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 |