" \t "http://baike.baidu.com/_blank)方式和运行期动态代理实现程序功能的统一维护的一种技术。在这一模块，我们通过AOP来对用户的行为进行监听。每当用户对某知识页面进行访问，我们便记录session中存储的用户信息与当前访问页面的信息，并将其存入数据库中。系统中根据此处的检索历史作为用户兴趣推荐模块的依据。

### 1.3.2 具体流程



图 1-2 AOP用户监听流程图

# 第2章 系统所用算法

## 2.1 Mahout智能算法

### 2.1.1 改进的PageRank算法--PeopleRank

PeopleRank算法是google的PageRank算法的变形。在该算法中，需要构建一个专家之间关联关系的矩阵，此关联关系是通过专家在同领域，论文有过合作等得到的。获取关联矩阵后，对该矩阵进行约10次的同源矩阵的计算，计算出每个专家的PeopleRank值，从而确定某领域的核心专家。

### 2.1.2 基于用户的协同过滤算法

基于用户的协同过滤算法可以推荐给用户一些与该用户品味相似的其他用户喜欢的内容。在该算法中需要四个主要的模块：首先，创建数据模型（DataModel），然后定义用户的相似度算法（UserSimilarity），接下来定义用户近邻算法（UserNeighborhood），最后是推荐算法（Recommender）完成计算过程。

## 2.2 算法的应用

本系统中PeopleRank算法主要应用在专家网络中，通过获得到的专家关联信息，在Hadoop集群中运行PeopleRank算法，获得相应的PeopleRank值，以此判断同领域专家的权重地位，实现领域权威专家的推送。最终在前台以专家网络图谱的形式进行显示，通过图谱中专家图形的大小为用户展示其在同等领域中的核心程度。

基于用户的协同过滤算法主要用在用户推荐中，通过用户的关注信息和收藏信息和检索历史，在hadoop集群中运行该算法，获得用户相似的邻居用户，然后根据邻居用户喜欢的内容进行推荐，并在前台显示。

# 第3章 系统数据结构设计

## 3.1 数据字典

表3-1 系统数据库列表

|  |  |  |
| --- | --- | --- |
| **序号** | **表名** | **表名描述** |
| 1 | sys\_user | 用户信息表 |
| 2 | sys\_role | 角色信息表 |
| 3 | acl\_permission | 权限信息表 |
| 4 | sys\_user\_role | 用户角色关联表 |
| 5 | acl\_role\_permission | 角色权限关联表 |
| 6 | theme\_parent\_type | 主题类主表 |
| 7 | theme\_type | 主题类细表 |
| 8 | file\_type | 文件类型编号表 |
| 9 | file\_info | 文件信息列表 |
| 10 | theme\_relation | 主题主细关联表 |
| 11 | file\_theme\_relation | 文件主题关联表 |
| 12 | file\_relation | 文件关联表 |
| 13 | expert\_info | 专家信息列表 |
| 14 | user\_theme\_relation | 用户主题关联表 |
| 15 | expert\_theme\_relation | 专家主题关联表 |
| 16 | user\_focus | 用户关注表 |
| 17 | expert\_file\_relation | 专家文件关联表 |
| 18 | file\_general | 用户常用表 |
| 19 | user\_history | 用户检索历史表 |
| 20 | user\_recommend | 用户推荐表 |
| 21 | user\_evaluation | 用户评价表 |
| 22 | theme\_hot | 热门主题表 |
| 23 | people\_rank | 主题专家排名表 |

## 3.2 数据结构详细列表

### 3.2.1 sys\_user（用户信息表）

表 3-2 用户信息表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | user\_id | 用户id | int | 11 | 是/是 |
| 2 | login\_name | 登录名称 | varchar | 50 | 否 |
| 3 | user\_name | 用户名称 | vachar | 50 | 否 |
| 4 | password | 用户密码 | varchar | 50 | 否 |
| 5 | email | 用户邮箱 | varchar | 45 | 否 |

### 3.2.2 sys\_role（角色信息表）

表3-3 角色信息表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | role\_id | 角色id | Int | 11 | 是/是 |
| 2 | role\_name | 角色名称 | varchar | 50 | 否 |
| 3 | role\_description | 角色描述 | varchar | 200 | 否 |

### 3.2.3 acl\_permission（权限信息表）

表3-4 权限信息表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | perm\_id | 权限id | Int | 11 | 是 |
| 2 | perm\_token | 权限名称 | varchar | 100 | 否 |
| 3 | description | 权限描述 | varchar | 100 | 否 |

### 3.2.4 sys\_user\_role（用户角色关联表）

表3-5 用户角色关联表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | user\_id | 用户id | Int | 11 | 是/是 |
| 2 | role\_id | 角色id | Int | 11 | 是/是 |

### 3.2.5 acl\_role\_permission（角色权限关联表）

表3-6 角色权限关联表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | role\_id | 角色id | Int | 11 | 是/是 |
| 2 | perm\_id | 权限id | int | 11 | 是/是 |
| 3 | perm\_token | 权限名称 | varchar | 100 | 否 |

### 3.2.6 theme\_parent\_type（主题类主表）

表3-7 主题类主表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | id | 主题编号 | Int | 11 | 是/否 |
| 2 | theme\_name | 父类主题名称 | varchar | 45 | 否 |

### 3.2.7 theme\_type（主题类细表）

表3-8 主题类细表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | Id | 子类主题编号 | Int | 11 | 是/否 |
| 2 | theme\_name | 子类主题名称 | varchar | 99 | 否 |
| 3 | key | 关键字 | text | 非空 | 否 |

### 3.2.8 file\_type（文件类型编号表）

表3-9 文件类型编号表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | id | 表序列号 | Int | 11 | 是/否 |
| 2 | type\_name | MIME（文件类型）名称 | varchar | 200 |  |
| 3 | type\_id | 类型编号 | Int | 11 |  |
| 4 | type\_general | 类型通俗名 | varchar | 50 |  |

### 3.2.9 file\_info（文件信息表）

表3-10 文件信息表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | file\_id | 文件编号 | Int | 11 | 是/否 |
| 2 | file\_name | 文件名 | varchar | 50 |  |
| 3 | type\_id | 文件类型 | Int | 11 | 否/是 |
| 4 | file\_abstract | 文件摘要 | text |  |  |
| 5 | file\_time | 文件上传时间 | datatime |  |  |
| 6 | file\_visit | 浏览量 | int | 11 |  |
| 7 | file\_down | 下载量 | Int | 11 |  |
| 8 | file\_url | 文件地址 | varchar | 200 |  |
| 9 | file\_keywords | 文件关键字 | text |  |  |
| 10 | file\_size | 文件大小 | varchar | 50 |  |
| 11 | file\_imgurl | 文件缩略地址 | varchar | 200 |  |
| 12 | flag | 文件是否发布 | Int | 11 |  |
| 13 | file\_downurl | 文件下载地址 | varchar | 200 |  |
| 14 | file\_unique | 文件唯一编号 | Int | 200 |  |
| 15 | user\_id | 上传用户 | Int | 11 | 否/是 |

### 3.2.10 theme\_relation（主题主细关联表）

表3-11 主题主细关联表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | id | 表顺序号 | Int | 11 | 是/否 |
| 2 | parent\_id | 父主题id | Int | 11 | 否/是 |
| 3 | child\_id | 子主题id | Int | 11 | 否/是 |

### 3.2.11 file\_theme\_relation（文件主题关联表）

表3-12 文件主题关联表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | file\_id | 文件id | Int | 11 | 是/是 |
| 2 | theme\_id | 子主题id | Int | 11 | 是/是 |

### 3.2.12 file\_relation（文件关联表）

表3-13 文件关联表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | id | 表顺序号 | Int | 11 | 是/否 |
| 2 | file\_id1 | 被关联文件编号 | Int | 11 | 否/是 |
| 3 | file\_id2 | 关联文件编号 | Int | 11 | 否/是 |

### 3.2.13 expert\_info（专家信息列表）

表3-14 专家信息列表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | id | 专家id | Int | 11 | 是/否 |
| 2 | expert\_name | 专家姓名 | varchar | 15 |  |
| 3 | expert\_org | 专家所属机构 | varchar | 153 |  |

### 3.2.14 user\_theme\_relation（用户主题关联表）

表3-15 用户主题关联表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | id | 表顺序号 | Int | 11 | 是/否 |
| 2 | user\_id | 用户id | Int | 11 | 否/是 |
| 3 | theme\_type\_id | 主题id | Int | 11 | 否/是 |

### 3.2.15 expert\_theme\_relation（专家主题关联表）

表3-16 专家主题关联表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | id | 表顺序号 | Int | 11 | 是/否 |
| 2 | expert\_id | 专家id | Int | 11 | 否/是 |
| 3 | theme\_type\_id | 子主题id | Int | 11 | 否/是 |
| 4 | expert\_score | 专家评分 | Int | 11 | 否 |

### 3.2.16 user\_focus（用户关注表）

表3-17 用户关注表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | Id | 表顺序号 | Int | 11 | 是/否 |
| 2 | focus\_themes | 关注主题 | text | 非空 |  |
| 3 | focus\_experts | 关注专家 | text | 非空 |  |

### 3.2.17 expert\_file\_relation（专家文件关联表）

表3-18 专家文件关联表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | expert\_id | 专家id | Int | 11 | 否/是 |
| 2 | file\_id | 文件id | Int | 11 | 否/是 |
| 3 | time | 专家创作文件时间 | datatime |  |  |

### 3.2.18 file\_general（用户常用表）

表3-19 用户常用表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | Id | 表序列号 | Int | 11 | 是/否 |
| 2 | file\_id | 文件id | Int | 11 | 否/是 |
| 3 | user\_id | 用户id | Int | 11 | 否/是 |
| 4 | time | 收藏时间 | datatime |  |  |

### **3.2.19 user\_history（用户检索历史表）**

表3-20 用户检索历史表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | Id | 表序列号 | Int | 11 | 是/否 |
| 2 | file\_id | 文件id | Int | 11 | 否/是 |
| 3 | user\_id | 用户id | Int | 11 | 否/是 |
| 4 | time | 浏览时间 | datatime |  |  |

### 3.2.20 user\_recommend（用户推荐表）

表3-21 用户推荐表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | Id | 表顺序号 | Int | 11 | 是/否 |
| 2 | user\_id | 用户id | Int | 11 | 否/是 |
| 3 | file\_id | 文件id | Int | 11 | 否/是 |
| 4 | expert\_id | 专家id | Int | 11 | 否/是 |

### 3.2.21 user\_evaluation（用户评价表）

表3-22 用户评价表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | Id | 表顺序号 | Int | 11 | 是/否 |
| 2 | user\_id | 用户id | Int | 11 | 否/是 |
| 3 | file\_id | 文件id | Int | 11 | 否/是 |
| 4 | evalution | 用户评价 | text |  |  |
| 5 | evalution\_time | 评价时间 | datatime |  |  |
| 6 | evalution\_score | 文件评价得分 | Int | 11 |  |

### 3.2.22 theme\_hot（热门主题表）

表3-23 热门主题表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | Id | 表顺序号 | Int | 11 | 是/否 |
| 2 | user\_visit | 用户访问量 | Int | 11 |  |
| 3 | file\_num | 专家数量 | Int | 11 |  |
| 4 | score | 主题评分 | double |  |  |

### 3.2.23 people\_rank（主题专家排名表）

表3-24 主题专家排名表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 字段名 | 字段描述 | 类型 | 长度 | 是否主键/外键 |
| 1 | people\_name | 专家姓名 | varchar | 50 |  |
| 2 | theme\_id | 主题id | Int | 11 |  |
| 3 | rank | 专家rank | double |  |  |
| 4 | rank\_id | 专家在某主题的id | Int | 11 |  |

# 第4章 系统模块详细设计

## 4.1 主题模块

### 4.1.1 主题知识浏览展示模块

1. 功能描述

系统中将用户可浏览的功能模块设置在Header选项卡处，便于用户清晰浏览系统各个模块。该部分主要包括对系统中，主题模块、全局检索、知识分类、可视化展示、用户中心、知识问答模块的浏览功能。点击主题模块，系统进入主题相关内容，页面列出所有的主题，点击每一个主题，可以查看里面的知识信息，包括文件、图片、视频、音频。

2. 流程图



图 4-1 主题知识展示流程图

3.主题知识界面



图4-2 主题知识界面

### 4.1.2 热门主题推荐

1. 功能描述

基于本系统的动态性要求，系统将以天为周期，对各个主题涉及的专家人数，用户关注程度，知识数量进行统计评分，从而为用户推送系统中的当下最为热门的行业信息。

2. 评分规则

某一主题知识数目/知识总量+专家数目/专家总数+用户关注数目/6\*总访问量

3. 流程图



图 4-3 热门主题推荐流程图

4.热门主题界面



图4-4 热门主题界面

## 4.2检索模块

### 4.2.1 高级检索

1.功能描述

在系统Header部分点击全局检索，进入检索页面，打开高级搜索，可以选择搜索范围（文件、图片、视频、音频），搜索字段（标题、摘要、关键字、时间、作者），搜索精度（精确、模糊），将这些信息进行综合搜索，搜索效率更高。

2.流程图

如图4-5所示：

3.全局检索界面

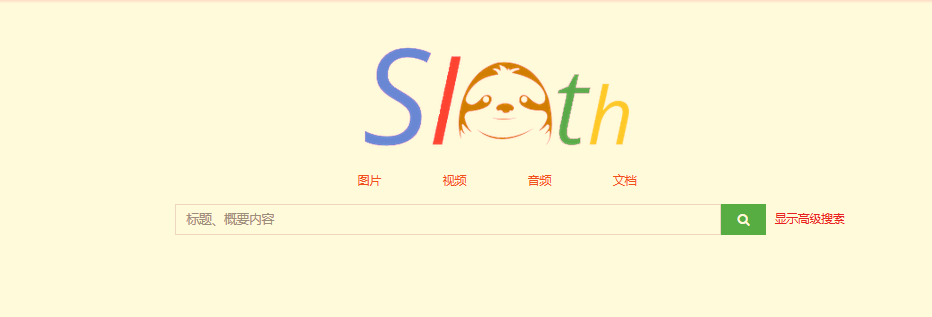


图4-6 全局检索界面



图4-5 高级检索流程图

4.高级检索界面

点击高级检索界面如下：



图4-7 高级检索界面

### 4.2.2 分类检索

1. 功能描述

用户通过Header部分的标签选项卡，选择知识分类，可以进入文件页面、图片页面、视频页面、音频页面，可以浏览每个页面所拥有的知识，并且可以进行检索。

2. 流程图

如图4-8所示。



图 4-8 分类检索流程图

## 4.3 知识分类模块

1.功能描述

知识分类模块主要用于对文件、图片、视频、音频的管理，可以查看各类知识的内容，进行检索，对知识进行下载、收藏、预览、评论，还可以查看每个知识的关联知识，对知识结构化的管理，方便用户寻找和管理公共知识，对于用户喜欢的知识主题等，可以进行关注，发送到常用文档，不用每次重新查找，甚至下载到本地。用户还可以对知识进行简单地评论，给出分数，系统可以根据评论分数给用户进行个性化推荐。

2. 流程图

如图4-9所示。



图 4-9 知识分类管理流程图

## 4.4 可视化展示模块

### 4.4.1 主题专家网络

1.功能描述

该部分构建以某一领域为中心的专家图谱，通过搜集专家的合作关系等数据，发掘专家之间的关联，并且通过在Hadoop集群上执行peopleRank算法得到各个专家的peopleRank值，以此为依据获得各主题网中的核心专家，系统通过echarts图谱的形式在前台展示，使数据中的秘密直观展现出来，便于用户直观的了解该领域的专家关系。

2. 流程图



图 4-10 主题专家网络流程图

3.专家网络界面



图4-11 专家网络界面

### 4.4.2 主题知识图谱

1.功能描述

该模块主要以图形化的形式对知识从不同维度进行统计展示，首先主题知识网络，主要是将各个主题下的小领域，以及领域下的知识类型，每种类型下的知识进行展示，主题图谱主要是将各个主题下的所有知识进行统计，并计算各个主题所占的比重，以图形化的形式直观的表现出来。

2. 流程图



图 4-12 主题知识图谱流程图

3.主题知识图谱界面

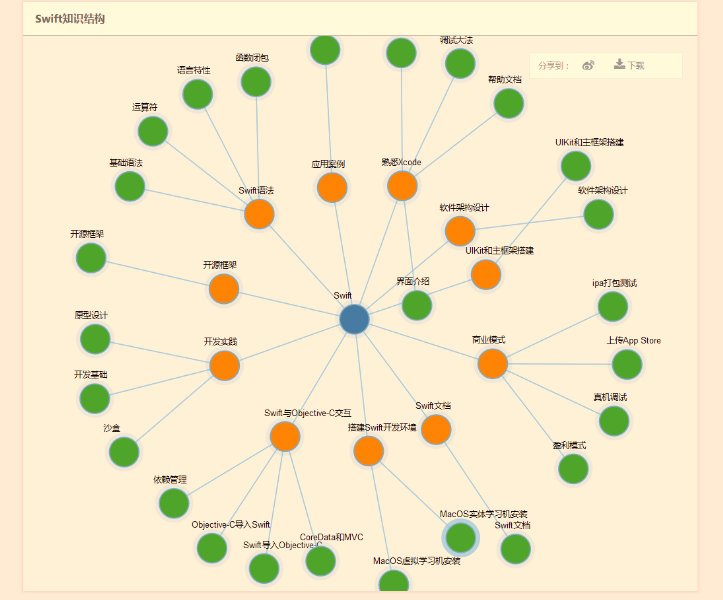


图4-13 主题知识图谱界面

## 4.5 知识问答模块

1.功能描述

系统中该部分主要体现在自动问答模块。首先由用户检索相应问题，系统会通过分词器对相应问题进行解析，同时对关键字进行提取，对所提出的问题进行问题分类。根据问题类别，在答案索引库进行搜索，得到的结果按照其在主题领域中的权重进行排序，最终推送给用户。

2.流程图



图 4-14 知识问答流程图

3.知识问答界面



图4-15 知识问答界面

## 4.6 用户中心模块

用户注册后会获得相应的登录权限，通过登录认证，用户可以进入个人中心，完善个人资料，修改登录密码，查看检索记录，用户可以上传文件，查看收藏的文件，查看关注的主题，并可以享用系统专门为用户提供的智能推荐功能，另外用户可以对上传的文件进行编辑，添加标识，并且可以批量设置关联。本系统通过此部分来打造更适合用户的知识库管理系统。



图 4-16 用户功能模块图

### 4.6.1 用户注册模块

1. 功能描述

游客需要通过填写基本信息注册账号，成为本系统的用户，系统将会为用户打造专属的个人空间，并且方便系统为用户推荐相关知识。

2. 流程图



图 4-17 用户注册流程图

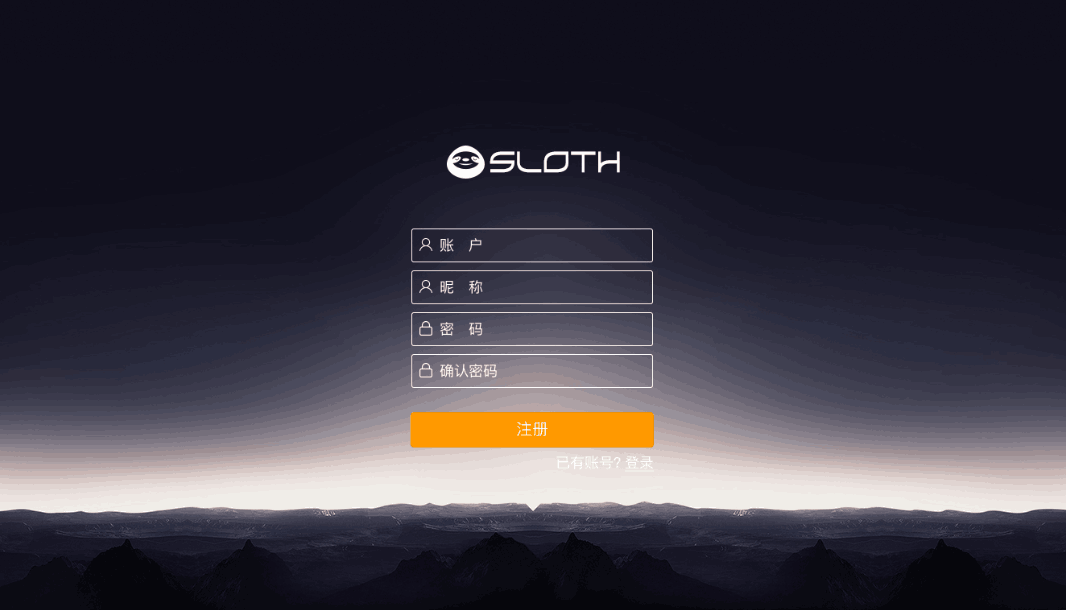
3. 用户注册界面

图4-18 用户注册界面

### 4.6.2 用户登录模块

1.功能描述

用户在未登录之前可以浏览基本信息，但不能享受系统提供的智能推荐功能，用户不能上传文件，同时也不能下载、收藏、评论文件。用户若要使用这些功能，需要登录系统进入用户个人空间，进行相应的功能操作。

2. 流程图

如图4-19所示。

3. 用户登录界面

图4-20 用户登录界面



图 4-19 用户登录流程图

### 4.6.3 个人资料管理

1.个人信息管理

用户可以查看、修改个人的基本信息，并管理自己关注的主题。



图 4-21 个人资料模块图

2. 修改密码

用户进入个人中心，可以通过此选项卡对密码进行修改。如图4-22所示：



图4-22 密码修改流程图

3.个人信息界面



图4-23 个人信息界面

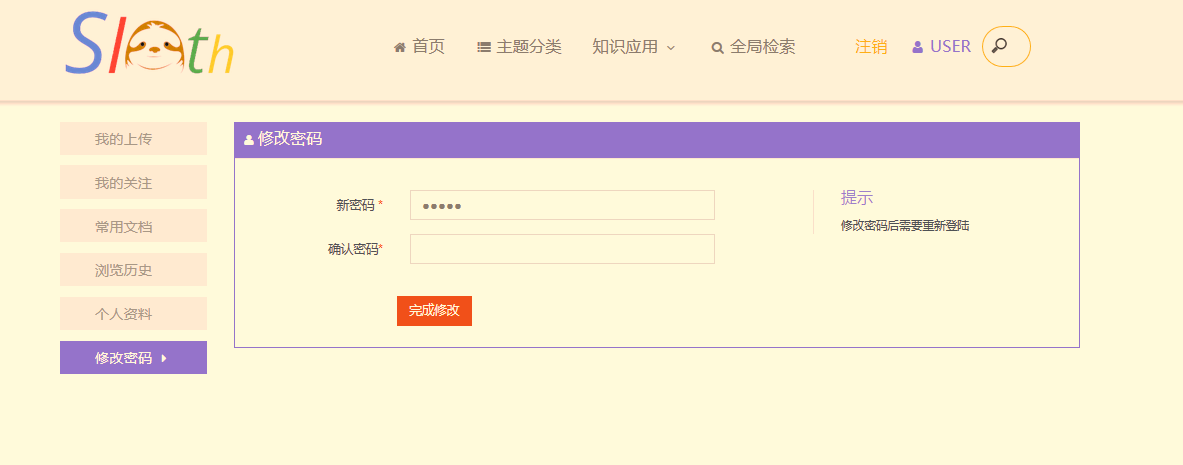
4.密码修改界面

图4-24 密码修改界面

### 4.6.4 检索历史

1.功能描述

通过AOP技术对用户的行为进行监听。每当用户对某一知识进行浏览，我们便记录session中存储的用户信息与当前访问知识的信息，并将其存入数据库中。记录用户每次检索足迹，并以此为依据为用户推荐相关的知识。

2.流程图



图4-25 用户检索历史监听流程图

3.检索历史界面



图4-26 检索历史界面

### 4.6.5 用户推荐

根据用户的关注信息，以及系统用户信息，对每一个用户计算出他的邻居用户即相似用户，然后推荐邻居用户关注的知识，具体流程如图4-27所示：



图 4-27 用户推荐流程图

### 4.6.6 用户上传

1.功能描述

用户登录个人中心，可以对自己拥有的知识进行上传，达到共享的状态，用户上传知识的同时，为了其他用户能更好的找到，需要添加一些介绍，比如标题，关键字等。对于上传的方式，用户可以选择单个上传，也可以选择批量上传，甚至可以拖拽上传。上传文件的类型包括word，excel，pdf，MP3，MP4，jpg，png，avi，zip等常见的文件格式。

2.流程图

具体流程如图4-28所示：



图 4-28 用户上传流程图

3.用户上传界面

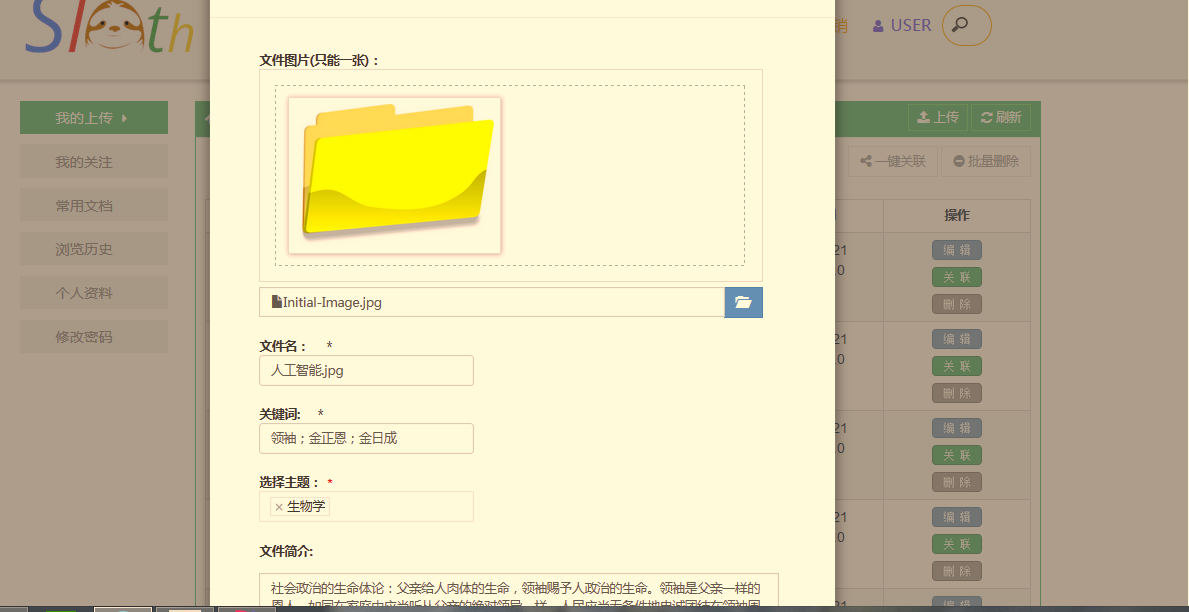


图4-29 用户上传界面

### 4.6.7 用户关注和用户收藏

用户关注模块记录用户关注的主题，用户收藏模块记录的是用户收藏的知识，用户可以查看关注、收藏，同时也可以取消关注，删除收藏。具体流程图如图3-30所示：



图 4-30 用户关注/用户收藏流程图

3.用户收藏界面

用户点击收藏，系统会将知识发送到常用文档。



图4-31 用户收藏界面

### 4.6.8 知识关联模块

知识关联是指将各类相关知识进行连接在一起，用户可以一键查看关联知识，可以更大范围的找到自己需要的知识。知识关联通过唯一编号进行关联，可以单个关联，也可以批量关联。具体流程图如图4-32所示：



图 4-32 知识关联流程图

3.手动知识关联界面



图4-33 手动知识关联界面

4.自动关联界面



图4-34 自动关联界面

## 4.7 后台管理模块

管理员通过前台登录用户名和密码，通过验证后将跳转到后台管理员模块。此模块主要用于管理用户，文件管理、图片管理、视频管理、音频管理，对知识进行删除、查询、更新等操作。实现知识库管理系统的动态化更新效果。



图 4-35 管理员功能模块图

# 第5章 系统开发阶段列表

## 5.1 基本功能开发列表

表 5-1 基本功能开发表

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  | | --- | --- | --- | --- | | **时间** | **任务** | | **人员** | | **概况** | **详情** | | 3月1日-3月7日 | 赛题分析 | 抽取基本功能点 | 全员 | | 设计创新点 | | 确定关键技术 | | 3月8日-3月12日 | 数据库设计 | 解析数据库 | 唐东凯 | | 设计数据库 | 刘泽豪 | | 建立数据库 | | 3月13日-4月1日 | 文件存储 | 单个上传 | 刘泽豪 | | 批量上传 | | 拖拽上传 | | 各种文件类型上传 | | 文档链接 | 发送至常用文档 | 孙境棋 | | 发送至目录 | | 文件收藏 | | 关联文档 | 批量关联 | 刘泽豪 | | 查看关联 | | 增加摘要及缩略图 | | 规则使用 | 规则设计 | 唐东凯 | | 规则实现 | 刘泽豪 | | 智能检索 | 全局检索 标题+摘要 | 刘泽豪 | | | | 分类检索 文件检索+ 图片检索 + 视频检索 + 音频检索 | 唐东凯  孙境棋 | | | | | 高级检索 标题+摘要+关键字+时间+作者 精确检索 + 模糊检索+词频检索 | | | | | 预览 | 文件预览 图片预览 | 刘泽豪 孙境棋 | | 视频播放 音频播放 | |  |  |  |

## 5.2 创新功能开发列表

表5-2 创新功能开发表

|  |  |  |  |
| --- | --- | --- | --- |
| **时间** | **任务** | | **人员** |
| **概况** | **详情** |
| 4月14日-4月18日 | 前期准备工作 | UI设计 | 全员 |
| 系统模型建立 |
| 数据收集 |  |
| 4月19日-4月26日 | 主题模块 | 确定主题分类 | 刘泽豪 |
| 主题知识填充 |
| 主题展示及知识展示 |
| 热门主题推荐 |
| 4月27日-5月3日 | 知识分类展示 | 文件管理 | 孙境棋 |
| 图片管理 |
| 视频管理 |
| 音频管理 |
| 下载 |
| 收藏 |
| 评论 |
| 5月4日-5月15日 | 用户中心 | 个人信息管理 | 刘泽豪 |
| 密码管理 |
| 用户关注 |
| 检索历史 |
| 用户收藏 |
| 用户上传 |
| 添加关联 |
| 5月16日-5月22日 | 用户推荐 | 查找算法 | 刘泽豪  唐东凯 |
| 算法实现 |
| 知识推荐展示 |
| 知识问答 | 关键字匹配检索 |
| 5月23日-5月30日 | 可视化展示 | 主题专家网络 | 孙境棋  唐东凯 |
| 主题知识网络 |
| 主题图谱 |
| Mahout智能算法 |
| Echarts显示 |
| 后台管理 | 用户管理 | 孙境棋 |
| 文件管理 |
| 图片管理 |
| 视频管理 |
| 音频管理 |
| 5月31日-6月8日 | 各功能完善 + 页面优化 | | 各自完成未完成的部分 |
| 6月9日-6月15日 | 测试+汇总文档+整理+录制视频 | | 全员 |