

# 个人总结报告

姓名：沈韵风

学号：3200104392

## 1 组内分工

### 编档工作

- 负责了 README 文档、接口文档和其他文档部分内容的撰写工作。
- 绘制了系统模块架构图、网络拓扑图等多个文档插图。

### 编码工作

负责使用 Flask 框架完成后端项目的编码工作，产出了：

- 合并且格式规范化的初始数据文件（.csv 格式）
- 自动化建表并导入初始数据的脚本文件
- 详细的接口文档
- 功能完备的后端仓库

### 展示工作

负责了结题报告的展示工作。

## 2 心得体会

### 2.1 创新

#### 项目创新

在老师介绍完“中国工程科技知识中心”这一项目后，我便开始了思考：凭借我们所掌握的知识，有什么是可以实现的？由于同组同学均没有图像处理与机器学习方面的基础，而大多数人都有数据可视化以及图形化用户界面方面的编程经验。因此，我们将讨论的目标聚焦于“数据可视化”方面。

但单纯的“可视化”是缺乏市场竞争力的，而目前市面上也存在很多实现完备的开源解决方案。通过对各个产业的相关情况进行查询，我们发现：作为我国的主要产业之一的“渔业”缺乏统一的数据统计与可视化平台，其数据来源分散，并且绝大多数数据表存在大量缺项的情况。因此，“渔业数据可视化”平台是存在市场空缺的。此外，渔业数据的地域性分布和随时间推移而变化的特征也与可视化方向相契合，具有可实现性。

## 深挖

### 1. 数据源整合与导入导出

进行数据可视化的前提是拥有准确的源数据，且我国渔业相关的统计数据分布零散。因此，如何整合的出一份完备的渔业统计数据库成为了我首先考虑的问题。在负责编写网页爬虫的同学完成数据爬取工具后，我使用了 MySQL 数据库将来自国家统计局与各地统计局的相关数据进行了合并，并保存为 `.csv` 格式，便于进行项目初始数据导入，也便于用户以 EXCEL 表格的形式查看。

此外，用户通过后台系统进行的修改将被实时同步到数据库中，不同的统计项将单独成表，用户可以边界的进行导入导出工作。

### 2. 可视化及成果导出

通过调研相关的竞品，我们发现：现有的数据可视化平台往往将数据选取与可视化呈现结果分置于两个标签页下，且转化过程中需要多次手动确认，操作较为繁琐。因此，我们采用了独特的左右分栏设计，使得用户在筛选表格的同时可以查看相关的可视化成果。这一操作固然会增加一定的服务器处理压力，但由于渔业数据表的单一筛选结果数量较少（每张表在1000条左右），这样的代价是可以接受的，且能给用户带来更好的使用体验。

此外，为了方便用户复用分析成果，我们也为用户提供了将可视化结果导出为图片的功能。

### 3. 数据预测

由于本平台的目的是方标渔业从业/研究群体对数据进行分析，进而做出决策。因此，我们为平台增加了基于历史数据对次年数据进行预测的功能。此外，随着对话模型的兴起，我们可以在后续版本中接入 `ChatGPT` 等对话机器人，为用户决策提供参考性建议。

## 2.2 组内协作

### 组织创新

相较于之前课程的组队模式，本次组队的方式增加了团队成员间相互了解的成本，但也同时模仿了实际工作环境中的不确定性。另外，5人的小组规模也使得先前课程中极小规模团队（2-3人）的管理模式受到挑战。

为此，我们采用了自荐的方式选举产生了小组长，负责日常组织小组例会、检查项目总体推荐程度。而每一次的编档 / 编码任务都会由不同的负责人进行汇总，使得每一个人都有机会接触团队工作的全貌，同时减轻组长压力。

此外，本次的团队模式是完全扁平化的。组长和负责人除组织会议、汇总相应材料外并无特权，都需要全程参与编码、编档工作，提升了团队工作效率与协作的紧密程度。

### 团队协作

#### 1. 版本管理

本项目的采用 git 进行版本管理，前后端项目分别托管在 GitHub 和 Gitee 两个平台上（在最后进行了汇总），前后端的实现细节相互隐藏，细节通过接口文档进行对接。

项目文档使用 markdown 进行编写，这一纯文本文件格式使得我们可以便捷的使用 diff 命令查看文档的版本差异（也可以通过 GitHub 中的相关服务进行直观对比）。相较于汇总 Word 文档显著减少了格式调整的时间成本。

#### 2. 文档

受大环境影响，本次的编码工作难以在线下统一完成——这使得前后端同学缺少面对面直观进行交流的问题。为减少因信息和理解偏差导致的对接错误，我撰写了较为完善的接口文档 - 按照不同的功能模块进行划分、编号，并注明了路由、请求方式、请求格式与返回值示例，提高了工作成果。

这些文档均以协作文档的形式托管在语雀上，使得修改可以即时同步，也使得开发进度与成果对全体组员透明，有效激发了开发的积极性。

### 3. 成本管理

基于课上所学的知识，我们对成本进行了预算与实际统计。结果表明：在使用甘特图统计实际开发进度前，组员的提交往往集中在小组 ddl 前一天左右的时间段内，而使用共享文档的甘特图对开发耗时进行统计后，提交的分布变得均匀，平均耗时也较为明显的缩短了。从总不难看出：使用合理、科学的方法对团队进行管理有助于提高团队效率、降低开发成本，这也提醒我们 - “软件工程”绝不仅仅是 code work，还综合了科学的管理方法与创新思维。