# 系统概要设计说明书

文档变更记录

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 变更说明 | 作者 | 版本号 | 日期 |
| 1 | 新建文档 | 黄俊远 | 1.0.0 | 2020.8.4 |
|  |  |  |  |  |

目录

[系统概要设计说明书 1](#_Toc47482168)

[1. 引言 1](#_Toc47482169)

[1.1 编写目的和范围 1](#_Toc47482170)

[1.2 术语表 1](#_Toc47482171)

[1.3 参考资料 1](#_Toc47482172)

[1.4 使用的工具 1](#_Toc47482173)

[2. 总体设计 2](#_Toc47482174)

[2.1 概述 2](#_Toc47482176)

[2.2 设计思想 2](#_Toc47482177)

[3. 模块设计 4](#_Toc47482178)

[3.1 爬虫功能 4](#_Toc47482182)

[3.2 文章比较功能 5](#_Toc47482183)

[4. 接口设计 6](#_Toc47482184)

[4.1 内部接口 6](#_Toc47482189)

[4.2 外部接口 6](#_Toc47482190)

[5. 数据库设计 6](#_Toc47482191)

[6. 系统安全设计 6](#_Toc47482192)

[7. 系统性能设计 7](#_Toc47482193)

[8. 系统容错处理 7](#_Toc47482194)

## 引言

### 编写目的和范围

本详细设计说明书编写的目的是说明程序模块的设计考虑，包括程序描述、输入输出、算法和流程逻辑等，为软件编程和系统维护提供基础。本说明书的预期读者为系统设计人员、软件开发人员、软件测试人员和项目评审人员。

### 术语表

|  |  |  |
| --- | --- | --- |
| 序号 | 术语或缩略语 | 说明性定义 |
| 1 | 爬虫 | 网络爬虫(又被称为网页蜘蛛,网络机器人),是一种按照一定的规则,自动地抓取互联网信息的程序或者脚本。另外一些不常使用的名字还有蚂蚁、自动索引、模拟程序或者蠕虫。 |
|  |  |  |

### 参考资料

暂无

### 使用的工具

文字处理工具：word

绘图工具 edraw

## 总体设计



### 概述

#### 功能描述

系统包含两大模块：爬虫模块和文章比较模块

#### 运行环境

#### 硬件环境

个人PC

CPU intel

内存 8G

#### 软件环境

操作系统 Win10/linux

数据库 mysql

其他 Python2.7

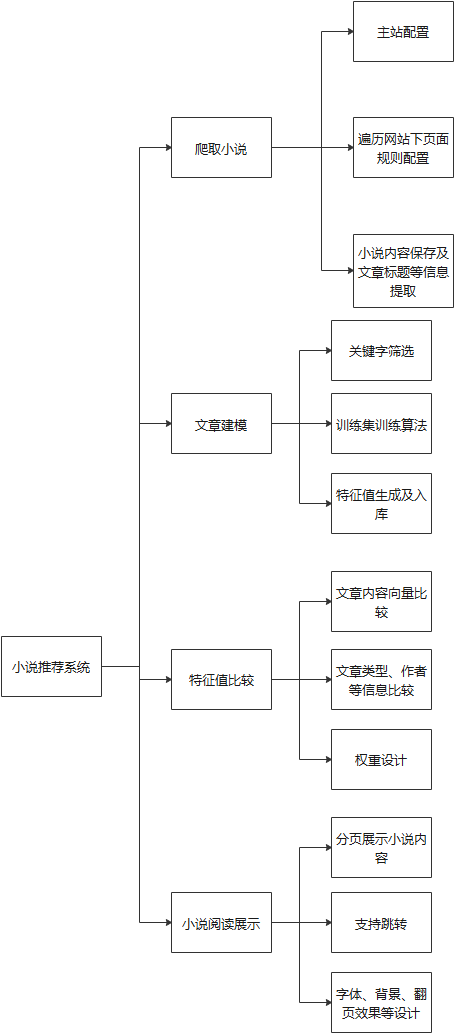
#### 开发环境

同上

### 设计思想

#### 系统架构

模块设计图：



#### 关键技术与算法

#### 关键数据结构

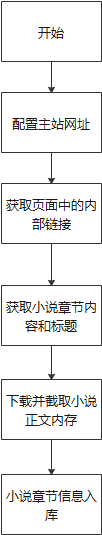
Mysql数据库，于数据库设计中列出

## 模块设计



### 爬虫功能

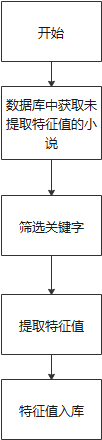
流程图：



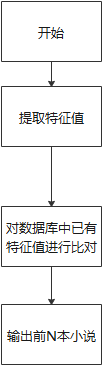
### 文章比较功能

流程图：

1、小说分析过程



2、小说比对功能



## 接口设计



### 内部接口

暂无

### 外部接口

## 数据库设计

NovelInfo表 包含id 小说名，小说作者，类型，简介，完结状态

NovelDetailInfo表 小说id 章节名 文件路径 大小

NovelModel表 小说id 特征值 提取特征值状态

## 系统安全设计

无

## 系统性能设计

无

## 系统容错处理

无