

浙江大学

垂直搜索引擎

“梦奇电影旅行”项目计划书



学生姓名：	<u>胡洋凡</u>	学号：	<u>3180103167</u>
学生姓名：	<u>蔡灿宇</u>	学号：	<u>3180101972</u>
学生姓名：	<u>刘轩铭</u>	学号：	<u>3180106071</u>
学生姓名：	<u>杨凌霄</u>	学号：	<u>3180103608</u>
学生姓名：	<u>王子腾</u>	学号：	<u>3180102173</u>
学生姓名：	<u>王绍兴</u>	学号：	<u>3180106074</u>

2021 年 07 月 09 日

Verse:1.1

版本变更历史

版本号	作者	修订日期	审核者	审核日期	修订原因
1.0	胡洋凡、蔡灿宇、 王子腾、王绍兴、 杨凌霄、刘轩铭	2021.07.09	蔡灿宇	2021.07.09	原始文档
2.0	胡洋凡、蔡灿宇、 王子腾、王绍兴、 杨凌霄、刘轩铭	2021.07.10	蔡灿宇	2021.07.10	样式的统一和部分修改

目录

1. 引言	5
1.1 编写目的	5
1.2 软件系统名称	5
1.3 相关人员及项目支持	5
1.4 读者对象	6
2. 项目概述	6
2.1 项目背景	6
2.2 项目介绍	7
2.3 名词定义	7
2.4 参考资料	8
2.5 产品成果	8
2.5.1 产品	8
2.5.2 手册	9
2.5.3 服务	9
2.6 项目相关信息	9
3. 开发进度计划	10
4. 角色分配	11
5. 软硬件资源说明	12
6. 专题计划要点	12
6.1 开发人员培训计划	12
6.2 质量保证计划	13
6.3 测试计划	13
6.3.1 静态测试	13
6.3.2 单元测试	15
6.3.3 集成测试	15
6.3.4 系统测试	16
6.3.5 确认测试	16

7. 项目可行性	16
7.1 经济可行性	16
7.2 技术可行性	17
7.3 用户使用可行性	17
8. 技术分析	18
8.1 前端技术.....	18
8.1.1 HTML 语言.....	18
8.1.2 JavaScript 语言.....	18
8.1.3 Vue.js 框架.....	18
8.2 后端技术	18
8.2.1 SpringBoot 框架.....	18
8.2.2 MongoDB 数据库.....	19
8.2.3 Scrapy 框架.....	19
8.3 服务器	19
8.3.1 Apache	19
8.3.2 Nginx	19
8.3.3 Elasticsearch	20

1. 引言

1.1 编写目的

项目管理与软件需求，作为软件工程当中最为重要的组成几个部分，已经引起了业内人士的高度重视。项目管理和需求工程概念的提出，就是为了把软件工程化，以更有效地开发需求，开发软件并实现有效的管理。

而本计划书旨在说明“梦奇”电影旅行搜索引擎系统在整个项目实施过程中的任务分配、人员分工、时间轴安排、过程规范和测试需要等，为后续的软件工程实施与管理奠定计划基础，有利于全局的发展与调控安排。

另外本计划书还将通过项目管理的角度展示本小组在项目的开发实施过程中的计划与组织方式。

1.2 软件系统名称

梦奇电影旅行搜索引擎

1.3 相关人员及项目支持

- 任务提出者
浙江大学项目实训任课老师-邵健
- 开发者
浙江大学项目实训课程 G14 小组
- 实现该软件的计算机网络
基于公网 IP 的服务器和数据库

- 该软件系统同其他系统或其他机构的基本的相互来往关系
- 项目组可使用一些开源数据平台接口获取电影有关信息。

1.4 读者对象

1. 软件客户代表
2. 项目经理
3. 项目开发人员
4. 软件质量分析员
5. 软件维护人员
6. 其它需要了解本项目质量管理情况的人员都可参阅

2. 项目概述

2.1 项目背景

该项目开发的软件为一个可以为电影和旅行爱好者使用的电影相关信息搜索引擎。

梦奇电影旅行是一个影视旅行主题应用，通过它可以沿着某部影视剧或电影了解到其取景地点或者配乐的情况，也能通过选择地点和音乐，了解到曾在此处取景的影视剧的取景片段及剧情发生的时间点或者以该音乐为配乐的众多音乐。

喜欢电影的人应该会有相同的感觉：遇到某部特别喜欢的电影，就很想去了解这部电影的取景地，或者这部电影的配乐信息，甚至想要跟随电影的主人公，走过他走过的路，身临其境，听着熟悉的音乐，感受电影和现实交织在一起的美妙，然后再重温一遍电影，会觉得是一场非常奇妙的体验

梦奇电影旅行就是一款献给热爱电影和旅行的朋友的一款指南应用，用户可以在应用中搜索熟悉的地点和音乐，或者是电影本身，了解到这些事物背后的电影故事。

2.2 项目介绍

本项目旨在让用户能够通过系统，输入音乐、地点、电影名等信息，搜索到相关电影，这些电影可能以输入的音乐作为配乐，可能以输入的地点作为取景地，或者是电影本身。项目希望用户搜索到相应信息后，了解电影的美好背景，在应用中进行一场美妙的旅行。

2.3 名词定义

➤ IntelliJ IDEA

java 编程语言开发的集成环境。IntelliJ 在业界被公认为最好的 java 开发工具，尤其在智能代码助手、代码自动提示、重构、JavaEE 支持、各类版本工具(git、svn 等)、JUnit、CVS 整合、代码分析、创新的 GUI 设计等方面的功能可以说是超常的。

➤ Chrome

又称谷歌浏览器，是一个由 Google(谷歌)公司开发的开放源代码网页浏览器。本软件的程式码是基于其他开放源代码软件所撰写，包括 WebKit 和 Mozilla，目标是提升稳定性、速度和安全性，并创造出简单且有效率的使用者界面。

➤ HTML

超文本标记语言 (Hyper Text Markup Language)，是标准通用标记语言下的一个应用，用于描述因特网上的网页文档。

➤ CSS

层叠样式表 (Cascading Style Sheets)，是一种用来表现 HTML 等文件样式的计算机语言，在网页中能够对网页中元素位置的排版进行像素级精确控制。

➤ UML

统一建模语言 (Unified Modeling Language)，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。

➤ ElasticSearch

ElasticSearch 是一个基于 Lucene 的搜索服务器。它提供了一个分布式多用户能力的全文搜索引擎，基于 RESTful web 接口。ElasticSearch 是用 Java 语言开发的，并作为 Apache 许可条款下的开放源码发布，是一种流行的企业级搜索引擎。

➤ MongoDB

MongoDB 是一个基于分布式文件存储的数据库。由 C++ 语言编写。旨在为 WEB 应用提供可扩展的高性能数据存储解决方案。MongoDB 是一个介于关系数据库和非关系数据库之间的产品，是非关系数据库当中功能最丰富，最像关系数据库的。

➤ MySQL

MySQL 是最流行的关系型数据库系统，在 WEB 应用方 MySQL 是最好的 RDBMS (关系数据库管理系统) 应用软件之一。

2.4 参考资料

1. 《软件设计文档国家标准》
2. 需求工程计划书、可行性分析报告
3. 《软件需求》刘伟琴、刘洪涛译
4. 《Software Requirements edition2》Karl.E.Wieger

2.5 产品成果

2.5.1 产品

最终成果是一个电影配乐、取景地、电影基本信息搜索引擎，包括：

- ①前后端代码
- ②数据库建立代码
- ③系统开发、维护文档
- ④网站所需资源文件

2.5.2 手册

- 验收报告以及交接手册
- 用户操作手册

对于系统的界面、功能以及操作说明进行描述，使用户对于软件的操作可以快速上手。

- 应用配置手册

提供给有一定技术的用户进行阅读，对于应用的搭建方法进行说明，对于相关参数的用途进行说明，使有一定技术的用户得已自己解决多数问题。

- 软件维护手册

主要包括系统说明、程序模块说明、操作环境、支持软件的说明、维护过程的说明，便于维护人员对于系统的维护。

2.5.3 服务

提供使用教程以及咨询服务，使得用户尽快上手使用系统。

提供后续一定年限的维护服务，当软件出现问题时提供技术支持。

提供使用相应的在线客服支持。

2.6 项目相关信息

项目批准者：项目实训课程老师

项目批准日期：2021 年 7 月 9 日

项目截止日期：2021 年 7 月 22 日项目验收前

3. 开发进度计划

项目阶段	开始时间	结束时间	负责人	主要工作	里程碑事件
项目启动	2021. 07. 09	2021. 07. 10	全体成员	进行项目可行性分析，指定项目总体计划和项目章程	完成《项目总体计划》
需求分析	2021. 07. 10	2021. 07. 12	全体成员	确定系统运行环境，建立系统逻辑模型，确定系统功能及性能要求，完成需求分析	完成《软件需求规格说明书》
系统设计	2021. 07. 13	2021. 07. 15	全体成员	进行系统架构及数据库设计	完成《详细设计说明书》
编程实现	2021. 07. 16	2021. 07. 18	全体成员	完成系统编码集成	完成可运行的系统
系统测试	2021. 07. 19	2021. 07. 20	全体成员	完成系统测试，完成工程部署进行项目总结	完成《测试报告》

表 1 开发进度计划

4. 角色分配

项目角色	角色分配	角色描述
产品定位	全体成员	负责市场调查并根据用户的需求确定开发何种产品，选择何种技术根据产品的生命周期，协调研发、营销、运营确定和组织实施相应的产品策略
总体设计	全体成员	建立系统框架；数据库设计；概要设计；参加技术评审；
服务器端	刘轩铭 胡洋凡	组织编写测试计划和测试方案，组织系统测试；参加技术评审；
前端	王绍兴 蔡灿宇	设计前端，提升网站美感，并编码实现
软件质量监督组	全体	是对质量经理以及项目经理提供项目进度与项目实际开发时的差异报告，提出差异原因和改进方法。
开发人员	全体	负责进行编码工作与单元测试，进行系统集成，及时解决测试时出现的 bug
测试人员	王子腾 杨凌霄	进行系统测试，向开发组反馈 bug

表 2 角色分配

5. 软硬件资源说明

软硬件资源名称	级别	详细配置	获取方式	使用说明
服务器	关键	CPU: 1 核 内存: 2G OS:ubuntu 16.04 带宽: 1Mbps	阿里云购买	开发阶段使用
数据库存储	关键	MySQL	阿里云购买	全程使用 自带备份功能
IntelliJ IDEA	关键	版本: 2020.2.3, Ultimate Ver	教育版	开发用IDE
Chrome, edge	关键		已经存在	测试浏览器兼容 本机调试
Git	关键	版本: 2.9.0	已经存在	配置管理 代码管理
Apache Jemeter	关键	版本: 5.3	免费下载	压力测试使用

表 3 软硬件资源说明

6. 专题计划要点

6.1 开发人员培训计划

➤ 培训内容

- 数据库基础，主要为 MySQL 或 MangoDB
- 编程语言培训，主要为 Java，JavaScript
- 编程工具的培训

针对需求，进一步加强软件开发规范和软件开发过程质量，了解客户的真正需求，提高开发人员开发素质，使开发人员成为高质量的软件工程师。结合项目进行案例讨论，根据讨论过程进行评估。通过上述培训，进一步提高开发人员技术水平，扩展开发人员的技术知识面和视野。

6.2 质量保证计划

项目的质量保证工作由质量经理、质量保证人员及项目负责人负责，质量保证人员在项目启动后制定适用于单个项目的《质量保证计划》，并检查和督促计划的实施。

应按照质量保证计划的规定进行项目生存周期的各项活动，加强各阶段质量的早期评审与检查工作，项目质量保证人员参加所有的评审与检查活动。

在项目生存周期过程期间，各阶段责任人应按项目进展情况完成相应阶段的管理及技术文档，还需按规范及其他相关文件的要求填写相应记录，质量保证人员通过对各阶段输出的文档及记录的检查发现存在的质量问题。

6.3 测试计划

6.3.1 静态测试

静态测试是指不执行程序而找出程序存在的错误，这种方法以人工的、非形式化的方法对程序进行分析和测试。在测试中，找出程序中的语法错误，编程所用 IDE 可以从很大程度上解决这个问题。

除此以外，我们可以通过下面的清单来检查程序的评测效果。

➤ 数据引用的错误

- 数组元素下标越界
- 指针变量访问的内存空间非法
- 访问了不存在的空间

➤ 数据说明错误

- 数组或字符串初始化不正确
- 变量类型错误
- 变量名错误
- 变量改变错误

➤ 计算错误

- 不同类型的变量混合计算
- 结果溢出、超出范围
- 运算符或者运算次序不对
- 整形变量使用错误或者算式非法

➤ 逻辑错误

- 控制流程错误
- 循环不终止，即死循环
- 入口条件不正确

➤ 接口错误

- 调用模块以及预备调用模块之间的参数个数、顺序、类型不匹配
- 模块之间的传送的变量的类型、范围、单位定义不一样
- 函数、过程和子程序调用的实参和形参个数、类型、顺序不匹配
- 子程序结果返回错误

➤ 输入输出错误

- 输入非法语句
- 忘记检测输入有效性
- 非法文件输入
- 结果输出错误

➤ 其他错误

-程序功能遗漏

-程序逻辑错误

-防止注入型攻击

经过和以上错误清单的比较，我们不断的修改源程序的语法和逻辑错误，进一步完善程序功能，这一阶段，我们还要规范程序的格式，添加必要的注释，是程序的结构变的清晰，层次分明，以便于其他开发人员的理解、开发、维护和修改。

6.3.2 单元测试

单元测试就是指被测试程序是单个子程序、过程的逻辑测试，我们的任务实验整体功能和接口说明是否有不符合规定的情况，以及编码是否有错。经过静态测试后，要注意测试每一个单元，在输入输出的测试案例上进行调整，这一次的测试重要对于单元进行整理，而非系统整体测试。在这一块测试中，需要对于各单位做内部逻辑检测，分模块进行，具体请见测试案例。同时我们需要审视模块的内聚程度是否足够高；审视模块的对外接口是否合乎设计。

6.3.3 集成测试

单元测试之后，就需要在单元测试的基础上，把每个单元按照设计要求逐步连接起来，进行联合测试，搭建整体的教学管理系统，我们主要是将各模块拼接起来中，进行接口的数据的类型和顺序的匹配。

重点关注：穿越各个模块的数据是否会丢失；各个子功能组合起来，能否达到预期的要求的功能；测试一个模块的功能是否会对另一个模块的功能产生不利影响；全聚德数据结构是否有问题，会不会被意外地修改；单个模块的误差积累起来，是否会放大而达到不可接受的程度。

6.3.4 系统测试

系统不仅仅包括软件本身，还包括计算机硬件及其外围设备、实际运行时大批量数据、非正常操作（比如黑客攻击），主要包括一下几个方面：

- 压力测试：测试服务器能够承受多大的并发访问请求，下载请求等。
- 容量测试：预先分析出系统能承受的某项指标的极限值。对于并发数，确定后可以提前建立线程池，减小反应时间。
- 安全测试：检查系统对非法入侵的防范能力。必须要能防范一般的攻击，如中间人攻击，回放攻击，DDOS 攻击，缓冲区溢出攻击、SQL 注入攻击等。
- 容错测试：检查系统的适度退化的处理能力。

6.3.5 确认测试

此阶段实行一系列黑盒测试，根据测试计划进行 α ， β 测试，其中最好在 α 测试中邀请一定数量的终端用户进行一系列测试。

7. 项目可行性

7.1 经济可行性

经济方面要考虑的是项目的开发运营成本及其经济收益。作为一个基于电影取景、配乐等信息检索的垂直搜索引擎，项目本身涉及多种技术，在开发和学习上需要投入一定的时间探索，但开发难度并不高，只需要我们一个小团队就能完成，在成本上的投入较为低廉。组内成员全体拥有开发计算机，且校内已经充分部署了网络环境，教师所在部门也已经都配备有计算机，同时服务器也可以通过学校的现有资源解决或租赁第三方服务器进行使用，所以在硬件方面不需要较大投入，不至于造成负担。对于信息的存储，如电影描述信息、取景地点信息、相关音频信息之类的数据存储量并不大，用租赁的数据库可以解决。因此在经济上是可行的。

7.2 技术可行性

校园中的网络正常运行。项目组的开发人员具备基本的开发能力，通过开发前对本次项目中使用的技术与工具的培训，可以使项目组成员达到项目开发的要求。应用所采用的架构是被广泛采用的，经受过时间的考验。

在软件开发需要掌握的知识体系有：Unix/Linux 平台技术、企业级数据库技术、Java 语言核心技术、软件工程和设计模式、计算机网络核心技术、超文本标记语言(Hyper Text Markup Language)、统一建模语言(Unified Modeling Language)等，以上技术开发人员在之前的课程中都已经进行过学习，所以技术上是可行的。

7.3 用户使用可行性

随着互联网和浏览器的普及，用户对计算机的基本操作和访问网站的基本使用方法都有了足够的了解，不存在计算机使用上的困难。对于搜索引擎应用本身来说，用户通过以往使用其他垂直搜索引擎，如今日头条、豆瓣、美团等应用的经验可以套用到本应用上来。另外，我们所开发的垂直搜索引擎能够提供帮助功能，用户可从中快速上手。经过统筹分析，我们得出结论：本项目在用户使用上是具有高度可行性的。

8. 技术分析

8.1 前端技术

8.1.1 HTML 语言

超文本标记语言，标准通用标记语言下的一个应用，也是一种规范，一种标准，超文本标记语言它通过标记符号来标记要显示的网页中的各个部分。网页文件本身是一种文本文件，通过在文本文件中添加标记符，可以告诉浏览器如何显示其中的内容。

8.1.2 JavaScript 语言

JavaScript 一种直译式脚本语言，是一种动态类型、弱类型、基于原型的语言，内置支持类型。它的解释器被称为 JavaScript 引擎，为浏览器的一部分，广泛用于客户端的脚本语言，最早是在 HTML（标准通用标记语言下的一个应用）网页上使用，用来给 HTML 网页增加动态功能。

8.1.3 Vue.js 框架

Vue 是一套用于构建用户界面的渐进式 JavaScript 框架。它与其它大型框架不同的是，Vue 被设计为可以自底向上逐层应用。Vue 的核心库只关注视图层，不仅易于上手，还便于与第三方库或既有项目整合。另一方面，当与现代化的工具链以及各种支持类库结合使用时，Vue 也完全能够为复杂的单页应用（SPA）提供驱动。

8.2 后端技术

8.2.1 SpringBoot 框架

SpringBoot 是全新开源的 Java 轻量级框架。它基于 Spring4.0 设计，不仅继承了 Spring 框架原有的优秀特性，而且还通过简化配置来进一步简化了 Spring 应用的整个搭建和开发过程。另外 SpringBoot 通过集成大量的框架使得依赖包的版本冲突，以及引用的不稳定性等问题得到了很好的解决。

8.2.2 MongoDB 数据库

MongoDB 是由 C++ 语言编写的, 是一个基于分布式文件存储的开源数据库系统。在高负载的情况下, 添加更多的节点, 可以保证服务器性能。MongoDB 旨在为 WEB 应用提供可扩展的高性能数据存储解决方案。MongoDB 将数据存储为一个文档, 数据结构由键值 (key=>value) 对组成。MongoDB 文档类似于 JSON 对象。字段值可以包含其他文档, 数组及文档数组。

8.2.3 Scrapy 框架

Scrapy 是一个适用爬取网站数据、提取结构性数据的应用程序框架, 它可以应用在广泛领域: Scrapy 常应用在包括数据挖掘, 信息处理或存储历史数据等一系列的程序中。通常我们可以很简单的通过 Scrapy 框架实现一个爬虫, 抓取指定网站的内容或图片。

尽管 Scrapy 原本是设计用来屏幕抓取 (更精确的说, 是网络抓取), 但它也可以用来访问 API 来提取数据。

8.3 服务器

8.3.1 Apache

Apache 是世界使用排名第一的 Web 服务器软件。它可以运行在几乎所有广泛使用的计算机平台上, 由于其跨平台和安全性被广泛使用。它快速、可靠并且可通过简单的 API 扩充, 将 Perl/Python 等解释器编译到服务器中。

8.3.2 Nginx

Nginx 是一个高性能的 HTTP 和反向代理服务器, 也是一个 IMAP/POP3/SMTP 服务器。这是一款轻量级的 Web 服务器/反向代理服务器及电子邮件代理服务器, 并在一个 BSD-like 协议下发行。其特点是占有内存少, 并发能力强, 事实上 Nginx 的并发能力确实在同类型的网页服务器中表现较好, 中国大陆使用 Nginx 网站用户有: 百度、京东、新浪、网易、腾讯、淘宝等。

8.3.3 Elasticsearch

Elasticsearch 是一个基于 Lucene 的搜索服务器。它提供了一个分布式多用户能力的全文搜索引擎，基于 RESTful web 接口。Elasticsearch 是用 Java 语言开发的，并作为 Apache 许可条款下的开放源码发布，是一种流行的企业级搜索引擎。Elasticsearch 用于云计算中，能够达到实时搜索，稳定，可靠，快速，安装使用方便。官方客户端在 Java、.NET (C#)、PHP、Python、Apache Groovy、Ruby 和许多其他语言中都是可用的。根据 DB-Engines 的排名显示，Elasticsearch 是最受欢迎的企业搜索引擎，其次是 Apache Solr，也是基于 Lucene。

