会议记录

|  |  |
| --- | --- |
| 时间 | 记录 |
| 2020.3.6 | 1.首次开组会，组员自我介绍，确定组长为赵正阳  2.根据组员的开发经验和使用的编程语言，经投票决定选择的开源项目为Scrapy |
| 2020.3.9 | 1.讨论汇报ppt的内容  2.初步确定组员分工：赵正阳负责做ppt，沈一聪负责汇报，其他人做相关调研工作 |
| 2020.3.13 | 1.确定项目计划书分工：赵正阳——成员介绍、项目简介；沈一聪、梁远志、郭浩隆——工作内容；宋冰晨——整体计划甘特图和贡献率计算公式  2.从下周开始每周五所有组员将自己的本周工作日志提交到GitHub  3.进一步确定组员分工：宋冰晨、赵正阳——Web UI及前后端交互；梁远志、郭浩隆——分布式；沈一聪——反爬虫 |
| 2020.3.16 | 1.讨论项目计划书和项目大体方向：  （1）最终完成一个针对特定网站的爬虫程序；  （2）根据具体网站决定要实现的反爬扩展功能，例如登录、动态JavaScript加载等；  （3）使用Scrapyd实现分布式功能，具体的用户参与程度（例如是否需要用户上传控制脚本）留待后续商定；  （4）使用ScrapydWeb等可视化工具实现监控功能（例如运行时间、爬取数据量等）  2.确定周五下午1点之前上传本周工作日志 |
| 2020.3.20 | 1.讨论项目具体功能：  （1）用户选择爬虫模板（爬取特定网站的特定内容）  （2）根据模板自动生成爬虫脚本  （3）通过Web界面进行爬虫监控和管理  得出项目扩展功能用例图  2.统一绘图软件：用例图使用StarUML，RUCM图使用Eclipse插件  3.确定需求规格说明书初稿分工：  赵正阳：绘制Scrapy框架本身的用例图和RUCM图  宋冰晨、郭浩隆：根据讨论形成的项目功能的用例图绘制RUCM图  沈一聪、梁远志：绘制分布式功能的用例图和RUCM图 |
| 2020.3.23 | 本周汇报内容的分工：赵正阳——Scrapy，宋冰晨——Web UI（以豆瓣热门电影模板为例讲解系统使用流程），沈一聪——后端系统和反爬虫，梁远志——分布式，郭浩隆——项目计划书改进 |
| 2020.3.27 | 1.需求规格说明书改进：  （1）补充假定和约束——沈一聪  （2）补充用例图和RUCM图的文字描述——每个人负责的部分  （3）增加Scrapy框架和后端的业务需求——每个人负责的部分  （4）补充Scrapy框架的非功能需求——赵正阳  （5）调整目录结构——赵正阳  2.文件命名规范：H-{文件名}v{版本号}  3.需求评审表格设计——沈一聪  4.实验8统计表格设计——赵正阳  5.贡献率计算公式改进，增加主观评价部分——宋冰晨 |
| 2020.3.30 | 1.确定编码部分大致分工：  Web UI（HTML页面和JavaScript）——郭浩隆、宋冰晨  网站后端——赵正阳、沈一聪  调度程序(Scrapyd)——梁远志、沈一聪  2.确定下次汇报ppt分工：  （1）需求规格说明书改进——赵正阳（评价意见-改进对照）  （2）需求评审表单——沈一聪  （3）实验6——宋冰晨  （4）实验7——梁远志  （5）实验8——赵正阳（统计表格数据，计算贡献率）  3.贡献率计算公式增加主观评价部分：  个人贡献率=(个人工作量/全组工作量)\*w1+(评价得分/∑评价得分)\*w2 |
| 2020.4.7 | 1.讨论每个人提出的F组和G组评审问题及汇总  2.确定评审意见反馈及汇报的分工：每个人负责自己撰写部分对应问题的修改和ppt |
| 2020.4.14 | 1.讨论每个人提出的A组合I组评审问题及汇总  2.确定复评审意见反馈分工：每个人负责自己撰写部分对应问题的修改，未完全接受或有争议的问题和评审组对应的报告人交流 |
| 2020.4.17 | 1.对被评审组的检查：每个人检查自己提出的问题  2.实验2的工作量由每个人提出和修改的意见个数按工时的计算公式统计，实验6的工作量由宋冰晨每周工作记录中维护MPP文件的工时统计，实验7的工作量暂不统计，实验8的工作量由工作量统计分析报告的字数统计  3.初步开发工作：宋冰晨、郭浩隆——尝试前端页面的开发，选择合适的前端框架；沈一聪——尝试根据爬虫模板写科技网的爬虫；梁远志——调度器部分的初步开发；赵正阳——网站后端的初步开发和数据库的创建  4.与GitHub推送和Pull Request等方面有关的问题和梁远志确认  5.下周三晚上讨论组间评分和汇报的问题 |
| 2020.4.22 | 1.确定评审的4个组的文档质量和反馈质量排序方法  2.讨论设计和开发方面的细节 |
| 2020.4.25 | 1.讨论一些设计方面的细节问题：前端界面改为使用Django模板、任务的“准备运行”和“取消”状态定义、新增模板功能使用自定义的界面、模板详情中的图片保存在服务器本地而不是Django项目的静态文件  2.下周开发工作目标：宋冰晨、郭浩隆——将前端HTML页面修改为Django模板；赵正阳——根据讨论结果修改网站后端接口；沈一聪——完善豆瓣爬虫模板；梁远志——给调度器增加上传egg文件功能  3.开发过程中遇到需要协调的问题及时沟通，暂定下周四（4月30日）再次开组会讨论开发进展 |
| 2020.4.29 | 参加组员：赵正阳、郭浩隆、宋冰晨  讨论Web UI设计问题：  （1）补充模板相关的实体类字段  （2）修改现有的用户登录注册逻辑，改用Django提供的功能  （3）Django项目按功能模块拆分为不同的app，由相关负责人分别开发 |