

**基于Scrapy的模板化爬虫程序管理平台**

**变更与管理分析报告**

v1.0

赵正阳 SY1906428

郭浩隆 SY1906430

沈一聪 SY1906510

梁远志 SY1906503

宋冰晨 SY1906429

2020年6月4日

**版本变更记录**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本 | 日期 | 修改人 | 修改章节 | 备注 |
| 1.0 | 2020/6/4 | 梁远志 | / | 初版 |

目录

[1. 实验目的 1](#_Toc42176547)

[2. 方法与工具 1](#_Toc42176548)

[3. 配置管理计划 1](#_Toc42176549)

[3.1 文档管理 1](#_Toc42176550)

[3.2 代码管理 2](#_Toc42176551)

[3.3 管理责任分配与监督 3](#_Toc42176552)

[4. 变更控制和版本管理 4](#_Toc42176553)

[4.1 文档版本管理 4](#_Toc42176554)

[4.2 代码变更控制 21](#_Toc42176555)

[4.3 分析与总结 28](#_Toc42176556)

# 1. 实验目的

在项目开发过程中，对文档和代码使用高效的版本管理方法和工具来维护版本变更和记录变更信息，便于项目开发人员协同工作、掌控进度，从而使项目开发更加敏捷高效。

# 2. 方法与工具

本实验的主要工具是Git，Git是一个分布式版本控制软件，最初由Linus Torvalds开发并使用在Linux内核的开发过程。为了实验方便，本实验使用了GitHub作为版本控制的软件源代码托管服务平台。本组也使用了各类GitHub专有的持续集成功能来进行配置管理。

# 3. 配置管理计划

本章主要介绍本组对文档管理和代码管理的计划。

## 3.1 文档管理

文档的管理本组按实验要求使用了GitHub的仓库，仓库地址为：<https://github.com/bhsei/20_H>

本组为文档定义了标准的命名规范，以本文档为例“H-EasySpider-变更与管理分析报告v1.0”，其中“H”是本组组号，“EasySpider”是本组开发项目的名称，“变更与管理分析报告”是本文档的标准名称（参照《软件工程综合实验 概述》），“v1.0”是文档的版本号。

文档的版本号视文档的变更频繁情况（需要评审的文档会更为频繁），对于变更频繁的文档使用三段的版本编号（如“v1.0.1”），其他文档使用两段的版本编号（如“v1.0”）。除了文件名外，文档的封面会注明文档的版本号以及最后修订时间。

每一个文档在封面后会放置一个“版本变更记录”表格，由对文档进行变更的组员进行填写，表格表头如下表所示，包括文档版本号、文档修订日期、修订人姓名、修改的章节范围以及备注。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本 | 日期 | 修改人 | 修改章节 | 备注 |

Git工具能够记录每一次变更的完整信息，因此也能够回退还原某一个特定提交的版本，因此在文件树中仅保留最新的文件版本。在Git中的目录树如下图所示。

├───1.会议记录

├───2.评审意见

├───3.项目PPT

├───4.项目提交制品

│ ├───4.0实验准备

│ ├───4.1软件需求分析

│ │ └───图片

│ │ ├───RUCM

│ │ └───用例图

│ ├───4.2软件需求评审

│ │ ├───初评审

│ │ └───复评审

│ ├───4.3软件设计与实现

│ ├───4.4软件测试

│ │ ├───1-单元测试

│ │ ├───2-集成测试

│ │ ├───3-场景测试

│ │ ├───4-功能测试

│ │ ├───5-非功能测试

│ │ └───归档

│ │ └───测试用例

│ ├───4.5软件测试评审

│ │ ├───初评审

│ │ └───复评审

│ ├───4.6软件项目计划与监控

│ │ ├───工作日志

│ │ └───项目进度计划（甘特图）

│ ├───4.7软件配置管理

│ ├───4.8软件工程实验追踪与分析

│ └───4.9实验总结

└───5.资料

## 3.2 代码管理

由于Git仓库不适用于管理文档等二进制的格式的文档的版本变更，文档仓库仅用于文档上传和说明用途。代码仓库将另外创建，可以使用submodule功能在文档仓库提供一个链接。代码仓库需要自主的控制权限以使用GitHub提供的Pull Request和Action功能分别用于变更管理和自动构建，测试或部署。

预计创建的代码仓库如下表所示。

|  |  |  |
| --- | --- | --- |
| 仓库名称 | 内容 | 仓库地址 |
| easyspider\_web | Django项目 | https://github.com/easy-spider/easyspider-web |
| easyspider\_scheduler | 调度程序项目 | https://github.com/easy-spider/easyspider-scheduler |
| easyspider\_template | 爬虫程序的基础模板 | https://github.com/easy-spider/easyspider-template |
| easyspider\_site\_douban | 针对豆瓣的爬虫模板 | https://github.com/easy-spider/easyspider-site-douban |
| easyspider\_site\_... | 更多模板 | N/A |

本组计划使用Pull Request来进行代码的变更合并，通过设置GitHub的分支保护，拒绝向master分支的直接push，要求所有修改都经过Pull Request流程。并且要求Pull Request合并时至少有另一名成员进行Review。设置如下图所示。

A screenshot of a cell phone

Description automatically generated

在版本号上，计划使用Tag功能管理版本，当功能稳定后创建分支stable，分支保持和最后一个Release的Tag相同，并用这个分支进行部署。tag名称参考格式“v4.0”。产生一个特定版本，以月份作为主版本号，月内次版本号递增，约一周一个次版本，并产生一个Release。在撰写本文的时候这一计划并没有实施。

在提交信息（commit message）上，计划约定使用一个统一的风格，其具体要求如下：

1. 代码仓库要求使用英文Message
2. 可以使用带有正文（不止一行的Message），主题和正文隔一行
3. 主题小于50字符
4. 主题首字母大写
5. 主题末尾不加句号
6. 主题用祈使句
7. 正文小于72词
8. 正文描述做了什么以及为什么

这一风格来源于网上的一些讨论，而上述规则来自<https://chris.beams.io/posts/git-commit/>。在撰写本文的时候这一计划并没有实施。

在持续集成上，使用GitHub的Actions，计划增加一个对Django项目的自动测试，以及对Django项目和调度器项目的自动部署。

在代码风格上，本项目参照Python的代码风格指导 PEP8（<https://www.python.org/dev/peps/pep-0008/>）代码提交前需要使用IDE对代码进行格式化。

## 3.3 管理责任分配与监督

本组在文档方面使用集中的管理方式，在集中编写文档时，会对每一个文档指派一个组员主持编写工作，并进行分工以及整合；在需要对文档进行修订时，会对每一个文档指派一个组员主持修订，在需要多个组员需要对文档进行修改的时候，该主持者会安排一个修订顺序，每一个需要修订的组员在上一个进行修订的组员完成修订后，获取上一个组员修订的版本，进行修订工作后增加版本后上传到GitHub，以此类推。本组并没有对文档设置明确的监督过程。

本组在代码方面使用Pull Request（PR）过程的Review作为监督过程，组员需要对其他组员的PR进行代码审查监督。对于其他代码管理配置上的问题，也指派了负责实验七的组员进行集中的配置和管理。

# 4. 变更控制和版本管理

本章节主要介绍代码的变更控制和文档的版本管理的统计和分析。

## 4.1 文档版本管理

在文档仓库上，共有177次的Commit记录（排除Merge），使用git log --pretty=format:"%h|%an|%ad|%s" --date=iso命令打印记录如下表所示。

|  |  |  |  |
| --- | --- | --- | --- |
| Hash | 作者 | 时间 | 提交消息 |
| aa8c4ef | 宋冰晨 | 2020-06-01 22:58:06 | 提交进度报告 |
| 2fe33cc | 宋冰晨 | 2020-05-29 23:37:58 | 提交MPP12.0 |
| 5561875 | 赵正阳 | 2020-05-29 20:28:48 | 会议记录更新至5.29 |
| 10afd5c | 赵正阳 | 2020-05-29 18:04:34 | 提交复评审意见反馈；工作量统计分析报告v1.4；测试需求复评审汇报 |
| cc06280 | 梁远志 | 2020-05-29 16:24:44 | Create 第14周\_梁远志.md |
| eff537b | 赵正阳 | 2020-05-29 12:13:48 | 提交赵正阳第14周工作日志 |
| 72db0dc | 郭浩隆 | 2020-05-29 12:08:20 | Create 第14周\_郭浩隆.md |
| 3397fa2 | 沈一聪 | 2020-05-29 11:14:41 | 提交沈一聪第14周工作日志 |
| d331508 | 沈一聪 | 2020-05-29 09:29:35 | update 性能测试报告 |
| 2d4b3ae | 梁远志 | 2020-05-28 23:41:57 | Add files via upload |
| 7c04117 | 梁远志 | 2020-05-28 23:41:30 | Delete H-EasySpider-被测软件执行结果分析报告v1.2.docx |
| 2b1cccb | 宋冰晨 | 2020-05-28 23:32:00 | 提交测试需求文档v1.3.6 |
| 2e74de9 | 赵正阳 | 2020-05-28 22:58:28 | 提交E组评审意见；测试需求规格说明书v1.3.5 |
| 194a3f4 | 梁远志 | 2020-05-28 19:51:14 | Delete H-EasySpider-测试需求规格说明书v1.3.3.docx |
| 929e148 | 梁远志 | 2020-05-28 19:50:47 | Add H-EasySpider-测试需求规格说明书v1.3.4.docx |
| 35cf503 | 郭浩隆 | 2020-05-28 19:25:48 | 测试需求规格说明书v1.3.3 |
| f2df4de | 宋冰晨 | 2020-05-28 19:12:12 | 提交测试需求文档v1.3.2 |
| ba1040c | 赵正阳 | 2020-05-28 18:06:56 | 提交测试需求规格说明书v1.3.1 |
| cce7dd0 | 赵正阳 | 2020-05-28 17:43:22 | 提交D组的评审意见、复评审意见反馈表 |
| 5995ca4 | 郭浩隆 | 2020-05-28 11:29:35 | Create H-B组软件问题汇总.docx |
| ef71aef | 郭浩隆 | 2020-05-28 00:01:03 | B、C组测试评审汇总 |
| 469525f | 梁远志 | 2020-05-24 14:43:50 | Delete H-EasySpider-软件问题报告v1.1.docx |
| c690da8 | 梁远志 | 2020-05-24 14:43:33 | Add files via upload |
| f4a031f | 赵正阳 | 2020-05-22 20:57:41 | 提交测试需求评审意见反馈；测试需求评审汇报ppt |
| 8f41e34 | 宋冰晨 | 2020-05-22 17:26:14 | 提交宋冰晨第13周工作日志；提交MPP11.0 |
| 46a3e3f | 宋冰晨 | 2020-05-22 13:24:12 | Merge branch 'master' of github.com:bhsei/20\_H update |
| 83d5a3f | 宋冰晨 | 2020-05-22 13:15:14 | update 可用性报告 |
| f91402d | 梁远志 | 2020-05-22 11:08:31 | Create 第13周\_梁远志.md |
| 1c1214a | 沈一聪 | 2020-05-22 10:54:45 | Merge branch 'master' of https://github.com/bhsei/20\_H |
| 5a4461d | 沈一聪 | 2020-05-22 10:54:15 | 沈一聪提交第13周工作总结 |
| 0c111c5 | 郭浩隆 | 2020-05-22 09:54:46 | 提交第13周工作日志 |
| 71f2f41 | 沈一聪 | 2020-05-22 09:22:58 | 合并测试需求规格说明书v1.2.2和v1.2.3 |
| 0cc0082 | 沈一聪 | 2020-05-22 01:49:51 | Merge branch 'master' of https://github.com/bhsei/20\_H |
| 792b18a | 沈一聪 | 2020-05-22 01:48:57 | 补充更改 H-EasySpider-测试需求规格说明书v1.2.2 |
| 75d728f | 宋冰晨 | 2020-05-22 00:20:19 | Merge branch 'master' of github.com:bhsei/20\_H update |
| 94711d1 | 宋冰晨 | 2020-05-22 00:18:22 | 更新测试需求文档1.2.3 |
| 6a5a846 | 赵正阳 | 2020-05-22 00:16:32 | 提交赵正阳第13周工作日志 |
| 6a3c215 | 宋冰晨 | 2020-05-22 00:04:15 | 更新部分测试文档 |
| 2980618 | 沈一聪 | 2020-05-21 23:47:10 | 提交H-EasySpider-性能测试报告v1.1.docx |
| 925fe5d | 沈一聪 | 2020-05-21 23:38:39 | merge |
| bff8877 | 沈一聪 | 2020-05-21 23:35:12 | 提交H-EasySpider-测试需求规格说明书v0.2.2.docx |
| 76f4e33 | 沈一聪 | 2020-05-21 23:35:12 | 提交H-EasySpider-测试需求规格说明书v1.2.2.docx |
| 14d3412 | 宋冰晨 | 2020-05-21 22:57:48 | update |
| 6994c11 | 宋冰晨 | 2020-05-21 22:52:00 | Merge branch 'master' of github.com:bhsei/20\_H merge |
| d0fea03 | 宋冰晨 | 2020-05-21 22:49:11 | update |
| a92865a | 郭浩隆 | 2020-05-21 20:51:17 | 提交软件测试计划v1.1 |
| 5f44700 | 赵正阳 | 2020-05-21 20:33:49 | 会议记录更新至5.21 |
| 29ee27d | 赵正阳 | 2020-05-21 19:27:13 | 提交测试需求规格说明书v1.2.1 |
| 3125ae3 | 梁远志 | 2020-05-21 17:12:21 | Delete H-EasySpider-软件问题报告v1.0.docx |
| c63b7c3 | 梁远志 | 2020-05-21 17:12:09 | Delete H-EasySpider-被测软件执行结果分析报告v1.0.docx |
| f7e50ac | 梁远志 | 2020-05-21 17:11:35 | Add files via upload |
| a80d11a | 梁远志 | 2020-05-21 17:07:51 | Add files via upload |
| 8470ea3 | 赵正阳 | 2020-05-21 15:37:25 | 提交B组和C组的评审意见、评审意见反馈表 |
| 8ffa473 | 赵正阳 | 2020-05-21 11:07:26 | 修改E组测试需求评审 |
| 370b0fe | 郭浩隆 | 2020-05-20 20:31:28 | 上传对E、D组评审结果 |
| 756046e | 梁远志 | 2020-05-20 11:25:22 | Add test result |
| b1bb3da | 宋冰晨 | 2020-05-19 20:24:09 | delete test requirements document 1.0 |
| 783fcf5 | 宋冰晨 | 2020-05-19 20:22:19 | 修改覆盖表 |
| ffed831 | 赵正阳 | 2020-05-19 14:09:27 | 修正测试需求规格说明书表标题 |
| 7a1addc | 赵正阳 | 2020-05-19 11:23:12 | 修改Scrapy单元测试编号 |
| 077d5f9 | 赵正阳 | 2020-05-18 23:35:07 | 提交测试需求规格说明书v1.1；工作量统计分析报告v1.3 |
| 7dfa129 | 郭浩隆 | 2020-05-18 21:52:12 | 上传测试需求表、软件问题表 |
| 0b5564f | 赵正阳 | 2020-05-18 12:49:46 | 添加项目访问地址 |
| 4b9b734 | 宋冰晨 | 2020-05-16 15:36:58 | 调整覆盖表格式 |
| edb185a | 宋冰晨 | 2020-05-16 14:27:58 | Merge branch 'master' of github.com:bhsei/20\_H merge |
| c55413e | 宋冰晨 | 2020-05-16 14:27:36 | 提交MPP10.0 |
| 9f207f1 | 赵正阳 | 2020-05-15 21:30:17 | 会议记录更新至5.15；更新测试需求分析汇报ppt |
| 3a6ac19 | 赵正阳 | 2020-05-15 15:42:01 | 提交测试需求分析汇报ppt；添加代码仓库链接 |
| 734eab0 | 梁远志 | 2020-05-15 13:49:43 | Create 第12周\_梁远志.md |
| 53ad3b5 | 沈一聪 | 2020-05-15 13:45:12 | Update README.md |
| a06f92e | 沈一聪 | 2020-05-15 13:43:59 | 沈一聪每周小结、教师意见提交 |
| 2bb597e | 沈一聪 | 2020-05-15 13:24:26 | 创建H-EasySpider-测试需求规格说明书v1.0.docx，测试用例和测试覆盖表文件夹 |
| 6468156 | 宋冰晨 | 2020-05-15 11:28:00 | 提交宋冰晨第12周工作日志 |
| bd88cfd | 赵正阳 | 2020-05-15 11:10:09 | 提交赵正阳第12周工作日志 |
| 8ddd833 | 郭浩隆 | 2020-05-15 09:12:58 | Create 第12周\_郭浩隆.md |
| 5b89c7a | 赵正阳 | 2020-05-13 21:41:36 | 会议记录更新至5.13 |
| 1a54a9c | 赵正阳 | 2020-05-12 00:02:15 | Merge branch 'master' of github.com:bhsei/20\_H |
| 981be66 | 赵正阳 | 2020-05-12 00:01:38 | 会议记录更新至5.11 |
| 6e4a2a6 | 宋冰晨 | 2020-05-08 23:35:48 | 提交MPP9.0 |
| c261a75 | 赵正阳 | 2020-05-08 21:58:58 | 修改第11周汇报ppt；会议记录更新至5.8 |
| d7987df | 郭浩隆 | 2020-05-08 13:16:18 | Create 第11周\_郭浩隆.md |
| 38184f0 | 赵正阳 | 2020-05-08 12:35:08 | 提交实现方案v1.0；工作量统计分析报告更新至v1.2；提交第11周汇报ppt |
| 6117302 | 赵正阳 | 2020-05-08 12:13:52 | 提交赵正阳第11周工作日志 |
| 8cff8d8 | 梁远志 | 2020-05-08 11:21:35 | Rename 第11周\_梁远志 to 第11周\_梁远志.md |
| b775533 | 沈一聪 | 2020-05-08 11:16:52 | 沈一聪提交第11周工作日志 |
| c7cbd42 | 梁远志 | 2020-05-08 10:48:47 | Create 第11周\_梁远志 |
| 659062 | 宋冰晨 | 2020-05-06 23:29:28 | 提交宋冰晨第11周工作日志 |
| bf67de9 | 赵正阳 | 2020-05-06 20:20:41 | Merge branch 'master' of github.com:bhsei/20\_H |
| cd9585d | 赵正阳 | 2020-05-06 20:19:49 | 会议记录更新至5.6 |
| 35162b6 | 宋冰晨 | 2020-05-03 10:26:15 | 提交MPP8.0 |
| 1fed2a2 | 沈一聪 | 2020-05-02 23:57:45 | 沈一聪第第十周工作日志 |
| 03216d2 | 赵正阳 | 2020-05-02 22:35:21 | 会议记录更新至5.2 |
| a4cb82b | 梁远志 | 2020-05-02 13:29:21 | Create 第10周\_梁远志.md |
| a29cde4 | 宋冰晨 | 2020-05-02 13:20:45 | 提交宋冰晨第10周工作日志 |
| ea7692c | 郭浩隆 | 2020-05-02 11:57:15 | Merge branch 'master' of https://github.com/bhsei/20\_H |
| c24368d | 郭浩隆 | 2020-05-02 11:55:51 | Create 第10周\_郭浩隆.md |
| 6cd0d17 | 赵正阳 | 2020-05-02 11:40:56 | 提交赵正阳第10周工作日志 |
| 1af0104 | 赵正阳 | 2020-04-29 23:09:12 | 会议记录更新至4.29 |
| 399e798 | 宋冰晨 | 2020-04-27 18:43:00 | 提交MPP7.0 |
| e396b4a | 赵正阳 | 2020-04-25 23:39:10 | 会议记录更新至4.25 |
| f09235b | 赵正阳 | 2020-04-24 18:28:09 | 修改第9周阶段性总结汇报ppt；工作量统计分析报告更新至v1.1 |
| aba3ead | 宋冰晨 | 2020-04-24 12:55:21 | 提交宋冰晨第9周工作日志 |
| 27b5b64 | 郭浩隆 | 2020-04-24 11:10:35 | 提交郭浩隆第九周工作日志 |
| fa57b35 | 沈一聪 | 2020-04-24 10:52:08 | Create 第9周\_沈一聪.md |
| 52ac05d | 梁远志 | 2020-04-24 10:30:01 | Create 第9周\_梁远志.md |
| 18a7515 | 赵正阳 | 2020-04-24 10:26:04 | 提交赵正阳第9周工作日志；README实验2标记完成 |
| 1a736e2 | 赵正阳 | 2020-04-22 21:34:50 | 会议记录更新至4.22 |
| 8957e81 | 宋冰晨 | 2020-04-17 23:53:14 | 提交MPP6.0 |
| ec3d061 | 赵正阳 | 2020-04-17 21:35:09 | 会议记录更新至4.17 |
| 6a9f15e | 赵正阳 | 2020-04-17 14:00:11 | 提交赵正阳第8周工作日志、需求复评审汇报ppt |
| a3d988a | 郭浩隆 | 2020-04-17 11:10:12 | 提交第八周工作日志 |
| 2ff7a92 | 沈一聪 | 2020-04-17 10:13:04 | 沈一聪提交第8周工作日志 |
| c5eae63 | 梁远志 | 2020-04-17 09:40:13 | Create 第8周\_梁远志.md |
| f7a5d0d | 宋冰晨 | 2020-04-16 20:01:00 | 提交需求复评审意见反馈文档 |
| ad67873 | 宋冰晨 | 2020-04-16 15:56:49 | 提交宋冰晨第8周工作日志 |
| 586c0e2 | 沈一聪 | 2020-04-16 15:45:40 | Update H-EasySpider-软件需求规格说明书v1.3.3.docx |
| a664d0a | 沈一聪 | 2020-04-16 15:44:16 | Update H-EasySpider-软件需求规格说明书v1.3.3.docx |
| 00efe49 | 宋冰晨 | 2020-04-16 15:34:34 | 提交需求分析文档v1.3.3 |
| cb438a9 | 梁远志 | 2020-04-16 13:46:08 | Delete H-EasySpider-软件需求规格说明书v1.3.1.docx |
| 00d4e1e | 梁远志 | 2020-04-16 13:45:29 | Add H-EasySpider-软件需求规格说明书v1.3.2.docx |
| 3a77aa4 | 梁远志 | 2020-04-16 13:36:25 | Update scrapyd.rucm |
| 7f91de1 | 沈一聪 | 2020-04-16 12:19:13 | 沈一聪更新需求规格说明书到v1.3.1，需求评审单中问题5需要讨论 |
| 5aafff8 | 沈一聪 | 2020-04-16 11:39:48 | Create 需求分析评审\_评审意见整理.docx |
| 7a012bc | 赵正阳 | 2020-04-16 10:39:44 | 提交F组和G组的复评审意见；需求规格说明书更新至v1.3.0；更新Scrapy用例图和RUCM图 |
| 97ce602 | 宋冰晨 | 2020-04-16 09:54:12 | 提交需求规格说明书v1.3.3 |
| 1b33542 | 赵正阳 | 2020-04-14 20:38:34 | 提交需求复评审意见汇总；会议记录更新至4.14 |
| 9e41d59 | 赵正阳 | 2020-04-12 20:59:18 | 提交需求分析初评审后需求规格说明书定稿v1.2.9 |
| 0d5702a | 沈一聪 | 2020-04-12 17:26:24 | Delete ~$EasySpider-软件需求规格说明书v1.2.8.docx |
| 26d156c | 沈一聪 | 2020-04-12 17:25:00 | 项目需求分析修改\_沈一聪\_4.12 |
| 25f567d | 梁远志 | 2020-04-10 23:07:31 | Delete H-EasySpider-软件需求规格说明书v1.2.6.docx |
| 016a823 | 梁远志 | 2020-04-10 23:06:55 | 修订软件需求规格说明书v1.2.7 |
| 46511e2 | 赵正阳 | 2020-04-10 22:46:30 | 提交需求评审汇报ppt；更新评审意见反馈；需求规格说明书更新至v1.2.6 |
| 12fe9f9 | 宋冰晨 | 2020-04-10 16:22:50 | 提交MPP文件5.0 |
| b922329 | 郭浩隆 | 2020-04-10 14:40:10 | Create 第7周\_郭浩隆.md |
| cb0b1f2 | 梁远志 | 2020-04-10 11:01:56 | Create 第7周\_梁远志.md |
| eb88b94 | 赵正阳 | 2020-04-10 09:41:57 | 提交赵正阳第7周工作日志 |
| 230e9a6 | 沈一聪 | 2020-04-10 09:34:10 | Rename 第七周\_沈一聪.md to 第7周\_沈一聪.md |
| e792083 | 沈一聪 | 2020-04-10 09:32:31 | Create 第七周\_沈一聪.md |
| 240ecbe | 宋冰晨 | 2020-04-09 21:37:27 | 提交宋冰晨第7周工作日志 |
| 0b35421 | 赵正阳 | 2020-04-09 16:08:19 | 调整需求规格说明书目录格式 |
| f343924 | 沈一聪 | 2020-04-09 15:54:12 | Delete H-EasySpider-软件需求规格说明书v1.2.4.docx |
| a668d71 | 沈一聪 | 2020-04-09 15:53:48 | H-EasySpider-软件需求规格说明书v1.2.4->v1.2.5.docx |
| 734c50a | 沈一聪 | 2020-04-09 15:40:46 | 完成H-EasySpider-评审意见反馈的初稿 |
| 0b23fdd | 沈一聪 | 2020-04-09 15:39:55 | Delete H-EasySpider-评审意见反馈 .docx |
| a1f1ba5 | 沈一聪 | 2020-04-09 15:39:15 | 完成H-EasySpider-评审意见反馈的初稿 |
| 5a0ca38 | 沈一聪 | 2020-04-09 15:30:56 | 修改了“管理服务器节点”用例中的actor |
| a7b25e9 | 梁远志 | 2020-04-09 14:24:11 | 修订需求规格 |
| d62053b | 梁远志 | 2020-04-09 14:04:38 | Update scrapyd.rucm |
| bb0aa0c | 梁远志 | 2020-04-09 13:52:46 | 添加评审意见反馈 |
| fa32fb4 | 郭浩隆 | 2020-04-09 12:53:32 | Create Web UI.rucm |
| 62fe6e4 | 宋冰晨 | 2020-04-09 12:24:14 | 更新需求规格说明书至v1.2.2和反馈文档 |
| 3bc5388 | 郭浩隆 | 2020-04-09 10:45:22 | Update H-EasySpider-评审意见反馈.docx |
| c8d015a | 赵正阳 | 2020-04-08 23:32:51 | 提交A组和I组的评审意见、评审意见反馈赵正阳部分、需求规格说明书v1.2.1；更新Scrapy用例图和RUCM图 |
| f1f70e8 | 赵正阳 | 2020-04-08 15:24:23 | 提交需求评审汇总；需求分析评审表v1.1；会议记录更新至4.7 |
| 69be9d0 | 赵正阳 | 2020-04-04 00:17:50 | 修改需求分析改进汇报ppt；README实验1标记完成 |
| ff5eb06 | 宋冰晨 | 2020-04-03 22:43:29 | 对制品文档重命名;提交MPP4.0 |
| 986cae4 | 沈一聪 | 2020-04-03 15:44:44 | Create 第6周\_沈一聪.md |
| 7c0363a | 梁远志 | 2020-04-03 11:55:06 | Create 第6周\_梁远志.md |
| 59ff775 | 郭浩隆 | 2020-04-03 09:33:41 | Create 第6周\_郭浩隆.md |
| d360e12 | 赵正阳 | 2020-04-02 16:44:52 | 会议记录更新至3.30；项目计划书v1.4.1；需求规格说明书v1.1；提交工作量统计分析报告v1.0、需求分析改进汇报ppt、赵正阳第6周工作日志 |
| 6826874 | 宋冰晨 | 2020-04-01 15:48:28 | 提交第6周工作日志 |
| 04c6d39 | 沈一聪 | 2020-03-30 17:29:20 | 需求分析评审表初稿 |
| 2784ccb | 沈一聪 | 2020-03-30 15:50:43 | week5需求规格说明书的评审意见 |
| 8a6d0ae | 宋冰晨 | 2020-03-28 17:45:30 | 提交MPP3.0 |
| bc4f1b1 | 赵正阳 | 2020-03-27 22:34:11 | 会议记录更新至3.27；修改需求分析汇报ppt；重命名需求规格说明书 |
| 9ee9e37 | 郭浩隆 | 2020-03-27 15:51:30 | Create 第5周\_郭浩隆.md |
| 61abef9 | 沈一聪 | 2020-03-27 15:40:00 | Create 第5周\_沈一聪.md |
| f6ac061 | 沈一聪 | 2020-03-27 15:39:12 | Delete 第五周\_沈一聪.md |
| 580c1d5 | 沈一聪 | 2020-03-27 15:38:46 | Create 第五周\_沈一聪.md |
| 9d70fdf | 宋冰晨 | 2020-03-27 11:44:43 | 提交宋冰晨第五周工作日志 |
| 8289aab | 梁远志 | 2020-03-27 11:08:46 | Add work log for week 5 |
| 959eed0 | 赵正阳 | 2020-03-27 00:00:42 | 提交赵正阳第5周工作日志；提交需求分析汇报ppt；更新readme；删除目录空格 |
| ba8f6d5 | 郭浩隆 | 2020-03-26 12:20:38 | Delete H-项目计划书v1.3.docx |
| 0476f5b | 郭浩隆 | 2020-03-26 12:20:34 | Create H-项目计划书v1.4.docx |
| 81296c3 | 郭浩隆 | 2020-03-26 12:20:30 | Create H-项目计划书v1.3.docx |
| 36af3ca | 沈一聪 | 2020-03-25 21:20:12 | Add files via upload |
| 7a71a07 | 赵正阳 | 2020-03-25 20:52:38 | 会议记录更新至3.23；修正工作日志的周数；提交实验1需求规格说明书初稿 |
| d9e4ab4 | 宋冰晨 | 2020-03-21 20:23:48 | 修正工作日志模板CSS显示问题、MPP更新至2.0、目录结构重建 |
| c309d10 | 沈一聪 | 2020-03-20 11:56:27 | Rename 第三周\_沈一聪 to 第3周\_沈一聪.md |
| 89926b3 | 沈一聪 | 2020-03-20 11:55:58 | Create 第三周\_沈一聪 |
| 40b93c3 | 赵正阳 | 2020-03-20 10:20:13 | 提交赵正阳第3周工作日志 |
| 063df99 | 梁远志 | 2020-03-20 09:44:03 | Create 第3周\_梁远志.md |
| 4a0d8bd | 郭浩隆 | 2020-03-20 09:08:21 | Create 第3周\_郭浩隆.md |
| 7ad161a | 赵正阳 | 2020-03-18 22:52:29 | 增加了原型爬虫程序；readme增加了实验内容 |
| b575298 | 赵正阳 | 2020-03-16 23:29:48 | 添加了项目计划书；会议记录更新至3.16 |
| 9e22bd4 | 宋冰晨 | 2020-03-16 17:21:19 | 提交工作日志模板、评审意见、MPP文件及版本管理日志 |
| 9357f29 | 宋冰晨 | 2020-03-14 15:04:11 | version 3 |
| 175ce7a | 赵正阳 | 2020-03-13 20:17:42 | Merge branch 'master' of github.com:bhsei/20\_H |
| ac810b8 | 赵正阳 | 2020-03-13 20:15:31 | 会议记录更新至3.13 |
| f31b9f8 | 沈一聪 | 2020-03-13 17:29:54 | -version 2.0 |
| 34c740d | 赵正阳 | 2020-03-10 21:51:00 | 添加了会议记录和汇报ppt |
| 804b4a2 | bhsei | 2020-03-10 14:44:32 | Initial commit |

本组在撰写本文时的文档仓库的目录树如下，其中荧光标记的是统计的制品文档。

├───1.会议记录

│ 会议记录.docx

│

├───2.评审意见

│ 实验准备汇报PPT\_评审意见整理.docx

│ 需求分析评审\_评审意见整理.docx

│ 需求规格说明书\_评审意见整理 .docx

│ 项目计划书\_评审意见整理 .docx

│ 项目设计实验部分\_评审意见整理.docx

│

├───3.项目PPT

│ 产品演示.mp4

│ 第11周\_开发进展汇报.pptx

│ 第12周\_测试需求分析汇报.pptx

│ 第13周\_测试需求评审汇报.pptx

│ 第14周\_测试需求复评审汇报.pptx

│ 第3周\_实验准备汇报.pptx

│ 第4周\_项目计划书汇报.pptx

│ 第5周\_需求分析汇报.pptx

│ 第6周\_需求分析改进汇报.pptx

│ 第7周\_需求评审汇报.pptx

│ 第8周\_需求复评审汇报.pptx

│ 第9周\_阶段性总结汇报.pptx

│

├───4.项目提交制品

│ ├───4.0实验准备

│ │ H-EasySpider-项目计划书v1.4.1.docx

│ │

│ ├───4.1软件需求分析

│ │ │ H-EasySpider-软件需求规格说明书v1.3.3.docx

│ │ │

│ │ └───图片

│ │ │ DFD.vsdx

│ │ │ ER.vsdx

│ │ │ 系统架构.png

│ │ │ 系统架构.vsdx

│ │ │

│ │ ├───RUCM

│ │ │ scrapyd.rucm

│ │ │ Scrapy框架.rucm

│ │ │ Web UI.rucm

│ │ │ 网站后端及调度程序服务器.rucm

│ │ │

│ │ └───用例图

│ │ scrapyd框架.mdj

│ │ Scrapy框架.mdj

│ │ webUi+后端总图.svg

│ │ 模板化爬虫程序管理平台.mdj

│ │

│ ├───4.2软件需求评审

│ │ │ H-EasySpider-需求分析评审表v1.1.docx

│ │ │ 需求分析评审标准.docx

│ │ │

│ │ ├───初评审

│ │ │ A评H\_评审检查单.docx

│ │ │ H-EasySpider-评审意见反馈.docx

│ │ │ H-F组需求分析评审汇总.docx

│ │ │ H-G组需求分析评审汇总.docx

│ │ │ I\_Node-RED\_200408\_I组对H组的评审结果.docx

│ │ │ 评审问题分配.docx

│ │ │

│ │ └───复评审

│ │ 20G对20H\_需求规格说明书评审表单.docx

│ │ F组对H组评审意见汇总.docx

│ │ H-A组需求分析评审汇总.docx

│ │ H-EasySpider-复评审意见反馈.docx

│ │ H-I组需求分析评审汇总.docx

│ │ 复评审问题分配.docx

│ │

│ ├───4.3软件设计与实现

│ │ │ H-EasySpider-实现方案v1.0.docx

│ │ │ README.md

│ │ │

│ │ └───图片

│ │ channelsHTTP.vsdx

│ │ djangoHTTP.vsdx

│ │ ER图.vsdx

│ │ MVT.vsdx

│ │ 任务操作.vsdx

│ │ 数据下载.vsdx

│ │ 模板浏览.vsdx

│ │ 状态图.vsdx

│ │ 系统组件图.vsdx

│ │ 调度器组件图.vsdx

│ │ 顺序图.vsdx

│ │

│ ├───4.4软件测试

│ │ │ H-EasySpider-性能测试报告v2.0.docx

│ │ │ H-EasySpider-测试需求规格说明书v1.3.6.docx

│ │ │ H-EasySpider-被测软件执行结果分析报告v1.3.docx

│ │ │ H-EasySpider-软件可用性报告v1.1.docx

│ │ │ H-EasySpider-软件测试计划v1.2.docx

│ │ │ H-EasySpider-软件问题报告v1.2.docx

│ │ │

│ │ ├───1-单元测试

│ │ │ ├───Django后端

│ │ │ │ │ Django单元测试报告.html

│ │ │ │ └───django\_coverage

│ │ │ │

│ │ │ ├───Scrapy

│ │ │ │ │ Scrapy单元测试报告1.html

│ │ │ │ │ Scrapy单元测试报告2.html

│ │ │ │ │ Scrapy单元测试报告3.html

│ │ │ │ │ Scrapy单元测试报告4.html

│ │ │ │ │ Scrapy单元测试报告5.html

│ │ │ │ └───scrapy\_coverage

│ │ │ │

│ │ │ └───调度器

│ │ │ │ 调度器测试报告.html

│ │ │ │

│ │ │ └───scheduler\_coverage

│ │ │

│ │ ├───2-集成测试

│ │ │ 集成测试覆盖表.xlsx

│ │ │

│ │ ├───3-场景测试

│ │ │ 场景测试覆盖表.xlsx

│ │ │

│ │ ├───4-功能测试

│ │ │ 测试覆盖表-梁远志.xlsx

│ │ │ 测试覆盖表-赵正阳.xlsx

│ │ │ 测试覆盖表\_宋冰晨.xlsx

│ │ │ 测试覆盖表\_沈一聪.xlsx

│ │ │ 测试覆盖表\_郭浩隆.xlsx

│ │ │

│ │ ├───5-非功能测试

│ │ │ 测试覆盖表\_非功能\_沈一聪.xlsx

│ │ │

│ │ └───归档

│ │ │ 功能性测试范围.xlsx

│ │ │ 缺陷跟踪流程.vsdx

│ │ │

│ │ └───测试用例

│ │ 功能性需求测试用例-宋冰晨.docx

│ │ 功能性需求测试用例-赵正阳.docx

│ │ 功能性需求测试用例\_梁远志.docx

│ │ 功能性需求测试用例\_郭浩隆.docx

│ │ 非功能需求测试用例及性能测试方案\_沈一聪.docx

│ │

│ ├───4.5软件测试评审

│ │ │ H-EasySpider-测试需求评审表v1.0.docx

│ │ │ H-EasySpider-软件问题表v1.0.docx

│ │ │ 测试需求规格说明书检查单.docx

│ │ │

│ │ ├───初评审

│ │ │ B对H组测试评审报告.docx

│ │ │ B对H组软件问题报告.docx

│ │ │ C\_汇总\_对H组的软件测试文档评审表单\_ver1.0.0.docx

│ │ │ C\_汇总\_对H组的软件问题报告\_ver1.0.0.docx

│ │ │ H-D组测试结果评审汇总.docx

│ │ │ H-D组测试需求评审汇总.docx

│ │ │ H-D组软件问题汇总.docx

│ │ │ H-EasySpider-评审意见反馈.docx

│ │ │ H-E组测试需求评审汇总.docx

│ │ │ H-E组软件问题汇总.docx

│ │ │

│ │ └───复评审

│ │ 20D对20H\_测试评审表格.docx

│ │ E组对H组评审单.docx

│ │ E组对H组软件问题汇总.docx

│ │ H-B组测试需求评审汇总.docx

│ │ H-B组软件问题汇总.docx

│ │ H-C组测试需求评审汇总.docx

│ │ H-C组软件问题15相关视频.mp4

│ │ H-C组软件问题汇总.docx

│ │ H-EasySpider-复评审意见反馈.docx

│ │

│ ├───4.6软件项目计划与监控

│ │ │ H-EasySpider-进度控制分析报告v1.0.docx

│ │ │

│ │ ├───工作日志

│ │ │ 工作日志模板.md

│ │ │ 第10周\_宋冰晨.md

│ │ │ 第10周\_梁远志.md

│ │ │ 第10周\_沈一聪.md

│ │ │ 第10周\_赵正阳.md

│ │ │ 第10周\_郭浩隆.md

│ │ │ 第11周\_宋冰晨.md

│ │ │ 第11周\_梁远志.md

│ │ │ 第11周\_沈一聪.md

│ │ │ 第11周\_赵正阳.md

│ │ │ 第11周\_郭浩隆.md

│ │ │ 第12周\_宋冰晨.md

│ │ │ 第12周\_梁远志.md

│ │ │ 第12周\_沈一聪.md

│ │ │ 第12周\_赵正阳.md

│ │ │ 第12周\_郭浩隆.md

│ │ │ 第13周\_宋冰晨.md

│ │ │ 第13周\_梁远志.md

│ │ │ 第13周\_沈一聪.md

│ │ │ 第13周\_赵正阳.md

│ │ │ 第13周\_郭浩隆.md

│ │ │ 第14周\_宋冰晨.md

│ │ │ 第14周\_梁远志.md

│ │ │ 第14周\_沈一聪.md

│ │ │ 第14周\_赵正阳.md

│ │ │ 第14周\_郭浩隆.md

│ │ │ 第4周\_宋冰晨.md

│ │ │ 第4周\_梁远志.md

│ │ │ 第4周\_沈一聪.md

│ │ │ 第4周\_赵正阳.md

│ │ │ 第4周\_郭浩隆.md

│ │ │ 第5周\_宋冰晨.md

│ │ │ 第5周\_梁远志.md

│ │ │ 第5周\_沈一聪.md

│ │ │ 第5周\_赵正阳.md

│ │ │ 第5周\_郭浩隆.md

│ │ │ 第6周\_宋冰晨.md

│ │ │ 第6周\_梁远志.md

│ │ │ 第6周\_沈一聪.md

│ │ │ 第6周\_赵正阳.md

│ │ │ 第6周\_郭浩隆.md

│ │ │ 第7周\_宋冰晨.md

│ │ │ 第7周\_梁远志.md

│ │ │ 第7周\_沈一聪.md

│ │ │ 第7周\_赵正阳.md

│ │ │ 第7周\_郭浩隆.md

│ │ │ 第8周\_宋冰晨.md

│ │ │ 第8周\_梁远志.md

│ │ │ 第8周\_沈一聪.md

│ │ │ 第8周\_赵正阳.md

│ │ │ 第8周\_郭浩隆.md

│ │ │ 第9周\_宋冰晨.md

│ │ │ 第9周\_梁远志.md

│ │ │ 第9周\_沈一聪.md

│ │ │ 第9周\_赵正阳.md

│ │ │ 第9周\_郭浩隆.md

│ │ │

│ │ └───项目进度计划（甘特图）

│ │ 版本管理日志.md

│ │ 项目进度计划V1.0.mpp

│ │ 项目进度计划V1.1.mpp

│ │ 项目进度计划v10.0.mpp

│ │ 项目进度计划v11.0.mpp

│ │ 项目进度计划v12.0.mpp

│ │ 项目进度计划V2.0.mpp

│ │ 项目进度计划V3.0.mpp

│ │ 项目进度计划v3.1.mpp

│ │ 项目进度计划v3.2.mpp

│ │ 项目进度计划v4.0.mpp

│ │ 项目进度计划v4.1.mpp

│ │ 项目进度计划v5.0.mpp

│ │ 项目进度计划v5.1.mpp

│ │ 项目进度计划v5.2.mpp

│ │ 项目进度计划v6.0.mpp

│ │ 项目进度计划v6.1.mpp

│ │ 项目进度计划v7.0.mpp

│ │ 项目进度计划v8.0.mpp

│ │ 项目进度计划v8.1.mpp

│ │ 项目进度计划v9.0.mpp

│ │ 项目进度计划v9.1.mpp

│ │

│ ├───4.7软件配置管理

│ │ README.md

│ │

│ ├───4.8软件工程实验追踪与分析

│ │ H-EasySpider-工作量统计分析报告v1.4.docx

│ │ H-EasySpider-贡献率统计.xlsx

│ │ 代码行数统计命令.txt

│ │

│ └───4.9实验总结

│ README.md

│

└───5.资料

README.md

本文共统计了12个文档的版本变更情况，如下表所示。包括每一个制品文档的版本数量以及各个版本的修订日期、修订者和修订内容。

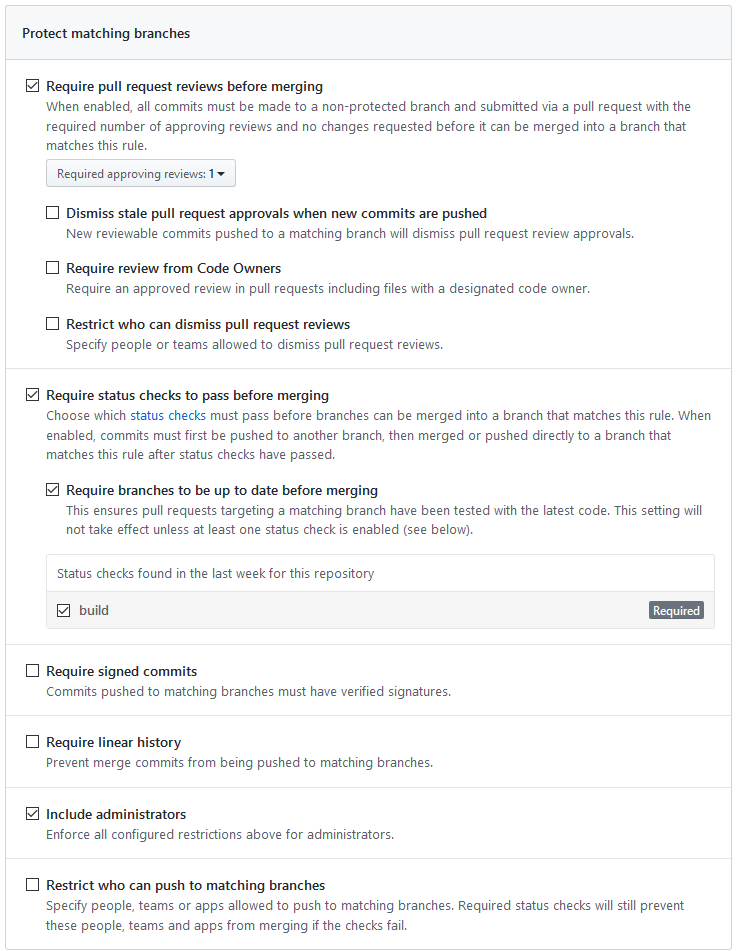
|  |  |
| --- | --- |
| 制品名称 | 版本数 |
| 软件项目计划 | 6   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/3/18 | 赵正阳 | 建立初始项目计划书文档 | | 1.1 | 2020/3/20 | 宋冰晨 | 建立版本变更表格、更正页码 | | 1.2 | 2020/3/20 | 赵正阳 | 更新组员分工 | | 1.3 | 2020/3/21 | 宋冰晨 | 修正甘特图 | | 1.4 | 2020/3/26 | 郭浩隆 | 添加项目目标、项目标准、修改Web UI计划、添加项目时间表 | | 1.4.1 | 2020/4/2 | 赵正阳 | 修改贡献率计算公式，增加主观评价 | |
| 软件需求规格说明书 | 15   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/3/25 | 赵正阳 | 完成引言，系统模型-Scrapy框架，Scrapy框架的需求分析、用例图和RUCM图，输入和输出，故障处理，运行环境，整合初稿 | | 郭浩隆 | 绘制Web UI的RUCM图 | | 沈一聪 | 绘制Web UI和后端的用例图、后端的RUCM图、非功能需求反爬虫 | | 梁远志 | 绘制EasySpider系统架构图，Scrapyd的需求分析、用例图和RUCM图，数据库特性 | | 宋冰晨 | 完成Web UI功能需求和非功能需求的文字描述 | | 1.1 | 2020/3/31 | 赵正阳 | 补充RUCM图文字描述、Scrapy框架非功能需求，整合及调整目录结构 | | 郭浩隆 | 补充RUCM图文字描述 | | 沈一聪 | 补充后端业务需求、RUCM图文字描述 | | 梁远志 | 补充RUCM图文字描述 | | 宋冰晨 | 补充项目背景、用户特点、假定和约束 | | 1.2.1 | 2020/4/8 | 赵正阳 | 补充了一些术语和缩略语；规范参考资料格式；修正标点；修正用例图和RUCM图；补充支持软件的版本号 | | 1.2.2 | 2020/4/9 | 宋冰晨 | 根据A组和I组的部分建议进行修改，另扩充了假定和约束，以及调整增加了Web UI部分的非功能需求 | | 1.2.3 | 2020/4/9 | 郭浩隆 | 修改了Web UI用例图，根据建议修改了部分RUCM图 | | 1.2.4 | 2020/4/9 | 梁远志 | 根据评审意见作出修改，包括：两处图注、一处冒号、一处语句表达、一处RUCM描述语句表达、删除了部分安全保密的描述 | | 1.2.5 | 2020/4/9 | 沈一聪 | 修改了若干部分错字；新增3.1.3参与者说明；修改3.2.3.4用例图的actor；更新目录 | | 1.2.6 | 2020/4/10 | 赵正阳 | 省略的RUCM图明确指出；对每个RUCM图添加文字引用；修改非功能需求反爬虫部分的表述；修改故障处理部分的表述 | | 1.2.7 | 2020/4/10 | 梁远志 | 修改数据库特性章节标题为数据存储，删除其中MySQL和MongoDB相关描述；将安全和保密章节移至非功能需求下（3.3.4）；删除需求分析不必要的服务器描述、支持软件描述、接口描述等 | | 1.2.8 | 2020/4/12 | 沈一聪 | EasySpider现在只存在Web UI和后端两部分，EasySpider的参与者说明中去除了scheduler了角色（对应scrapyd）；后端部分修改了用例图、添加和修改了RUCM用例，修改了部分文字说明；调整了第三章的部分图片编号；调整了全文的目录 | | 1.2.9 | 2020/4/12 | 赵正阳 | EasySpider系统架构删除“Scrapyd上部署的爬虫程序”；Scrapyd框架的功能需求移至Web UI之前 | | 1.3.0 | 2020/4/16 | 赵正阳 | 补充术语；增加对Scrapy框架本身的参与者的说明；删除“系统设置”用例；调整RUCM图的大小；修正其他字词标点问题 | | 1.3.1 | 2020/4/16 | 沈一聪 | 统一了各个框架和模型的actor名称（使用者、编程者）；修正3.2.3.2错字；删除3.3.3反爬虫中的出现的“解决方案”字样，修改标点，修改图片说明；  TODO：问题5待议，涉及多个方面问题 | | 1.3.2 | 2020/4/16 | 梁远志 | 添加了JSON API的解释及JSON的术语解释；修订了数据存储章节的标点符号；删除了不通顺的表达；将RUCM图中的用例名称改为中文；对数据存储章节的“状态”进行进一步说明 | | 1.3.3 | 2020/4/16 | 宋冰晨 | 根据A组和I组的部分建议对涉及章节进行修改（规范了文字描述）；合并3.1.1和3.1.2；合并3.3.2和3.3.3；将安全和保密章节移至3.7 | |
| 实现方案 | 1   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/5/7 | 赵正阳 | 完成概述，绘制系统组件图和ER图，整合初稿 | | 宋冰晨 | 完成Web UI的模板浏览、任务操作和数据下载模块的流程图 | | 郭浩隆 | 完成Web UI的登录注册和模板管理模块的流程图 | | 梁远志 | 绘制系统状态图，完成系统调度器部分 | | 沈一聪 | 完成爬虫模板部分 | |
| 测试计划 | 3   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/5/17 | 宋冰晨 | 建立初稿 | | 2020/5/18 | 梁远志 | 补充调度器的单元测试假设 | | 2020/5/18 | 沈一聪 | 补充除可用性测试之外，其他非功能性测试的假设和细节 | | 2020/5/19 | 赵正阳 | 补充功能性测试范围工作表 | | 1.1 | 2020/5/21 | 郭浩隆 | 补充图表标题、修正错误格式、补充部分图表说明 | | 1.2 | 2020/5/22 | 宋冰晨 | 更正非功能性测试部分内容 | |
| 测试规格说明书 | 11   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/5/14 | 赵正阳 |  | | 郭浩隆 |  | | 沈一聪 | 整合文档 | | 梁远志 |  | | 宋冰晨 |  | | 1.1 | 2020/5/18 | 赵正阳 | 增加Django后端单元测试，文档整合 | | 沈一聪 | 增加Scrapy单元测试和集成测试 | | 梁远志 | 增加调度器单元测试和集成测试 | | 宋冰晨 | 增加场景测试 | | 1.2.1 | 2020/5/21 | 赵正阳 | 场景测试“编辑后重新运行任务”补充任务状态为等待运行和正在运行的情况；修正其他文字问题 | | 1.2.2 | 2020/5/21 | 沈一聪 | 更新“测试相关文档”、“用户密码强度要求”用例明确了“预期结果”；修正“性能测试方案”出现两次“豆瓣电影”的描述并修改表格格式； 修正“负载/压力测试”的标点符号和字体 | | 1.2.3 | 2020/5/21 | 宋冰晨 | 补充可用性测试相关内容 | | 1.3.1 | 2020/5/28 | 赵正阳 | 根据D组评审意见修改 | | 1.3.2 | 2020/5/28 | 宋冰晨 | 根据D组评审意见完成修改，更正测试前提、补充测试数据、调整部分测试步骤的顺序 | | 1.3.3 | 2020/5/28 | 郭浩隆 | 根据D组评审意见完成修改 | | 1.3.4 | 2020/5/28 | 梁远志 | 根据D组评审意见完成修改 | | 1.3.5 | 2020/5/28 | 赵正阳 | 根据E组评审意见完成修改 | | 1.3.6 | 2020/5/28 | 宋冰晨 | 根据E组评审意见完成修改：修正场景测试定义相关措辞 | |
| 被测软件执行结果分析报告 | 3   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/5/19 | 梁远志 | 初版 | | 1.1 | 2020/05/21 | 梁远志 | 统一了目录序号字重；  将表格的文字说明移至在表格前。 | | 1.2 | 2020/5/21 | 宋冰晨 | 补充可用性测试结果章节 | | 1.3 | 2020/5/28 | 梁远志 | 参照E组评审意见修改  更新附录索引文档文件名 | |
| 软件可用性报告 | 3   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/5/18 | 宋冰晨 | 建立初稿 | | 1.1 | 2020/5/21 | 宋冰晨 | 根据c组的评审意见，补充表格题注，统一段落格式 | |
| 性能测试报告 | 3   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/5/19 | 沈一聪 | 初版 | | 1.1 | 2020/5/21 | 沈一聪 | 添加变更记录；修改表一样式和错误引用；统一Easyspider表述；调整图2缩进；修改表表头3.4字体 | | 2.0 | 2020/5/29 | 沈一聪 | 添加目录 | |
| 软件问题报告 | 3   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/5/19 | 梁远志 | 初版 | | 1.1 | 2020/5/21 | 梁远志 | 更新目录 | | 1.2 | 2020/05/24 | 梁远志 | 增加评审报告的问题 | |
| 进度控制分析报告 | 1 |
| 变更与管理分析报告 | 1 |
| 工作量统计分析报告 | 5   |  |  |  |  | | --- | --- | --- | --- | | 1.0 | 2020/3/30 | 赵正阳 | 完成初稿、统计实验准备和实验1的数据 | | 1.1 | 2020/4/24 | 赵正阳 | 完成实验2、实验6和实验8截至第8周的数据统计 | | 1.2 | 2020/5/8 | 赵正阳 | 完成实验3、实验6和实验8截至第10周的数据统计 | | 1.3 | 2020/5/18 | 赵正阳 | 完成实验3、4、6、8截至第12周的数据统计 | | 1.4 | 2020/5/29 | 赵正阳 | 完成实验4、5、6、8截至第13周的数据统计 | |

## 4.2 代码变更控制

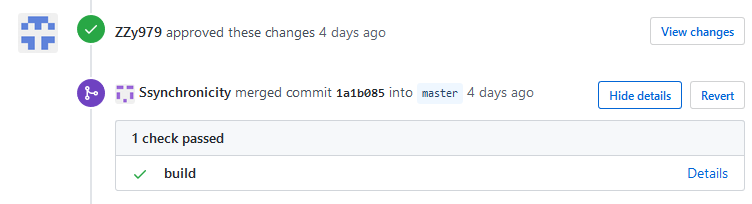
在代码变更控制上，本组使用的流程如下：

1. 开发前拉取主分支
2. 从主分支创建开发分支
3. 在开发分支上进行开发
4. 推送到GitHub
5. 创建一个Pull Request
6. 其他组员Review
7. Merge到主分支

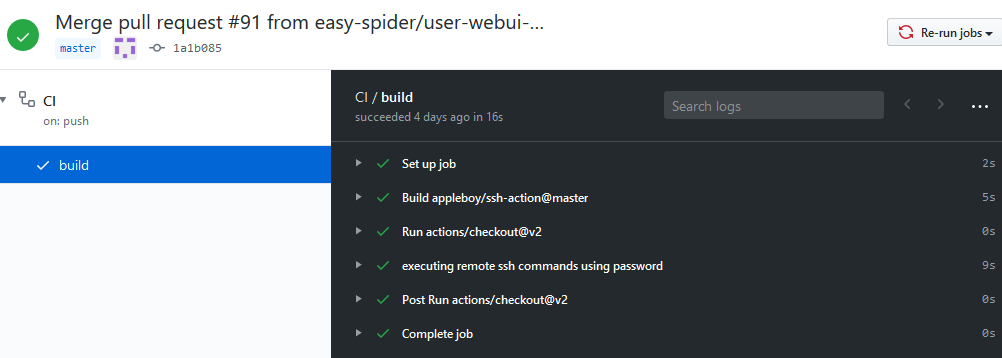
本组还要求在Pull Request合并前通过一个状态检查，如下图所示。本组在测试环节中设计了针对Django的单元测试，因而本组编写的一个GitHub Action在PR创建时运行这些Django测试。仅有在测试通过后同时要求有其他组员Review后才可合并。



下图展示了一个Pull Request的Review和自动测试。



在Pull Request合并到主分支后，会触发另外一个GitHub Action，如下图所示。本组编写的这一Action可以将新的主分支部署到生产服务器上，进而实现了自动部署。这一持续集成技术的应用使得项目的开发可以更加敏捷，同时也减少了服务器运维的时间开销。



本组的Django项目（easyspider-web）和调度器（easyspider-scheduler）项目使用了Pull Request机制和自动部署技术，Django项目使用了自动测试技术。

Django项目（easyspider-web）的变更情况如下表所示。统计的项目包括GitHub PR序号、PR作者、PR的审查人（reviewer）、状态（是否合并）、提交信息。共有82个PR，其中72个已合并，10个已关闭。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| PR | 作者 | 审查人 | 状态 | 提交信息 |
| 1 | 赵正阳 | 梁远志 | 已合并 | Add basic CRUD for user, template and task |
| 2 | 赵正阳 | 梁远志 | 已合并 | Change models Template and Job; Exclude settings.py |
| 3 | 宋冰晨 | 梁远志 | 已合并 | 提交普通用户WebUI |
| 4 | 郭浩隆 | 赵正阳 | 已关闭 | 提交部分Web UI页面 |
| 5 | 郭浩隆 | 赵正阳 | 已合并 | 提交部分WEB UI |
| 6 | 郭浩隆 | 宋冰晨 | 已合并 | 统一页面风格 |
| 7 | 梁远志 | 赵正阳 | 已合并 | Add JSON api for scheduler |
| 8 | 梁远志 | 赵正阳 | 已合并 | Add node manage page |
| 9 | 梁远志 | 赵正阳 | 已合并 | Add scrapyd-related operations |
| 10 | 郭浩隆 | 赵正阳 | 已合并 | 增加css和js，上传html至templates |
| 11 | 宋冰晨 | 赵正阳 | 已合并 | 创建userAPP |
| 12 | 宋冰晨 | 赵正阳 | 已合并 | User webui dev |
| 13 | 宋冰晨 | 赵正阳 | 已合并 | 创建template APP、增加AJAX |
| 14 | 郭浩隆 | 赵正阳 | 已合并 | 提交user相关的static，修改html |
| 15 | 赵正阳 | 宋冰晨 | 已合并 | Finish login/logout views for app user; Create models for app spiderTemplate; Delete useless models and views from app easyspider |
| 16 | 赵正阳 | 宋冰晨 | 已合并 | Create app pic |
| 17 | 宋冰晨 | 赵正阳 | 已合并 | 创建task APP |
| 18 | 梁远志 | 赵正阳 | 已合并 | Move scheduler API out |
| 19 | 赵正阳 | 宋冰晨 | 已合并 | Finish views for app spiderTemplate (except creating task) |
| 20 | 梁远志 | 赵正阳 | 已关闭 | Add action for auto test |
| 21 | 宋冰晨 | 赵正阳 | 已合并 | User webui dev |
| 22 | 赵正阳 | 宋冰晨 | 已合并 | Fix search and filter problems in sites\_view |
| 23 | 宋冰晨 | 赵正阳 | 已合并 | fix templateFirst.html |
| 24 | 宋冰晨 | 赵正阳 | 已合并 | 增加channels相关代码 |
| 25 | 赵正阳 | 梁远志 | 已合并 | Add tests for app spiderTemplate; Finish create\_task view and add tests |
| 26 | 梁远志 | 赵正阳 | 已合并 | Fix scheduler api and setup Django CI |
| 27 | 赵正阳 | 宋冰晨 | 已合并 | Finish task and data related views |
| 28 | 郭浩隆 | 赵正阳 | 已合并 | 修改个人信息页面 |
| 29 | 赵正阳 | 郭浩隆 | 已合并 | Finish views user\_profile and reset\_password |
| 30 | 郭浩隆 | 赵正阳 | 已合并 | 修改管理员部分代码 |
| 31 | 赵正阳 | 郭浩隆 | 已合并 | Finish views for creating template and node management |
| 32 | 宋冰晨 | 赵正阳 | 已关闭 | 完善task APP前端 |
| 33 | 宋冰晨 | 赵正阳 | 已关闭 | 完善task app前端 |
| 34 | 宋冰晨 | 赵正阳 | 已关闭 | User webui dev |
| 35 | 宋冰晨 | 赵正阳 | 已关闭 | User webui dev |
| 36 | 宋冰晨 | 赵正阳 | 已合并 | User webui dev |
| 37 | 宋冰晨 | 赵正阳 | 已合并 | User webui dev |
| 38 | 郭浩隆 | 赵正阳 | 已合并 | 修改管理员部分页面 |
| 39 | 宋冰晨 | 赵正阳 | 已合并 | User webui dev |
| 40 | 宋冰晨 | 赵正阳 | 已合并 | fix delete post problem |
| 41 | 赵正阳 | 宋冰晨 | 已合并 | Change task operation views to returning JsonResponse; Add view restart\_task; Send message to frontend when task status is updated; Change sorting rule of "hot sites" to view times |
| 42 | 梁远志 | 赵正阳 | 已合并 | Add delete and inspect node, fix urls |
| 43 | 宋冰晨 | 赵正阳 | 已合并 | 完成task app下html的修改 |
| 44 | 赵正阳 | 宋冰晨 | 已合并 | Bug fix: ZeroDivisionError in Task.progress() |
| 45 | 宋冰晨 | 赵正阳 | 已合并 | User webui dev |
| 46 | 赵正阳 | 宋冰晨 | 已合并 | Fix minor problems |
| 47 | 郭浩隆 | 宋冰晨 | 已合并 | fix index bug ,change templateUpload html and js |
| 48 | 赵正阳 | 郭浩隆 | 已合并 | Add 3 limit fields for Param; Finish views modify\_template and delete\_template |
| 49 | 赵正阳 | 宋冰晨 | 已合并 | Bug fix: get\_recent\_tasks return tasks of other users; Change: rename\_task and delete\_task returns top 5 recent tasks |
| 50 | 郭浩隆 | 赵正阳 | 已合并 | complate templateUpload templateSearchAdmin and templateModify |
| 51 | 郭浩隆 | 赵正阳 | 已合并 | fix bugs |
| 52 | 宋冰晨 | 赵正阳 | 已合并 | fix recent edit bug |
| 53 | 宋冰晨 | 赵正阳 | 已合并 | 将easyspider重命名为administrator |
| 54 | 郭浩隆 | 赵正阳 | 已合并 | bug fixed |
| 55 | 赵正阳 | 宋冰晨 | 已合并 | Change: rename\_task and restart\_task return create\_time |
| 56 | 赵正阳 | 宋冰晨 | 已合并 | Bug fix: create\_time is not correctly returned in rename\_task and restart\_task |
| 57 | 宋冰晨 | 赵正阳 | 已合并 | add loading animation and block other operations; add escape; fix edit time bug; fix websocket bug; |
| 58 | 宋冰晨 | 赵正阳 | 已合并 | fix dashboardBase bug |
| 59 | 赵正阳 | 宋冰晨 | 已合并 | Bug fix: finish\_time and create\_time is not converted into correct time zone |
| 60 | 赵正阳 | 宋冰晨 | 已合并 | Change: rename\_task no longer updates or returns create\_time of task |
| 61 | 宋冰晨 | 赵正阳 | 已合并 | remove edit time in rename task |
| 62 | 宋冰晨 | 赵正阳 | 已合并 | User webui dev |
| 63 | 郭浩隆 | 赵正阳 | 已合并 | change logo css |
| 64 | 郭浩隆 | 赵正阳 | 已合并 | fix bugs |
| 65 | 郭浩隆 | 赵正阳 | 已关闭 | Update templateUpload.html |
| 66 | 郭浩隆 | 宋冰晨 | 已合并 | fix bug |
| 67 | 郭浩隆 | 赵正阳 | 已合并 | Update templateModify.html |
| 68 | 郭浩隆 | 赵正阳 | 已合并 | fix bugs |
| 69 | 赵正阳 | 郭浩隆 | 已合并 | Bug fix: pictures are uploaded without ".jpg" in create\_template |
| 70 | 宋冰晨 | 赵正阳 | 已合并 | User webui dev |
| 72 | 宋冰晨 | 赵正阳 | 已合并 | fix datatable thead not aligned with tbody |
| 73 | 宋冰晨 | 梁远志 | 已合并 | User webui dev |
| 76 | 梁远志 | 赵正阳 | 已合并 | Fix display of node status and amount of finished job |
| 77 | 宋冰晨 | 赵正阳 | 已合并 | fix dataDownload.html overflow bug |
| 80 | 赵正阳 | 郭浩隆 | 已合并 | Change text of data preview page; Bug fix: invalid email is accepted in register and user\_profile views |
| 83 | 宋冰晨 | 赵正阳 | 已关闭 | fix pause bug |
| 84 | 宋冰晨 | 赵正阳 | 已合并 | fix pause bug |
| 86 | 赵正阳 | 宋冰晨 | 已合并 | Change response of restart\_task, change\_task\_status and clear\_data on disallowed operation to JSON |
| 87 | 宋冰晨 | 赵正阳 | 已关闭 | Fix editRestart BUG |
| 88 | 宋冰晨 | 赵正阳 | 已合并 | fix editRestart bug |
| 90 | 宋冰晨 | 赵正阳 | 已关闭 | delete deleteTask frontend check |
| 91 | 宋冰晨 | 赵正阳 | 已合并 | delete deleteTask status check |

调度器项目（easysipder-scheduler）的变更情况如下表所示。共有8个PR，均已合并。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| PR | 作者 | 审查人 | 状态 | 提交信息 |
| 1 | 梁远志 | 沈一聪 | 已合并 | Add basic node operations |
| 2 | 梁远志 | 沈一聪 | 已合并 | Add basic job operations |
| 3 | 梁远志 | 沈一聪 | 已合并 | Move database operations to Django |
| 4 | 梁远志 | 沈一聪 | 已合并 | Make scheduler ready to deploy |
| 5 | 梁远志 | 沈一聪 | 已合并 | Update following django API |
| 6 | 梁远志 | 赵正阳 | 已合并 | Update following django backend |
| 7 | 梁远志 | 沈一聪 | 已合并 | Tweak scheduler |
| 8 | 梁远志 | 沈一聪 | 已合并 | Add tests |

豆瓣网站爬虫项目（easyspider-site-douban）未使用PR，由沈一聪单独开发，共53个Commit（不含merge）。

|  |  |
| --- | --- |
| Hash | 提交信息 |
| 4781a6d | update egg and unitteest result |
| d209643 | fix spider bugs |
| 46dc599 | fix booksearch&musicsearch when field appears in content |
| 7071eee | update egg for master |
| 21a2791 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| db71236 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| b2562f0 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| c4ebdcd | remove track to htmlcov |
| c8e19fe | Update README.md |
| f9b9648 | support unittest,README,requirement.txt |
| cb2d4f3 | update egg |
| 6d14a10 | fix run.py;update testbench.py and support logs now;update proxypool support;optimize settings;update README |
| 6574789 | revise run.py;add testbench;fix miss 1 item each time run a spider |
| b92ecfc | complete merge |
| 2f0050f | apply proxy pool for all spiders;update log info;update README;fix example pics name |
| e409c9a | apply proxy pool for all spiders;update log info;update README;fix example pics name |
| 7b5b915 | finish proxypool for moviesearch spider;upgrade dealing with anti crawl stratagies |
| c86981b | support ip proxypoll(not stable) |
| 9dd67d8 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| 6fc5e94 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| 9bdf10b | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| 66b5c64 | update example for douban\_music, douban\_book |
| 40337d4 | update example for douban\_music, douban\_book |
| 9a03057 | update README |
| 22b8ddc | rebuild for deployment |
| a70f420 | support douban\_booksearch spider;update readme, TODO;optimize spider code |
| 9646b28 | globel replace novelsearchspider->booksearchspider |
| 9b529b2 | add track to release folder(.egg) |
| 7a44967 | support douban\_musicsearch spider |
| 1fd5e6f | update example |
| d3f5f05 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| 549e841 | update example |
| 3158d8d | Update README.md |
| abb6e22 | Update README.md |
| 0ef4495 | add examples for web page display;update README;update TODO |
| 5395be1 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| 37bfee5 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| 2214587 | Merge branch 'master' of https://github.com/easy-spider/easyspider-site-douban |
| 43b76e3 | support chrome headless;fix moviesearch description;remove some selenium settings from custom\_settings.py;update extract useragent to useragent.py for reuse;update README.md;add new spiders(not working);update TODO; |
| 7aa9afb | support chrome headless;fix moviesearch description;remove some selenium settings from custom\_settings.py;update extract useragent to useragent.py for reuse;update README.md;add new spiders(not working);update TODO; |
| 63cf12b | update README |
| f81ea65 | update README |
| f4dc1f2 | fmt;apply custom\_settings;update todo;update build.sh |
| 2f2ac34 | merge |
| 3612ce1 | accomplish spider 'moviesearch';accomplish SeleniumMiddleware;;update some setting and extract some part to customesetting;update README.md |
| 8d2bf66 | revise some term in settngs.py ot better work around anti-crawl |
| 930181f | Update README.md |
| 700c80e | finish crawl an arbitary movie page |
| 0e8c8e1 | finish crawl a certain movie search result;add the item for moviesearch spider;add log requirement with todo and update todo info |
| ad9e913 | Add user\_agent middleware |
| f86cfcd | Make README complete;update build.sh |
| dcbb598 | update todo & add custom\_settings |
| 561bd29 | add build.sh |
| 0434a8e | finish design mongodb interface and now can apply argus to pipeline;update run.py |
| 084e488 | add new spider moviesearch and finish convey argus;add run.py |
| 954584a | add todo;delete run.py;add build.py for deployment |
| 93583a1 | save data to local mongodb; reformat |
| 66af2c0 | Merge branch 'master' of github.com:icessspark/easyspider\_DouBan |
| a06ad58 | Initial commit |
| de68007 | crwal douban top 250 movies info and print |

爬虫模板项目（easyspider-template）未使用PR，由沈一聪单独开发，共4个Commit（不含merge）。

|  |  |
| --- | --- |
| Hash | 提交信息 |
| 4f94727 | Merge branch 'master' of github.com:easy-spider/easyspider-template |
| a7d1773 | update example dir;fix some spell |
| e2da684 | update example dir;fix some spell |
| 97ce0d2 | Finish first version of tempalte |
| 3a22049 | Merge branch 'master' of github.com:easy-spider/easyspider-template |
| 384d68d | init template repo |
| e68c104 | Initial commit |

## 4.3 分析与总结

本组主要按每个组员的Git的Commit数量等统计指标进行变更和管理分析。使用命令git shortlog -s -n --all --no-merges计算出各个仓库中每个人的Commit数量（不含merge）。

其中文档仓库的Commit数量由低到高排列如下表所示。组长赵正阳有明显的领先。

|  |  |
| --- | --- |
| 组员 | Commit数 |
| 赵正阳 | 52 |
| 沈一聪 | 39 |
| 宋冰晨 | 33 |
| 梁远志 | 31 |
| 郭浩隆 | 22 |

在文档仓库的提交信息中，因为没有事先的管理约定，因此略显杂乱。但大部分的提交信息中包括修改的文件名。值得注意的是组长赵正阳的提交信息会更为详细。就在具体提交的粒度和提交信息的风格上，本组还需要反思和改进。

在文档内的版本变更记录中，组员都能较详细的记录变更修改的内容。但是除了第一个制品文档外，其他文档因为本组的文档修订方针的变化，不再明确设立审查者对变更进行审查，而是使用一种自由组织的方式进行审查，这一个也可以是本组能够改进的地方。

在代码仓库上，本组的主要仓库使用了Pull Request且有多个组员进行开发。Pull Request机制有效的减少了代码合并的冲突问题，同时因为使用了自动测试技术，能够一定程度保证本组的软件制品的质量。

对于Django项目（easyspider-web）的各个指标如下表所示。宋冰晨提出了最多的PR，组长赵正阳负责了多数的PR的审查。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 组员 | Commit数 | PR数 | Review数 | Python行 | Html行 | CSS行 | JS行 |
| 赵正阳 | 32 | 22 | 45 | 2208 | 54 | 0 | 0 |
| 梁远志 | 20 | 7 | 5 | 231 | 311 | 0 | 0 |
| 宋冰晨 | 41 | 27 | 17 | 264 | 2505 | 350 | 237 |
| 郭浩隆 | 25 | 16 | 5 | 9 | 1268 | 768 | 405 |

对于调度器项目（easyspider-scheduler）的各个指标如下表所示。梁远志提出了所有的PR，沈一聪负责了多数的PR的审查。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 组员 | Commit数 | PR数 | Review数 | Python行 |
| 赵正阳 | 0 | 0 | 1 | 0 |
| 梁远志 | 11 | 8 | 0 | 508 |
| 沈一聪 | 0 | 0 | 7 | 0 |

对于豆瓣网站爬虫项目（easyspider-site-douban）的Commit数量如下表所示。沈一聪完成了全部的开发工作。

|  |  |  |
| --- | --- | --- |
| 组员 | Commit数 | Python行 |
| 沈一聪 | 53 | 1837 |

对于豆瓣网站爬虫项目（easyspider-template）的Commit数量如下表所示。沈一聪完成了全部的开发工作。

|  |  |  |
| --- | --- | --- |
| 组员 | Commit数 | Python行 |
| 沈一聪 | 4 | 443 |

这些统计结果也间接反映了本组的代码开发工作的分工。受制于分工安排，并没有能够在所有的项目中使用Pull Request，这也是本组能够改进的地方。

总体来说，本组有效地利用了版本管理工具Git，不论是在文档管理还是在代码管理上，工具的使用带来了很大的好处。此外GitHub的持续集成工具的使用也同样保证了本项目的开发工作能够顺畅高效地进行。过程中也发现了许多能够改进的地方，这也对本组成员日后进行开发工作提供了有益的启示。