|  |  |
| --- | --- |
|  | **<项目编号>**  **<项目名称>** |
| 分 类:  <模板>  使用者:  <项目组>  文档编号:  HD-PP-310 | **项目计划**  **<版本号>**  项 目 承 担 部 门：大数据第二组开发计划  撰 写 人（签名）：  完 成 日 期：  本文档 使 用部门： □主管领导 □项目组  □客户（市场） □维护人员 □用户  评审负责人（签名）：  评 审 日 期： |
|  |

**文档信息**

|  |
| --- |
| 标题: 大数据第二组开发计划 |
| 作者: 杨国杰、解沁林、马韵佳、陈真如、董晨露、李晖茜 |
| 创建日期: 20200611 |
| 上次更新日期: |
| 版本: |
|  |
| 部门名称: 大数据第二组 |

修订文档历史记录

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **版本** | **说明** | **作者** |
| 20200611 | 1.0 | 初步进行项目开发计划的规划 |  |

**目录**

1. 前言 5

1.1 目的 5

1.2 术语与缩略语 5

2. 项目概述 5

2.1 项目背景和目标 5

2.2 项目范围 5

2.3 交付的产品 5

2.4 约束和假设 5

3. 角色和职责 6

3.1 利益相关人角色和职责 6

3.2 有关的利益相关人介入计划 6

4. 项目约定 6

5. 项目的已定义过程 6

5.1 项目的生命周期选择 6

5.2 项目阶段划分及主要工作产品 6

5.3 本项目采用的过程 7

5.4 裁剪结论 7

6. 工作任务分解 7

7. 项目估计 7

8. 项目所需技能和培训计划 8

8.1 项目所需技能 8

8.2 项目培训计划 8

9. 项目相关计划 8

10. 开发计划 8

10.1 项目监控计划 8

10.1.1 活动列表 8

10.1.2 改进指标 8

10.1.3 偏差控制（项目计划变更与重估计约定） 9

10.2 风险管理计划 9

10.3 度量与分析计划 9

10.4 质量保证计划 10

10.5 配置管理计划 10

10.6 系统测试计划 10

10.7 介入计划 10

10.8 数据管理计划 10

10.9 项目培训计划 11

10.10 需求管理计划 11

11. 工作环境 11

11.1 开发环境 11

11.1.1 硬件设备 11

11.1.2 支持工具和软件环境 11

11.2 测试环境 12

11.2.1 硬件设备 12

11.2.2 支持工具和软件环境 12

12. 其他 12

**项目开发计划**

# **前言**

## **目的**

为有需求的广告主提供一个对于广告投放的渠道参数方面的数据可视化，将投放对象的各种参数化成简单易懂的可视化表格，提供各项参数的考量，优化广告投放的成本，达到更好的投放效果。

## **术语与缩略语**

B站：[www.bilibili.com](http://www.bilibili.com)

UP主：进行视频制作并且上传视频的用户

# **项目概述**

## **项目背景和目标**

项目名称：B站2019年度百大UP视频数据分析平台

任务提出者：华迪公司

用户：具有在B站投放广告需求的广告主

开发者：大数据实训第二组

项目提出的背景：广告主对于广告投放渠道没有简单易懂的数据模型对于投放的广告进行投放效率方面的评估

## **项目范围**

《B站2019年度百大UP视频数据分析平台项目》立足建立专业性的广告主数据分析平台。该平台UP主的视频播放相关数据为主要参数，为广告主建立标准化的分析平台

需求服务。本网络平台细分企业和个人用户，对UP相关的视频播放量数据进行模块化的分析并且利用可视化技术进行简单可懂的数据图表的展示，突出使用过程的高效率、易用性特点，做到既符合国际主流标准又满足广告投放方和UP主的需要。

本系统基于互联网技术，可以使企业和用户方便的使用本系统，高效的查找出需要的信息数据。

本系统的应用目标是全国的B站的广告投放主和B站成百上千的UP主

## **交付的产品**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **产品名称（包括服务）** | **介质** | **交付日期** | **接收方** | **接受准则** |
| 01 | B站2019年度百大UP视频数据分析平台 | 文件（多种交付方式） |  | 华迪 | 数据可视化  平台 |

## **约束和假设**

需求改动较小，并且小组成员能够完成相应的开发计划

# **角色和职责**

## **利益相关人角色和职责**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **序号** | **角色** | **成员姓名** | **职责** | **备注** |
|  | 项目经理 | 杨国杰 | 推动项目进展，拉动项目的完成度，为小组成员分配相应的任务 |  |
|  | 数据采集和爬取 | 解沁林、李晖茜 | 爬取相应的数据并且生成相关的csv表格 |  |
|  | 数据分析 | 杨国杰、董晨露 | 数据分析，为可视化铺好相应的基础 |  |
|  | 数据可视化 | 马韵佳、陈真如 | 制作相应的可视化图形 |  |

## **有关的利益相关人介入计划**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **角色** | **成员姓名** | **职责** | **介入时间** | **具体任务** | **备注** |
| 1. | 陶建兵 | 系统验收 | 系统完成进行接收阶段之后 | 对于项目进行评审和相关的测试验收 |  |

# **项目的已定义过程**

## **项目的生命周期选择**

本项目的生命周期，采用瀑布模型括计划，需求分析、设计、编码、测试、运行维护六个阶段，阶段自上而下，相互衔接，次序固定。

## **项目阶段划分及主要工作产品**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **阶段划分** | **阶段目标** | **使用的工具和方法** | **主要活动及工作产品** | **阶段周期** |
|  | 项目策划 | 完成项目开发计划 | word文档 | 进行项目探讨，确定开发人员，完成项目开发计划 | 6.3-6.6 |
|  | 需求分析 | 完成数据需求规约 | Word文档 | 确定开发方向，对1用户原始需求进行详细的功能需求分析和数据需求分析 | 6.7-6.8 |
|  | 分析设计 | 完成数据平台系统分析和设计模型 | IBM RationalRose工件 | 通过探讨确定开发工具，设计系统模型 | 6.8-6.13 |
| 完成数据平台系统架构设计说明书 | Word文档 | 对所用开发环境进行设计，进行Python系统架构设计 |
| 完成数据库设计说明书 | Word文档 | 进行数据库设计，研究将数据传入数据库存储 |
| 4. | 编码测试 | 软件代码工程 | Pycharm软件,Mysql | 进行代码设计，并进行测试 | 6.14-6.15 |
| 5. | 系统测试 | 完成软件测试报告 | Word文档 | 进行系统整合，调整界面，检测开发平台是否符合需求 | 6.16 |
| 6. | 项目结项 | 发布工程 | HTML文档 | 对数据平台进行展示 | 6.17 |
| 数据平台安装和使用手册 | Word文档 | 总结数据平台使用手册 | 6.18 |
| 项目开发总结报告 | Word文档 | 对整个项目开发进行总结 | 6.18 |
| 7. | 整个开发过程 | 项目阶段评审报告 | Word文档 |  |  |

## **本项目采用的过程**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **工程Engineering** | **度量与分析MA** | **过程管理ProcessManage** | **项目管理ProjectManage** | **支持Sustain** |
| 需求管理（REQM） | 度量与分析（MA） | 组织过程定义（OPD） | 项目策划（PP） | 过程和产品质量保证（PPQA） |
| 需求开发（RD） |  | 组织过程焦点（OPF） | 项目监督和控制（PMC） | 配置管理（CM） |
| 技术解决方案（TS） |  | 组织级培训（OT） | 风险管理（RSKM） | 决策分析和解决方案（DAR） |
| 产品集成（PI） |  |  | 集成项目管理（IPM） |  |
| 验证（VER） |  |  |  |  |
| 确认（VAL） |  |  |  |  |

# **工作任务分解**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **任务项** | **负责人** | **参加人员** | **交付项，应提交文档** | **预计时间** |
| 项目策划 | 杨国杰 | 全组成员 | 《项目开发计划书》  《项目审批表》 | **6.6** |
| 需求分析 | 杨国杰 | 全组成员 | 《需求分析规格说明书》 | **6.8** |
| 分析设计 | 解沁林 | 解沁林、李晖茜 | 《数据平台系统架构设计说明书》 | **6.13** |
| **数据库设计** | 董晨露 | 杨国杰、董晨露 | 《数据库设计说明书》 | **6.13** |
| **可视化页面设计** | 马韵佳 | 马韵佳、陈真如 | **《可视化页面设计说明书》** | **6.13** |
| **代码测试** | 李晖茜 | 解沁林、李晖茜、杨国杰、 | 《源代码及设计文稿》 | **6.15** |
| **系统测试** | 陈真如 | 董晨露、马韵佳、陈真如 | 《系统（验收）测试计划》  《系统（验收）测试报告》 | **6.16** |
| 项目结项 | 杨国杰 | 全组成员 | 《产品使用说明书》《项目开发总结报告》 |  |

# **项目估计**

|  |  |
| --- | --- |
| **估计类别** | **条件设置** |
| 代码规模估计 | 语言：pyhton/HTML+css  计量单位：行  项目代码估计：2000左右 |
| 工作量估计 | 计量单位：人天  项目人均生产率：50行/人天 |
| 进度估计 | 需求阶段占项目活动工作量比例：20%  设计阶段占项目活动工作比例：20%  编码阶段占项目活动工作比例：40%  测试阶段占项目活动工作比例：20%  公式：项目活动总工作量\*各阶段比例 |

# **项目所需技能和培训计划**

## **项目所需技能**

|  |  |  |  |
| --- | --- | --- | --- |
| **项目组组成人员** | **在项目中角色** | **预计人数** | **所需技能** |
| 项目经理 | 1 | 能够对项目有完整的认识和把控，对小组成员工作有相当的了解。 |
| 前端web | 2 | 能够运用HTML+css进行页面编写；能够使用echarts对数据进行整合分析并可视化。 |
| 爬虫 | 2 | 熟练使用python相关操作对网页数据进行爬取。 |
| hadoop后端 | 2 | 能够在完全分布式环境下进行Hadoop编程，分析、存储和清洗数据。 |

## **项目培训计划**

根据课程相关要求进行相关学习。

# **项目相关计划**

**8.1 需求管理计划**

参见需求规格说明书

**8.2 进度控制计划**

采用在项目生命周期每次迭代结束时对项目的规模、进度、质量、风险进行跟踪、评审和审计的方式对项目过程跟踪和控制；

实行的机制：

每周开始前由项目经理给组员发放周工作计划清单；

每周由项目组成员填写周工作量统计表；

每周由项目经理汇总项目总工作量、评估项目实际进展以及关键路径，并完成《项目进度跟踪表》；

每周由项目经理组织评估项目风险、关键依赖关系，并完成《项目风险跟踪表》；

每月由项目经理组织项目状态评估，并完成《项目状态评估报告》；若项目进度、规模、成本超过一定界限（参见《中央研究院软件项目跟踪与监控规范》、《中央研究院软件项目跟踪与监控指南》要求），经项目SCCB讨论，由项目经理根据规范和指南要求做相应项目的重新评估和计划变更。

迭代完成后，由项目经理完成《迭代评估报告》；

*由项目监控部组织对迭代里程碑进行评审。*

**8.3 质量控制计划**

参见各阶段迭代计划中包含的评审计划以及测试计划。

**8.4 报告计划**

每周由项目经理提交周项目进度跟踪报告、周风险跟踪报告、下周工作计划；

每周由项目需求管理经理提交需求管理跟踪报告；

每月由项目经理提交项目状态评估报告；

每次迭代结束时，由项目经理提交项目迭代评估报告。

**8.5 评测计划**

每月由项目经理按项目状态报告要求提交度量指标，如需求分析时间、需求管理时间、计划工作时间、项目组工作时间、组间协调问题解决、提交状态报告的数量等。

# **开发计划**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 里程碑标识 | 任务名称 | 资源名称 | 开始时间 | 完成时间 | 工期 | 前置任务 | 提交成果 | 备注 |
| 决定项目产品 | 项目策划 | 无 | 6/3 | 6/5 | 2天 | 无 | 项目产品决策 |  |
| 完成项目分工 | 项目分工 | 项目产品所需工作清单 | 6/6 | 6/7 | 1天 | 项目策划 | 项目分工名单 |  |
| 完成数据收集 | 爬虫 | 无 | 6/10 | 6/14 | 3天 | 无 | 数据收集csv文档 |  |
| 完成数据处理 | hadoop后端 | 数据收集csv文档 | 6/14 | 6/20 | 6天 | 爬虫 | 数据处理后的csv文档 |  |
| 完成完整的产品网页 | 前端开发 | 处理后数据文档、数据收集文档 | 6/14 | 6/24 | 10天 | 爬虫、hadoop后端 | 产品网页成品 |  |
| 完成产品测试报告 | 测试 | 产品网页成品 | 6/24 | 6/26 | 2天 | 前端开发 | 产品测试报告 |  |
| 完成项目改进 | 项目改进 | 产品网页成品 | 6/27 | 6/28 | 1天 | 测试 | 产品网页成品改进版，各项报告文档 |  |
| 完成项目汇报 | 项目汇报 | 项目相关的所有资料 | 6/29 | 6/29 | 半天 | 以上所有 | 项目汇报 |  |

## **9.1 项目监控计划**

### *活动列表*

|  |  |  |  |
| --- | --- | --- | --- |
| **跟踪活动** | **周期** | **形式或工具** | **结果或记录** |
| 项目策划 | 2天 | 腾讯会议 | 项目产品决策 |
| 项目分工 | 1天 | 会议投票 | 项目分工名单 |
| 爬虫 | 3天 | python爬虫 | 数据收集csv文档 |
| hadoop后端 | 6天 | hadoop | 数据处理后的csv文档 |
| 前端开发 | 10天 | JavaScript，css，echarts | 产品网页成品 |
| 测试 | 2天 | 黑白盒测试 | 产品测试报告 |
| 项目改进 | 1天 | python爬虫，hadoop，JavaScript，css，echarts | 产品网页成品改进版，各项报告文档 |
| 项目汇报 | 半天 | 腾讯会议 | 项目汇报 |

### *改进指标*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **度量元类型** | **度量元** | **经验数据** | **改进目标数据** | **改进目标制定理由** |
| **生产率** | 代码平均生产率（KLOC/人时） | 0.13 | 0.20 |  |
| **成本** | 总工作量（人日） | 15.38368 | 17.6638 |  |
| **进度** | 项目周期（日） | 27 | 27 |  |
| **质量** | 缺陷清除率（%） | 86 | 95 |  |
| 缺陷密度（个/KLOC） | 0.8 | 0.5 |  |
| 残留缺陷（个） | 4 | 2 |  |

### *偏差控制（项目计划变更与重估计约定）*

|  |  |  |
| --- | --- | --- |
| **监控参数** | **控制值** | **行动** |
| **工作量** | 阈 值：15.39  预警值：5.387 |  |
| **进度** | 关键路径  阈 值：28.93  预警值：4.339  非关键路径  阈值：4.641 |  |
| **规模** | 阈 值：118.3  预警值：41.3 |  |
| **质量** | 控制阈值在相应的测试计划中逐步体现。 |  |

## **风险管理计划**

**9.2.1 风险摘要**

该项目需要爬取哔哩哔哩上的各种数据信息，多次爬取访问可能导致IP地址被哔哩哔哩平台拒绝访问。

该项目爬取的数据是截止到2020年6月1日的数据，后续如果出现数据变动可能无法被分析。

项目涉及的风险包括：技术风险、进度风险

**9.2.2 风险管理任务**

有关风险管理工作程序（风险的识别、风险优先级的确定、确定风险防范降低减弱的措施、风险应急计划、风险跟踪、风险度量等）参见中央研究院《软件项目风险管理工作规范》。

**9.2.3 组织和职责**

项目风险管理活动涉及项目组中：

项目经理——组织项目风险管理活动，参与讨论、解决并跟踪等；

分析设计人员、测试人员——参与风险讨论、状态变化及解决已知风险，并整理书面材料；

SCM人员——承担风险跟踪，对风险进行度量和分析、维护风险跟踪表和报告，并向项目经理提供项目风险汇总信息；

SQA人员——抽查风险管理活动的开展及审计活动是否符合中央研究院《软件项目风险管理工作规范》要求；

**9.3 度量与分析计划**

采用在项目生命周期每次活动任务结束时对项目的规模、进度、质量、风险进行跟踪、评审和审计的方式对项目过程跟踪和控制；

实行的机制：

每个任务开始前由项目经理给组员发放周工作计划清单；

每个任务由项目组成员填写周工作量统计表；

每个任务由项目经理汇总项目总工作量、评估项目实际进展以及关键路径，并完成《项目进度跟踪表》；

每个任务完成后由项目经理组织项目状态评估，并完成《项目状态评估报告》；若项目进度、规模、成本超过一定界限（参见《中央研究院软件项目跟踪与监控规范》、《中央研究院软件项目跟踪与监控指南》要求），经项目SCCB讨论，由项目经理根据规范和指南要求做相应项目的重新评估和计划变更。

迭代完成后，由项目经理完成《迭代评估报告》；

## **9.4 质量保证计划**

参见各活动任务阶段的评审计划和最终的《项目测试分析表》

## **9.5 配置管理计划**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **项**  **目**  **组**  **组**  **成**  **人**  **员** | **在项目中角色** | **预计人数** | **到位时间** | **技能/经验** |
| 项目经理 | 1 | 06/03 | Software Project Management |
| 系统分析员 | 1 | 06/07 | 需求分析,需求管理， UML, Clear Case |
| 构架设计师 | 1 | 06/07 | Rational Rose, UML, Clear Case |
| 实施员 | 6 | 06/07 | pycharm，SQL， Database, JavaScript  ，echarts |
| 测试 | 3 | 06/07 | Test Manage, Clear Case |
| 文档 | 6 | 06/07 | Microsoft Word、编写文档 |

## **系统测试计划**

**9.6.1 测试类型**

功能测试：确保功能测试需求项以及用例场景能够实现。采用黑盒测试技术设计测试用例。要保证所有测试用例使用完毕，所有功能都被测试到。

性能测试：针对系统的响应时间、并发性、吞吐量等方面制定测试用例，分析结果确定系统性能。采用黑盒测试对每个不同性能进行测试。

强度测试：测试在极端情况下可能出现的问题，保证在资源少和共享资源竞争的情况下，无重大问题发生。

容量测试：使用大量数据给软件系统以考验，以确定达到限制时不会引发软件失败。通常对其正常值以增加一个数量级，即\*10来确定“大量数据”的值。

安全性测试：确保本系统权限分明，不会有越权情况出现。

配置测试：确保本系统在有其他软件运行的情况下不受太大影响。

**9.6.2 时间安排和人员安排**

|  |  |  |
| --- | --- | --- |
| **制定内容** | **时间安排** | **参与人** |
| 测试计划 | 2020/6/12-2020/6/13 | 杨国杰 |
| 测试用例 | 2020/6/14-2020/6/16 | 全体组员 |
| 测试日记 | 每个阶段完成后编写 | 全体组员 |
| 测试通知单 | 相应测试完成后 | 杨国杰 |
| 测试分析报告 | 2020/6/25-2020/6/28 | 杨国杰 |

## **介入计划**

具体参见3.2 有关的利益相关人介入计划。

## **数据管理计划**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **计划收集数据** | **格式** | **存储**  **介质** | **保密**  **要求** | **来源** | **使用限制** | **收集**  **时间** | **收集人** | **检查频率** | **其他** |
| 项目开发过程中生的工作产品（包括各种开发文档、报告、纪要等） | 电子 | 腾讯在线文档、群文件 | 公开 | 项目开发中产生 | 组织级和本项目组使用，对于可供参考的内容，根据组织得的需要，可作为公共资源使用。 | 项目开发中各个阶段 | 根据配置管理计划的要求采集 | 根据配置管理计划的要求检查。 |  |
| 项目组收集的参考资料 | 电子 | 腾讯在线文档、群文件 | 小组内部 | 从网上下载或其他人处获得 | 供小组内部所有人员使用 | 项目开发中各个阶段 | 项目组成员 | 根据配置管理计划的要求检查。 | 分类存放 |

## **项目培训计划**

|  |  |  |
| --- | --- | --- |
| **培训内容** | **培训时间** | **指导工程师** |
| python编程与网络数据采集 | 2020/6/2-2020/6/3 | 陶建兵 |
| linux基础 | 2020/6/4 | 许富强 |
| Hadoop与Spark基础与环境搭建 | 2020/6/4-2020/6/5 | 许富强 |
| Spark RDD编程 | 2020/6/6 | 陶建兵 |
| Spark SQL编程 | 2020/6/7 | 陶建兵 |
| 大数据开发 | 2020/6/8-2020/6/9 | 陶建兵 |

## **需求管理计划**

6月2日至6月29日的工作日每天上午9:00-11:50，下午2:20-5:10为固定项目开发时间其余时间自行分配。

# **工作环境**

## **开发环境**

### *硬件设备*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **设备名称／类别** | **配置** | **数量** | **获取方式** | **负责人** | **到位时间** | **费用预算（万元）** | **备注** |
| Laptop | intel core i5 | 1 | 自带 | 李晖茜 | 2020/6/2 | 0 |  |
| Laptop | intel core i5 | 1 | 自带 | 董晨露 | 2020/6/2 | 0 |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Hadoop集群 |  | 1 | 华迪提供 | 杨国杰 | 2020/6/5 | 0 |  |

### *支持工具和软件环境*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **软件／工具** | **版本** | **获取方式** | **负责人** | **到位时间** | **费用预算（万元）** | **备注** |
| Anaconda | 3 | 老师提供 | 无 | 2020/6/2 | 0 |  |
| pycharm | 2019 | 老师提供 | 无 | 2020/6/2 | 0 |  |
| VSCode | 1.46.0 | 网上下载 | 无 | 2020/6/2 | 0 |  |

## **测试环境**

### *硬件设备*

具体参见10.1.1 硬件设备。

### *支持工具和软件环境*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **软件／工具** | **版本** | **获取方式** | **负责人** | **到位时间** | **费用预算（万元）** | **备注** |
| UnitTest | 5 | 网上下载 | 杨国杰 | 预计2020/6/20 | 0 |  |