

海外藏中国文物知识管理与服务平台

项目管理计划

1. 简介

1.1 项目概述

本项目是一个文物知识图谱系统的开发项目，旨在通过构建知识图谱，实现海量文物信息的整合、普及和交流。该项目的目标是将各地博物馆、文物机构的文物信息进行汇总，构建一个全面、系统的文物知识图谱，并开发相应的查询和可视化工具，提供更加方便、直观、全面的文物信息查询和研究服务。

1.2 项目范围说明

本项目的主要内容包括文物知识图谱构建子系统、海外文物知识服务子系统和文物数据爬取子系统。其中，文物知识图谱构建子系统负责构建文物知识图谱，文物数据爬取子系统负责爬取各地博物馆、文物机构的文物数据，海外文物知识服务子系统负责提供查询和可视化服务。

1.3 软件项目计划书的演化

软件项目计划书将根据实际开发情况不断演化，包括需求定义、软件设计、开发、测试、交付等各个阶段的计划和执行情况的更新。

2. 项目组织管理

2.1 过程模型

本项目采用线性模型瀑布模型作为过程模型，按5.1项目进度分阶段进行，每个阶段完成后进行组内讨论和团内审核。

2.2 团队的分工与合作

团号 组号	题目	组长	电话	学号	姓名	班级
团2组 1	知识图谱构建子 系统	宋玉 佳	15386964756	2022040200	宋玉 佳	计科 2202
				2022040204	宋朝 威	计科 2202
				2022040041	陈旭 东	计科 2202
				2022040207	林瀚 之	计科 2202
				2021040066	杨博 超	计科 2202

团号 组号	题目	组长	电话	学号	姓名	班级
				2022040210	徐承昊	计科 2202
组号	题目	组长	电话	学号	姓名	班级
团2组 2	海外文物知识服 务子系统	龙源	17773384332	2022040110	王纤惠	计科 2202
				2022040136	王皓宇	计科 2202
				2022040069	张馨敏	计科 2202
				2022040279	张云菲	计科 2202
				2022040257	陈芊桦	计科 2202
				22022040087	龙源	计科 2202
组号	题目	组长	电话	学号	姓名	班级
团2组 3	知识问答子系统	许楷 烨	19835228872	2022040172	许楷 烨	计科 2202
				2022040295	熊家 怡	计科 2202
				2022040278	杨忆 非	计科 2202
				2022040173	李卓	计科 2202
				2022040073	苗少 凯	计科 2202
组号	题目	组长	电话	学号	姓名	班级
团2组 4	掌上博物馆	李典 坤	15844425379	2022040307	李典 坤	计科 2202
				2022040042	张琪 伟	计科 2202
				2022040321	彭秀 涛	计科 2202
				2022040209	段志 豪	计科 2202

团号 组号	题目	组长	电话	学号	姓名	班级
				2022040286	赵一 泽	计科 2202
组号	题目	组长	电话	学号	姓名	班级
团2组 5	后台管理子系统	罗时 尚	15805240790	2022040341	罗时 尚	计科 2202
				2022040337	张天 瑜	计科 2202
				2022040346	黄岩	计科 2202
				2022040345	徐嘉 聪	计科 2202
				2021010233	杨子 鉴	计科 2202
				2020030020	陈佳 艺	计科 2202

3. 管理过程

3.1 管理目标及优先级

在本项目中，管理目标的重点是实现高质量的知识图谱构建和数据爬取，并为用户提供海外文物知识服务。以下是该项目的管理目标和优先级：

- 构建高质量的知识图谱，实现知识的结构化和标准化。
- 实现高效、准确、稳定的数据爬取，保证数据的完整性和可靠性。
- 提供海外文物知识服务，包括文物信息检索、时间轴展示、关系图谱可视化等功能，为用户提供便捷、丰富的文物知识服务。
- 确保项目开发的可持续性和可扩展性，以满足未来的需求变化。

3.2 项目风险管理

在项目开发过程中，可能会遇到一些风险，例如技术难题、人员流失、进度延误等，这些风险可能会对项目的进度和质量造成影响。因此，需要对这些风险进行管理和控制。以下是本项目的主要风险和应对措施

在该项目中，有一些潜在的风险，包括但不限于以下几种类型：

- 数据来源网站更新导致数据结构变更
- 爬虫被网站识别并限制访问
- 数据量较大导致爬取和处理时间过长
- 人员变动和沟通不畅

为了降低风险的发生概率和减轻风险的影响，我们将采取以下措施：

- 监测网站的更新，及时调整爬虫代码以适应变化。
- 采用IP代理池技术，轮流使用多个IP地址模拟不同的访问来源，以规避网站限制。
- 优化爬虫代码，增加多线程、异步请求等功能，以提高爬取效率。
- 建立良好的沟通渠道，定期开会讨论进展和问题，及时解决人员变动和沟通不畅带来的问题。

3.3 项目沟通管理

为了保证项目的顺利进行，我们将采取以下措施：

- 建立团队内部沟通机制，定期开会、讨论进展和问题，及时解决团队内部的问题。
- 建立和数据来源方的沟通渠道，了解数据更新情况，及时调整爬虫代码以适应数据变化。
- 建立和客户的沟通渠道，及时反馈项目进展和问题，并听取客户的反馈意见。
- 建立项目文档共享平台，包括代码仓库、文档库、会议记录等，以方便团队成员之间的沟通和协作。

4 技术过程

4.1 开发工具、方法和技术

我们将采用以下开发工具、方法和技术：

编程语言：Python、Java, HTML, JavaScript, SQL等

爬虫框架：requests, BeautifulSoup

前端框架：Vue.js、ElementUI、D3.js等

数据库管理系统：MySQL

代码托管平台：GitHub

代码管理工具：Git

使用阿里云服务器管理所有后台数据，包括爬取下来的博物馆数据和用户数据等

主机号：123.56.94.39

端口号：3306

4.2 软件需交付的文档

交付物	阶段
项目管理计划	在项目初始阶段形成，在后续每一个阶段更新
需求规格说明	在需求分析阶段结束形成，在后续的设计和开发阶段更新
设计报告	在设计阶段形成，在后续开发阶段补充和更新
测试报告	在软件开发阶段形成，在后续的测试阶段更新
用户手册	在需求分析阶段结束形成，在后续阶段更新，在系统交付阶段交付
每周会议纪要	每周会议记录，包括讨论内容、本周开发情况和小结。
源代码	在开发阶段形成，每个团的代码在github/Gitee/GitLink等平台上建立一个库，至少每周更行一次。
可执行系统	在开发阶段形成，在测试阶段更新。手机端提交apk文件。

5. 项目进度及成本管理

5.1 进度描述

本项目的整体进度计划如下：

关键时间	任务	要求
第7周	确定分组、制定项目管理计划，确定技术路线	分组名单提交给助教；在 GitHub 上建立项目库，发布管理计划。
第8周	撰写需求规格说明初稿、编写程序	发布需求规格说明书。
第9周	撰写设计报告初稿、编写程序	发布设计报告。
第10-11周	编写程序	建议编程后期每周增加会议次数。
第12周	子系统运行检查	以小组为单位，全部成员参加。每人介绍自己完成部分。与助教老师预约时间。
第13周	总结汇报演讲	以团队为单位，按抽签顺序进行。
第14周	系统集成后的运行检查	提交文档电子版，源代码和可执行系统，汇报 PPT 和演示视频。

5.2 开发过程中的资源需求

- 人员：本班各小组成员
- 支持软件：Visual Studio Code, MySQL, PyCharm
- 实验设备：个人笔记本、实验楼机房PC

5.3 软件管理过程中预算及资源分配

- a) 该项目属于课程项目，不涉及经济预算或分配
- b) 资源分配为团内成员个人拥有的所有设备

6. 各个项目具体进度和管理

6.1 知识图谱构建子系统

技术栈：Neo4j, requests , BeautifulSoup, Python

人员分工：

任务	成员
大都会博物馆 宾夕法尼亚大学考古与人类学博物馆 数据爬取	宋玉佳
克利夫兰 博物馆数据爬取	杨博超
数据清洗	陈旭东
数据翻译	宋朝威 徐承昊
三元组转换及知识图谱制作	林瀚之主要负责 其他同学辅助林瀚之完成
服务器购买及各种环境配置以及数据上传	宋玉佳

目前数据已经清理完毕，已经上传至云端数据库，后续完成知识图谱的构建

6.2 海外文物知识服务子系统

技术栈：

- 数据库：Neo4j
- 前端：Vue.js
- 后端：Flask

项目的目标是使用户能够浏览、查询和可视化展示文物信息。提供多种查看文物的方式，包括详细信息展示和高级搜索选项，并允许用户注册以使用更多个性化功能，如收藏和评论等。

项目进展计划：

时间	项目进度计划
第八周	需求分析阶段 - 详细理解并分析项目需求，设计系统架构和数据库设计 - 完成原型设计 - 同时学习 Neo4j、vue、flask 等相关技术，配置统一的环境
第九周	设计阶段 - 撰写需求规格说明，设计报告，用户手册初稿 继续学习技术，开始着手程序编写： - 在数据浏览功能上，实现文物信息的多种展示形式、筛选排序功能 - 文物详情查看；在数据查询功能上，实现简单查询和高级查询功能。
第十周	开发阶段 - 更新需求规格说明，设计报告，用户手册 - 实现文物知识图谱的可视化展示，包括力导向图、时间轴等 - 实现用户注册登录、信息设置、浏览记录、收藏、评论等功能（选做部分根据进度选择实现）
第十一周	开发阶段 - 撰写测试报告初稿，更新用户手册，完成需求规格说明，设计报告 - 继续完成代码，集成各模块
第十二周	测试阶段 - 更新测试报告，完成用户手册 - 继续优化代码，进行系统测试，修复发现的问题

6.3 知识问答子系统

技术栈：

- 数据库：Neo4j
- 前端：Vue.js
- 后端：Flask, Python

项目目的是通过知识图谱等数据，使用RAG增强大模型的路线，实现知识问答服务。包括单实体属性的问答以及闲聊等功能。

小组分工：

- 许楷烨：组长，负责总体协调把握，并实现后端api接口接入
- 熊家怡：后端代码实现，并实现附加功能
- 李卓：前端代码实现，需求、接口等相关文档撰写
- 苗少凯：协助整理相关文档
- 杨忆非：负责功能测试、PPT制作等

项目进展计划：

时间	项目进度计划
第八周	需求分析阶段 - 详细理解并分析项目需求，设计系统架构和数据库设计 - 完成原型设计 - 同时学习 Neo4j、vue、flask、大模型等相关技术，配置统一的环境
第九周	设计阶段 - 撰写需求规格说明，设计报告，用户手册初稿，继续学习技术，开始着手程序编写： - 确定接口文档，实现基础的本地问答，同时前端开始编写代码。
第十周	开发阶段 - 更新需求规格说明，设计报告，用户手册 - 实现问答系统的前后端连接实现基础的问答 - 实现用户注册登录注册等功能
第十一周	开发阶段 - 撰写测试报告初稿，更新用户手册，完成需求规格说明，设计报告 - 继续完成代码，进行闲聊选做功能技术路线确定
第十二周	测试阶段 - 更新测试报告，完成用户手册 - 继续优化代码，进行系统测试，修复发现的问题

6.4 掌上博物馆

技术栈：

- **开发平台**：华为HarmonyOS
- **开发工具**：华为DevEco Studio
- **前端框架**：ArkUI
- **后端**：Spring Boot
- **接口**：RESTful API

人员分工

任务	成员
需求规格说明书，设计报告	段志豪
测试报告，用户使用手册	彭秀涛
汇报ppt，接口文档	赵一泽
后端代码	李典坤
前端代码	张琪伟

时间安排

关键时间	进度计划
第8周	- 在github上建立项目库，发布管理计划； - 完成原型设计 - 确定分组、制定项目管理计划、确定技术路线。
第9周	- 撰写需求规格说明初稿，发布需求规格说明书； - 撰写设计报告初稿，发布设计报告。 - 学习代码相关语法和工具使用等，同时开始编写程序。
第10周	- 开始编写用户手册； - 正式开始开发主要功能； - 更新设计报告和需求规格说明。
第11周	- 完成主要功能开发，根据进度继续开发选做部分； - 测试主要功能，编写测试报告； - 更新用户手册。
第12周	- 继续进行系统测试，更新测试报告，完善各个文档； - 优化代码，修复问题。

6.5 后台管理子系统

技术栈

- 后端：springboot
- 前端：Vue.js (JavaScript框架)

人员安排

1. 罗时尚 (组长)
 - 负责总体协调把握项目的进度。
 - 修改其中出现的疑难杂症和测试项目编写登录页面，注册页面。
每周编写markdown文档记录工作进度
2. 张天瑜
与徐嘉诚完成审核模块的前后端编写
3. 徐嘉聪
完成了数据库的设计工作，与张天瑜完成审核模块的前后端编写，同时也编写注册，登录页面，找回密码页面等功能
4. 黄岩
完成日志记录的模块的前后端代码编写，数据库恢复等页面功能的编写
5. 杨子鉴
对人员管理的crud操作的编写。
6. 陈佳艺
对人员管理的crud操作的编写。

项目进展计划：

第八周：分配任务，与大团队协商。初步确定技术方案。

第九周：完成基础的crud任务如人员管理，数据管理等。

第十周：完成登录，注册，找回密码等任务的完成。

第十一周：完成日志管理，以及数据库恢复以及持久化等任务。

第十二周：测试子系统的功能。

第十三周：验收。