**Project Plan**

**1 Summary of the project**

1.1 Project Description

With the development of the times, the search engine built into the City University website is increasingly unable to meet the needs of students' searching.

In order to help CityU students make better use of the resources provided by the Computer Science department, we have developed this software and provided two efficient and practical sorting methods for sorting results.

1.2 Stakeholders

Related stakeholders of the software:

|  |  |
| --- | --- |
| Stakeholders | Description & Responsibility |
| Project Manager | Project Manager |
| Project Team | Developing the Software for Users. |
| Users  (CityU students) | Users could input several words and select a sorting method between TF-IDF and PageRank and view the sorted searching results. |

1.3 Objectives of the project

The purpose of developing this software is to help CityU students and teachers to better search, use and arrange various resources provided by the Department of Computer Science, such as servers, computers, VPNs, etc.

This search engine is quite efficient and practical. The built-in two sorting methods (TF-IDF and PageRank) would help users to find the required web resources faster and more accurately. In other words, you could avoid wasting time on finding resource in CityU website by hand or through the inefficient engine

**2 Summary of Methodology**

2.1 Software Development Methodology

Concurrent Engineering Model is used for planning software development since nearly all classes (components) could work without any other ones. To promise the low coupling, we create so many dataset files (in real practice, we would use database) which store just one kind of information of the URLs and many well-designed interfaces are used.

As shown in the following picture, nearly all components could be developed individually. By using divide and conquer principle, each team member would focus on several components by spiral or evolutionary approach.

图片包含 文字, 地图



自动生成的说明

2.2 Project Team Organization

|  |  |  |
| --- | --- | --- |
| Name | Title | Job Description |
| LI Haoran | Project Manager | Management, Development, Testing |
| ZHAO Zinan | Software Engineer | Development, Testing |
| HY C Matthew | Software Engineer | Development, Testing |
| LIU Jinchao | Software Engineer | Development, Testing |
| GUO Yuhan | Software Engineer | Development, Testing |

2.3 Development Tools

2.3.1 Development Tools

1. Eclipse IDE for Java Developers Version: 2018-09 (4.9.0)

2.3.2 Development Platform

1. Mac OS X 10.14.1
2. Windows 10 1607

2.3.3 Test Cases

1. Junit 4
   * 1. Coverage Checking

a.

2.3.5 Testing Platform

1. Mac OS X 10.14.1
2. Windows 10 1607

2.3.6 Documentation

1. Visual Paradigm Version 15.1 Build 20181103: Sequence Diagrams, Class Diagrams and Use Case Diagrams
2. Javadoc: Generating Java manual
3. Bugzilla: Bug report
4. Microsoft Word: Reports

2.3.7 Project Management

1. GitHub
2. Google Drive

2.3.8 Presentation

1. Microsoft PowerPoint
2. Google Slides

2.4 Configuration Management

**GitHub** is used to host various versions of source code of this program, as it allows users from the same team to cooperate by working on different parts, other team members can sync the program and know the status of those works. If necessary, they can help modify and commit the modified part to GitHub. As it keeps the record of each commit, version control can be done by tracing commits.

**Google Drive** is used to manage documents as team members can edit on a document in the same time, the work can be reviewed in real time. Also, the record of each changes made by different members can be reviewed, which is convenient to trace the versions.

2.5 Work Breakdown Structure

We choose the process type WBS and the work is distributed by top-down method, which the assignation of sub-tasks is done according to a milestone.

1.0 stage0

1.1 Project Management

1.2 Requirements Gathering

1.3 Analysis & Design

2.0 stage1 Release 1.0

2.1 Software Development

2.1.1 Query Part (Implement Query and 1 sorting method)

2.1.1.1 Main

2.1.1.2 Common Query

2.1.1.3 TF-IDF Rank

2.1.2 Preprocess Part (Generate Forward and Inverted method)

2.1.2.1 Forward Index Module

2.1.2.2 Inverted Index Module

2.1.2.3 String Map Module

2.1.2.4 Extract Body Module

2.2 Testing

2.2.1 Unit Testing

2.2.2 Other Testing

2.3 Debugging & Refactoring

3.0 stage2 Release 2.0

3.1 Software Development

3.1.1 Query Part

3.1.1.1 Main (Display more information)

3.1.1.2 PageRank

3.1.2 Preprocess Part (Generate Forward and Inverted method)

3.1.2.1 Extract Link Module

3.1.2.2 Extract Title Module

3.2 Testing

3.2.1 Unit Testing

3.2.2 Other Testing

3.3 Debugging & Refactoring

3 Project Schedule

Project Schedule can help us make the best use of time. We plan to complete a release in about 3-4 weeks, including 1 to 2 weeks for software development, 1 week for unit testing and 1 week for debugging and refactoring.

Here is the screenshot of project schedule:

图片包含 屏幕截图



自动生成的说明

4 Future Planning

In the future, we would like to implement functions in order to enhance the application.

1. Use cookies or other to provide more personalized search results

May be the search engine could store what the user have clicked and analysis their personal preference.

1. Provide other available sorting methods
2. Increase the probability that certain webpages will be searched

For some important webpages which contain some important or useful information or resources, we should improve their search rankings.