

后端知识库搜索引擎系统

项目章程

小组：G08

组长：黄亦非

组员：陈鑫、傅诤哲、胡瀚丹

2019/7/5

目录

1. 概要	4
1.1. 项目背景	4
1.2. 项目名称	4
1.3. 项目目标	4
1.3.1. 总体目标	4
1.3.2. 具体目标	4
2. 项目组织	6
2.1. 项目经理	6
2.2. 项目成员	6
3. 项目范围	6
3.1. 功能范围	6
3.2. 实体范围	6
4. 项目总体计划	6
4.1. 项目阶段划分	6
4.2. 预计开始时间	6
4.3. 预计结束时间	6
4.4. 预计时间表和里程碑	6
5. 项目主要干系人	8
5.1. 项目用户方简介	8
5.2. 项目开发方简介	8
6. 项目假设和约束条件	8
7. 项目风险管理	8
7.1. 实施周期延期的风险	8
7.2. 实施范围风险	8

7.2.1.	前期项目描述的不准确性导致各相关部门之间的分歧。	8
7.2.2.	开发过程中需求外延导致周期延长。	9
7.2.3.	临近项目交付， 用户突然提出新的需求。	9
8.	项目变更管理	9
8.1.	提出方申请	9
8.2.	接收方响应	10
8.3.	申请方认可	10
9.	验收标准	10
9.1.	验收方式	10
9.2.	验收标准	10
10.	文件签署	12

1. 概要

1.1. 项目背景

Techhub 后端知识库搜索引擎系统主要是面向在校大学生用户，致力于在互联网上学习后端的相关技术，能够帮助用户快速的搜索相关领域的技术文档，包括该技术的描述、使用手册、教学视频、使用当中的相关问题等方面。对于用户而言，可以根据输入的问题或者关键字，准确的定位用户的意图，并且返回最准确的结果给用户。同时用户也可以根据提供的几个过滤条件对返回的结果进行对应的过滤，实现更加精确的结果定位。

同时系统对于数据的来源和处理也非常的重视，致力于提供最准确、最完整、最精确的知识库架构。对于数据我们会进行相关的过滤、去重、结构化信息的提取等等数据处理的动作，保证系统数据的稳定和准确。

1.2. 项目名称

TechHub 后端搜索引擎系统

1.3. 项目目标

1.3.1. 总体目标

21 世纪是以网络的全面深入运用为特征的世纪。网络环境下的教育不仅是教育信息化的必然产物，也是教育发展的必然走向。通过因特网或其他数字化内容进行学习交流与教学的活动即网络化学习（e-learning），可以充分利用现代信息技术所提供的、具有全新沟通机制与丰富资源的学习环境，实现一种全新的学习交流方式。

基于这种方式，我们将设计一个后端的知识库搜索引擎，以期在海量的信息下，网站用户能够快速的获取到学习某门新技术的完整路线。

具体目标：

该网站的潜在用户可以简单分为 3 类，分别为网站用户、管理员。

网站用户（实际使用该网站进行查询的用户）功能实现的主要目标有：

- 用户在该网站上面可以输入自己想要搜索和查找的内容，然后网站进行查询相关的内容。
- 同时在一定的响应阈值时间之内，网站能够展示相对应的搜索结果。
- 同时网站能对搜索结果进行内部排序然后返回显示，同时用户能够重置搜索搜索结果。
- 点击具体的条目时候显示条目的具体内容，并且可以链接到原网站当中。
- 支持分页显示，使得页面的显示更加的美观。
- 同时还支持用户使用不同的搜索条件进行查询相对应的更加准确的信息。

管理员功能实现的主要目标有：

- 支持数据的手动导入，在后续还有数据更新的时候能够手动导入更新后台的数据库支持。
- 同时也支持数据的定时 import，这样为了方便管理员不需要随时随刻的去手动导入搜索引擎当中，当数据库当中有变动的时候，可以被自动的导入到搜索引擎当中。
- 管理员进行信息爬取，这个是系统的最基本的需求，需要在各个网站上爬取相关的信息，作为后台数据库的支持。
- 信息自动滤重功能，网上爬取的各种信息可能会存在很多的重复，所以需要对爬取的数据进行滤重的动作。
- 信息自动分类和摘要，这个对于信息是比较重要的一个后处理，通过对信息自动摘要和分类，方便后续用户进行条件搜索。
- 信息自动爬取，当用户搜索结果小于我们设定的阈值的时候，能够支持实时的去各大搜索引擎当中爬取新的信息并且导入我们的后台数据库当中。

2. 项目组织

2.1. 项目经理

黄亦非

2.2. 项目成员

陈鑫、傅诤哲、胡瀚丹

3. 项目范围

3.1. 功能范围

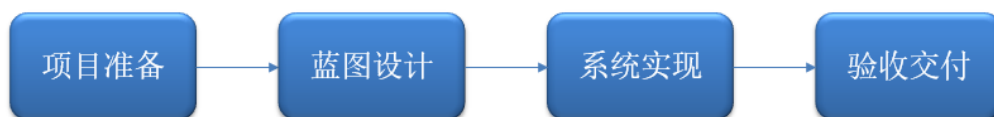
这个网站的实现方法将和其他的网站一样，没有特殊的技术。网站的范围是：1.搜索信息，2.展示信息，3.信息获取。不再另外开设可供网站用户使用的邮箱，如有邮件都将使用个人自己在其他网站上的邮箱。

3.2. 实体范围

网站用户、网站管理员、游客

4. 项目总体计划

4.1. 项目阶段划分



4.2. 预计开始时间

2019 年 7 月 5 日。

4.3. 预计结束时间

2019 年 7 月 19 日。

4.4. 预计时间表和里程碑

教学网站项目工程时间表

项目阶段	里程碑	计划时间
项目准备	项目可行性报告	2019.7.5
	项目章程	
	项目总体计划	
蓝图设计	前景与范围	2019.7.11
	质量保证计划	
	需求工程计划	
	软件需求规格说明书	
	系统设计计划	
	需求变更控制会规程	
系统实现	系统编码与实现计划	2019.7.15
	测试计划	
	需求变更控制文档	
验收交付	用户手册	2019.7.18
	软件需求规格书更新版	
	软件概要设计说明书	
	测试报告	
	工程部署计划	
	培训计划	
	系统维护计划	
	项目总结报告	

5. 项目主要干系人

5.1. 项目用户方简介

网站用户：指预期使用该搜索引擎系统的网站用户。

管理员：指该网站的后台管理员

项目开发方简介：

2019 学年小学期项目实训课程第 8 小组。

6. 项目假设和约束条件

对于成功开发该项目，需要如下几项假设条件：

- 相关教师（学院）的支持和认可，包括技术支持和设备支持；
- 参与教师和项目小组成员的高度配合；
- 可用的编程电脑，服务器（若后期有需求），Webstorm、Pycharm 等开发平台；
- 团队的合作精神，良好的工作能力，充足的空余时间。

本网站要求提供对外服务的能力，保证至少 300 名同学同时进行搜索的要求。包括数据存储能力，网络服务吞吐能力，数据安全特性等。

7. 项目风险管理

实施过程中的风险不可避免，需要提前针对可能的风险做出合理和高效的应对措施。

7.1. 实施周期延期的风险

周期延期，是指开发人员因为各种因素而未按预期时间完成开发阶段而致使项目交付延期。

应对方案：定期召开会议，沟通项目进展和遇到的困难。针对个人实际情况对项目分工和时间线做出合理和灵活调整。

7.2. 实施范围风险

7.2.1. 前期项目描述的不准确性导致各相关部门之间的分歧。

应对方案：以项目的总体目标和具体目标为准绳，充分尊重需求工程的成果。

7.2.2. 开发过程中需求外延导致周期延长。

应对方案：明确规范各开发阶段满足的需求和具体功能，外延的需求可以在下一发行版实现。

7.2.3. 临近项目交付，用户突然提出新的需求。

应对方案：与用户充分沟通，探讨新需求实现的可行性和合理性，权衡交付的时效性和新需求的重要性，尽量满足用户的需要。

8. 项目变更管理

变更控制是通过有序地管理变更来稳定开发过程、减少项目风险。本程序的制定是为了检查所有的变更请求，决定哪些需要实施、哪些需要推延、哪些需要否决。在得到对方的认可后，进度和成本将相应地做出调整。一个有效的变更控制程序对于避免项目延期和超支是必要的。

8.1. 提出方申请

申请方在提出变更前需要内部讨论并且最终总结意见为项目变更申请表，并由申请方项目经理交付接收方项目经理。该申请表的功能是使双方的每一个成员都能充分理解变更的后果。

变更申请表

变更申请表
(系统名称) 变更申请序号#：
申请人：
日期：
申请变更内容：
申请变更原因：
变更类型：
A:功能方面 B：运行性能方面 C：文档方面
授权人签字：

日期：

8.2. 接收方响应

接收方项目经理在接受到变更申请表后在 24 小时确认收讫，并开始评估变更请求的技术可行性、成本并将其与项目业务需求和资源约束对齐。接收方经理将在项目组内部讨论后最终决定是否批准或驳回该变更。

8.3. 申请方认可

申请方经理对于变更申请书进行确认，并且提交双方不能决定的争议部分至项目领导小组。

若变更申请涉及对合同和费用的修改，亦需要报告领导小组批准。

9. 验收标准

9.1. 验收方式

本项目采用分阶段提交成果和验收的方法。在得到本阶段成果的确认以后，再开始下一阶段的实施工作。以保证项目始终在实施双方意见一致的前提下进行。

项目阶段验收将根据双方确认的本阶段实施目标，工作计划和提交的阶段工作完成报告作出结论。

在开发方书面提出验收申请之日起，浙江大学项目经理（或其授权人）应在 5 个工作日内，书面签署确认报告或向项目组提出优化的建议。

9.2. 验收标准

功能模块	功能	实现程度
爬虫系统	数据获取	
	数据过滤	
	关键字提取	

	摘要生成	
前端页面	支持搜索内容输入	
	展示搜索结果	
	具体的结果显示	
	根据相关条件进行过滤	
后端系统	搜索接口	
	过滤接口	

10. 文件签署

该章程已审阅并予批准。

签字确认

甲方：

年月日

乙方：

年月日