



需求规格说明书

——BookWise 书籍垂直搜索引擎

《软件工程管理》课程 G13 小组

组长：王伟杰

组员：安俊霖、包越、陈华杰、黄
乐诚、刘逸杰

2023.11.25

修改历史

修订日期	版本号	修改人	修改内容	审核人
2023-11-14	Version1.0	全体组员	初稿	王伟杰
2023-11-25	Version2.0	全体组员	需求分析	王伟杰
2023-11-27	Version3.0	全体组员	终稿	王伟杰

目录

1 引言	7
1.1 编写目的.....	7
1.2 软件项目背景	7
1.3 目标读者和阅读建议	8
1.3.1 软件需求规格说明书的组成部分	8
1.3.2 软件需求规格说明书面向的读者对象	10
1.4 项目目标	10
1.5 参考文献与资料	11
2 总体描述.....	12
2.1 产品基本信息	12
2.2 产品前景.....	12
2.3 用户类及其特征	13
2.4 产品功能概述	14
2.4.1 搜索词条模块	14
2.4.2 搜索结果模块	15
2.4.3 内容展示模块	15
2.4.4 用户注册模块	17
2.4.5 用户登录模块	18

2.4.6 个人信息模块	19
2.5 设计和实现上的约束	19
2.5.1 数据存储	20
2.5.2 网络服务吞吐	20
2.5.3 数据安全	20
2.6 假设和依赖	22
2.7 开发、测试和运行环境	23
2.8 项目文档.....	23
2.8.1 描述类文档	24
2.8.2 过程类文档	24
2.8.3 参考类文档	24
2.9 产品开发项目设计人员	24
3 系统需求分析概述	26
3.1 系统总用例图	26
3.2 需求优先级分析	27
3.3 系统模块架构图	28
4 用户场景.....	28
4.1 关键词搜索	28
4.2 用户注册.....	29
4.3 用户登录.....	30
4.4 搜索结果展示	31
4.5 模糊搜索.....	33

4.6 搜索历史.....	34
4.7 热门搜索	35
4.8 结果相关度展示	36
4.9 结果详情展示	37
5 数据流图.....	38
5.1 搜索引擎子系统数据流图.....	38
5.2 网站维护子系统数据流图.....	38
6 接口需求.....	39
6.1 用户接口	39
6.2 客户端接口	39
6.2.1 客户端硬件需求.....	39
6.2.1 客户端软件需求.....	40
6.3 服务器接口:	40
6.3.1 硬件	40
6.3.2 软件	40
6.4 网络接口	41
9 非功能性需求.....	41
9.1 性能需求	41
9.2 界面需求	42
9.3 易用性	43
9.4 安全性	44
9.5 系统的完整性	45

9.6 可移植性	45
9.7 可扩展性	45
9.8 可维护性	46
9.9 可靠性	46
10 数据字典.....	47
10.1 数据项.....	47
10.2 数据存储	48
10.3 数据流.....	49
10.4 处理逻辑	50
11 业务规则与业务算法	53
11.1 业务规则.....	53
11.2 业务算法.....	55
12. DM 图.....	56
13.上下文图	58
14.附录.....	58
14.1 术语与缩写解释	58
14.2 待确定问题清单	60

1 引言

1.1 编写目的

(1) 定义系统的范围和边界，包括系统的功能、非功能需求以及用户界面和交互设计等方面的要求。

(2) 明确系统的输入和输出，以及与外部系统或组件的接口和交互方式，确保系统能够与其他系统或组件进行有效的集成和交互。

(3) 指导开发团队在软件设计和实施过程中，明确系统的架构和设计约束，以便保证系统的稳定性、可靠性和安全性。

(4) 帮助测试团队理解系统的功能和性能需求，从而能够有效地进行测试计划和测试用例的编写和执行。

(5) 帮助项目管理团队进行进度和成本的控制，明确系统开发的目标和里程碑，以便能够合理安排资源和时间。

(6) 作为后续系统维护和升级的依据，帮助后续团队理解系统的设计和实现，从而能够进行有效的维护和升级工作。

1.2 软件项目背景

本项目是浙江大学 2023 学年《软件工程管理》课程的课程项目，目标是实现一个书籍垂直搜索引擎。

随着互联网的发展，人们获取信息的方式发生了巨大的变化。特别是在图书领域，读者需要更加高效和便捷地获取他们感兴趣的书籍

信息。传统的搜索引擎虽然可以提供大量的信息，但是在特定领域的深度搜索和精准推荐方面仍然存在不足。因此，开发一款专注于书籍领域的垂直搜索引擎成为了当下的需求。

该书籍垂直搜索引擎项目旨在为读者提供一个专注于书籍的搜索平台，通过整合各种图书信息资源，包括线上线下书店、图书馆、电子书平台等，为用户提供更加全面、深入的书籍搜索和推荐服务。这将有助于读者更快速、更准确地找到他们感兴趣的书籍，推动图书行业的数字化和信息化发展。

项目团队将致力于开发一款功能强大、用户体验友好的书籍垂直搜索引擎，通过技术手段提高图书信息的获取效率，为用户提供更加便捷的阅读体验。同时，该搜索引擎还将为图书出版商、书店和图书馆等机构提供更广泛的推广和宣传渠道，促进图书产业的发展和繁荣。

综上所述，书籍垂直搜索引擎项目的开发将填补当前书籍领域搜索服务的空白，满足用户对于高效获取图书信息的需求，促进图书产业的数字化和信息化发展。

1.3 目标读者和阅读建议

1.3.1 软件需求规格说明书的组成部分

书籍垂直搜索引擎项目的需求规格书是对该项目的功能、性能、界面设计、数据存储、安全性等方面的详细描述和规划。其主要包括以下内容：

- ✓ 项目背景和目标：介绍书籍垂直搜索引擎项目的背景和目标，包括项目的意义、目标用户群体、预期效果等。
- ✓ 功能需求：详细描述书籍垂直搜索引擎的功能需求，包括搜索功能、图书信息展示、个性化推荐、用户管理、数据整合等方面的功能。
- ✓ 性能需求：描述书籍垂直搜索引擎的性能需求，包括系统响应时间、并发访问量、数据处理能力等方面的性能要求。
- ✓ 界面设计：对书籍垂直搜索引擎的界面设计进行详细描述，包括搜索界面、图书信息展示界面、用户个人中心界面等方面的设计要求。
- ✓ 数据存储：描述书籍垂直搜索引擎的数据存储需求，包括对图书信息、用户信息、搜索记录等数据的存储和管理要求。
- ✓ 安全性需求：描述书籍垂直搜索引擎的安全性需求，包括用户数据保护、系统安全防护、数据传输加密等方面的安全要求。
- ✓ 部署和维护需求：描述书籍垂直搜索引擎的部署和维护需求，包括系统部署环境、备份与恢复策略、系统维护计划等方面的要求。

- ✓ 需求确认和验收标准：明确书籍垂直搜索引擎项目的需求确认和验收标准，以便在项目开发和交付过程中进行验收和确认。

1.3.2 软件需求规格说明书面向的读者对象

- ✓ 开发团队：开发团队需要了解软件需求规格说明书中的功能和性能要求，以便根据规格书进行软件开发和测试。
- ✓ 测试团队：测试团队需要了解软件需求规格说明书中的功能和性能要求，以便根据规格书进行软件测试，确保软件的质量和稳定性。
- ✓ 项目管理人员：项目管理人员需要了解软件需求规格说明书中的项目进度和交付要求，以便进行项目管理和进度跟踪。
- ✓ 客户和用户：客户和用户需要了解软件需求规格说明书中的功能和性能要求，以便了解软件的功能和使用方式。同时，他们也可以通过规格书来确认软件是否符合他们的需求和期望。

1.4 项目目标

- ✓ 提供用户一个专注于图书的搜索引擎平台，让用户能够更快速、准确地找到他们想要的图书信息。
- ✓ 为用户提供一个可以针对特定图书类别、作者、出版社等进行精确搜索的工具，帮助他们更好地找到所需的图书。
- ✓ 收集并整合各大图书商、出版社的图书信息，为用户提供一个全面的图书搜索平台，让用户能够在一个地方找到他们需要的图书。
- ✓ 提供个性化的推荐功能，根据用户的搜索历史和偏好，推荐相关的图书给用户，提高用户体验。
- ✓ 不断改进搜索算法和用户界面，提高搜索的准确性和用户体验，成为用户信赖的图书搜索工具。

1.5 参考文献与资料

- 《软件需求规格说明书》IEEE 标准 830-1998
- 《软件需求》Karl Wieggers, Joy Beatty
- 《敏捷软件需求》Dean Leffingwell
- 《软件需求工程》（第三版）Ian Sommerville
- 《需求管理实践指南》（第二版）Karl Wieggers, Joy Beatty

- [G13] “书籍垂直搜索引擎”项目计划书

2 总体描述

2.1 产品基本信息

产品名称

BookWise 书籍垂直搜索引擎

任务提出者

浙江大学软件工程管理课程组——任课教师：邵健、邹强；

助教：李昊泽、苏展、俞乃铖

开发者

浙江大学软件工程管理 G13 小组

12 月 3 日前完成初步开发，1 月 5 日前完成最终系统。

2.2 产品前景

书籍垂直搜索系统具有广阔的产品前景，主要体现在以下几个方面：

- ✓ 用户需求：随着人们对知识和信息的需求不断增长，书籍垂直搜索系统能够满足用户对特定书籍的准确、快速搜索需求，提高用

户获取信息的效率和便利性。

- ✓ 电子书市场增长：随着电子书市场的不断扩大，用户对于电子书的需求也在增加。书籍垂直搜索系统可以帮助用户在海量的电子书中快速找到所需内容，满足用户对电子书的搜索和获取需求。
- ✓ 个性化推荐：书籍垂直搜索系统可以根据用户的搜索历史和兴趣，提供个性化的书籍推荐服务，增加用户粘性，提高用户满意度。
- ✓ 数据挖掘和分析：通过书籍垂直搜索系统收集的用户搜索行为和偏好数据，可以进行数据挖掘和分析，为出版社、书店等提供市场分析和用户行为分析的数据支持。
- ✓ 跨境电商：随着跨境电商的发展，书籍垂直搜索系统可以为用户提供跨境图书购买的便利，拓展了图书市场的边界。
- ✓ 教育和学术研究：对于学生、教师和研究人员来说，书籍垂直搜索系统可以为其提供丰富的学术资源和研究文献，满足其学术和研究需求。

综上所述，书籍垂直搜索系统具有巨大的市场潜力和发展前景，能够满足用户需求，为出版商、电商平台等提供数据支持，同时也能够服务于教育、学术等领域，是一个具有广阔前景的产品。

2.3 用户类及其特征

我们将产品使用方定义为一个用户类，即对书籍信息有检索需求的人。用

户类的具体特征如下表所示

使用者	主要特征	备注信息
用户	可以通过关键词进行搜索	支持高级搜索、模糊搜索、拼音搜索等多项搜索功能
	可以查看搜索结果	搜索结果默认按相关度排序，允许用户根据不同条件对搜索结果进行排序
	可以查看书籍具体信息	包括书籍封面、简介、作者、出版信息等
	可以查看热搜	根据搜索频次排列热搜
	可以获取个性化推荐	根据用户的兴趣和历史行为提供个性化的书籍推荐。
	可以进行账号注册	用户可以在系统填写信息注册自己的账号
	可以进行账号登录	用户可以登录系统
	可以修改个人信息	用户可以进入自己的个人信息界面，对自己的信息进行修改

表 2-3-1 用户类及其特征

2.4 产品功能概述

2.4.1 搜索词条模块

该模块的基本功能为：用户输入关键字，从而对书籍的名称、作者、主题、出版日期、ISBN、书籍语言等内容进行搜索。搜索将支持模糊搜索、精确搜索。

以下为部分功能释义：

模糊搜索：系统能够允许被搜索信息和搜索提问之间存在一定的差异，这种差异就是“模糊”在搜索中的含义。SQL 模糊查询是指使用类似 SQL 语句中的“LIKE”关键字进行模糊查询的技术。它可以帮助我们查找特定字符串的数据，而不需要确切地知道要查找的字符串，只要知道一些关键字即可。

2.4.2 搜索结果模块

用户输入关键字并点击搜索按钮后，系统会将搜索结果呈现给用户。搜索结果将以卡片列表的形式进行呈现，呈现内容为相关的书籍词条。用户可以点击卡片，跳转页面查看详细的内容，包括书籍的内容、简介、作者信息、背景等等。

搜索内容将按照搜索的相关度和点击量进行排序。同时也支持用户自主选择条件来对搜索内容进行筛选和排序。

此外，系统将会根据用户的搜索内容，对用户进行相关书籍和作者的个性化推荐，展示一些同类的词条方便用户跳转。而系统同样也会展现一些搜索热词为用户提供搜索参考。

2.4.3 内容展示模块

用户对搜索结果进行点击后，将进行具体内容的展示。

对于书籍主要包括：

- ✓ **书籍信息展示：**展示搜索结果中符合条件的书籍的基本信息，包括书名、作者、出版社、出版日期、ISBN 号码等。用户可以通过这些信息快速了解到搜索到的书籍的基本情况。用户可在此界面选择收藏该书籍。
- ✓ **书籍封面展示：**搜索结果页面通常会展示搜索到的书籍的封面图片，这有助于用户更直观地识别和辨别搜索结果中的书籍。
- ✓ **书籍摘要和简介：**搜索结果页面可能会展示书籍的摘要或简介内容，这有助于用户快速了解书籍的主要内容和特点，从而决定是否进一步查看或购买。
- ✓ **评价和评分信息：**一些书籍垂直搜索系统会展示用户对于搜索到的书籍的评价和评分，这有助于用户了解书籍的受欢迎程度和质量。同时用户可以在评价界面对书籍进行评分与评价。
- ✓ **相关推荐书籍：**搜索结果页面可能会展示与用户搜索书籍相关的其他推荐书籍，这有助于用户进一步扩展阅读范围，发现更多感兴趣的书籍。
- ✓ **作者信息：**对于搜索结果中的书籍，系统可能会展示相关作者的基本信息、其他作品等内容，同时设置跳转链接，点击可展示作者详细信息界面，帮助用户更全面地了解作者的背景和作品。

对于作者主要包括：

- ✓ **作者简介：**作者信息模块通常会展示作者的简介，包括作者的生平、教育背景、职业经历等内容，帮助用户更全面地了解作者的背景和经历。用户可在此界面选择订阅作者。
- ✓ **代表作品：**作者信息模块通常会列举作者的代表作品，方便用户快速了解作者的作品风格和特点。
- ✓ **其他作品：**除了代表作品外，作者信息模块还可能会列举作者的其他作品，帮助用户更全面地了解作者的创作成果。
- ✓ **评价和评分：**一些书籍垂直搜索系统会展示用户对于作者的评价和评分，这有助于用户了解作者的受欢迎程度和质量。
- ✓ **相关推荐作者：**作者信息模块可能会推荐与搜索到的作者相关的其他作者，帮助用户进一步扩展阅读范围，发现更多感兴趣的作者。

2.4.4 用户注册模块

用于新账户的注册，利于系统管理用户账户，同时可以确保用户信息的准确性与安全性，主要包括以下功能：

- ✓ **用户信息收集：**用户注册模块会收集用户的基本信息，如用户名、密码、电子邮件地址等。有些注册模块还会要求用户提供其他信

息，如姓名、性别、出生日期等。

- ✓ **验证和安全性：**注册模块通常会包括验证用户输入的信息的功能，以确保用户提供的信息是准确和有效的。它还会包括安全性功能，如密码加密和验证码，以保护用户账户的安全。
- ✓ **用户协议和隐私政策：**注册模块可能会要求用户同意网站或应用程序的用户协议和隐私政策，以确保用户了解和同意相关的条款和条件。
- ✓ **邮件确认和激活：**一些注册模块会要求用户通过电子邮件确认其注册信息，并激活其账户，以确保用户提供的电子邮件地址是有效的。
- ✓ **错误处理和反馈：**注册模块通常会包括错误处理和反馈功能，以帮助用户解决注册过程中可能遇到的问题。

2.4.5 用户登录模块

用户登录模块是一个用于验证用户身份并授权访问系统资源的功能模块，是系统安全性的重要组成部分，它能够有效地保护系统资源免受未经授权的访问。

它通常包括以下功能：

- ✓ **用户身份验证：**用户输入用户名和密码进行身份验证，系统通过验证用户提供的凭据来确认用户身份。
- ✓ **访问控制：**一旦用户身份验证成功，系统会授予用户访问特定资源的权限。这可以通过角色和权限管理来实现，确保用户只能访问其被授权的资源。
- ✓ **安全性：**用户登录模块也负责确保用户登录的安全性，例如通过使用加密技术来保护用户的凭据和会话管理来避免未经授权的访问。
- ✓ **记录和审计：**用户登录模块通常也会记录用户的登录活动，以便系统管理员能够进行审计和监控用户的访问行为。

2.4.6 个人信息模块

个人信息模块用于存储和管理用户的个人基本信息。它包括用户名、年龄、性别、联系方式、地址、账号、密码等基本信息。用户可以通过个人信息模块方便地管理自己的个人信息，并在需要进行更新和修改。个人信息模块还可以用于身份验证和安全管理，确保用户信息的安全性和隐私保护。

2.5 设计和实现上的约束

系统的设计、编码以及维护将遵照文档严格进行。

在具体设计和实现上，按照以下约束进行。

2.5.1 数据存储

考虑到大部分开发人员都熟悉 MySQL 数据库系统，因此统一选择使用 MySQL 作为项目的数据库系统。

2.5.2 网络服务吞吐

网络服务的吞吐量是指网络服务在单位时间内处理的请求或数据量。它是衡量网络服务性能的重要指标之一，对于评估网络服务的性能和容量具有重要意义。根据项目要求，本项目将可以提供对外服务的能力，确保至少 1000 名用户可以同时使用本系统。

2.5.3 数据安全

保证数据安全是网络服务开发中非常重要的一环。以下是一些常用的方法来确保数据安全：

- ✓ 数据加密：对于敏感数据，应该使用加密算法进行加密存储，以防止数据在存储或传输过程中被窃取。常用的加密算法包括对称加密算法（如 AES）和非对称加密算法（如 RSA）。
- ✓ 访问控制：通过访问控制列表（ACL）或者基于角色的访问控制

（RBAC）等方法，限制对数据的访问权限，确保只有授权的用户或系统可以访问特定的数据。

- ✓ 输入验证：在接收用户输入数据时，进行有效性验证和过滤，以防止恶意输入或注入攻击。例如，对于 Web 应用，需要对用户输入的数据进行 XSS（跨站脚本攻击）和 SQL 注入等安全性检查。
- ✓ 安全传输：使用安全的传输协议，如 HTTPS，来保证数据在网络传输过程中的安全性，防止数据被中间人攻击窃取或篡改。
- ✓ 安全存储：对于数据库中的数据，需要采取安全存储的措施，如对数据库进行加密、定期备份、灾难恢复计划等。
- ✓ 安全审计：记录和审计数据访问和操作日志，以便在发生安全事件时进行追踪和调查。
- ✓ 安全更新：及时更新和修复系统和组件的安全漏洞，确保系统的安全性。
- ✓ 安全培训：对开发人员和系统管理员进行安全培训，提高其对安全问题的认识和应对能力。

- ✓ 安全监控：部署安全监控系统，及时发现和应对安全事件。

2.6 假设和依赖

假设与依赖编号	假设与依赖描述
AS-1	不考虑一些不可控的突发性意外
AS-2	系统需要依赖于网络数据获取，包括从数据源获取书籍信息、更新数据等，需要确保网络连接的可用性和稳定性。
AS-3	系统可能假设数据源提供的书籍信息质量较高，包括准确性、完整性、更新性等，需要确保数据质量能够满足搜索引擎的需求。
AS-4	系统可能对书籍信息的格式和结构有一定的假设，包括书名、作者、出版日期、ISBN 码等信息，需要确保数据源提供的信息符合搜索引擎的预期格式和结构。
AS-5	系统可以确保能够获取到准确、全面的书籍信息。
AS-6	系统可能基于特定的搜索算法进行书籍信息的检索和排序，需要确保搜索算法的有效性和性能。
AS-7	系统需要确保书籍信息的安全性和用户隐私的保护，需要

	依赖于安全措施和合规性要求。
AS-8	系统多轮的迭代开发过程有相应的人员负责维护与更新

表 2-6-1 假设和依赖

2.7 开发、测试和运行环境

开发环境

- a. JetBrains IDEA IntelliJ（教育版免费）
- b. Visual Studio Code（开源免费）

测试工具

- a. Postman（社区版免费）

运行环境

- a. 联网 PC
- b. 浏览器

2.8 项目文档

所有的项目文档将分为三类：关于系统功能和使用的描述类文档、在系统开发过程中产出的过程类文档以及为开发提供方向和依据的

参考类文档。这些文档 将为整个开发过程以及后续的系统使用提供依据。

2.8.1 描述类文档

描述性文档会阐明软件或设备的功能性、使用方法、技术规范、界面设计等信息，帮助用户理解和使用产品或工具，主要用于帮助用户理解如何使用 BookWise 书籍垂直搜索引擎，其典型代表为用户手册。

2.8.2 过程类文档

过程类文档用于记录项目开发过程, 有助于确保特定的任务或过程能够按照一致的方法和步骤进行, 并提供指导和支持给执行者。这类文档的典型代表为测试报告。

2.8.3 参考类文档

参考类文档用于对系统的功能、接口、需求等设计进行明确约定与记录, 便于以团队协作的方式进行系统的迭代开发, 供开发人员参考执行。这类文档的典型代表为需求规格说明书。

2.9 产品开发项目设计人员

编号	角色	角色职责	具体负责人
1	项目经理	在预算范围内领导小组成员完成所有的工作	王伟杰
2	产品经理	负责进行市场调研，了解和设计具体的用户需求，并领导制定本项目的需求规格说明书	包越
3	设计总监	建立系统框架，负责系统的总体架构和具体模块的设计，参加技术评审	黄乐诚
4	测试团队经理	负责领导测试团队完成整个项目的测试	陈华杰
5	美术监督	设计网站原型，并负责维持网站的美术质量	安俊霖
6	质量经理	带领软件质量监督巡视组成员制定质量保证计划等文档，并对实际的文档和代码的质量进行监督评审	黄乐诚
7	开发团队	负责完成搜索引擎的具体开发和部署工作	全体成员
8	测试团队	负责完成整个系统的测试和维护	全体成员
9	软件质量监督师	负责实时对质量经理以及项目经理提供的项目进度和项目实际开发时的差异提出报告并	全体成员

表 2-9-1 产品开发项目设计人员

3 系统需求分析概述

3.1 系统总用例图

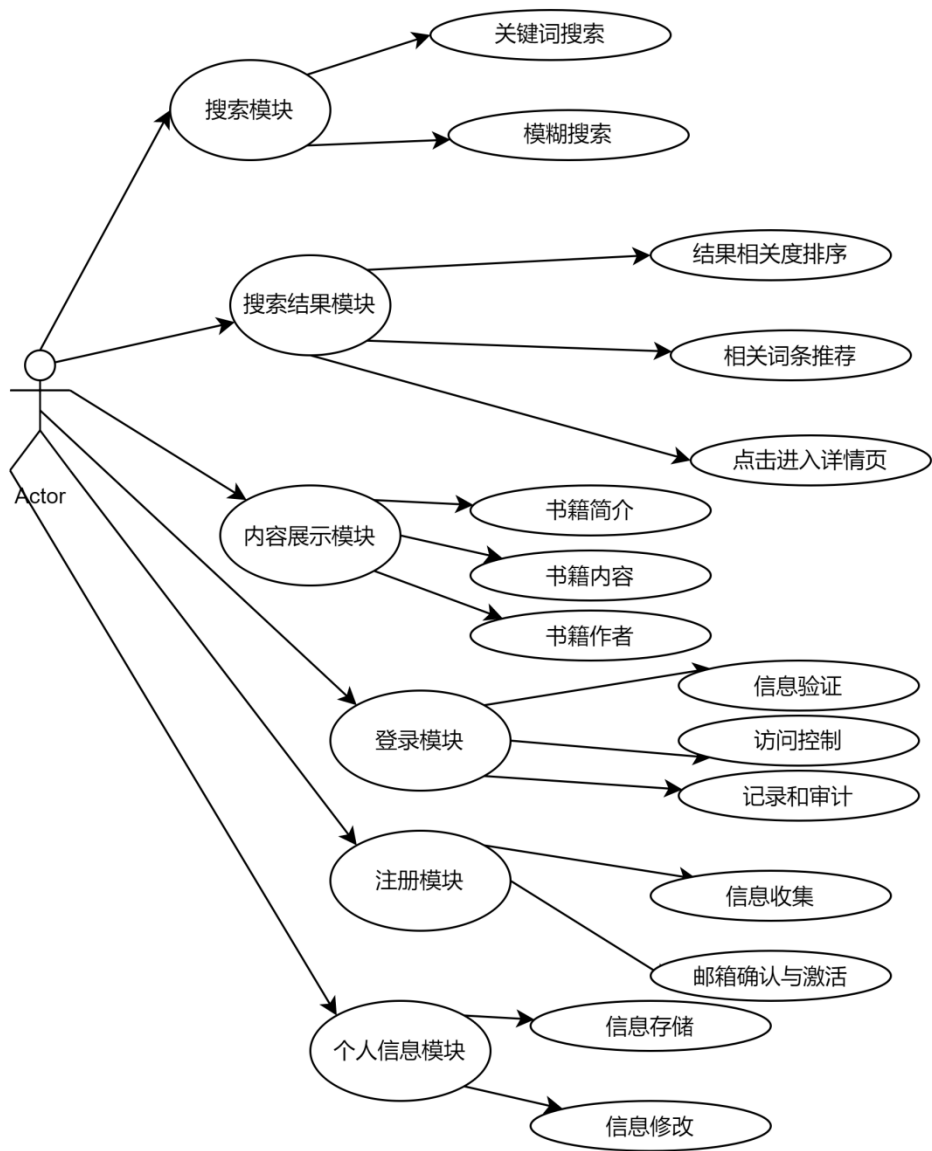


图 3-1-1 系统用例总图

3.2 需求优先级分析

本需求优先级分析模块依据用户量-发生频率四象限进行评估，优先解决大用户量的高频问题，保证基础体验，最后解决少量用户的低频问题，优化体验。

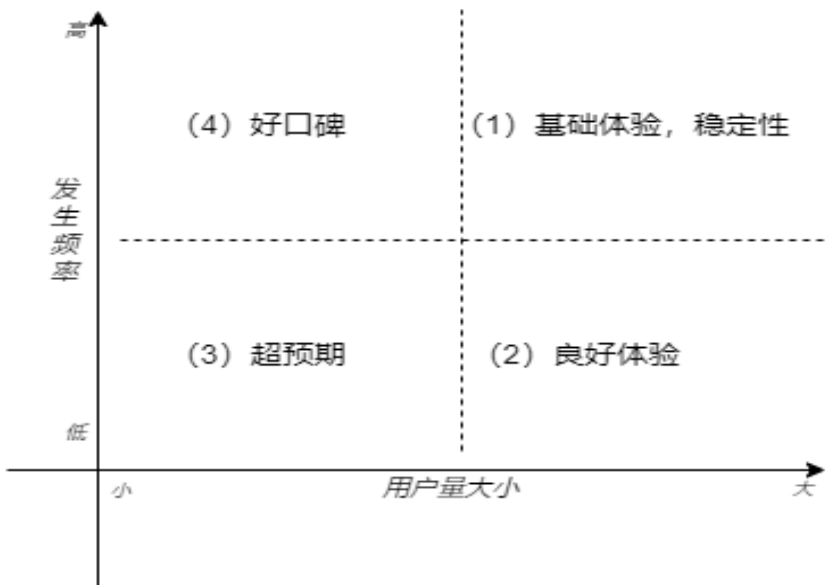


表 3-2-1 用户量-发生频率四象限

根据基本的用户需求分析，我们可以把定义的用户需求按照优先级作如下排序：

用例编号	用例名称	发生频率	用户量	优先级
SE-01	关键词搜索	高	大	高
SE-02	搜索结果展示	高	大	高
SE-03	模糊搜索	高	大	高
SE-04	搜索历史	中	大	中
SE-05	热门搜索	中	大	中

SE-06	结果相关度排序	高	大	中
SE-07	相关词条推荐	中	中	中
SE-08	账号注册	高	大	高
SE-09	账号登录	高	大	高
SE-10	个人信息修改	中	大	高

表 3-2-1 需求优先级分析

3.3 系统模块架构图

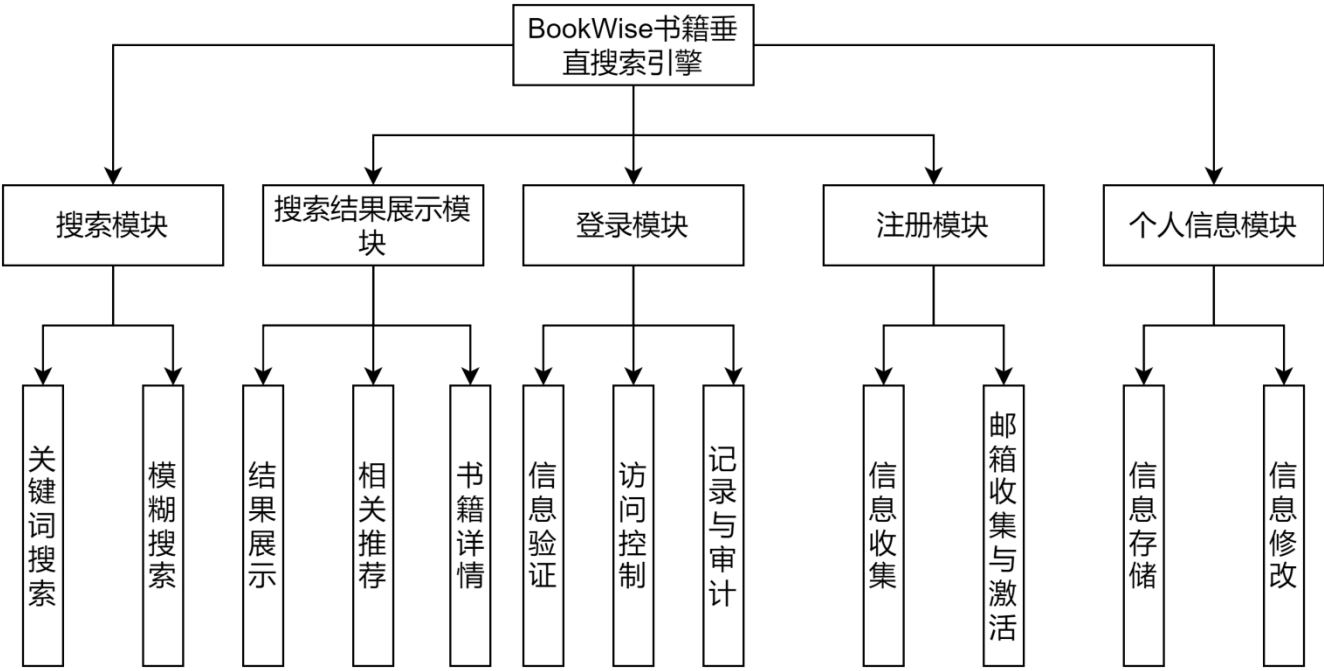


图 3-3-1 系统模块架构图

4 用户场景

4.1 关键词搜索

SE-01	关键词搜索
主要参与者	普通用户
目标	查询用户想要的书籍信息
前提条件	1. 系统正常运行。 2. 用户已进入搜索模块
触发器	用户输入书籍名称进行搜索
工作流程	1. 用户进入 BookWise 搜索引擎主页。 2. 用户输入关键词。 3. 用户点击搜索按钮。 4. 跳转搜索结果界面。
异常	1. 用户输入不合法信息。（发生于工作流程 2） 处理：系统提示“含不合法信息！请重新输入！” 2. 点击搜索后页面显示异常 处理：根据网站的联系方式联系系统管理员
优先级	高
使用频率	高
输入	用户想要搜索的关键词
输出	搜索结果页面

表 4-2-1 关键词过滤

4.2 用户注册

SE-02	用户注册
-------	------

主要参与者	普通用户
目标	用户进行系统内的账号注册
前提条件	1. 系统正常运行。 2. 用户有可用邮箱进行绑定。
触发器	用户点击注册。
工作流程	1. 用户进入 BookWise 搜索引擎主页。 2. 用户点击注册按钮进入注册页面。 3. 用户填写相关信息 4. 完成注册。
异常	页面显示异常 处理：根据网站的联系方式联系系统管理员
优先级	高
使用频率	高
输入	注册请求
输出	注册结果

4.3 用户登录

SE-03	用户登录
主要参与者	普通用户
目标	支持用户输入搜索“关键词+关键词”类语句

前提条件	1.系统正常运行
触发器	用户输入语句，点击搜索
工作流程	1.用户进入 BookWise 搜索引擎主页。 2.用户点击打开登陆界面。 3.用户输入账号密码。 4.系统进行认证。 5.登录成功或失败
异常	1.用户输入的账号密码不匹配。（ 处理：系统提示“账号密码输入有 误！请重新输入！” 2.点击搜索后页面显示异常 处理：根据网站的联系方式联系系 统管理员
优先级	高
使用频率	高
输入	账号密码、登录请求
输出	登录结果

4.4 搜索结果展示

SE-06	搜索结果展示
主要参与者	普通用户

目标	用户搜索后可以查看搜索结果
前提条件	<ol style="list-style-type: none"> 1.系统正常运行 2.用户可正常打开相关页面
触发器	用户输入语句，点击搜索，进入搜索页面
工作流程	<ol style="list-style-type: none"> 1.用户进入 BookWise 搜索引擎主页。 2.用户输入关键词。 3.用户点击搜索按钮。 4.跳转搜索结果界面，显示结果卡片列表，结果默认按相关度进行排序。
异常	<ol style="list-style-type: none"> 1.用户输入不合法信息。（发生于工作流程 2） 处理：系统提示“含不合法信息！请重新输入！” 2.点击搜索后页面显示异常 处理：根据网站的联系方式联系系统管理员
优先级	高
使用频率	高
输入	关键词、语句
输出	搜索结果显示页面

4.5 模糊搜索

SE-07	模糊搜索
主要参与者	普通用户
目标	系统允许被搜索信息和搜索提问之间存在一定的差异
前提条件	1.系统正常运行 2.用户可正常打开相关页面
触发器	用户输入语句，点击搜索，进入搜索页面
工作流程	1.用户进入 BookWise 搜索引擎主页。 2.用户输入关键词。 3.用户点击搜索按钮。 4.跳转搜索结果界面，显示结果卡片列表，结果默认按相关度进行排序。
异常	1.用户输入不合法信息。（发生于工作流程 2） 处理：系统提示“含不合法信息！请重新输入！” 2.点击搜索后页面显示异常 处理：根据网站的联系方式联系系统管理员
优先级	高
使用频率	高

输入	与完整关键词有差异的语句
输出	搜索结果显示页面

4.6 搜索历史

SE-10	搜索历史
主要参与者	普通用户
目标	记录并显示用户曾搜索过的词条
前提条件	1.系统正常运行 2.用户可正常打开相关页面 3.系统正确记录用户搜索历史
触发器	用户点开搜索首页
工作流程	1.用户进入 BookWise 搜索引擎主页。 2.搜索框下显示最近 5 条搜索记录
异常	无
优先级	高
使用频率	高
输入	无
输出	搜索历史展示界面

4.7 热门搜索

SE-11	热门搜索
主要参与者	普通用户
目标	显示热门搜索内容
前提条件	1.系统正常运行 2.用户可正常打开相关页面 3.系统根据用户搜索统计信息得到热门搜索词条列表。
触发器	用户点开搜索首页
工作流程	1.用户进入 BookWise 搜索引擎主页。 2.用户输入关键词。 3.用户点击搜索按钮。 4.跳转搜索结果界面，显示结果卡片列表，结果默认按相关度进行排序。
异常	无
优先级	高
使用频率	高
输入	无

输出

热门搜索展示界面

4.8 结果相关度展示

SE-12	结果相关度展示
主要参与者	普通用户
目标	系统默认按相关度排序显示搜索结果，用户可以点击相关度按钮，选择搜索结果按照相关度排序
前提条件	1.系统正常运行 2.用户可正常打开相关页面 3.系统正确计算相关度
触发器	用户输入关键词进行搜索
工作流程	1.用户进入 BookWise 搜索引擎主页。 2.搜索框下显示热门搜索内容
异常	无
优先级	高
使用频率	高
输入	无
输出	按照相关度排序的搜索结果页面

4.9 结果详情展示

SE-14	结果详情展示
主要参与者	普通用户
目标	用户点击某项特定搜索结果，进入书籍或作者详情界面。
前提条件	1.系统正常运行 2.用户可正常打开相关页面 3.搜索结果正常
触发器	用户搜索后点击某结果
工作流程	1. 用户完成搜索 2. 用户点击某结果词条 3. 系统展示该词条详细信息
异常	1. 无法显示书籍或作者详情。 处理：可能发生了数据更新或丢失，或由网络原因导致。 系统提示用户稍后刷新重试或联系管理员
优先级	高
使用频率	高
输入	鼠标点击
输出	结果详情界面

5 数据流图

5.1 搜索引擎子系统数据流图

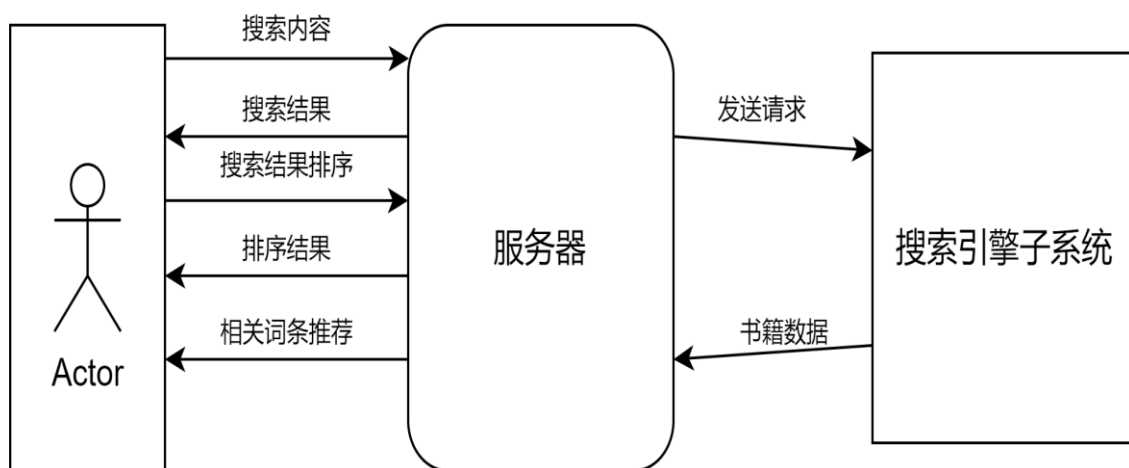


图 5-1-1 搜索引擎子系统数据流图

5.2 网站维护子系统数据流图

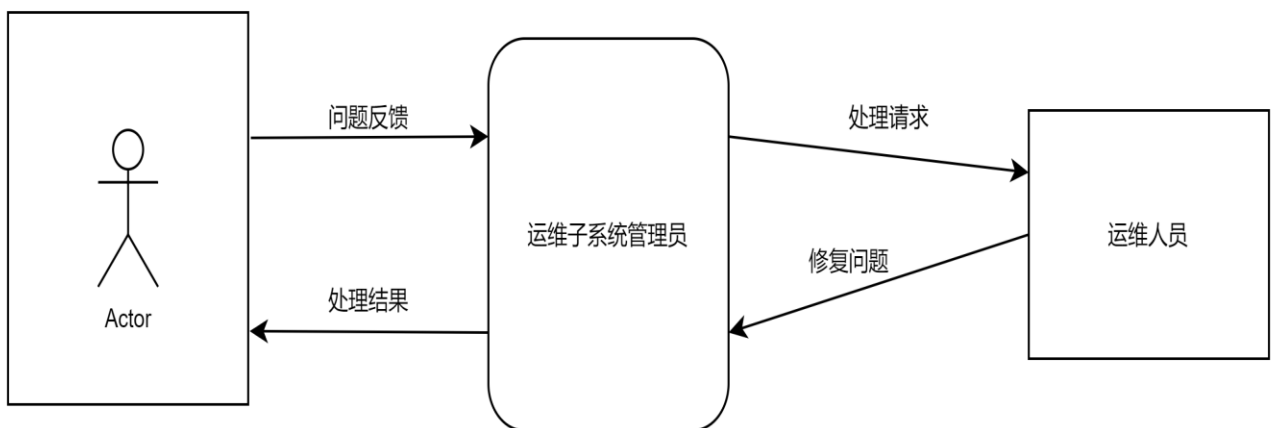


图 5-2-1 网站维护子系统数据流图

6 接口需求

6.1 用户接口

软件需要提供用户友好的界面。版面设计上需要做到重点突出、方便浏览，方便用户使用。图标和图形的选择遵照简单化、标准化的原则，使的所选择使用的图形和图标更容易被大众接受和认可。。用户接口的设计应该考虑用户的需求和习惯，以使用户能够方便、快捷地与程序或设备进行交互。用户接口设计的目标是提供良好的用户体验，使用户能够轻松地完成任务并享受使用过程。

6.2 客户端接口

6.2.1 客户端硬件需求

处理器	Intel i5 及以上
内存	2G 以上
存储	磁盘
网卡	速率 10Mbps 以上
备份	RAID 5

6.2.1 客户端软件需求

操作系统	Windows10 , Linux Ubuntu 18.04
浏览器	IE, Google Chrome 等主流浏览器

6.3 服务器接口：

6.3.1 硬件

处理器	pentium 2G
内存	3G 以上
存储	1T 以上
备份	RAID 5

6.3.2 软件

操作系统	Linux Ubuntu 18.04
服务器	Nginx
数据库	MySQL

6.4 网络接口

软件需要能够通过网络与其他设备或系统进行通信，包括局域网、互联网等。

9 非功能性需求

9.1 性能需求

- 响应时间：在 95%的情况下，一般时段响应时间不超过 1.5 秒，高峰时段不超过 4 秒；在网络畅通时，页面刷新时间不超过 10 秒；
- 并发数：系统需要能够同时处理大量用户的搜索请求，保证在高峰期也能保持良好的性能。系统可以同时满足 10,0 个用户请求；
- 资源使用率：CPU 占用率 \leq 50%，内存占用率 \leq 50%。
- 可用性：该系统应实现多 Web 浏览器支持：在大多数流行的 Web 浏览器中正确显示和执行，包括 Firefox、Chrome、Edge 等。
- 系统需要能够及时更新书籍数据库的信息，保证搜索结果的准确性和完整性。

- 精准搜索：系统需要能够准确地匹配用户输入的关键词，返回相关度高的书籍搜索结果。

9.2 界面需求

- 要将书籍信息用良好的可视化界面呈现。
- 对于一个书籍垂直搜索系统界面，需要做到设计简洁、功能明晰，使 80% 用户能够在看到该系统首页的 5 秒内意识到这是一个书籍信息搜索系统。
- 简约是一个被广泛认可的设计原则，即尽量避免过多的复杂元素和信息。一个干净、直观的界面能够让用户迅速理解应用的功能，减少用户混淆和迷失。在设计中，去除不必要的装饰和元素，专注于核心功能，有助于提供更好的用户体验。
- 一致性是用户界面设计中的关键原则，它指的是在整个应用中保持相似的设计和交互方式。无论用户在应用的哪个部分，他们都应该感到熟悉和舒适。保持一致性有助于用户快速学习应用的使用方法，减少混淆和困惑。使用相同的颜色、字体、按钮样式等，使界面的各个部分具有统一的外观。

- 用户界面的设计应该以用户为中心，注重提供简单、直观且易于操作的界面。可用性是用户体验的核心，设计师应该考虑用户的需求和行为模式。合理的布局、易于理解的标签和图标，以及易于点击的按钮等，都是提高可用性的方法。
- 导航系统是用户与应用交互的关键。一个简洁而直观的导航系统可以让用户轻松地浏览应用，找到他们所需的功能和信息。以下是一些设计简洁导航系统的建议：使用明确的导航标签，避免使用晦涩难懂的术语；采用层级结构，将相关功能和内容进行分组；提供一致的导航布局，使用户在不同页面间能够快速适应；使用图标和颜色来帮助用户快速识别不同的导航选项。

9.3 易用性

- 易理解性：使用文档易于理解，所有文档语言简练，内容应该与产品实际情况相一致，且所有文档中的语句无歧义，能很好地解释每一步骤的含义，用户一看便清楚。
- 易学习性：菜单选项易找到，各图标含义明确、简单易懂，操作步骤向导解释清楚、易懂，产品具有很好的引导性，即一个软件客户不用看说明书都能正确地使用。

- 易操作性：要求人机界面友好、界面设计科学合理、操作简单等。
操作简单的软件让用户可以直接根据窗口提示进行使用，不去过多地参考使用说明书。各项功能流程设计直接明了，尽量每一个操作都很简单的完成。在一个业务功能中可以关联了解其相关的业务数据，具有层次感。合理的默认值和可选项的预先设定，避免过多的手工操作。如果某个操作将产生失败的后果，该功能执行应是可逆的，或程序应给出该失败的明显警告，并且在执行该命令前要求确认。一旦出现操作失败，及时的信息反馈是非常重要的，没有处理结果或者是处理过程没有相关信息反馈的系统不是一个优秀的系统。流畅自然的操作感觉，来源于每一次操作都是最合理的设计。
- 易吸引力：拥有良好外观和界面设计，吸引客户的眼球。

9.4 安全性

- 严格权限访问控制，用户在经过身份认证后，只能访问其权限范围内的数据，只能进行其权限范围内的操作；
- 提供运行日志管理及安全审计功能，可追踪系统的历史使用情况；
- 能经受来自互联网的一般性恶意攻击。如病毒（包括木马）攻击、口令猜测攻击、黑客入侵等；
- 保证数据信息的安全，做到若有数据库误删除的情况，可以使用

撤销操作进行修复。并及时进行信息备份防止丢失；

- 系统应该保护未开放下载权限的资料不被下载。

9.5 系统的完整性

- 该系统包括数据备份、恢复、日志管理、垃圾数据清除等基本功能。
- 系统能够成功地执行各项功能，并提供正确的结果。

9.6 可移植性

- 系统是便携式的，从一个操作系统移动到其他操作系统不会产生任何问题。

9.7 可扩展性

- 可实现负载均衡，日后若信息量较大，系统可相应增加服务器实现扩展；
- 由于系统模块之间的低耦合性，在增加新功能模块时，对于旧的功能模块影响不大。
- 能够根据需求进行横向和纵向扩展，以应对日益增长的用户和数据量。

9.8 可维护性

- 从接到修改请求后，对于普通修改应在 1~2 天内完成；
- 对于评估后为重大需求或设计修改应在 1 周内完成；
- 90%的 BUG 修改时间不超过 1 个工作日，其他不超过 2 个工作日。
- 做到高内聚、低耦合的系统模块划分。开发者需要充分考虑模块内部结构的紧密型及模块间联系的独立性。便于后期对系统的修改与维护。
- 该系统需要具有良好的监控和日志记录功能，以便管理员能够及时发现和解决潜在的问题，并对系统运行情况进行及时跟踪和分析。
- 需要能够进行全面的测试，以确保其功能和性能符合预期，并且能够及时发现和修复潜在的问题。
- 搜索引擎的代码和架构应该易于理解和维护，以便开发人员能够快速诊断和修复问题。

9.9 可靠性

- 分布式系统在发生通信异常时会先暂时切断连接，等问题修复完成后重新连接，恢复软件的运行；
- 系统缺陷率每 1,000 小时最多发生 1 次故障；
- 因软件系统的失效而造成无法完成业务。

10 数据字典

10.1 数据项

编号	数据项名称	类型	取值范围	描述	举例
E1	搜索关键字	字符	长度范围:(0, 256)	用户输入的需要查询的关键字	围城
E2	书籍名称	字符	长度范围:(0, 256)	/	《围城》
E3	书籍介绍	字符	/	/	《围城》是钱钟书所著的长篇小说，是中国现代文学史上一部风格独特的讽刺小说。被誉为“新儒林外史”。故事主要写抗战初期知识分子的群像。

E4	书籍作者	字符	长度范围:(0, /	钱钟书
			256)	
E5	出版时间	时间	/	1947 年

10.2 数据存储

文件或数据库名	书籍信息
描述	书籍的信息集合
组成	书籍简介+书籍内容+书记作者+平均评分+评价集合+评分人数+出版时间+书籍类型
相关处理	用户查询书籍信息 用户对书籍评分评论 管理员添加/修改书籍信息

文件或数据库名	作者信息
描述	作者的信息集合
组成	作者简介+作者生平+代表作+作品集合+写作风格
相关处理	用户查询作者信息 管理员添加/修改作者信息

文件或数据库名	用户信息
描述	用户的信息集合
组成	用户名+账号+密码+邮箱
相关处理	用户注册 用户登录 用户修改个人信息

10.3 数据流

数据流名称	描述	组成	来源	终点
搜索请求	用户在搜索框中输入书籍信息搜索请求	搜索关键字	用户	搜索结果
搜索结果	系统根据用户的搜索请求返回书籍信息搜索结果	书籍信息，简介，作者，评分评价等	系统	搜索结果界面
注册请求	用户进行账	用户输入信	用户	注册结果

	号注册	息		
登录请求	用户进行账号登录	用户账号密码	用户	登录结果
修改个人信息请求	用户进行个人信息修改	用户要修改的信息	用户	修改结果

10.4 处理逻辑

处理名	用户搜索查询书籍信息
描述	用户输入关键词，系统返回相关书籍信息
输入数据流	用户输入的关键词
输出数据流	书籍信息结果
存取的数据库	书籍信息数据库
处理逻辑	<p>读取书籍信息数据库</p> <p>检索是否有与用户输入相关的书籍信息</p> <p>IF 有 THEN</p> <p>返回相关书籍信息结果</p> <p>ELSE</p> <p>返回无相关信息提示</p>

处理名	用户搜索查询作者信息
描述	用户输入关键词，系统返回相关作者信息
输入数据流	用户输入的关键词
输出数据流	作者信息结果
存取的数据库	作者信息数据库
处理逻辑	<p>读取作者信息数据库</p> <p>检索是否有与用户输入相关的作者信息</p> <p>IF 有 THEN</p> <p>返回相关作者信息结果</p> <p>ELSE</p> <p>返回无相关信息提示</p>

处理名	用户进行账号注册
描述	用户点开注册界面，输入信息，进行注册
输入数据流	填写的信息，注册请求
输出数据流	注册结果

存取的数据库	用户信息数据库
处理逻辑	<p>读取用户信息数据库</p> <p>检索数据库中是否有已有用户信息，</p> <p>IF 有 THEN</p> <p>返回请勿重复注册提示</p> <p>ELSE</p> <p>返回成功注册</p>

处理名	用户登录
描述	用户输入账号密码，进行验证与登录
输入数据流	账号密码，登录请求
输出数据流	登录结果
存取的数据库	个人信息数据库
处理逻辑	<p>读取个人信息数据库</p> <p>检索是否有用户输入的账号信息，以及其输入的账号密码是否匹配</p> <p>IF 无或不匹配 THEN</p> <p>返回账号密码错误提示</p>

	ELSE
	登录成功

处理名	用户修改个人信息
描述	用户点击个人信息界面，进行修改
输入数据流	填写的信息，个人信息修改申请
输出数据流	信息修改结果
存取的数据库	个人信息数据库
处理逻辑	读取个人信息数据库 更新信息

11 业务规则与业务算法

11.1 业务规则

标识符	规则定义	规则类型	静态或动态	来源
BR-01	用户仅能搜索相关书籍的信息	约束	静态	数据限制

BR-02	用户搜索关键词支持与或非逻辑运算	动作处罚规则	动态	用户
BR-03	用户搜索时语言仅限于中文(包括拼音)以及英文	约束	静态	管理策略
BR-04	用户在搜索后跳转到的搜索结果页面的列表中点击某项即可进入详情页面	动作触发规则	动态	用户
BR-05	搜索结果列表默认按照搜索的相关度排序	约束	静态	管理策略
BR-06	用户注册时填写用户名	约束	静态	数据限制

	有字数限制			
BR-07	用户仅能查看自己的个人信息界面	约束	静态	管理策略

表 11-1 业务规则

11.2 业务算法

- **检索算法：**根据用户输入的关键词，搜索引擎需要使用检索算法从图书数据库中找到相关的图书信息。这可能涉及到关键词匹配、相关性排序、过滤等技术。
- **排名算法：**搜索结果的排名通常需要考虑多个因素，包括关键词的相关性、图书的权威性、用户的搜索历史等。排名算法需要综合考虑这些因素，为用户提供最合适的搜索结果。
- **自然语言处理算法：**搜索引擎需要能够理解用户输入的自然语言，将用户的搜索意图转化为数据库查询。自然语言处理算法通常包括分词、词性标注、语义理解等技术。能正确识别处理输入中的“与”“或”“非”逻辑词。

- **数据挖掘算法：**搜索引擎可能会根据用户的搜索行为和图书信息，进行数据挖掘分析，发现用户的兴趣和行为规律，从而优化搜索结果和推荐结果。

12. DM 图

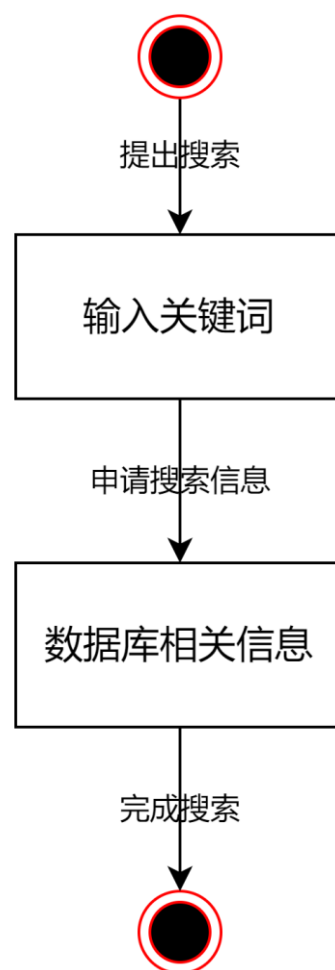


图 11.1 来自搜索信息的对话图

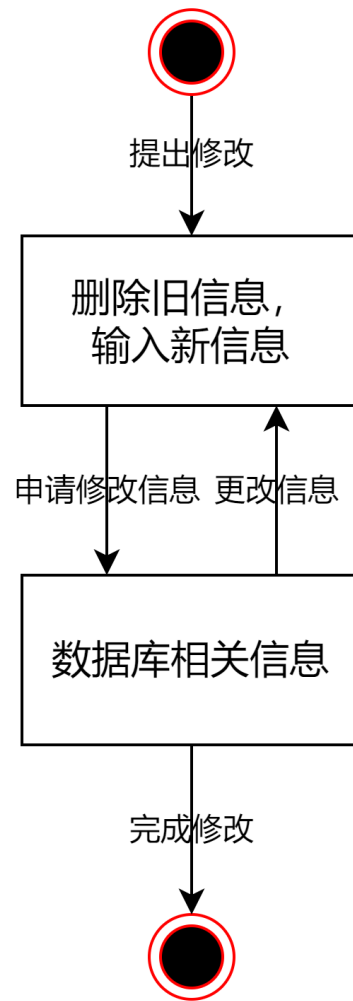
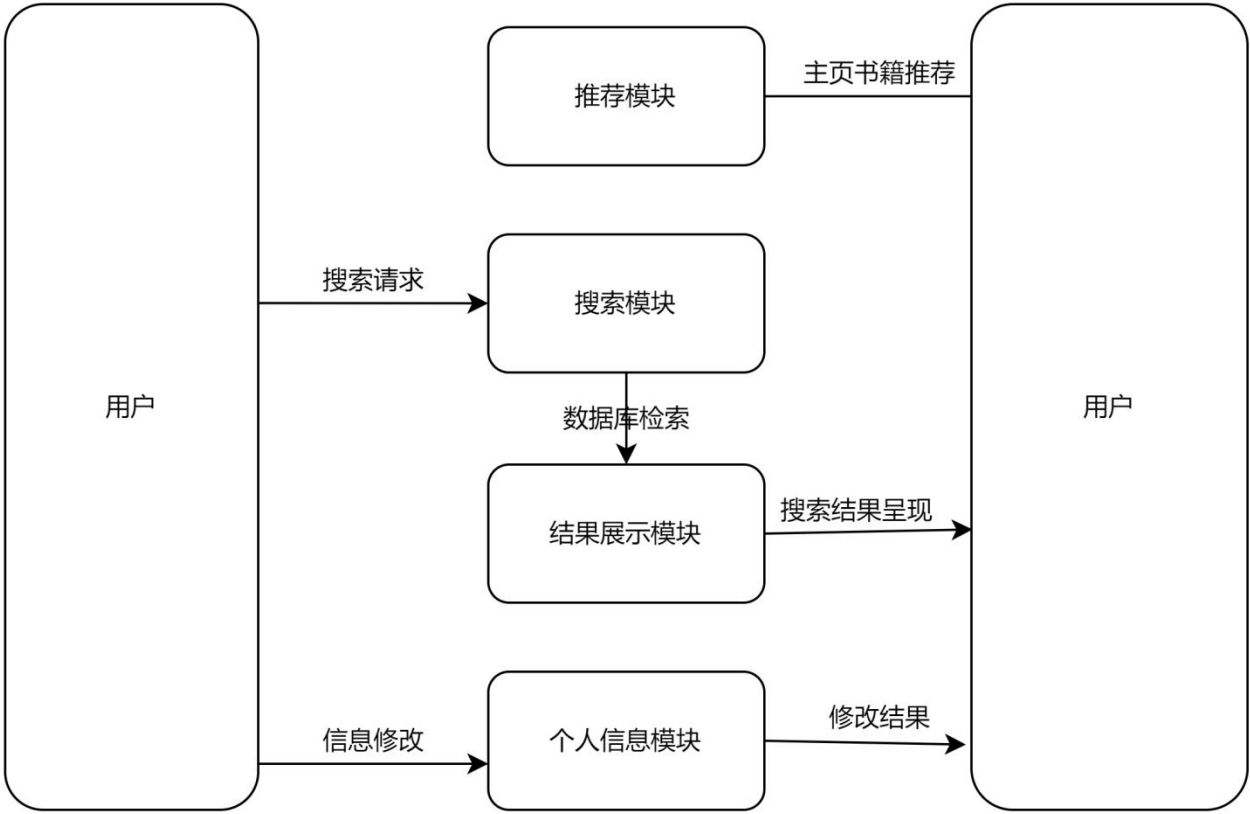


图 11.2 来自个人信息修改的对话图

13.上下文图



14.附录

14.1 术语与缩写解释

名词	解释
软件工程	把系统的、规范的、可度量的途径应用于软件开发、运行和维护过程，也就是把工程应用于软件

管理	通过计划、组织和控制等一系列活动，合理地配置和使用各种资源，以达到既定目标的过程。
软件生命周期	软件定义、软件开发和软件维护期。
需求分析	确定目标系统必须具备哪些功能。
软件	软件是程序、数据及相关文档的完整集合
软件过程	为了获得高质量软件所需要完成的一系列任务的框架，它规定了完成各项任务的工作步骤
UML	一种书写软件蓝图的标准语言
用例图	从用户的角度描述系统功能，并指出各个功能的操作者
顺序图	描述对象之间的动态交互关系，着重表现对象间消息传递的时间顺序
状态图	描述对象的所有可能状态及引起状态转换的事件
功能性需求	指定系统或产品应具有的特定功能或行为，它描述了软件应如何运行以满足预定目标。
非功能性需求	关乎软件系统的可用性，如果未予以精确定义，可能会影响终端用户的体验。
数据流	数据流（data stream）是一组有序，有起点和终点的字节的数据序列。
数据字典	数据字典是指对数据的数据项、数据结构、数据流、数据存储、处理逻辑

	等进行定义和描述，其目的是对数据流程图中的各个元素作出详细的说明，使用数据字典为简单的建模项目。
用户场景	用户场景通常称为用例，它提供了系统将如何被使用的描述。
业务规则	业务规则是指对业务定义和约束的描述，用于维持业务结构或控制和影响业务的行为。
数据库	数据库是“按照数据结构来组织、存储和管理数据的仓库”。是一个长期存储在计算机内的、有组织的、可共享的、统一管理的大量数据的集合。
数据挖掘	数据挖掘是指从大量的数据中通过算法搜索隐藏于其中信息的过程。数据挖掘通常与计算机科学有关，并通过统计、在线分析处理、情报检索、机器学习、专家系统（依靠过去的经验法则）和模式识别等诸多方法来实现上述目标。

14.2 待确定问题清单

暂无