编码：项目级-软件开发计划-Version\_1.0.0-alpha-20200812

文档等级：□发布级 ■项目级 □团队级

采纳标准：团队文档质量标准

网络文学大数据分析系统

**Big Data Analysis System of Network Literature**

软件开发计划

**Version 1.0.0-alpha**

编写人：阳玉洁

西工大-华迪实训第四小组

**All Rights Reserved**

**目录**

**[1 引言 3](#_Toc24492)**

[1.1 标识 3](#_Toc27918)

[1.2 系统概述 3](#_Toc12811)

[1.3 文档概述 3](#_Toc3157)

[1.4 基线 4](#_Toc28575)

**[2 引用文件 4](#_Toc25978)**

**[3 交付产品 4](#_Toc560)**

[3.1 程序 4](#_Toc22763)

[3.2 文档 4](#_Toc5906)

[3.3 非移交产品 7](#_Toc31430)

**[4 开发及文档编制要求 8](#_Toc1440)**

[4.1 约束和假设 8](#_Toc20965)

[4.2 对项目文档编制的需求和约束 9](#_Toc14985)

**[5 角色和职责 9](#_Toc7285)**

[5.1 利益相关人角色和职责 9](#_Toc22174)

**[6 项目的已定义过程 10](#_Toc11784)**

[6.1 项目阶段划分及主要工作产品 10](#_Toc25863)

[6.2 本项目采用的过程 11](#_Toc23431)

[6.3 部分过程规范 12](#_Toc27643)

**[7 工作任务分解 13](#_Toc19316)**

**[8 项目所需技能和培训技能 13](#_Toc29817)**

[8.1 项目所需技能 13](#_Toc30029)

[8.2 项目培训技能 14](#_Toc18581)

**[9 开发计划 14](#_Toc12574)**

**[10 开发环境 17](#_Toc25937)**

[10.1 开发环境 17](#_Toc125)

[10.1.1硬件设备 17](#_Toc17438)

[10.1.2支持工具和软件环境 17](#_Toc17537)

[10.2 测试环境 18](#_Toc370)

**[11 其他 18](#_Toc28802)**

[11.1 术语与缩略语 18](#_Toc20767)

# 

# 引言

## 标识

适用系统：网络文学大数据分析系统

中文名称：软件开发计划（SDP）

英文名称：Software Development Plan (SDP)

文档版本：1.0.0

文档编号：项目级-软件开发计划-Version\_1.0.0-alpha-20200812

## 系统概述

本文档适用于“网络文学大数据分析系统”项目的开发过程。网络文学大数据分析系统项目由西北工业大学软件学院提出，由西工大-华迪实训第四小组成员负责实施，该项目编号为项目级-软件开发计划-Version\_1.0.0-alpha-20200812，其软件产品版本号为“1.0”。

该项目基于B/S（Brower/Server）浏览器/服务器结构，采用传统的三层结构方式进行解耦开发。对当下最热门的网络文学阅读机构——“晋江文学”提供的各项数据进行可视化展示和分析。数据分析采用Hadoop和Spark。

## 文档概述

本文档依据《GB/T 8567-2006 计算机软件文档编制规范》制定，属于技术文档，仅限于项目相关人员阅读。

本文档用于制定软件开发计划，其内容包括工作任务分解；估算估算工作产品和活动的规模、工作量、成本以及所需资源；制定开发计划（工作进度表）、风险管理计划、质量保证计划、配置管理计划、度量计划等，经评审、批准后作为项目实施和完成的依据，项目跟踪和监控的基础。

本文档的保密范围如下：

①开发人员

②测试阶段人员

③对本文档进行评审的人员或机构

④项目组以及其他有权需要调用本文档的人员

## 基线

基线：GB/T 8567-2006

# 引用文件

《GB/T 8567-2006 计算机软件文档编制规范》，国家标准

《项目级-文档编写格式规范-Version\_1.0.0-alpha-20200812》，工作室编写

# 交付产品

## 程序

网络文学大数据分析系统项目计划提交的程序文件包括：

1. 系统各迭代版本源代码
2. 系统各迭代版本可执行文件

本系统最终展示为前端网站。

**3.2 文档**

网络文学大数据分析系统项目计划提交的文档包括：

1. 《项目级-软件开发计划-Version\_1.0.0-alpha-20200812》

《项目级-软件开发计划-Version\_X.X.X-release-202008XX》

《项目级-软件开发计划-Version\_X.X.X-signed-202008XX》

1. 《项目级-WBS-Version\_1.0.0-alpha-20200812》

《项目级-WBS-Version\_X.X.X-release-202008XX》

《项目级-WBS-Version\_X.X.X-signed-202008XX》

1. 《项目级-软件配置管理计划-Version\_1.0.0-alpha-20200812》

《项目级-软件配置管理计划-Version\_X.X.X-release-202008XX》

《项目级-软件配置管理计划-Version\_X.X.X-signed-202008XX》

1. 《项目级-配置库-Version\_1.0.0-alpha-20200812》

《项目级-配置库-Version\_X.X.X-release-202008XX》

《项目级-配置库-Version\_X.X.X-signed-202008XX》

1. 《项目级-项目问题跟踪表-Version\_1.0.0-alpha-20200812》

《项目级-项目问题跟踪表-Version\_X.X.X-release-202008XX》

《项目级-项目问题跟踪表-Version\_X.X.X-signed-202008XX》

1. 《项目级-软件需求规格说明-Version\_1.0.0-alpha-20200812》

《项目级-软件需求规格说明-Version\_X.X.X-release-202008XX》

《项目级-软件需求规格说明-Version\_X.X.X-signed-202008XX》

1. 《项目级-系统架构设计-Version\_1.0.0-alpha-202008XX》

《项目级-系统架构设计-Version\_X.X.X-release-202008XX》

《项目级-系统架构设计-Version\_X.X.X-signed-202008XX》

1. 《项目级-数据库系统设计说明-Version\_1.0.0-alpha-202008XX》

《项目级-数据库系统设计说明-Version\_X.X.X-release-202008XX》

《项目级-数据库系统设计说明-Version\_X.X.X-signed-202008XX》

1. 《项目级-界面原型设计-Version\_1.0.0-alpha-202008XX》

《项目级-界面原型设计-Version\_X.X.X-release-202008XX》

《项目级-界面原型设计-Version\_X.X.X-signed-202008XX》

1. 《项目级-配置状态报告-Version\_1.0.0-alpha-202008XX》

《项目级-配置状态报告-Version\_X.X.X-release-202008XX》

《项目级-配置状态报告-Version\_X.X.X-signed-202008XX》

1. 《项目级-软件测试计划-Version\_1.0.0-alpha-20200812》

《项目级-软件测试计划-Version\_X.X.X-release-202008XX》

《项目级-软件测试计划-Version\_X.X.X-signed-202008XX》

1. 《项目级-集成测试计划-Version\_1.0.0-alpha-20200812》

《项目级-集成测试计划-Version\_X.X.X-release-202008XX》

《项目级-集成测试计划-Version\_X.X.X-signed-202008XX》

1. 《项目级-单元测试报告1-Version\_1.0.0-alpha-20200812》

《项目级-单元测试报告1-Version\_X.X.X-release-202008XX》

《项目级-单元测试报告1-Version\_X.X.X-signed-202008XX》

1. 《项目级-单元测试报告2-Version\_1.0.0-alpha-202008XX》

《项目级-单元测试报告2-Version\_X.X.X-release-202008XX》

《项目级-单元测试报告2-Version\_X.X.X-signed-202008XX》

1. 《项目级-单元测试报告3-Version\_1.0.0-alpha-202008XX》

《项目级-单元测试报告3-Version\_X.X.X-release-202008XX》

《项目级-单元测试报告3-Version\_X.X.X-signed-202008XX》

1. 《项目级-单元测试报告4-Version\_1.0.0-alpha-202008XX》

《项目级-单元测试报告4-Version\_X.X.X-release-202008XX》

《项目级-单元测试报告4-Version\_X.X.X-signed-202008XX》

1. 《项目级-单元测试报告5-Version\_1.0.0-alpha-202008XX》

《项目级-单元测试报告5-Version\_X.X.X-release-202008XX》

《项目级-单元测试报告5-Version\_X.X.X-signed-202008XX》

1. 《项目级-系统测试1\_主界面-Version\_1.0.0-alpha-202008XX》

《项目级-系统测试1\_主界面-Version\_X.X.X-release-202008XX》

《项目级-系统测试1\_主界面-Version\_X.X.X-signed-202008XX》

1. 《项目级-系统测试2\_二级界面-Version\_1.0.0-alpha-202008XX》

《项目级-系统测试2\_二级界面-Version\_X.X.X-release-202008XX》

《项目级-系统测试2\_二级界面-Version\_X.X.X-signed-202008XX》

1. 《项目级-类图设计-Version\_1.0.0-alpha-202008XX》

《项目级-类图设计-Version\_X.X.X-release-202008XX》

《项目级-类图设计-Version\_X.X.X-signed-202008XX》

1. 《项目级-项目周报-Version\_1.0.0-signed-202008XX》
2. 《项目级-项目工作日志-Version\_1.0.0-signed-202008XX》

**3.3 非移交产品**

《项目级-文档编写格式规范-Version\_1.0.0-alpha-20200812》

# 开发及文档编制要求

## 约束和假设

约束：

1. 项目交付日期为2020年8月29日
2. 爬取数据量达到百万级别
3. 使用Hadoop集群资源管理系统管理存储数据
4. 使用Spark框架对数据清洗处理

假设：

1. 无人员变更成本
2. 因事因病请假次数在项目期间每人不超过一次
3. 项目进度能按照进程规划如期进行

## 对项目文档编制的需求和约束

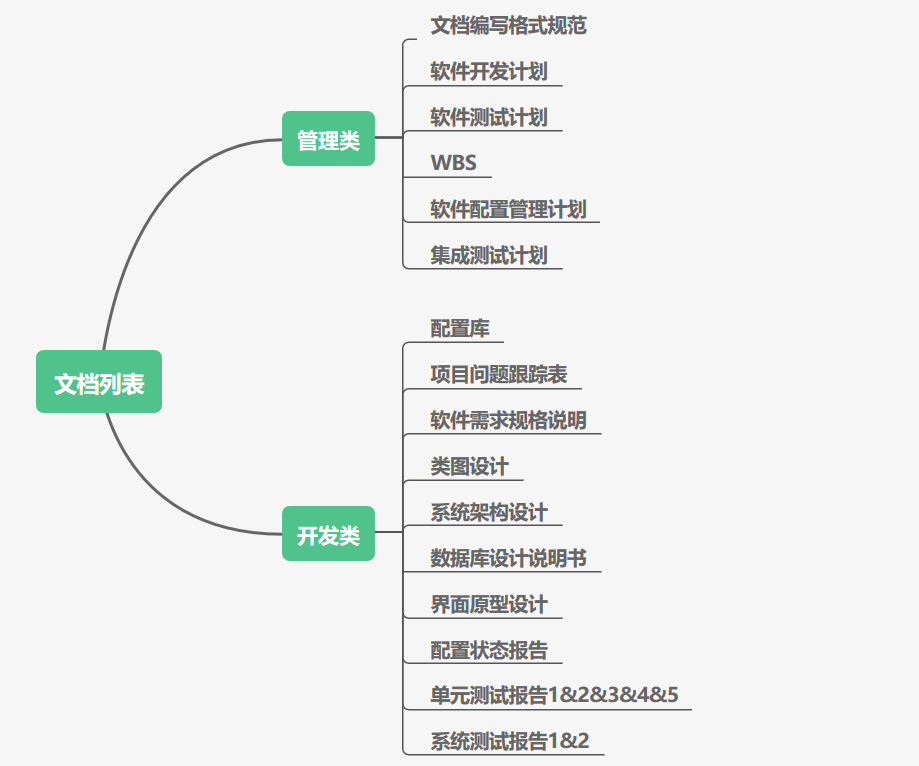


图1 文档归档方案

# 角色和职责

## 利益相关人角色和职责

表1 利益相关人角色和职责分配图

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 角色 | 成员 | 职责 | 备注 |
| 1 | 项目经理&系统分析师 | 焦晨阳 | 负责团队成员的任务调度与任务进程管理；负责了解需求，规划设计产品架 | 无 |
| 2 | 数据清洗师 | 阳玉洁 | 爬取数据并分类整理 | 无 |
| 3 | 前端设计师 | 张婷睿 | 前端界面的框架构建设计编码 | 无 |
| 4 | 后端设计师 | 全芷玥 | 后端代码编写，数据清洗算子设计 | 无 |
| 5 | 功能测试师 | 李明霞 | 对产品进行功能性测试并返回测试报告 | 无 |
| 6 | 数据库操作师 | 康艳晴 | 对数据进行清洗并整合入云数据库 | 无 |

# 项目的已定义过程

## 项目阶段划分及主要工作产品

表2 项目阶段计划表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 阶段划分 | 阶段目标 | 使用的工具和方法 | 主要产品及工作产品 | 阶段周期 |
| 1 | 需求分析 | 明确需求 | 头脑风暴 | 软件需求规格说明 | 8/12-8/14 |
| 2 | 获取数据与集群环境搭建 | 爬取所需信息 | PyCharm、  谷歌浏览器  Requests、BeautifulSoup | 编写网站爬虫爬取网文阅读信息并储存在本地 | 8/15-8/17 |
| 搭建Hadoop、Spark集群环境 | VmVare下配置虚拟机系统、使用pip命令安装 | 搭建集群，并为集群内每一台机器配置Hadoop与Spark环境 | 8/18-8/20 |
| 3 | 数据处理与构建数据库 | 完成数据清洗与入库 | Spark集群计算平台 | 编写Spark算子处理数据并将数据整合入数据库表格 | 8/21-8/23 |
| 4 | 前后端编写、数据处理与显示 | 完成项目后端 | PyCharm，Python语言编写 | 编写后端代码 | 8/24-8/27 |
| 完成前端 | PyCharm、Web storm | 编写前端代码 |
| 完成前后端对接 | Ajax | 编写交互代码 |
| 完成平台数据展示 | Echarts | 设计与美化前端界面，从数据库导入数据 |

## 本项目采用的过程

表3 依据CMMI3制定的适合于本项目的过程

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 工程 | 度量与分析 | 过程管理 | 项目管理 | 支持 |
| 需求开发（RD） | 度量与分析（MA） |  | 项目计划（PP） | 过程和产品质量保证（PPQA） |
| 需求管理（REQM） | 确认（VAL） |  | 项目监督和控制（PMC） | 配置管理（CM） |
| 技术解决方案（TS） | 验证（VER） | 组织级培训管理（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） |  |

## 部分过程规范

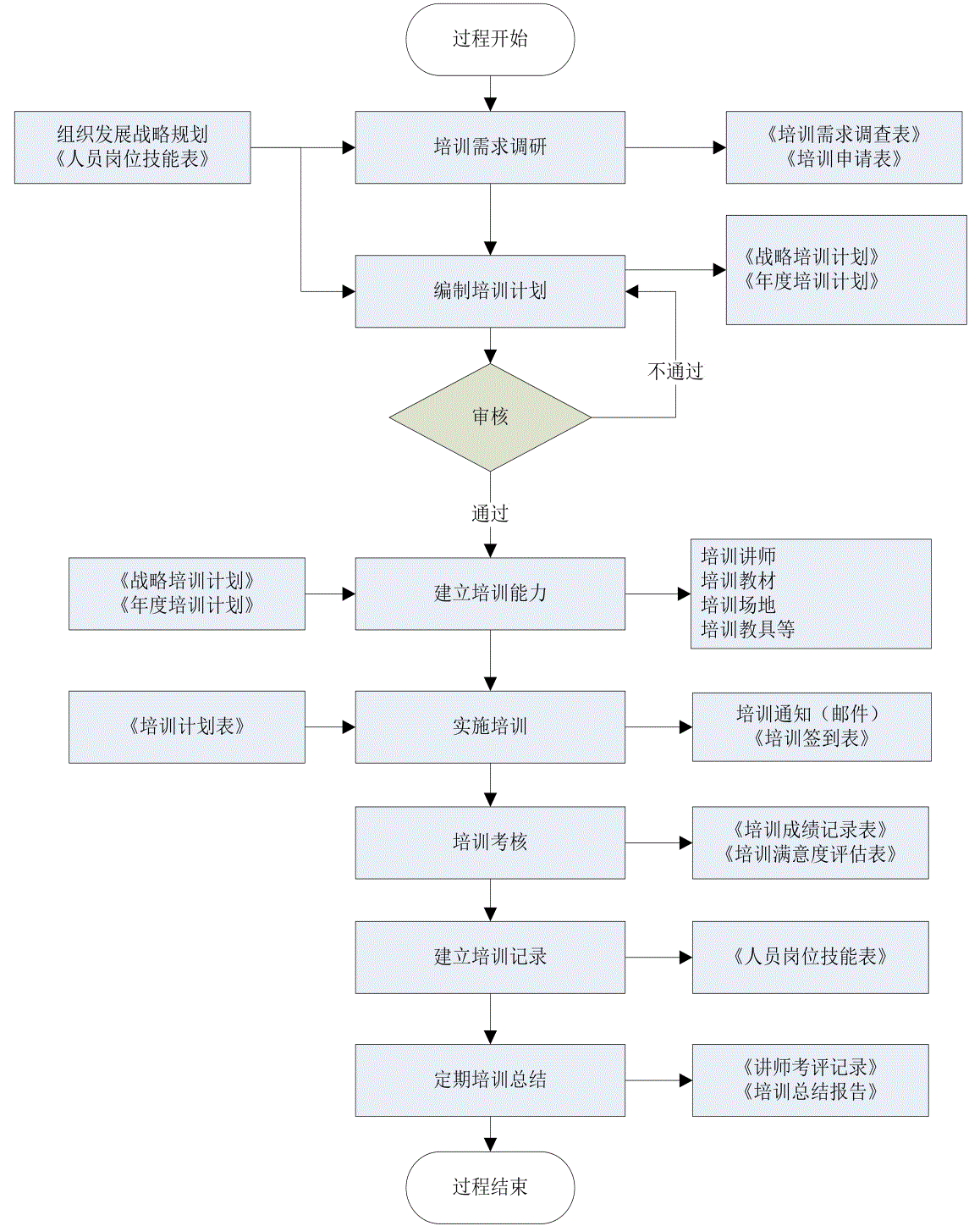


图2 组织级培训管理（OT）标准流程

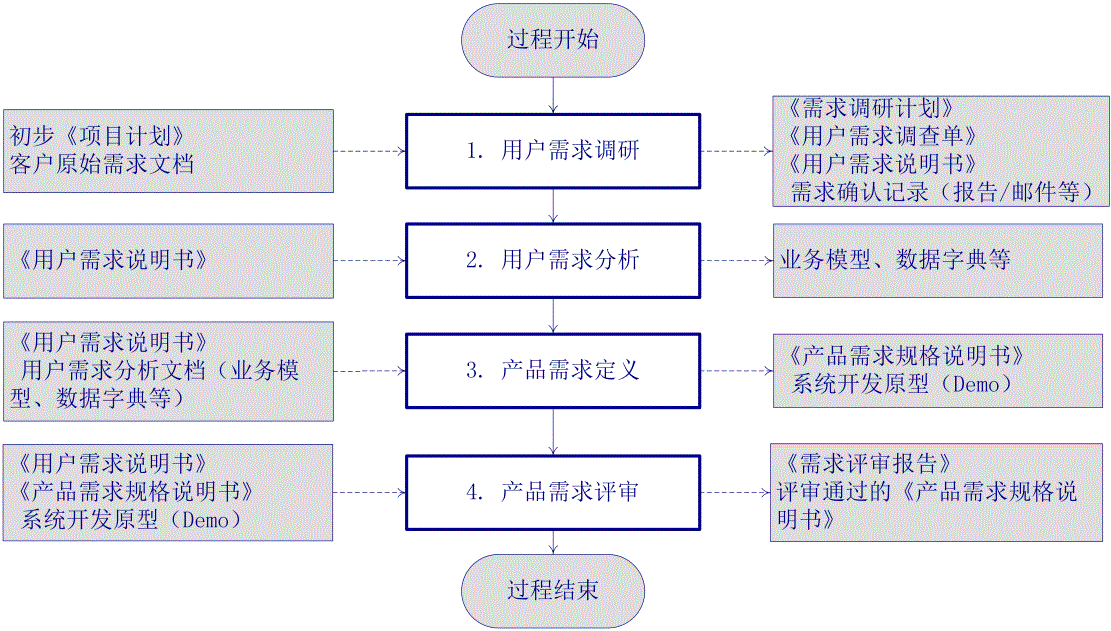


图3 需求开发（RD）标准流程图

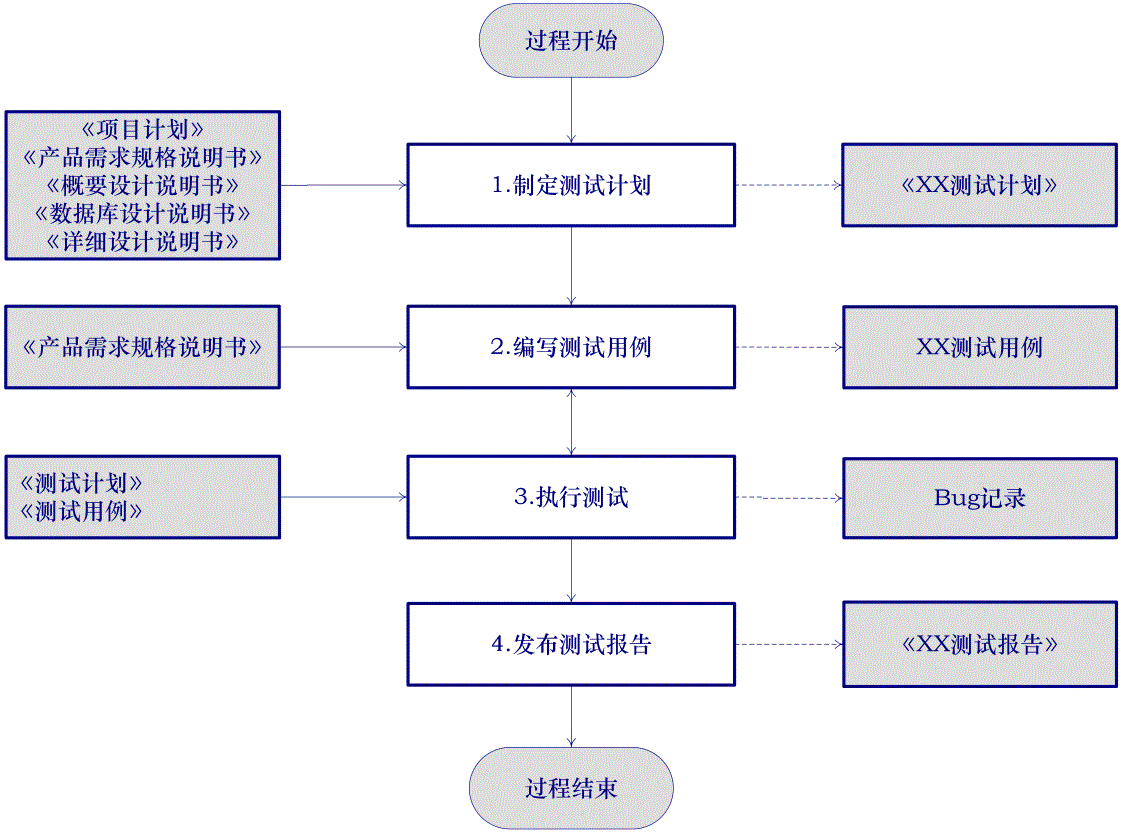


图4 验证（VER）标准流程图

# 工作任务分解

详情请参见：《项目级-WBS-Version\_1.0.0-alpha-20200812》

# 项目所需技能和培训技能

## 项目所需技能

管理技能：项目负责人对整体进度的调度调整，对人力资源进行合理分配的技能。

技术技能：网站爬虫、前后端实现、数据库交互、分布式集群计算。

业务技能：团队成员相互沟通的能力、解决开发过程中遇到的各种代码问题的能力等。

## 项目培训技能

1. Python语言基础运用
2. Requests、Scrapy爬虫框架实例教学
3. 虚拟机分布式集群环境搭建教程
4. 前端框架构建教程

# 开发计划

**9.1 项目监控计划**

**9.1.1 活动列表**

表4 活动列表

|  |  |  |  |
| --- | --- | --- | --- |
| 跟踪活动 | 周期 | 形式或工具 | 结果或记录 |
| 进度跟踪 | 每3天进行一次 | 人工检查+召开会议 | 全部按时完成 |

**9.1.2 偏差控制（项目计划变更与重估计约定）**

表5 偏差控制表

|  |  |  |
| --- | --- | --- |
| 监控参数 | 控制值 | 行动 |
| 工作量 | 阈 值：25%  预警值：35% | 对因事因病造成的工作量缺失采取空闲时间加班的方法 |
| 进度 | 关键路径：产品前端与后端编码  阈 值：20%  预警值：25%  非关键路径：爬取数据，清洗并入库  阈值：25% | 未能按进度完成的工作内容将会由项目经理重新统一调配，将超时任务优先分配给已处理完自身任务的成员 |
| 规模 | 阈 值：15%  预警值：20% | 扩大爬取规模，获得足够满足要求的数据量 |
| 质量 | 控制阈值在相应的测试计划中逐步体现。 | 对产品质量进行定期测试，记录测试结论并随后优化 |

**9.2风险管理计划**

1. 人员流动与请假风险：无流动可能、因事因病请假每人每周控制在8个公时。
2. 项目要求变更风险：采取敏捷开发模式，将产品功能模块化，高内聚低耦合，方便追加功能。
3. 技术风险计划：尽可能选用开发社区活跃、开发体系完整成熟的语言进行项目开发，统一组内成员开发环境。

**9.3度量与分析计划**

表6 度量与分析计划表

|  |  |  |  |
| --- | --- | --- | --- |
| 度量周期 | 度量任务描述 | 负责人 | 参与成员 |
| 每周 | 描述一周内工作执行情况与完成度 | 项目经理 | 全部小组成员 |
| 每周 | 各组员按照自己工作内容完成周报 | 各位成员 | 全部小组成员 |
| 系统测试完成后一天 | 测试人员生成《系统测试报告》进行反馈 | 项目经理 | 测试人员 |

**9.4质量保证计划**

1. 每阶段任务结束前进行质量保证测试，根据测试结果优化软件产品。
2. 质量保障内容：可修改性、易用性、可移植性。

**9.5配置管理计划**

1. 配置管理内容：源代码版本控制、开发环境配置、数据整合控制
2. 配置管理方案：定义各类配置项如库、出库的准则和操作流程；定义基线变更的准则和操作流程；明确配置库的备份及维护的方法，当出现异常后如何恢复的预案等；版本发布的准则、发布流程及发布计划，如测试版本、β版本等。
3. 管理报告内容：基线库的基线项的清单、基线项的名称、版本、存放位置。

**9.6系统测试计划**

详情请参见：《项目级-集成测试计划-Version\_1.0.0-alpha-20200812》

**9.6数据管理计划**

表7 数据管理计划表

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **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 | 硬盘（配置库） | 公司 | 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，可作为公司内部公共资源使用。 | 2020/08/13 | 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。 |  |
| 开发过程打印的纸质资料（包括各种开发文档、报告、纪要等） | 纸质 | 档案袋 | 公司 | 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，根据公司的需要，可作为公司内部公共资源使用。 | 2020/08/13 | 项目配置人员 | 每个开发周期开始前 | 分类存放 |
| 项目组收集的参考资料 | 电子 | 硬盘  （资料库） | 公共 | 开源社区等 | 供公司内部所有人员使用 | 2020/08/13 | 项目组成员 | 每当需要时 | 分类存放 |

# 10 开发环境

## 10.1 开发环境

### 10.1.1硬件设备

表8 硬件设备列表

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 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 |
| 个人笔记本 | 8G运存 | 6 | 组员自带 | 各组员 | 8/03 | 0 |  |

### **10.1.2支持工具和软件环境**

表9 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 |
| JetBrains PyCharm | 2020 | 官方下载 | 无 | 8/05 | 0 | 学生资格认证 |
| 虚拟机系统 | 7 | 国内镜像 | 无 | 8/06 | 0 |  |
| MySQL数据库管理 | 6.5 | 官方下载配置 | 无 | 8/11 | 0 |  |
| Redis | 5.0.9 | 官方下载配置 | 无 | 8/11 | 0 |  |

## 10.2 测试环境

详情请参见：《项目级-集成测试计划-Version\_1.0.0-alpha-20200812》

《项目级-软件测试计划-Version\_1.0.0-alpha-20200812》

# 11 其他

## 11.1 术语与缩略语

PyCharm：是一种Python IDE，带有一整套可以帮助用户在使用Python语言开发时提高其效率的工具。

Requests：用[Python](http://blog.csdn.net/shanzhizi" \t "_blank)语言编写，基于 urllib，采用 Apache2 Licensed 开源协议的 HTTP 库。

Hadoop：以分布式文件存储和MapRuduce为核心的基础架构，用户可以在不了解分布式底层细节的情况下，开发分布式程序。充分利用集群的威力进行高速运算和存储。

Apache Spark：是专为大规模数据处理而设计的快速通用的计算引擎项目概述。