

# 知识库管理系统

## Knowledge Management System

### 软件概要设计说明书

参赛学校： 河海大学常州校区

组 名： Vamos

指导老师： 丁波

队 长： 白宏熙

队 员： 耿玉霞、任玉倩

# 目录

1. 引言 .....	1
1.1 编写目的.....	1
1.2 项目背景.....	1
1.3 技术背景.....	1
1.4 定义.....	1
1.5 参考资料.....	1
2 . 任务概述.....	2
2.1 目标.....	2
2.2 运行环境.....	2
2.2.1 基本配置.....	2
2.2.2 其它配置.....	3
2.3 功能需求概述.....	3
2.3.1 文件集中存储上传.....	3
2.3.2 文档链接.....	3
2.3.3 关联文档.....	3
2.3.4 规则使用.....	4
2.3.5 智能检索.....	4
2.3.6 文件的基本管理.....	5
2.3.7 文本抽取.....	5
2.3.8 语言监测.....	5
2.3.9 元数据提取.....	5
2.3.10 判断文件真实类型.....	5
2.3.11 预览与编辑.....	5
2.3.12 自动补全.....	5
2.3.13 中文分词.....	5
2.3.14 全文检索.....	5
2.3.15 命名实体识别.....	6
2.3.16 文件动态.....	6

2.3.17 推荐系统.....	6
2.4 性能需求概述.....	6
2.4.1 资料存储容量.....	6
2.4.2 资料综合分析速率.....	6
2.4.3 文本处理速度.....	6
2.4.4 命名实体识别功能.....	6
2.4.5 文本抽取功能的范围.....	6
2.5 条件与限制.....	7
3 . 总体设计.....	7
3.1 平台设计原则与思想.....	7
3.1.1 设计原则.....	7
3.1.2 设计思想.....	7
3.2 系统总体设计方案.....	7
3.2.1 系统结构.....	7
3.2.2 功能分配.....	8
3.3 系统部署方案.....	11
3.4 系统界面设计.....	12
4 . 接口设计.....	21
4.1 外部接口.....	21
4.2 内部接口.....	21
5 . 数据库设计.....	22
5.1 数据库表列表 .....	22
5.2 逻辑结构设计.....	23
5.3 物理结构设计 .....	23
6. 出错处理设计.....	30
6.1 出错输出信息 .....	30
6.2 出错处理对策（如设置后备、性能降级、恢复及再启动等） .....	30
7. 安全保密设计.....	30
8. 维护设计.....	31

## 1. 引言

### 1.1 编写目的

为开发人员和用户提供软件系统的总体结构设计、数据库(包括数据结构)设计、外部接口设计、功能部件分配设计、部件之间的接口设计规范。

在系统需求分析的基础上,对需求分析中产生的功能模块进行过程描述,设计功能模块的内部细节,包括算法和详细数据结构,为编写源代码提供必要的说明,并作为测试和维护的参考文档。

### 1.2 项目背景

在日常的生活和工作中,每个人在各自的范围内不断累计,形成不同层次的过程资源财富,涵盖了文本、传真、图形、视频、音频等各个类型的知识财富载体。

在知识财富的不断累计过程中,由于知识财富的来源不同、用途不同、载体不同等,往往导致知识财富的时间链条错乱、关联知识无法有效检索、各类知识无法综合运用等问题的出现,伴随知识财富的持续积累,这种问题愈发突出,迫切需要行之有效的手段来录入、管理、分析、统计知识财富,持续有效的归纳利用知识财富,将助推个人及组织目标的实现。

### 1.3 技术背景

B/S 结构(Browser/Server,浏览器/服务器模式),是WEB兴起后的一种网络结构模式,WEB 浏览器是客户端最主要的应用软件。这种模式统一了客户端,将系统功能实现的核心部分集中到服务器上,简化了系统的开发、维护和使用。客户机上只要安装一个浏览器(Browser 英 ['braʊzə]美 ['braʊzə]),如Netscape Navigator 或 Internet Explorer,服务器安装 SQL Server、Oracle、MYSQL 等数据库。浏览器通过 Web Server 同数据库进行数据交互。

### 1.4 定义

“V\_kmos” 在本文档中表示为 “知识库管理系统”的简称

### 1.5 参考资料

[1] 《SQL 注入攻击与防御(第2版)》

- [2] 《响应式 Web 设计：HTML5 和 CSS3 实战》
- [3] 《CSS 权威指南（第三版）》
- [4] 《JavaScript 权威指南（第 6 版）》
- [5] 《疯狂 Ajax 讲义（第 3 版）》
- [6] 《Lucene 实战(第 2 版)》
- [7] 《推荐系统实践》

## 2 . 任务概述

### 2.1 目标

本系统旨在通过录入、管理、分析、统计知识财富，帮助用户实现持续有效的归纳利用知识财富，助推个人及组织目标的实现。解决在知识财富的不断累计过程中，由于知识财富的来源不同、用途不同、载体不同等导致知识财富的时间链条错乱、关联知识无法有效检索、各类知识无法综合运用等问题。

基本技术目标包括：

- （1）支持各种类型文件的批量上传及压缩包导入；
- （2）支持将文档发送至常用文档，用户无需每次在打开各个目录中查找；
- （3）支持将文档发送至目录，方便其他用户查阅拖拽式配置模拟器界面；
- （4）支持文档的批量关联，支持通过文件唯一编号，实现文档与相关文档、图纸、图片及其它格式附件进行关联，点击链接即可一键快速查看；
- （5）支持为图片、音、视频类文件增加摘要及缩略图；
- （6）支持为文件设定规则，指定动作（组合）、条件和操作，当动作触发符合设定的条件，系统则自动执行规则的操作；
- （7）支持智能检索，能快速从海量资料中精准检索所需文件；完成检索操作后对信息进行接受、判断、提取、分析和概括之后形成自己的知识；
- （8）支持各类格式的图片、视频、音频以及各类 Office 文件的在线预览。

### 2.2 运行环境

#### 2.2.1 基本配置

知识库管理系统所需的基本配置如下：

- （1）操作系统 Windows XP SP3 或 Win7；

- (2) 内存 512M 及以上
- (3) 硬盘空间 40G 及以上
- (4) 服务器操作系统: Windows 2003 server 64 位;

### 2.2.2 其它配置

运行知识库管理系统的开发端所需的其它配置如下:

- (1) tomcat/apache-php/IIS 至少其中之一 (tomcat 和 IIS 需启动 CGI 支持);
- (2) 开发工具: Eclipse 3.6 以上、JDK 1.6 以上;
- (3) Mysql 5.0 及以上或 oracle10g。

## 2.3 功能需求概述

知识库管理系统功能需求包括文件集中存储上传、文档链接、关联文档、规则使用、智能检索和文件的基本管理。

### 2.3.1 文件集中存储上传

在知识库管理系统中,点击本地上传可以加载数据源,实现文件的批量上传及压缩包导入,支持文件的在线压缩和解压缩,支持 Office 文档、PDF、图像、音视频和图纸等各类型文件。

除此之外知识库管理系统还提供文件下载的功能,用户可自由下载自己知识库中的文件,也可以下载其他用户共享的文件。

### 2.3.2 文档链接

文档链接功能支持将文档发送至常用文档,用户无需每次在打开各个目录中查找;知识库管理系统为用户提供桌面形式的界面,常用文档以桌面快捷方式的形式展现给用户,用户无需每次在打开各个目录中查找,只需点击该快捷方式即可查看相应文档。

另一方面,本系统支持将文档发送至目录,方便其他用户查阅。

### 2.3.3 关联文档

文档新建或上传后,系统会自动或手动为文件生成唯一编号,支持文档的批量关联。支持通过文件唯一编号,实现文档与相关文档、图纸、图片及其它格式附件进行关联,点击链接即可一键快速查看。

系统可以为视频文件生成缩略图，也可以对图片文件进行压缩，从而实现为图片、音、视频类文件增加缩略图。此外系统就还可以为各类文件增加摘要。

#### 2.3.4 规则使用

系统支持为文件设定规则，指定动作（组合）、条件和操作，当动作触发符合设定的条件，系统则自动执行规则的操作。目前，知识库管理系统所提供的主要规则有：

- (1) 按照 Windows 任务计划生成缩略图；
- (2) 按照用户自定义的规则，定时向指定邮箱发送指定文件，规则可包括时间、收信人、抄送人、发送文件、文字内容等；
- (3) 可以按照用户需求，将一长定时间内不用的文件放入知识库回收站；
- (4) 为上传的文件设置用户自定义分类标签或系统默认归档标签；
- (5) 文件加密隐藏，需保护的文件被放入一个加密的文件空间里，访问时需输入独立密码；
- (6) 文件关联，用户可自定义常用文件的打开方式、查看模式；
- (7) 用户可以选择是否自动备份，防止账户被恶意删除，可有效恢复文件；
- (8) 文件中转站，设定文件暂存周期，以定期清理节约空间，并能有效避免一次性文件的产生；
- (9) 设置文件生成链接时的默认访问规则，如访问权限和生存周期，增强用户知识的版权和安全性；
- (10) 自定义文件推荐功能的开启，帮助用户在检索时更快更精确的获取推荐知识；
- (11) 用户可自定义文件摘要的表现形式——属性、关键词、系统所提取的文件内容等方面，方便用户查找文件。

#### 2.3.5 智能检索

知识库管理系统能够根据文件类型、文件关键信息、时间以及用户使用各文件的频率等，通过全文检索、自动补全、中文分词、命名实体识别、跨语言信息检索、拼写检查等功能，能够快速从海量资料中精准检索所需文件进行智能搜索，并通过在检索操作后对信息进行接受、判断、提取、分析和概括之后形成用户的知识体系。

除此之外，系统支持在线预览各类格式的图片、播放视频、音频以及浏览各类 Office 文件，包括 Word、Excel、PowerPoint、WPS、Visio 等格式。

### 2.3.6 文件的基本管理

文件的基本管理功能可以实现对文件的基本操作：复制、粘贴、剪切、重命名、删除。还可以实现文本抽取并支持文本在线编辑。

### 2.3.7 文本抽取

文本抽取利用 Apache Lucene Tika，该功能支持 pdf、doc、docx、ppt、excel、txt、html、xml、zip、tar 等常见格式的文本文件。

### 2.3.8 语言监测

语言监测利用 Apache Lucene Tika，支持语言：在所有由 ISO639-1 标准的 184 标准语言，Tika 可检测 18 种语言，皆为常用语言。

### 2.3.9 元数据提取

元数据提取利用 Apache Lucene Tika，支持格式：支持常见文档、音视频和图像

### 2.3.10 判断文件真实类型

判断文件真实类型利用 Apache Lucene Tika，支持格式：XML、HTMLv、MS-OFFICE、ODF、PDF、EPUB、RTF、TXT、FEED、AUDIO、JPEG、VIDEO、ASM、MBOX、DWG、FONT、EXE 等常见文件格式。

### 2.3.11 预览与编辑

该功能通过使用 PageOFFICE 本地组件实现。

### 2.3.12 自动补全

结合历史记录和网络热搜在搜索是可以自动补全。

### 2.3.13 中文分词

IKAnalyzer 中文分词包

### 2.3.14 全文检索

Apache Lucene Solr 以数据库为源，建立索引库，查询速度达到百万条/毫秒



利用 TF-IDF 计算权重,检索结按权重智能排序,关键词高亮显示。

#### **2.3.15 命名实体识别**

命名实体识别是利用斯坦福 NER 包实现的。

#### **2.3.16 文件动态**

利用伪静态技术生成文件动态,可以帮助用户根据上传的时间地点等信息更快的查找所需文件。

#### **2.3.17 推荐系统**

有评分时,采用基于 items 的协同过滤(Mahout 单机实现),无评分时,感知上下文,采用基于文件内容和行为的神经网络预测(RapidMiner 5 java 开源支持),根据文件属性和行为为用户推荐文档。

### **2.4 性能需求概述**

#### **2.4.1 资料存储容量**

知识库管理系统预期的资料存储容量是管理的资料素材数目大于等于 1 百万。

#### **2.4.2 资料综合分析速率**

知识库管理系统预期的资料综合分析速率是 10 万条记录下,通联关系分析时间小于 3 分钟。

#### **2.4.3 文本处理速度**

知识库管理系统预期的文本处理速度是当用户输入的查询词小于 10 个时,全文检索的响应速度小于 5 秒。

#### **2.4.4 命名实体识别功能**

知识库管理系统预期的命名实体识别功能是能够支持中文,人名、组织机构名、时间、地名、目标类型、目标名称等。

#### **2.4.5 文本抽取功能的范围**

知识库管理系统预期的文本抽取功能的范围是能够对 pdf,office 2003/2007, html, email, chm, zip, rar 等各种类型的文本进行内容抽取。

## 2.5 条件与限制

知识库管理系统使用的条件与限制主要有：除了离线功能外，本作品的客户端功能均需要连接知识库管理系统的服务器。

## 3 . 总体设计

### 3.1 平台设计原则与思想

#### 3.1.1 设计原则

知识库管理系统设计遵循以下原则：

- (1) 方便的拖拽式用户操作界面；
- (2) 客户端分辨率自适应；
- (3) 跨平台特性；
- (4) 系统可扩展性；
- (5) 准确性，智能检索的正确率应达到 90%以上，误报率应低于 10%；
- (6) 快速性，资料综合分析速率和智能检索速率要达到系统性能需求；
- (7) 操作简便性。

#### 3.1.2 设计思想

拖拽式用户操作界面功能的目的在于让用户进行可视化预览，方便客户端上传文件。对于文件列表，采用了 HTML5 技术构建，使其有条目拖动浏览的效果。

为了使不同分辨率和不同浏览器的客户端拥有一致的显示效果和用户体验。知识库管理系统采用了响应式布局和 bootstrap 的框架，从而可以为不同终端的用户提供更加舒适的界面和更好的用户体验。如此，也就实现了高度的跨平台性。

为了使功能具有一定的可扩展性，平台拟定使用分层模式的思想，将系统分为三层：数据层、逻辑层、表现层。对于可能变化的文件内容，系统只需要修改数据层与新闻数据库的交互代码即可，而无需改动逻辑层和表现层。

根据用户的行为(如使用某文件的频率)与搜索本身的内容，进行关联分析，利用神经网络模型，从而实现亏阿苏、准确的智能检索。

### 3.2 系统总体设计方案

#### 3.2.1 系统结构

知识库管理系统结构如图 1 所示。

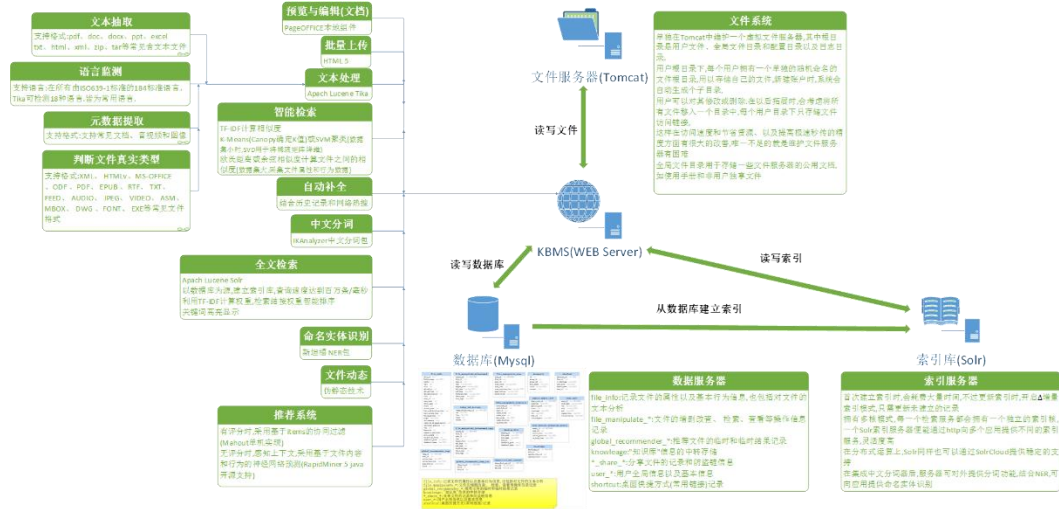


图 1 知识库管理系统结构

### 3.2.2 功能分配

#### 1、核心功能设计

知识库管理系统的核心功能如图 2 所示。

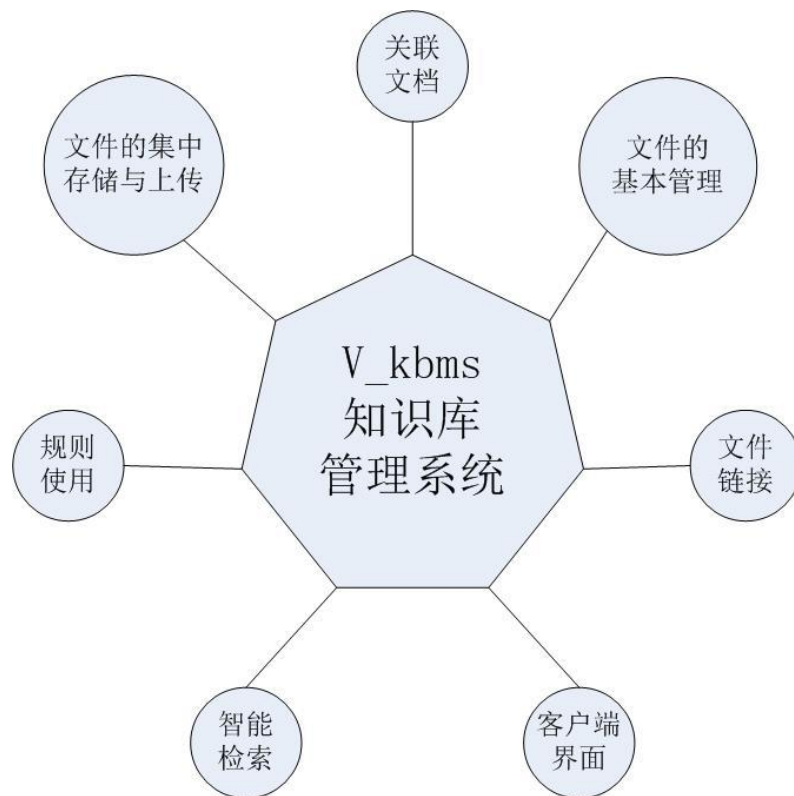


图 2 知识库管理系统核心功能设计

#### 2、分模块设计

知识库管理系统主要包括文件的集中存储与上传功能模块，文档链接功能

模块，关联文档功能模块，规则使用功能模块，智能检索功能模块和文件的基本管理功能模块。

(1) 文件的集中存储与上传功能模块包含了文件的批量上传及压缩包导入等功能。次模块设计如图 3 所示。

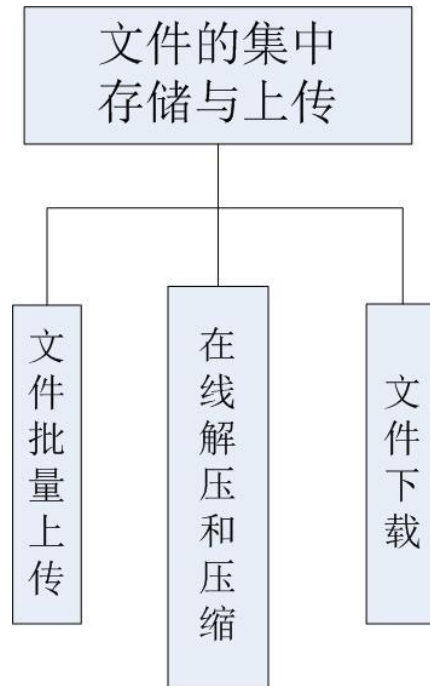


图 3 集中存储与上传功能模块

(2) 文档链接功能模块，文档链接功能模块主要包括将文档发送至常用文档，用户无需每次在打开各个目录中查找以及将文档发送至目录，方便其他用户查阅这两个主要功能。文档链接功能模块设计如图 4 所示。

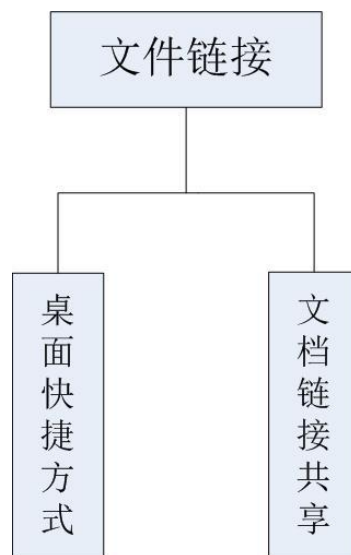


图 4 文档链接模块模块设计

(3) 关联文档功能模块，文档新建或上传后，系统会自动或手动为文件生成唯一编号，支持文档的批量关联。支持通过文件唯一编号，实现文档与相关文档、图纸、图片及其它格式附件进行关联，点击链接即可一键快速查看。该模块还包括为图片、音、视频类文件增加摘要及缩略图的功能。关联文档模块设计如图 5 所示。

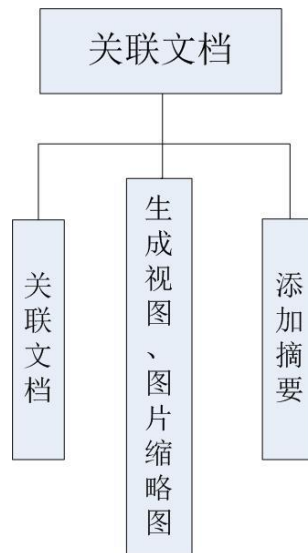


图 5 关联文档模块设计

(4) 规则使用功能模块主要包括按 Windows 任务计划生成缩略图、邮箱定时定向发送文件、文件定时放入回收站等功能，规则使用功能模块设计如图 6 所示。

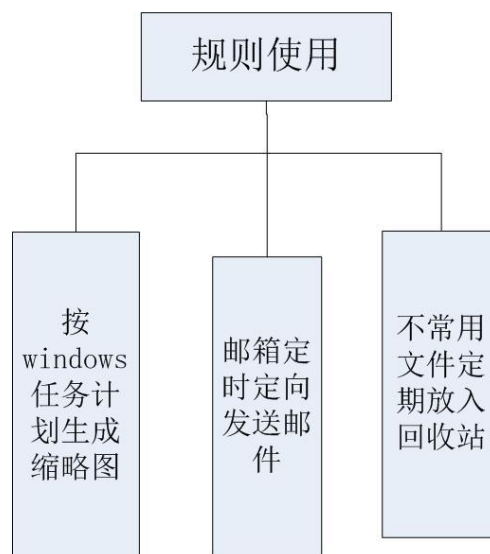


图 6 规则使用功能模块设计

(5) 智能检索功能模块主要包括智能检索的实现与各种类型文件的在线预览，智能检索功能模块设计如图 7 所示。

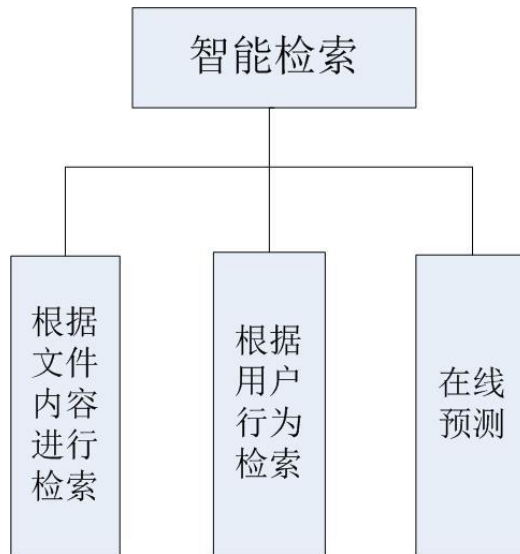


图 7 智能检索功能模块设计

(6) 文件的基本管理功能模块除了实现对文件的基本操作：复制、粘贴、剪切、重命名、删除，还包括文本抽取并支持文本在线编辑功能。基本管理功能模块设计如图 8 所示。

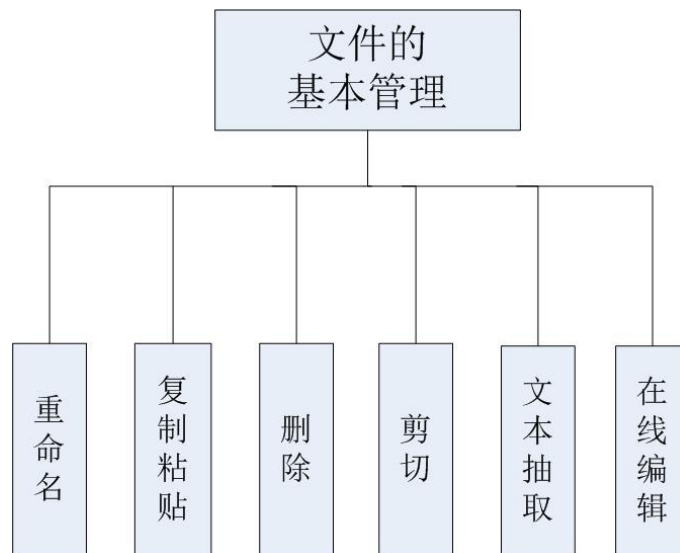


图 8 基本管理功能模块设计

### 3.3 系统部署方案

知识库管理系统的基础构架由个人计算机，应用服务器，数据库服务器，FTP 服务器，索引服务器以及之间相连的网络和防火墙构成。用户使用个人计算机，

访问外部网络，并从外部的服务器中获取文件数据，并进行远程操作。个人计算机通过带有防火墙安全设置的网络连接到应用服务器，并向应用服务器发送数据和操作请求。应用服务器与 FTP 服务器和数据库服务器直接相连，其根据个人计算机发送的请求，返回来自数据库服务器和 FTP 服务器的内容，或对数据库服务器和 FTP 服务器上的数据信息进行读写。用户的检索请求由索引服务器完成，应用服务器进行转发，其部署方案如图 9 所示。

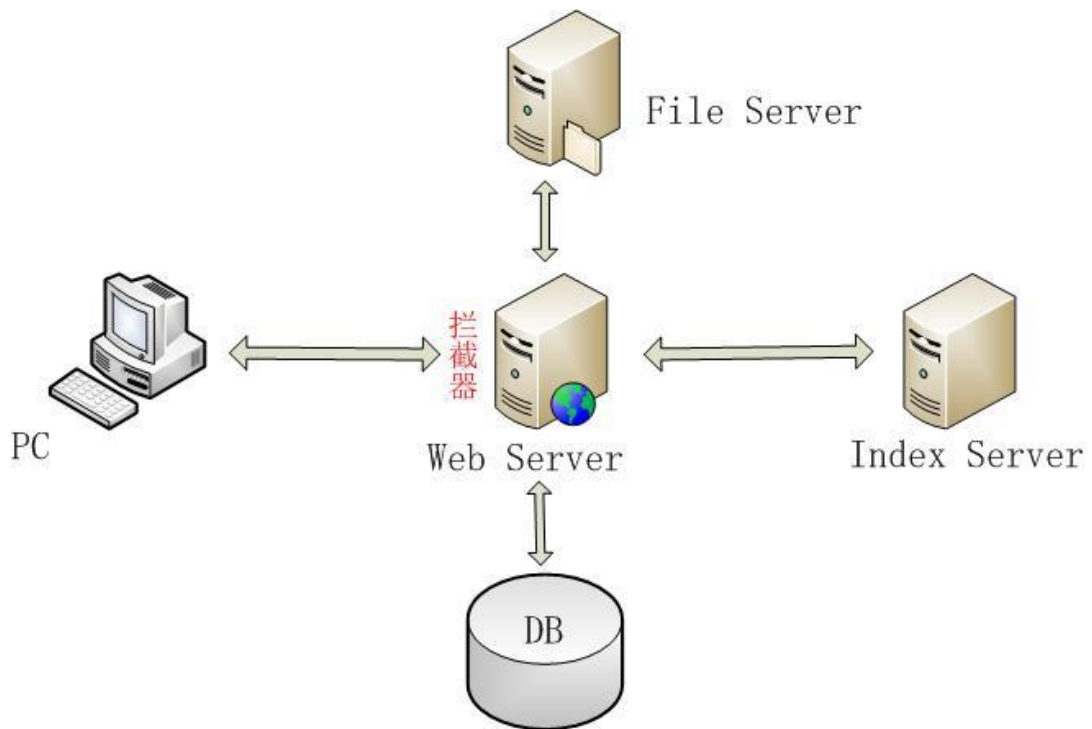


图 9 系统部署方案

### 3.4 系统界面设计

用户界面设计：

(1) 用户登录注册界面：

游客必须经过登录注册成为系统用户之后，方可使用系统，用户注册，首先以手机短信验证身份的真实性，随后填写密码等信息完成注册，登录注册成功后即可进入系统。



图 10 用户登录界面

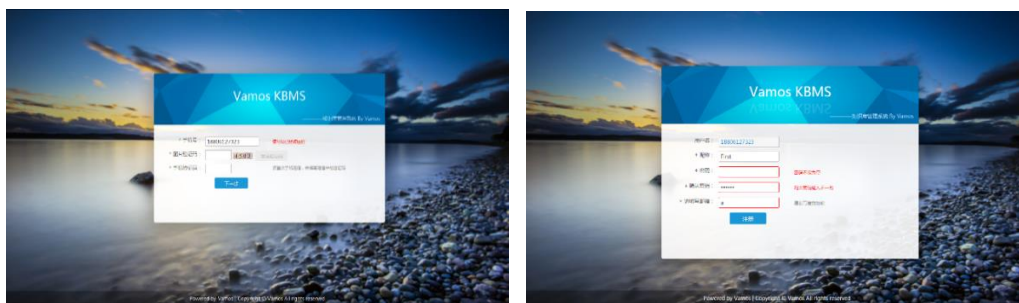


图 11 用户注册界面

## (2) 文件管理界面:

1、用户登入系统之后, 首先进入文件管理界面, 管理自己的知识。该界面主要由文件展示列表, 文件目录列表组成。左侧列表为系统在用户完成注册后根据用户信息自动创建的根目录列表, 为用户进行分类存储提供条件; 页面主视窗中主要展示用户目录下的文件, 将文件名、文件大小、文件修改时间等展示出来, 供用户对文件进行操作。



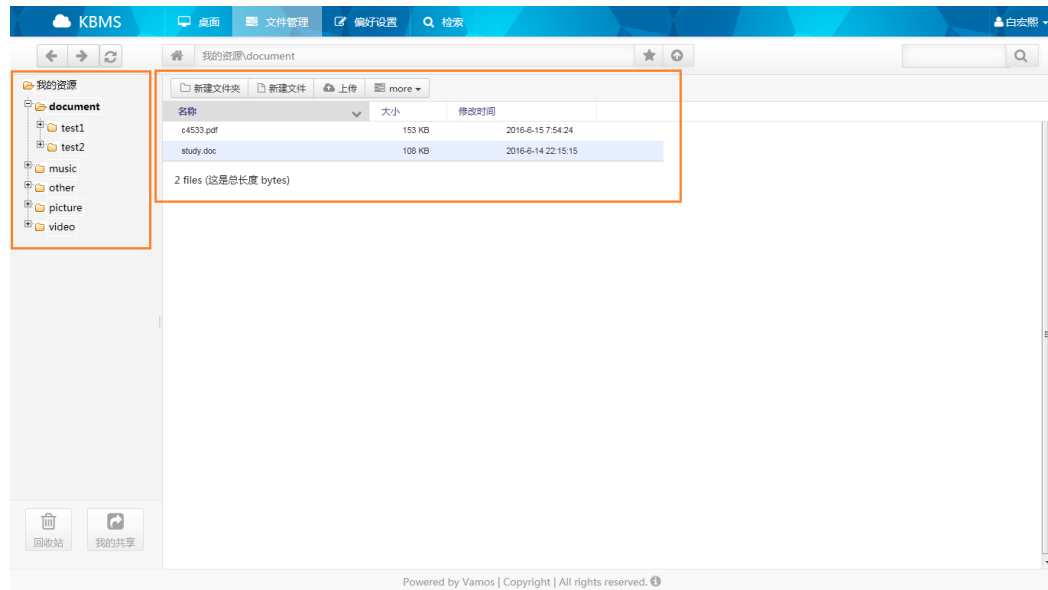


图 12 用户根目录列表界面

2、页面右上方为用户的基本管理信息设置，使得用户拥有更好的系统体验。页面左下方的“回收站”及“共享按钮”，可分别展示出用户的文件删除记录以及共享记录，供用户查看资料的“去向”。并且回收站中将收录用户的不常用知识，保证用户知识系统的灵活多变以及活跃性。

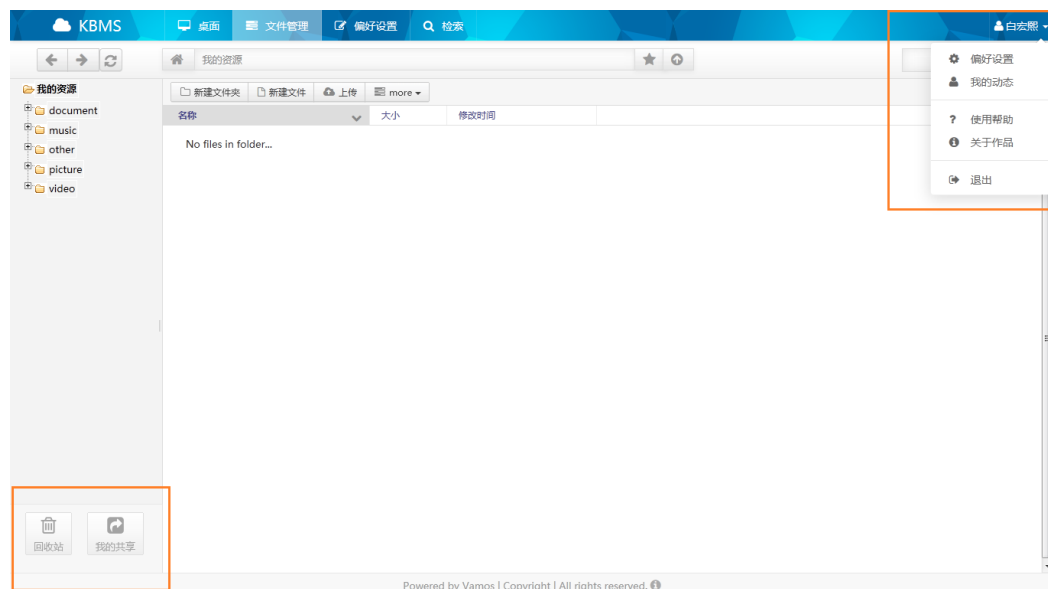


图 13 基本管理功能界面

3、在主视窗的菜单栏部分即为用户可对文件进行的各类操作，包括文件夹的新建，对文件进行的增删改查、在线预览、在线编辑等部分。亦可创建针对某一文件进行创建快捷方式至常用目录（系统桌面），用户即可方便快捷地对

常用文件进行查找。页面中的一些按钮，可以提供用户收藏、跳转、检索的链接。

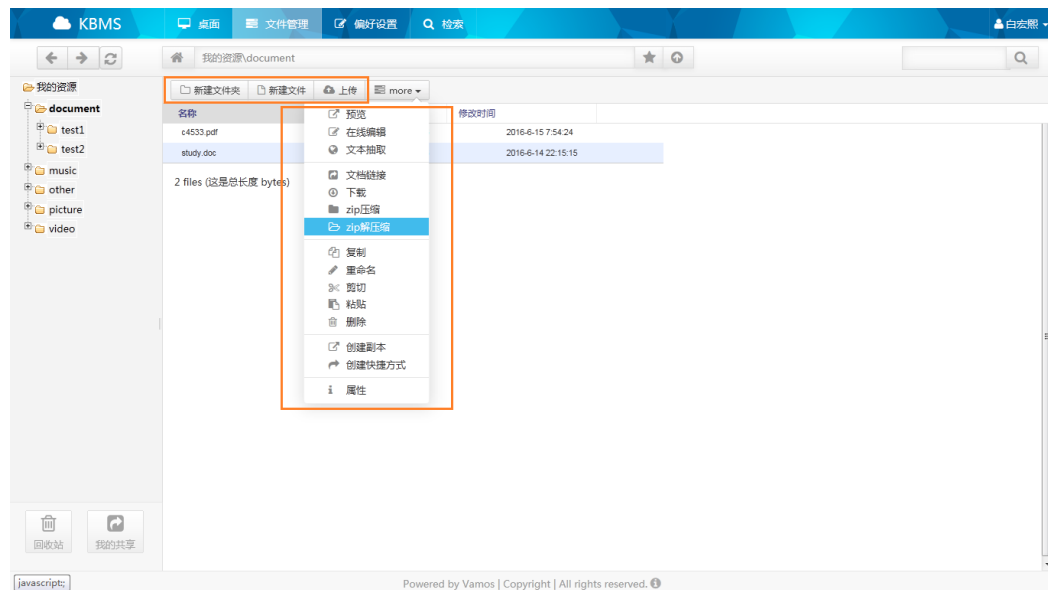


图 14 用户文件操作界面

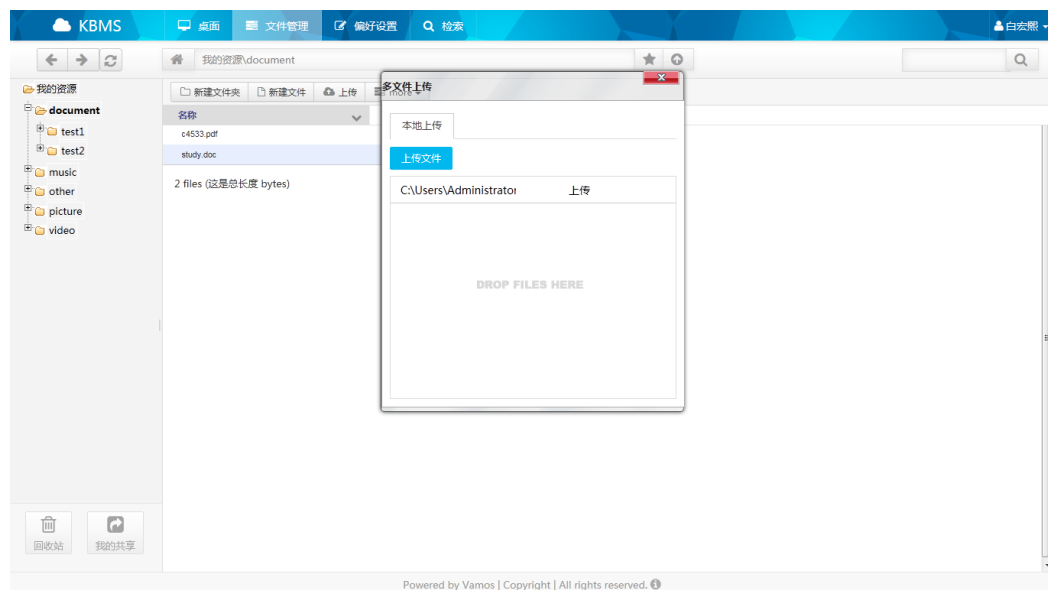


图 15 用户上传文件界面

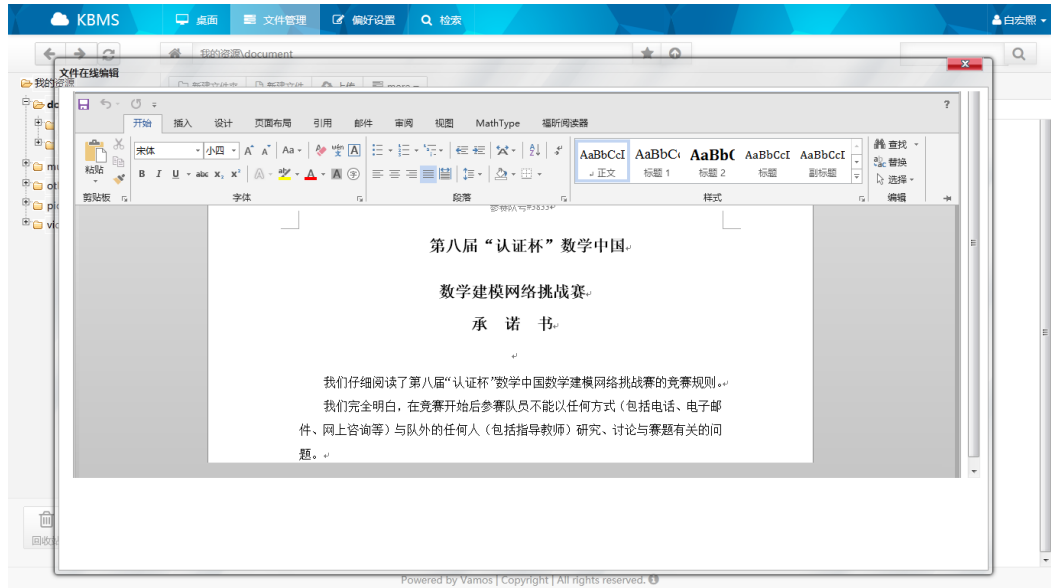


图 16 用户在线编辑文件界面

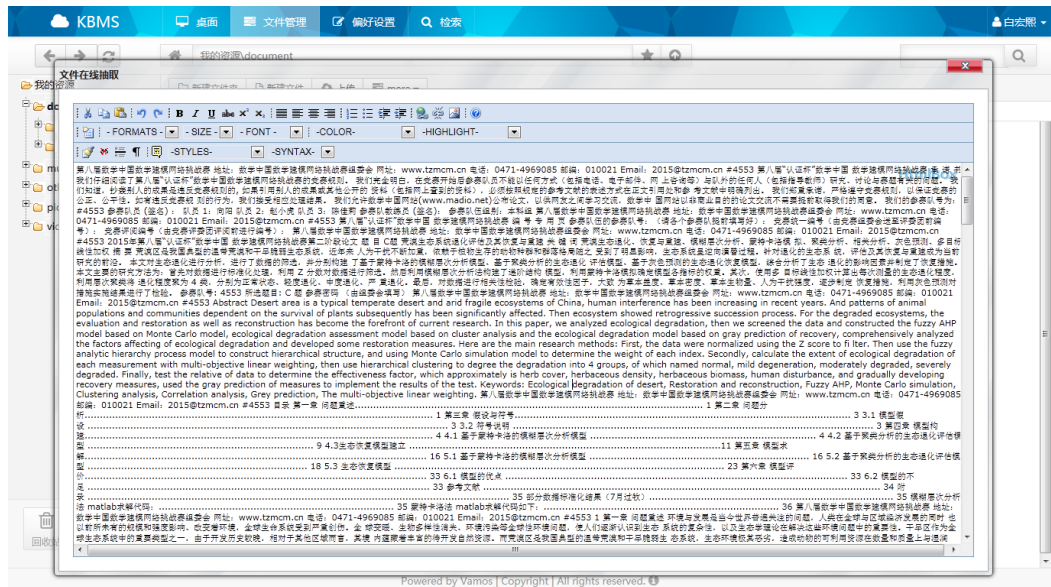


图 17 文档在线抽取界面

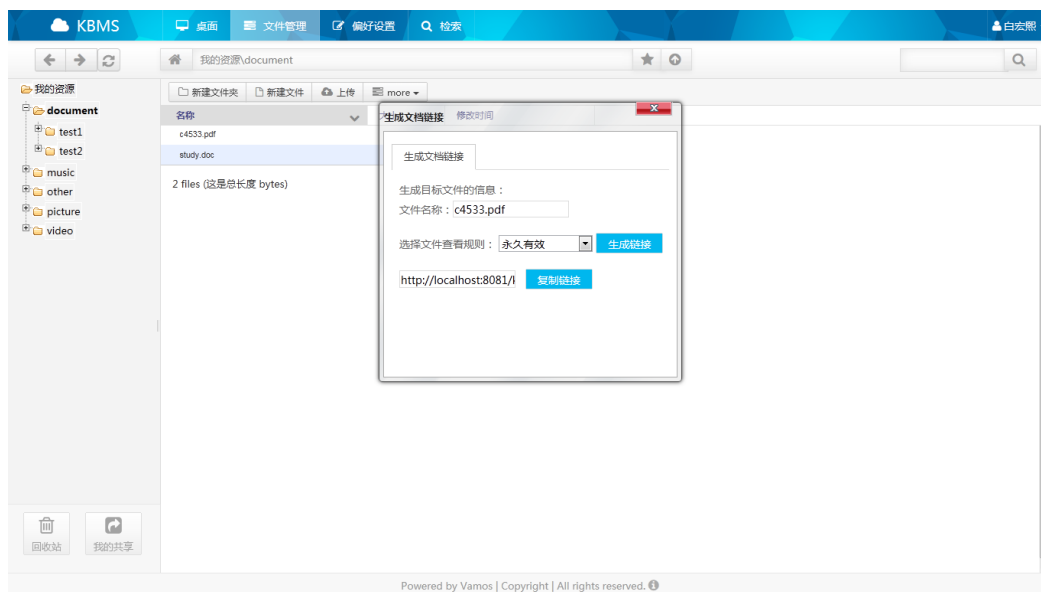


图 18 用户生成文件共享链接界面

### (3) 桌面系统页面

1、本页面收录用户的一些常用文件，根据用户的检索使用情况和自定义情况，提供给用户快速使用知识的通道。

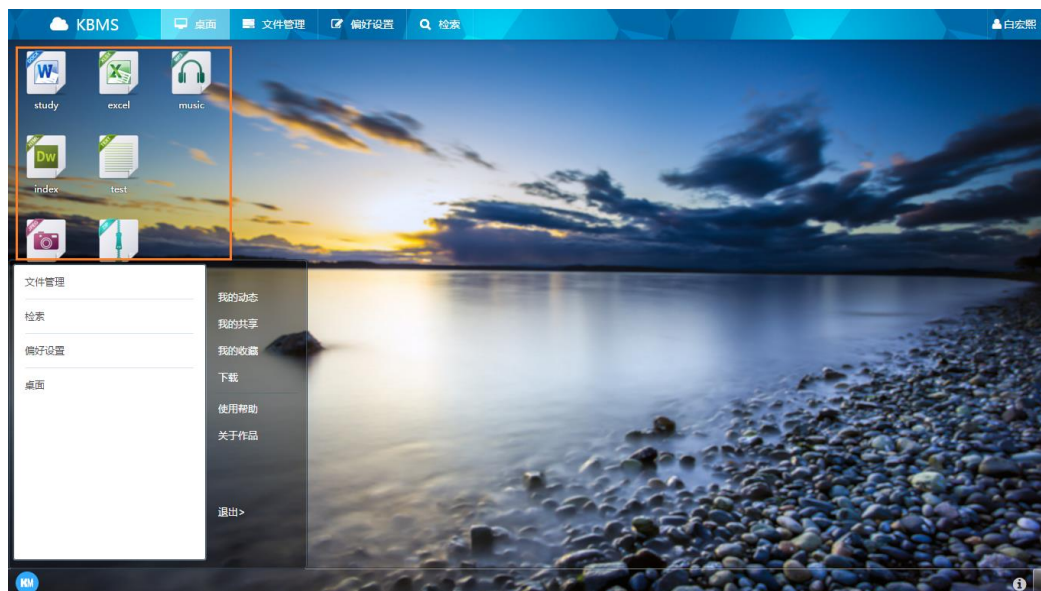


图 19 桌面系统界面

2、右键点击文件，可对文件进行一系列操作。

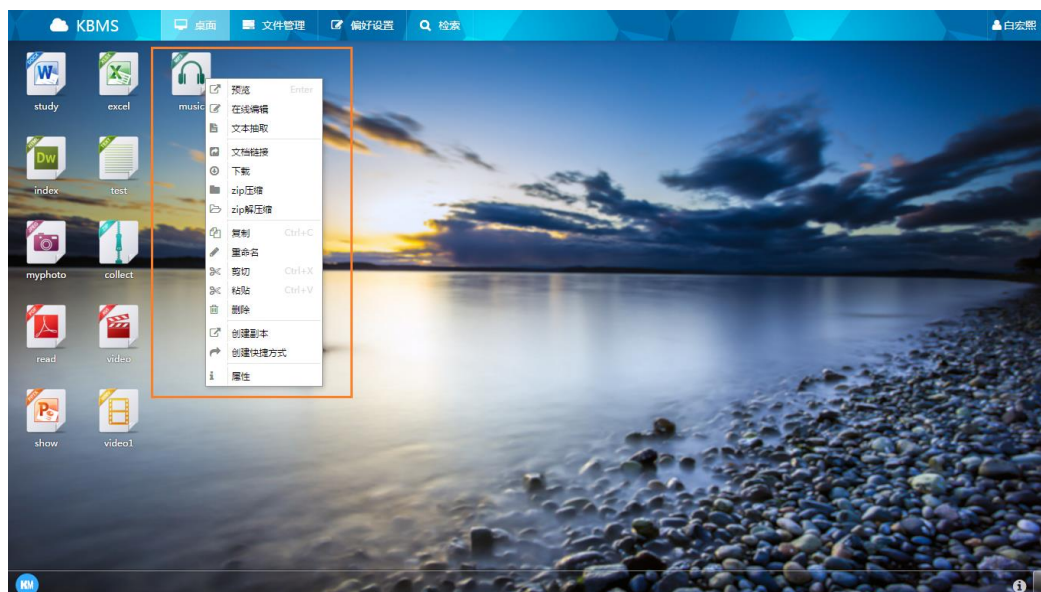


图 20 桌面系统操作文件界面

3、系统桌面部分主要提供文件的在线查看功能，供用户快速查找文档。

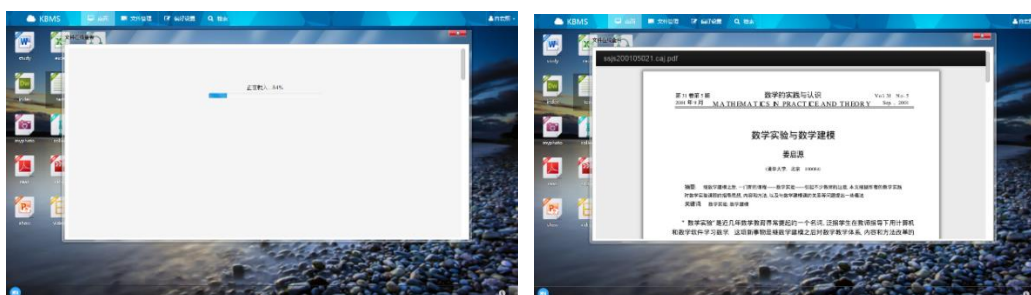


图 21 桌面系统预览文件界面

(4) 系统检索页面

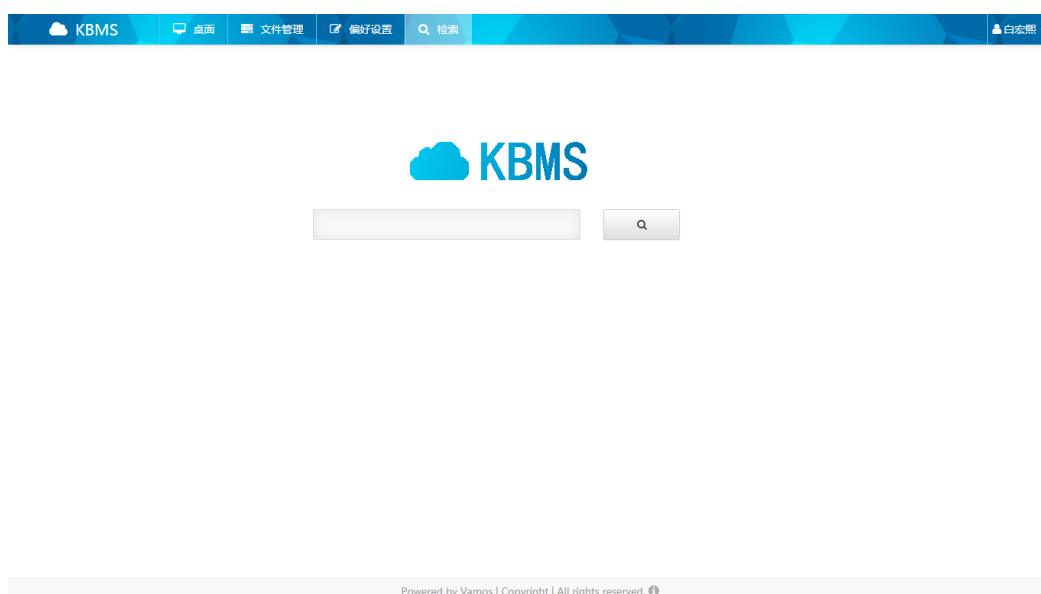


图 22 检索部分初始界面

将检索结果高亮显示，并显示检索的响应时间和检索时间，应用全文检索的方式，提高检索的速度，同时检索结果的列表显示，使得用户更清晰地查找到所需的知识。



图 23 检索结果显示界面

#### (5) 用户偏好设置页面

除去修改个人资料信息外，系统提供给用户更多个性化的选择，更改系统的“颜值”。并且对用户自身的知识操作做一定的管理，进行分享、收藏等操作，并且针对分享的操作可设置知识的“时限”，保证用户知识的安全。

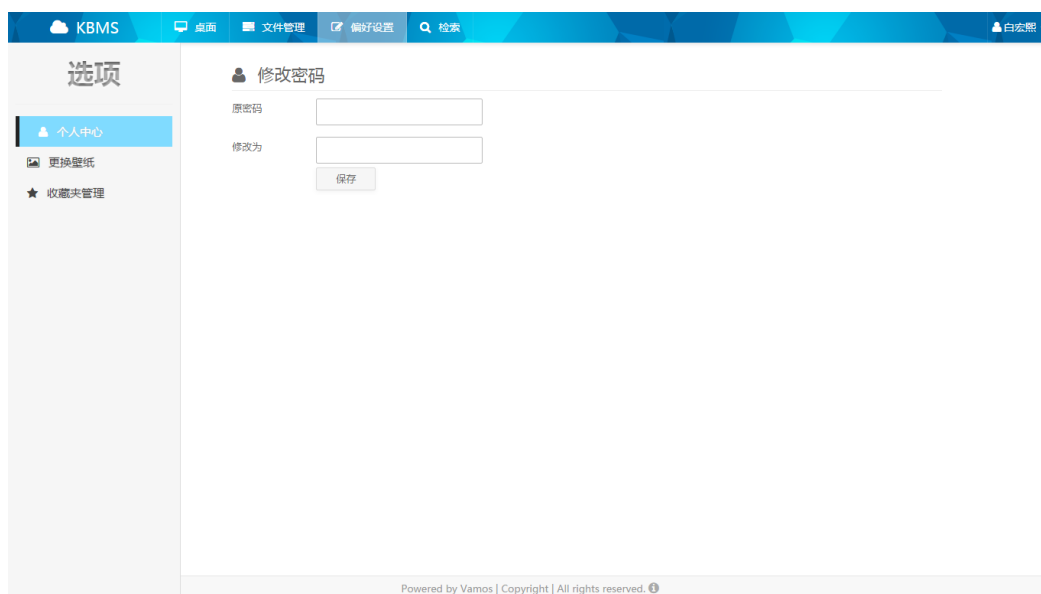


图 24 用户修改个人资料界面



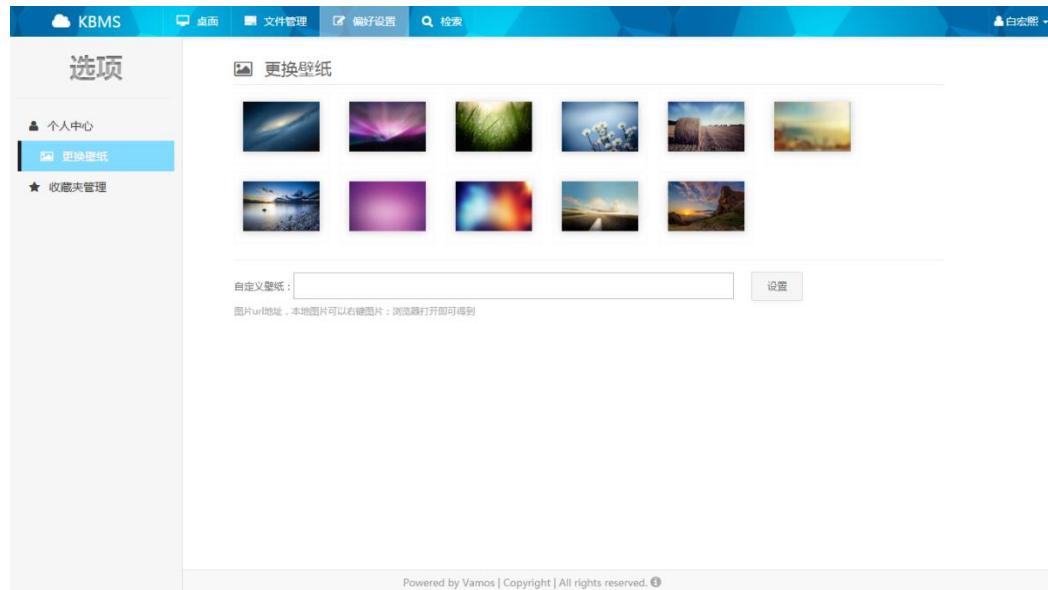


图 25 用户个性化界面

#### (6) 用户动态部分页面

将用户最近的知识操作以动态的方式展示，作为用户使用知识的纪录者，帮助用户查找及及时概括形成自己的知识。



图 26 用户动态展示部分界面

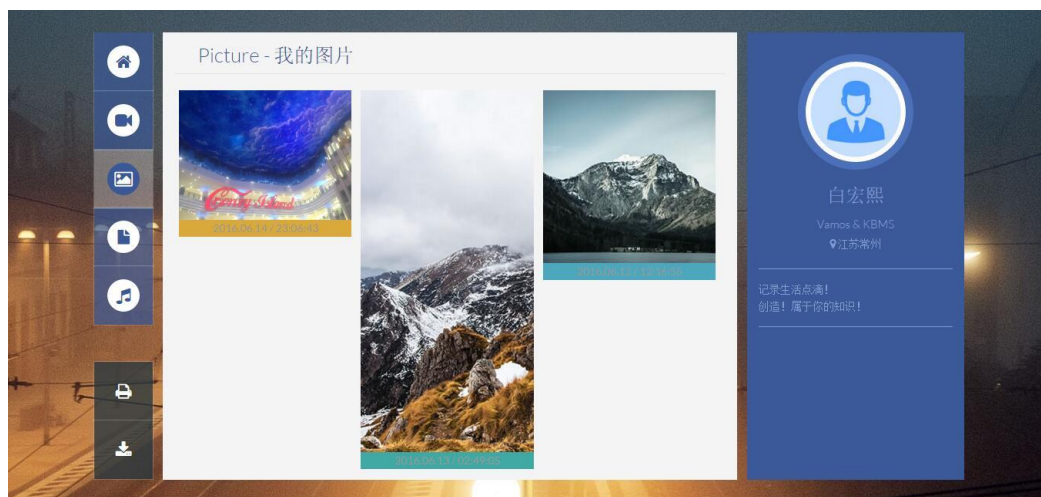


图 27 用户动态分类展示部分界面

## 4 . 接口设计

### 4.1 外部接口

知识库管理系统的短信验证和在线预览功能需调用外部接口

#### (1) SMS 短消息通信协议接口

客户端请求: HTTP-POST

服务器返回数据: Json XML 类返回

#### (2) IdocV 接口

客户端请求: HTTP-POST

服务器返回数据: Json 类返回

### 4.2 内部接口

客户端向知识管理系统服务器发送权限验证请求、数据访问请求或数据更新请求。服务器根据对应的请求进行应答, 将数据返回给客户端。

#### (1) Solr 实时检索接口

客户端请求: HTTP-GET

服务器返回数据: Json XML 类返回

#### (2) PageOffice 文档在线编辑接口

客户端请求: HTTP-GET

服务器返回数据: Json 类返回

平台的内部接口如图 28 所示



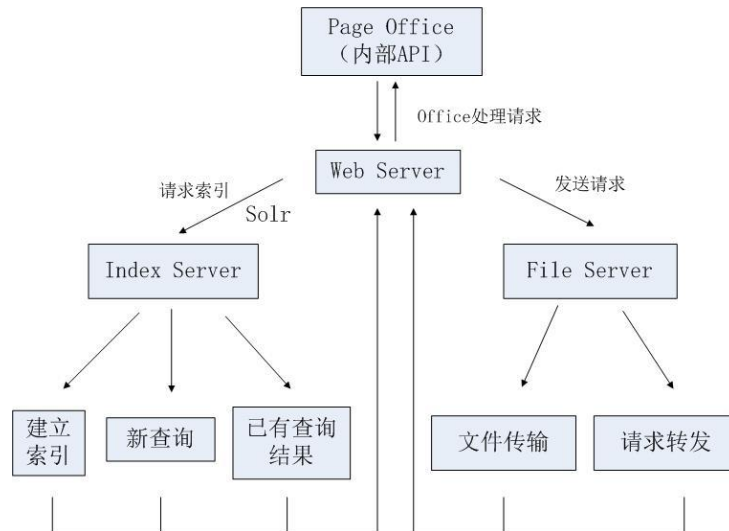


图 28 平台的内部接口

## 5 . 数据库设计

数据库名称：Kmos

### 5.1 数据库表列表

序号	表名称	表描述
1	desktop_file	存储生成桌面快捷方式文件的基本信息
2	documents	存储摘要关键字信息
3	file_association	文件关联打开方式字典表
4	file_info	存储文件的基本信息
5	file_manipulate_delmovemod	存储文件的行为信息
6	file_manipulate_retrieval	存储文件的检索记录
7	file_manipulate_scan	存储文件的浏览记录
8	file_type	存储文件类型信息
9	global_recommender_temp_res	存储临时推荐结果
10	global_recommender_temp	存储临时推荐中间过程信息
11	knowledge	存储推荐文件信息
12	public_share_list	存储分享文件信息
13	share_list_url_convert	存储分享链接转换信息
14	shortcut	存储缩略图信息



字段	说明	类型	可空	备注
user_id	ID	int	NO	主键，自增
nick_name	昵称	varchar	YES	
tel	电话	varchar	YES	
mail	邮件	varchar	YES	
pwd	密码	char	YES	
signin_time	注册时间	char	YES	
root_location	文件根目录	char	YES	

表 2 用户信息表

2、桌面快捷方式文件信息表（desktop\_file）如表 3 所示

字段	说明	类型	可空	备注
desktop_id	ID	int unsigned	NO	主键，自增
file_path	文件路径	char	YES	
file_name	文件名称	char	YES	
file_id	文件 ID	int	YES	
thumbnail_path	缩略图路径	char	YES	
thumbnail_name	缩略图名称	char	YES	
user_id	用户 ID	int	NO	主键

表 3 桌面快捷方式文件信息表

3、摘要信息表（documents）如表 4 所示

字段	说明	类型	可空	备注
id	ID	int	NO	主键，自增
group_id	组别 ID	Int	NO	
date_added	添加日期	datetime	NO	
title	标题	varchar	NO	
content	文本内容	text	NO	

表 4 摘要信息表

4、文件关联打开方式字典表（file\_association）如表 5 所示

字段	说明	类型	可空	备注
Association_id	ID	Int unsigned	NO	主键，自增
name	打开方式名称	Char	YES	

表 5 文件关联打开方式字典表

5、文件基本信息表（file\_info）如表 6 所示

字段	说明	类型	可空	备注
file_id	ID	int	NO	主键，自增
before_name	原始文件名	char	YES	
suffix	文件原始后缀	int	YES	
type	文件真实类型	int	YES	
size	文件大小	int	YES	
upload_ip	上传 IP	int	YES	
upload_time	上传时间	int	YES	
upload_location	上传地点	int	YES	
tool	上传工具	int	YES	
user_id	用户 ID	int	YES	
save_path	保存路径	int	YES	
after_name	系统存储名称	int	YES	
state	文件状态	int	YES	
create_time	文件创建时间	int	YES	
last_modification_time	文件最后一次 修改时间	int	YES	
last_access_time	最后一次访问 时间	int	YES	
belong_groups	文件所属组别	int	YES	
right0	文件权限	char	YES	
doc_summary	文件摘要	text	YES	
identity_of_named_entity	文件类别	int	YES	
sentiment_analisis	情感	int	YES	
tag	标签	text	YES	

keywords	关键字	int	YES	
senmatic_associate	文本关联	int	YES	
give_a_mark	评分	int	YES	
if_special_focus	是否特别关注	int	YES	
thumbnail_path	缩略图路径	char	YES	
thumbnail_name	缩略图名称	char	YES	
file_association	文件关联	char	YES	

表 6 文件基本信息表

6、文件行为信息表（file\_manipulate\_delmovemod）如表 7 所示

字段	说明	类型	可空	备注
events_id	ID	int	NO	主键，自增
file_id	文件 ID	int	YES	
user_id	用户 ID	int	YES	
type	文件类型	varchar	YES	
happen_time	行为发生时间	varchar	YES	
del_type	是否为删除行为	varchar	YES	
mod_type	是否为修改行为	char	YES	
download_type	是否为下载行为	varchar	YES	
share_type	是否为分享行为	char	YES	
move_type	是否为移动行为	char	YES	

表 7 文件行为信息表

7、文件检索记录表（file\_manipulate\_retrieval）如表 8 所示

字段	说明	类型	可空	备注
retrieval_id	ID	int	NO	主键，自增
goal_file_id	目标文件 ID	int	YES	
user_id	用户 ID	int	YES	
index_times	检索次数	int	YES	
index_time	检索时间	int	YES	
key_words	关键字	char	YES	

index_location	结果文件路径	int	YES	
on_hook_time	挂机时间	int	YES	

表 8 文件检索记录表

8、文件浏览记录表（file\_manipulate\_scan）如表 9 所示

字段	说明	类型	可空	备注
scan_id	ID	int	NO	主键，自增
file_id	文件 ID	int	YES	
user_id	用户 ID	int	YES	
scan_times	浏览次数	varchar	YES	
scan_time	浏览时间	int	YES	
scan_continue_time	浏览时长	varchar	YES	
page_rolling	页面滚动次数	varchar	YES	
mouse_track	鼠标轨迹	varchar	YES	

表 9 文件浏览记录表

9、文件类型信息表（file\_type）如表 10 所示

字段	说明	类型	可空	备注
type_id	ID	int	NO	主键，自增
name	文件后缀名	char	YES	
icon	图标名称	char	YES	

表 10 文件类型信息表

10、临时推荐结果记录表（global\_recommender\_temp\_res）如表 11 所示

字段	说明	类型	可空	备注
temp_res_id	ID	int	NO	主键，自增
thing	被推荐文件 ID	int	YES	
file_id	原文件 ID	int	YES	

表 11 临时推荐结果记录表

11、临时推荐中间过程信息记录表（global\_recommender\_temp）如表 12 所示

字段	说明	类型	可空	备注
----	----	----	----	----

temp_id	ID	int	NO	主键，自增
user_id	用户 ID	int	YES	
active_time	活动时间	char	YES	
active_position	活动位置	char	YES	
goal_file_id	推荐目标文件 ID	char	YES	
state	状态	int	YES	

表 12 临时推荐中间过程信息记录表

12、推荐文件信息表（knowledge）如表 13 所示

字段	说明	类型	可空	备注
knowledge_id	ID	int	NO	主键，自增
file_id	文件 ID	int	YES	
related_file_id	关联文件 ID	int	YES	
related_votes	关联评分	double	YES	

表 13 推荐文件信息表

13、分享文件信息表（public\_share\_list）如表 14 所示

字段	说明	类型	可空	备注
share_list_id	ID	int	NO	主键，自增
mdshcode	防伪码	char	YES	
physics_path	物理路径	char	YES	
state	分享状态	int	YES	
timestamp	分享时间戳	char	YES	
ttl	存活时间	char	YES	
user_id	分享人 ID	int	YES	
timestampdie	失效时间戳	char	YES	

表 14 分享文件信息表

14、分享链接转换信息表（share\_list\_url\_convert）如表 15 所示

字段	说明	类型	可空	备注
share_list_url_id	ID	int	NO	主键，自增

old_url	真实链接	varchar	YES	
new_url	短连接	varchar	YES	

表 15 分享链接转换信息表

15、缩略图信息表（shortcut）如表 16 所示

字段	说明	类型	可空	备注
shortcut_id	ID	int	NO	主键，自增
file_id	文件 ID	int	YES	
c_timestamp	创建时间戳	char	YES	
save_path	存储路径	char	YES	
before_name	文件名称	char	YES	
user_id	用户 ID	int	YES	

表 16 缩略图信息表

16、今日短信记录表（today\_tel\_history）如表 17 所示

字段	说明	类型	可空	备注
today_tel_history_id	ID	int	NO	主键，自增
tel	手机号码	char	YES	
time	发送时间	char	YES	
times	发送次数	int	YES	
states	发送状态	int	YES	

表 17 今日短信记录表

17、用户全局行为信息表（user\_overall\_situation\_events）如表 18 所示

字段	说明	类型	可空	备注
event_id	ID	int	NO	主键，自增
user_id	用户 ID	int	YES	
active_time	活动时间	char	YES	
active_position	活动地点	char	YES	
on_hook_time	挂机时间	char	YES	

表 18 用户全局行为信息表



## 6. 出错处理设计

### 6.1 出错输出信息

在用户使用错误的数据或访问没有权限的数据后，系统给出提示：“对不起，你非法使用数据，没有权限！”而且用户的密码管理可以允许用户修改自己的密码，不允许用户的匿名登录。

### 6.2 出错处理对策（如设置后备、性能降级、恢复及再启动等）

如：由于数据在数据库中已经有备份，故在系统出错后可以依靠数据库的恢复功能，并且依靠日志文件使系统再启动，就算系统崩溃用户数据也不会丢失或遭到破坏。但有可能占用更多的数据存储空间，权衡措施由用户来决定。

## 7. 安全保密设计

为提供系统安全的信息环境，知识库管理系统具有以下措施保证信息安全：

### （1）客户端权限设定

利用客户端身份验证，将用户登录账号和密码发送至应用服务器端，通过服务器端进行数据处理后在数据库服务器中进行匹配检索。匹配成功则发送登录成功验证信息给用户。

### （2）数据安全设定

为了保证数据库的安全性，将数据实际操作部分在应用服务器端完成。数据库端只能向应用服务器发送数据操作请求。

### （3）数据存储保密

对于需要保密的数据，提供了 UTF-8 转义存储的方式，利用密文进行数据存储，保证安全性。并在路径存储时采用 Base64 编码方式，在便于数据传输的同时，客户端与服务器端都能够进行编解码。

### （4）链接分享保护

对于用户分享的文件链接，我们对其进行不可逆加密，为其提供一个映射数据库内容的唯一标识，不会泄露文件的信息，以保护文件服务器不受攻击。

### （5）数据库保护

sql 注入攻击，采用预编译以免数据库遭受攻击；此外，在进行全文检索时，用正则表达式剔除敏感字符，以保护 solr 全文检索索引免受攻击。

## 8. 维护设计

为了方便程序的后期维护，知识库管理系统提供了以下措施：

- （1）在数据库设计过程中，使用基础数据结构，方便数据库的修改和移植。
- （2）采用增量索引，避免每次新数据导入重复建立索引，以提高效率。
- （3）提供了知识库管理系统智能分析在线服务平台，提供通联关系分析技术支持及智能推荐等服务。