

SpY计划文档

迭代一

阶段目标

- 1. 完成爬虫收集特定新闻网站的新闻数据
- 2. 能够将爬虫收集到的新闻数据持久化到数据库
- 3. 可对收集到的新闻数据进行界面展示
- 4. 可对收集到新闻数据进行关键字搜索
- 5. 可对收集到新闻数据进行增删改查

日程表

活动	参与人	计划完成时间
需求规格文档	王柏霏	2月28日
体系结构设计文档	尹天禹 赵培炎	2月28日
仓库建设，前端初始化	袁晨	2月28日
确定数据库表结构	杜周纬	2月29日
完成新浪爬虫	杜周纬	3月5日
完成前后端接口文档	尹天禹	3月2日
后端初始化构建	赵培炎	2月29日至3月2日
后端接口实现	尹天禹 赵培炎	3月3日至3月8日
前端 - 首页 详情页	袁晨	2月29日至3月4日
前端 - 登录，搜索，筛选	张瑞林	3月5日至3月12日
前端 - 后台	王柏霏	3月5日至3月12日

里程碑和里程碑标志日期

--	--

里程碑	日期
完成新浪爬虫	3月5日
完成后端实现	3月8日
完成前端实现	3月12日

计划执行检查表

检查内容	时间
需求规格文档，体系结构设计文档	2月29日
仓库建设，前端初始化，后端初始化	3月1日
完成前后端接口文档	3月2日
确定数据库表结构，完成新浪爬虫	3月5日
完成后端实现	3月8日
完成前端实现	3月12日

迭代二

阶段目标

- 1. 新闻分类
- 2. 数据统计
- 3. 变量统计
- 4. 知识图谱生成与可视化

日程表

活动	参与人	计划完成时间
需求文档	袁晨 张瑞林	3月22日
体系结构设计文档	尹天禹 赵培炎	3月22日
讨论前端页面及其大致实现思路	袁晨 张瑞林 王柏霏	3月22日
完成数据库设计	赵培炎 杜周纬	3月24日

完成实时爬虫添加到数据库	杜周纬	3月24日
完成CI/CD流水线构建	尹天禹 王柏霏	3月27日
完成接口文档	尹天禹 张瑞林 赵培炎 杜周纬	3月27日
完成neo4j数据库与实体关系提取	张瑞林 杜周纬	3月27日
完成云部署与大语言模型接入	尹天禹 杜周纬	3月27日
完成知识图谱设计	袁晨 杜周纬 张瑞林 王柏霏	3月31日
完成Elasticsearch模糊搜索	赵培炎	3月31日
后端接口实现	尹天禹 赵培炎 杜周纬 张瑞林	4月11日
前端实现	袁晨 王柏霏	4月17日

里程碑和里程碑标志日期

里程碑	日期
完成CI/CD流水线构建	3月27日
完成neo4j与知识图谱	3月31日
完成Elasticsearch模糊搜索	3月31日
完成后端实现	4月11日
完成前端实现	4月17日

计划执行检查表

检查内容	时间
需求规格文档，体系结构设计文档	3月22日
完成数据库设计，并实时爬虫添加到数据库	3月24日
完成CI/CD流水线构建	3月27日
完成知识图谱及可视化(neo4j、ECharts、大语言模型)	3月31日
完成后端实现	4月11日

迭代三

阶段目标

- 1. 自动化情报收集
- 2. 情报可信性分析
- 3. 提高迭代2中知识图谱生成的准确率

日程表

活动	参与人	计划完成时间
需求文档	袁晨	5月16日
计划文档	袁晨	5月16日
体系结构设计文档	尹天禹 赵培炎	5月16日
可信性分析调研	袁晨	5月18日
前端CICD完善	王柏霏	5月20日
fink环境搭建	尹天禹	5月20日
自动化爬虫与推送	杜周纬	5月20日
数据备份存储	赵培炎	5月20日
可信性分析	张瑞林	5月20日
知识图谱改善	杜周纬	5月24日
后端接口实现	尹天禹 赵培炎 杜周纬 张瑞林	5月28日
前端实现	袁晨 王柏霏	5月28日
前后端联调	尹天禹 赵培炎 杜周纬 张瑞林 袁晨 王柏霏	6月2日

里程碑和里程碑标志日期

里程碑	日期

完成可信性分析	5月20日
完成自动化情报收集	5月20日
完成知识图谱改善	5月24日
完成前后端联调	6月2日

计划执行检查表

检查内容	时间
需求规格文档，体系结构设计文档	5月16日
完成可信性分析	5月20日
完成自动化情报收集	5月20日
完成知识图谱改善	5月24日
完成前后端联调	6月2日