

浙江大学



Techhub 后端知识库搜索引擎系统 用户手册

小组: G08

组长: 黄亦非

组员: 傅诤哲、胡瀚丹、陈鑫

2019/7/18

目录

修订表.....	3
审批记录.....	4
1 引言.....	5
1.1 编写目的.....	5
1.2 背景.....	5
1.3 定义.....	5
1.4 参考资料.....	6
2. 用途.....	6
2.1 功能.....	6
2.2 性能.....	7
2.2.1 时间特性.....	7
2.2.2 灵活性.....	7
2.3 安全保密.....	8
3 运行环境.....	8
3.1 硬设备.....	8
3.2 支持软件.....	8
3.3 数据结构.....	9
4 使用过程.....	9
4.1 安装和初始化.....	9
4.2 输入.....	9
4.2.1 输入数据的现实背景.....	10
4.2.2 输入格式.....	10
4.2.3 输入举例.....	10
4.3 输出.....	10
4.3.1 输出数据的现实背景.....	11
4.3.2 输出格式.....	11
4.3.3 输出举例.....	11
4.4 文卷查询.....	12
4.5 出错处理和恢复.....	13
4.6 终端操作.....	13

修订表

编号	生成版本	修订人	修订章节和内容	修订日期
1	1.0	全体小组成员	所有章节	2019/7/18

审批记录

编号	版本	审批人	审批意见	审批日期
1	1.0		通过	

1 引言

1.1 编写目的

本文档的主要编写目的是为网站用户提供使用网站的指导。本文档的主要读者为使用本网站的用户，次要读者为开发人员、测试人员和项目经理。

1.2 背景

随着互联网信息技术的发展，越来越多学生和职场人员希望学习一门新的编程技术。在学习技术的过程中，我们大多会使用百度、Google 等网站搜索相关资源，但由于搜索结果较多较杂，用户需要耗费大量时间进行信息筛选，此外也存在学习资源质量低、路线不明确等问题。

因此，我们提出了一个面向技术学习的垂直搜索引擎，致力于帮助用户更高效地获取技术知识。本系统通过整合官方文档、技术博客、技术问答、视频教程等多源数据，构建一个专业的技术学习知识库，并通过信息抽取与搜索结果排序等手段，呈现出最符合用户需求的检索结果。

1.3 定义

事务流：数据进入模块后可能有多种路径进行处理

主键：数据库表中的关键域。主键中的值互不相同

外部主键：数据库表中和其他表主键关联的域

缩写：

系统：通常情况下未进行特殊说明，系统统称为该网站。

UML：统一建模语言。UML 是一套用来设计软件蓝图的标准建模语言，是一种从软件分析、设计到编写程序规范的标准化建模语言。

Spring Boot：微服务框架

React: Web 开发框架

MongoDB: 资讯库存储

1.4 参考资料

文档格式按照我国 GB/T8567-1988 国家标准和 IEEE/ANSI1830-1993 标准规范要求进行。包括以下文件：

- 1、项目描述
- 2、“软件需求分析与设计”教学与实践安排
- 3、软件工程项目开发文档范例
- 4、软件工程国家标准文档
- 5、软件需求说明书编写规范

书籍包括：

- 1、《软件项目管理》朱少民，韩莹编，人民邮电出版社
- 2、《软件项目管理》Rajeev T Shandilya 编，科学出版

2. 用途

2.1 功能

用户类	功能与说明
网站用户	<div>1. 主要用户</div> <div>2. 同时使用该网站的用户数目可能较多，主要服务的对象为想要找到特定领域的专业技术文档或者是相关问</div>

	<p>题解答。</p> <ol style="list-style-type: none"> 3. 要求能够返回足够精确和足够深度的结果 4. 能够根据用户的喜好进行返回数据的过滤 5. 支持根据用户的输入进入搜索引擎搜索，并且返回相关的内容到网站首页。
系统管理员	<ol style="list-style-type: none"> 1. 次要用户 2. 但是权限比较高，具有数据库的更新和审核等管理权限 3. 面向的用户数目比较多，要求能够提供方便并且快速的管理员接口进行管理 4. 操作频率相对较低，但是每一次的操作对于系统造成的影响比较大。

2.2 性能

2.2.1 时间特性

要求五百人同时使用时响应时间小于 0.5s。

2.2.2 灵活性

当用户需求如对操作方式、运行环境、结果精度、时间特性等的要求有某些变化时，本软件的适应能力一般，所以请严格按照提示使用本网站。

2.3 安全保密

本软件在安全、保密方面的设计考虑和实际达到的能力较好，能有效避免注入 SQL 注入等基本的恶意攻击方式，确保用户信息和系统、数据的安全。

3 运行环境

3.1 硬设备

需要一台联网的 PC 机。

3.2 支持软件

操作系统：Windows10/Mac OS

前端开发框架：React

后端开发：Eclipse、Pycharm

数据库平台：Mongodb

Web 服务器：Apache

搜索引擎：Solr

开发工具：能支持网页开发的工具均可，如 JetBrains WebStorm、Visual Studio Code

测试工具：请测试人员自行选择，推荐使用 LoadRunner

建模工具：根据项目情况，请自行选择合适方便的建模工具，《软件工程-实践者的研究方法》一书中介绍了很多建模工具，可对比参考。

办公软件：Microsoft Office 2016 系列产品

浏览器：Chrome、Firefox、IE11、Safari

3.3 数据结构

数据库平台：Mongodb

4 使用过程

4.1 安装和初始化

本软件系统为网站，所以不需安装和初始化，只需要确保自己的 PC 机上的浏览器版本符合要求。再打开浏览器，输入 `http:// 49.234.190.100:5000` 即可成功打开页面。



4.2 输入

相关的输入应按照网站指示进行输入工作。

4.2.1 输入数据的现实背景

输入数据的现实背景是教学实际情况中应输入的数据，相关资料可以参考上文所提到的功能和数据精度表。值得注意的是，现实背景应包括情况、情况出现的频度、情况来源、输入媒体、限制、质量管理和支配等方面。

4.2.2 输入格式

值得注意的是，输入格式的现实背景是用户希望得到的技术文档的关键词句，应注意长度、格式基准、标号、顺序、标点、词汇表、省略、重复和控制等方面。

4.2.3 输入举例

以搜索 spring 为例。可以输入诸如 Spring cloud，SpringCloud 等，系统支持接受一定的误差。



4.3 输出

本网站的输出为网站自己生成，不需要用户操心。

4.3.1 输出数据的现实背景

输出数据的现实背景是用户搜索的技术文档的查询结果。值得注意的是，现实背景应包括文档来源、作者，分类，关键词、关键句。

4.3.2 输出格式

本网站的输出格式为网站通过搜索引擎自己生成，不用用户手动调整。

4.3.3 输出举例

以下是用户在搜索 Spring 的输出举例。



搜索结果的 Tag 分布：

The screenshot shows the techhub search results for 'spring'. The header is purple with the techhub logo and a search bar containing 'spring'. Below the header, there's a sidebar on the left with '所有资源' (All Resources) and '分类' (Categories). The main content area displays '共计 14729 个技术标签:' (Total 14729 technical tags). A grid of tags is shown, each with a count: Docker (3874), 编程 (3039), apache (2468), Spring (2445), MongoDB (2023), 视频教程 (1933), Redis (1865), 学习 (1853), 程序员 (1764), Nginx (1640), Solr (1624), MySQL (1605), 分布式 (1453), 公开课 (1452), tomcat (1420), docker-compose (1396), RabbitMQ (1363), 架构师 (1310), 微服务 (1195), Korea相关 (1177), Neo4j (1165), Linux (1142), GraphQL (1136), 高并发 (1093). On the right, there's a '标签词云' (Tag Word Cloud) showing various terms like RESTful API, 公开课, Spring, tomcat, Redis, OAuth, MySQL, 视频教程, 编程, Nginx, Solr, Docker, MongoDB, 分布式, 学习, Neo4j, RabbitMQ, 架构师, 微服务, and Korea相关. At the bottom, it says 'made with ❤ by Group-8 @ZJU'.

搜索结果精细化选择:

This screenshot shows the techhub search results for 'spring' with refined filters. The header is purple with the techhub logo and a search bar containing 'spring'. The sidebar on the left has '所有资源' (All Resources) and '分类' (Categories). The main content area displays '[博客文章] - 显示 8 条最优搜索结果' (Blog Articles - Display 8 best search results). Two article cards are shown. The first card is titled 'Spring SpringMVC SpringBoot SpringCloud概念、关系及区别' (Spring SpringMVC SpringBoot SpringCloud concepts, relationships and differences) and includes a brief description, a source link, and tags for Spring, Spring MVC, and Spring Boot. The second card is titled '10 道 Spring Boot 面试题' (10 Spring Boot interview questions) and includes a brief description, a source link, and tags for Spring Boot and 程序员 (Programmer). On the right, there's a '标签词云' (Tag Word Cloud) showing various terms like 微服务 (Microservices), RabbitMQ, Neo4j, feign, MySQL, 程序员 (Programmer), PHP, tomcat, jwt, Nginx, Spring, Solr, MongoDB, 视频教程 (Video Tutorial), Docker, Linux, 学习 (Learning), Redis, 编程 (Programming), 分布式 (Distributed), 教育 (Education), apache, k8s, 架构师 (Architect), 公开课 (Open Course), and 课程 (Course).

4.4 文卷查询

使用支持 mongodb 的数据库连接方式比如 Studio 3T，在链接中进行如下的配置：

Connection name: TechHub

Connection group: [dropdown]

Server Authentication SSL SSH Tunnel Proxy Advanced

Connection Type: Standard Connection

Members: 49.234.90.146:37017
49.234.90.146:47017
49.234.90.146:57017

Add... Remove Discover

Replica Set Name: [text field]

Read Preference: Primary [dropdown] [info]

☐ Read-Only Lock [info]

From URI... Use this option to import connection details from a URI

To URI... Use this option to export complete connection details to a URI

Test Connection Cancel Save

配置成功即可连接到数据库。

4.5 出错处理和恢复

当软件出错时，可以重启系统或重启计算机再次尝试使用。但请不要擅自修改系统代码。当错误实在无法解决之时，请拨打售后服务热线 18868120866 或尝试联系系统维护人员。

4.6 终端操作

本系统不是多终端系统，因此终端操作板块在此略过。