比书网

需求规格说明书

华东理工大学信息学院软件工程项目组

版本历史信息

|  |  |  |  |
| --- | --- | --- | --- |
| 版本 | 修订日期 | 修订章节 | 修订原因 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

1. 引言
   1. 编写目的
   2. 开发背景
2. 任务概括
   1. 总体功能要求
   2. 运行环境
3. 需求评审
   1. 需求分析
   2. 确定原型
   3. 需求验证

4. 数据描述

4.1 数据的来源

4.2 数据存储

4.3 数据字典

5. 功能需求及任务优先级

5.1 用户及管理员用例

5.2 模块功能需求

5.2.1 获取商品信息模块

5.2.2 商品信息整合并存储模块

5.2.3用户查询模块

5.3 任务优先级

6. 性能需求

6.1时间性能需求

6.2并行性需求

6.3 健壮性

7. 系统测试及模块测试

7.1 系统测试

7.2 模块测试

8. 软件的交付准备

7.1 交付清单

7.2 验收具体内容

1. **引言**

**1.1 编写目的**

本指南用于指导软件开发者为商品比价的过程，通过规范项目开发过程达到提高软件质量，降低维护成本的目的。开发者应该根据此文档进行项目开发。本指南是对项目开发者的基本要求，通过该文档项目开发者可以了解该项目的基本需求和项目功能，通过对一些功能以及人员分工情况做以说明。

**1.2 项目背景**

项目的委托单位：2014级软件工程项目小组

1. **任务概括**

**2.1 任务目标**

开发者应在充分分析需求分析的基础上，选择采用B/S结构或者C/S结构。

本指南中没有规定开发者采用何种具体工具的开发方法，开发者可根据项目具体特点、自身特长来选择面向过程的方法、面向对象的方法。本项目的最终目标是把网络上不同网站同种的商品的信息存储在数据库中，然后根据用户的查询向用户反馈最终相关联的结果。

**2.2 运行环境**

**2.2.1软件接口**

操作系统： Windows 2000及以上

数据库：MySQL

**2.2.2硬件接口**

客户端：CPU在1GHz以上，内存2G以上。

服务端：CPU在4GHz以上，内存6G以上。

1. **需求评审**
   1. 需求分析

根据本项目所需要实现的相关功能，即实现对当当和京东图书的爬取并保存下来，然后对图书信息进行整合，最后实现用户对输入的关键字的查询返回结果。确定功能需求，性能需求，接口，约束及后期产品迭代。

* 1. 确定原型

根据上述确定的需求，确定本项目的原型，在项目实施过程中根据客户的反馈不断对项目原型进行适当的调整。

3.3需求验证

对项目原型进行验证，验证功能需求、项目范围、异常流程需求、约束条件。

1. **数据描述**

**4.1数据来源**

来自当前流行电商网站当当网和京东网的各类书籍信息

**4.2 数据存储**

商品的信息存储在MySQL数据库中

**4.3 数据词典**

书籍信息应当包括以下字段：书籍ID，书籍名称，书籍类别，书籍的出版社，版次，书籍的价格，出版时间，页数，书页大小，作者，收藏数，纸张属性，图片链接。

网页的基本信息应当包括以下字段：网站名称，书籍ID，书籍的url。

1. **功能需求**

此次项目的开发的目标是完成各个电商网站对同一商品的比价，因此需要获取各个电商网站的各种商品的信息（特别是商品的价格），通过也要存储下来便于用户查询，以便给出相对及时的商品价格信息，并将此种商品在各个网站的信息反馈给用户。因此本项目分为三个模块：获取商品信息模块、商品信息整合并存储模块和用户查询模块。

**5.1 用户及管理员用例**

****

UC1. 查看图书详细分类（用户）

用户可以通过移动鼠标指针到页面左侧分级栏查看图书分级

UC2. 请求搜索信息（用户）

用户在搜索框内填入相关搜索关键字，系统根据关键字返回相应的图书信息

UC3. 浏览搜索结果（用户）

系统根据关键字返回相应的图书信息，用户浏览返回显示的相关图书信息

UC4. 选择图书（用户）

用户根据页面显示的图书信息选择图书，系统返回图书的信息

UC5. 维护（管理员）

管理员定期对图书目录系统进行维护，更新相应的图书信息

**5.2模块功能需求**

**5.2.1获取商品信息模块**

网络爬虫的功能是从当前流行的电商网站中获取其中的书籍信息，并保存下来。其需求是：

* 1. 灵活性

当给定不同电商的URL时，能够根据URL获取相对应的电商网站上的商品信息

* 1. 输入输出要求

应该输入一个合法的初始URL作为初始入口

* 1. 故障的处理能力

当获取的数据过多时可能造成服务器负载过重，因此要保证系统的健壮性

* 1. 内容的准确性

在网页中获取的商品信息应该是具体商品的相关信息，这样就减少了获取的信息中不相关数据的数量，同时也直接减少了下面的数据清洗工作量。

**5.2.2 商品信息整合模块**

**5.2.2.1 数据清洗**

对模块一部分所传来的数据进行清洗，这需要对其所传递过来的数据格式有一定要求，需要求模块一传递过来如Json, XML等易解析的数据格式。数据的清洗应该得到准确的商品信息，为下一步数据的整合做好准备。

**5.2.2.2数据整合**

在上一步数据清洗的基础上，得到商品的具体信息，再利用商品的名称等属性将多个网站得到的商品进行匹配，找到相互对应的商品，建立映射关系，完成商品的融合，当然也要完整的同一个商品在各个网站的信息，以便做对比分析。

**5.2.2.3数据库存储**

首先要确定的是数据库的设计，需要与模块一、三的人员进行协商，确定下来表的数量及属性列。同时也要有组织有结构的去定义数据库中的表格，这样有利用下一步对数据库操作时减少访问时间，提高用户的体验。

**5.2.3用户查询模块**

用户输入需要查询的商品名称关键字,返回相关商品的具体信息，包括商品名称、商品规格、商品类型等等，这样就可以辅助用户买到心满意足的商品。

**5.2.3.1 用户查询界面**

用户查询界面应简洁大方，便于用户操作。

**5.2.3.2 搜索结果界面**

搜索结果展示界面应优雅美观，展示的信息应全面详细。

**5.3任务优先级**

|  |  |  |  |
| --- | --- | --- | --- |
| **一级**  **模块** | **二级**  **模块** | **三级模块** | **优先级** |
| **比书网** | **获取商品信息模块** |  | **P1** |
|  |  |
|  |  |
|  |  |
| **商品信息整合模块** |  |  |
|  |  |
|  |  |
| **用户查**  **询模块** |  |  |
|  |  |

**6． 性能需求**

**6.1 时间性能需求**

对于用户给出的查询要在一定的时间给出相关结果，时间要求上不能超过3s。

**6.2 并行性需求**

为了便于查询能够满足多用户的查询。

**6.3 健壮性**

要求系统能处理不合法的输入

**7 系统测试及模块测试**

**7.1 系统测试**

为了尽早发现软件系统中的错误，从而达到提高系统质量、降低系统维护的费用，开发者应在编码过程中对各个模块的程序代码进行单元测试，系统集成时进行集成测试，系统集成完成后对整个系统进行系统测试。单元测试是在系统开发过程中针对程序模块进行正确性检验。集成测试是在单元测试的基础上，将所有模块按照设计要求组装成系统或子系统，对模块组装过程和模块接口进行正确性检验。系统系统测试不仅是检测系统的整体行为表 现，从另一个侧面看，也是对系统开发设计的再确认。进行系统系统测试工作时。测试主要包括界面测试、可用性测试、功能测试、稳定性(强度)测试、性能测试、强壮性(恢复)测试、逻辑性测试等。

**7.2 模块测试**

第一步：设计基本功能测试用例，证明被测单元至少在某种正常情况下能够运行了；

第二步：设计功能正面测试用例，找出被测单元对于设计要求的正确输入可能做出的不正确处理；

第三步：设计功能反面测试用例，找出被测单元对于设计要求的错误输入可能做出的不正确处理；

第五步：设计性能测试用例，找出单元对于设计要求的性能可能做不到的错误。

第六步：设计[白盒测试](javascript:;)用例，找出单元内部控制结构和数据使用可能存在的问题。

**8 软件的交付准备**

8.1 交付清单

在软件测试证明软件达到要求后，软件开发者应该提交开发目标运行程序、数据库的数据字典、用户使用指南、需求报告、设计报告、测试报告等开发过程中相应的文档

8.2 验收具体内容

验收内容包括：合法性检查、文档检查、软件一致性检查、软件系统测试与测试结果评审等几项工作。