

# IRBL 迭代一：软件项目计划书

IRBL Software Project Scheme Specification

--基于信息检索的缺陷定位系统

khyyds 小组

2021年3月11日

## 1.项目介绍

### 1.1 项目概述

#### 1.1.1 目的

编写本项目开发计划的目的是把IRBL系统开发过程中对各项工作任务的责任人员、开发的进度、预期的功能、硬件和软件资源条件等问题所做的安排用文档的形式记载下来，以便根据本计划开展和检查项目开发工作，保证项目开发成功。

#### 1.1.2 项目背景

随着软件工程的发展，对系统缺陷的定位逐渐复杂也愈发重要。缺陷定位是最常见的开发实践之一，缺陷的存在会影响软件系统的正常使用甚至带来重大危害。本项目将使用基于信息缺陷的定位手段，将缺陷定位转化为以bug报告为查询，代码文件为搜索对象的搜索问题，为每个缺陷报告生成一份按照程序实体与缺陷相关度降序排序的列表，提升软件开发的效率。

#### 1.1.3 项目的范围与目标

##### 1.1.3.1 范围描述

首先，使用系统前应将需要进行缺陷定位的文件放置在特定文件夹，方便系统进行读取。

其次因为信息检索是面向特定项目文件的，IRBL系统应该进行文件读取与预处理。

在完成准备工作后，系统应对预处理过的报告进行进一步处理，通过TF-IDF、权重计算等方式找到与缺陷高相关性的文件。

系统应该具有一个交互性好，鲁棒性强的命令行作为用户输入界面，进行指令执行功能。

1.1.3.2 主要功能

- （一）terminal：有用户友善、稳定性高的命令行系统
- （二）预处理：对bugReport文件以及代码文件进行预处理
- （三）文件的读取与写入
- （四）word count计算及权重计算
- （五）TF-IDF计算
- （六）相似度计算
- （七）结果打印

1.2 项目可交付成果

1.2.1 需完成的软件

- 1. 源程序
- 2. 数据库对象创建语句
- 3. 可执行程序

1.2.2 需要完成的文档

序号	开发任务	交付成果	文档类型	参照标准
1	软件项目管理	软件项目计划书	管理文档	《GB8567—88计算机软件产品开发文件编制指南》
2	软件架构开发	需求规格说明文档	开发文档	事件驱动架构
3	软件系统开发	详细设计文档	开发文档	螺旋模型
4	软件系统测试	测试文档	验证与确认文档	《软件质量模型与度量》

### 1.3 有关定义和缩写说明

IRBL： Information Retrieval Based Bug Localization，基于信息检索的缺陷定位

terminal： 命令行窗口，用于接收命令和反馈结果

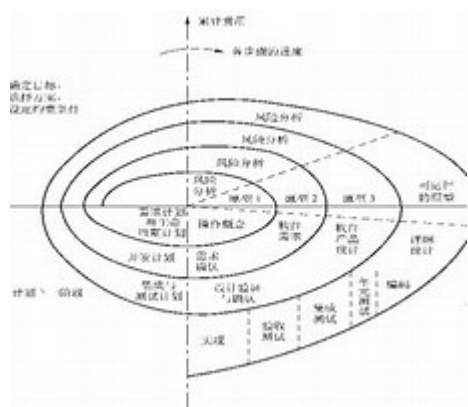
TF-IDF： 一种统计方法，用以评估一字词对于一个文件集或一个语料库中的其中一份文件的重要程度

word count： 词数统计

## 2. 项目组织

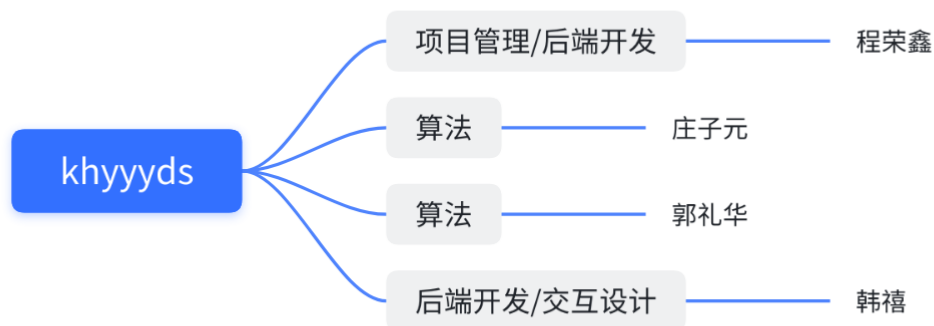
### 2.1 开发模型

在IRBL系统中，由于对软件质量有较高要求，且整个开发过程分为三期，需要通过迭代不断提升软件质量，并且涉及需求变更和某些模块的变化，采用螺旋模型可以兼顾了快速原型的迭代的特征以及瀑布模型的系统化与严格监控，并且引入了其他模型不具备的风险分析。



### 2.2 开发人员

系统开发项目组人员结构如下图所示：



### 3. 管理过程

#### 3.1 管理目标和优先级

##### 3.1.1 管理目标

基本管理原则：每位成员既是项目的决策者，又是项目的开发者，共同完成项目的开发、质量把控、问题处理以及交付流程。决策应在充分的讨论基础上由大家共同做出，一旦决策做出就必须被及时有效的执行。

目标1：按时按量完成项目的迭代过程，实现基本功能。

目标2：遵循规范的项目运作标准，文档严谨完整，代码注释充分，便于后续维护

目标3：项目的最终数据良好，在系统缺陷定位场景的作用可以优于现有结果

目标4：交互体验良好，用户使用方便，无学习成本

#### 3.2 风险管理

本计划主要针对项目开发涉及到的风险，包括在项目开发周期过程中可能出现的风险，以及项目实施过程中外部环境的变化可能引起的风险进行评估。

##### 3.3.1 风险条目

###### 3.3.1.1 管理风险

(一) 项目进展因突发性困难导致不符预期

(二) 不可抗力因素造成的影响

#### 3.3.1.2 技术风险

(一) 设计错误导致程序实现困难

(二) 缺乏质量跟踪

(三) 测试存在缺漏项

(四) 算法研究过程中遇到瓶颈

#### 3.3.1.3 开发环境风险

(一) 所使用开发软件的质量问题

(二) 开发环境不稳定

(三) 服务器不稳定或被攻击

## 4. 技术过程

### 4.1 方法、工具和技术

本小组的团队组织结构为民主团队架构，编程语言为 java, 采用面向对象的分析设计方法，利用 UML 进行系统建模，使用MySQL 5.7作为数据库管理系统，并采用Alibaba标准的文件命名方式、代码版式、注释等编码规范，编码人员对代码进行严格检查后再进行代码编译，测试人员根据测试文档进行单元测试，最后实现软件的交付。开发环境为：IntelliJ IDEA+Python3.8 + Anaconda3 + MySQL 5.7，

部署环境为：CentOS 7.6

### 4.2 计算机系统

团队内所使用的开发及运行系统有：

Windows 10 1909

Mac OS

Manjaro

CentOS 7.6（服务器环境）

多样化的系统开发环境可以保证软件的高可移植性、高稳定性。