目录

[第一章 方案概要 2](#_Toc6159562)

[1.1目标问题 2](#_Toc6159563)

[1.2总体思路 2](#_Toc6159564)

[1.3具体做法 2](#_Toc6159565)

[1.4创新亮点 2](#_Toc6159566)

[第二章 项目目标与服务模型 2](#_Toc6159567)

[2.1项目目标 2](#_Toc6159568)

[2.2服务模型 2](#_Toc6159569)

[2.3项目价值及创新点 2](#_Toc6159570)

[2.4项目创新点 2](#_Toc6159571)

[第三章 组织管理与业务分析 2](#_Toc6159572)

[3.1组织管理 2](#_Toc6159573)

[3.2业务分析 2](#_Toc6159574)

[第四章 技术路线级实现方法 2](#_Toc6159575)

[4.1技术开发框架 2](#_Toc6159576)

[4.2关键开发技术 2](#_Toc6159577)

[4.3开发环境搭建 3](#_Toc6159578)

[4.4数据库设计 3](#_Toc6159579)

[4.5服务器设计 3](#_Toc6159580)

[4.6测试 3](#_Toc6159581)

[4.7扩展性 3](#_Toc6159582)

[第五章 成本模型及可行性分析 3](#_Toc6159583)

[5.1成本模型 3](#_Toc6159584)

[5.2可行性分析 3](#_Toc6159585)

# 第一章 方案概要

## 1.1目标问题

## 1.2总体思路

## 1.3具体做法

## 1.4创新亮点

# 第二章 项目目标与服务模型

## 2.1项目目标

### 2.1.1 功能性需求分析

投诉线索智能筛查系统，旨在协助投诉处理工作人员从海量的投诉信息中，快速找到重要信息，自动合并同类投诉内容，并行成投诉线索，自动分发流程给各个行政部门，提供线索处置监督管理，最后完成投诉内容的处理和反馈；同时通过汇总一段时间投诉内容、投诉部门、事件影响、处理流程等信息，可以从更宏观角度跟踪和分析市民关心的热点问题和社会问题，为城市管理提供辅助决策功能。

需要做到的四个核心模块：

1. 投诉线索数据采集模块

从市政府12345在线投诉平台爬取投诉举报信息

1. 投诉线索数据分析处理模块

对抓取的投诉举报信息进行处理：格式化，结构化提取出实体（如：xxx超市等），地址信息，地理位置，行政区域，事件类型等信息。

1. 投诉线索智能筛查模块

对上一模块生成的结构化数据和关联关系，对投诉信息自动分类和合并，并根据投诉内容、类别、区域，自动将生成的投诉线索分发给各个行政单位。

1. 投诉线索统计分析模块

通过对一段时间内容投诉信息、投诉线索、投诉处置、行政区域、事件影响等信息汇总分析，完成（包括但不仅限于）：

1. 各个类别、行政区域、行政单位的投诉信息统计分析（被投诉主体，区域、单位、领域等统计分析）

2. 对各个热点领域社会投诉问题的跟踪和报警，并提供趋势分析等功能

### 2.1.2 需求分析方案

按照确定项目目标及范围→获取用户需求→分析用户需求→编写需求文档→评审需求文档→管理需求的流程，进行需求分析后，我们整理分析出了如下几个需要突破的难点：

1. **如何对投诉进行准确分类**
2. **将来源不同的同一件事合并**
3. **准确的进行热点分析和预测**

## 2.2服务模型

### 2.2.1初期支持

在用户（有关机关部门）使用站点初期，我们设置了多方面的许多直观的可视化数据，使用户可以一眼便了解到其可能想要获得的信息。

### 2.2.2后期支持

在用户（有关机关部门）使用站点的过程中，我们可以随时为用户解决相关问题，并为用户提出的新需求增设模块以及不同种类图表。

## 2.3项目价值及创新点

## 2.4项目创新点

# 第三章 组织管理与业务分析

## 3.1组织管理

## 3.2业务分析

# 第四章 技术路线及实现方法

## 4.1技术开发框架

## 4.2关键开发技术

### 4.2.1 算法

### 4.2.2 前端

整个站点采用基于HTML5 ★ BOILERPLATE的模版，可以构建快速，强大且适应性强的 Web应用程序或站点。它包含了jQuery和Modernizr以及Normalize.css，相较于传统的网页开发提高了开发效率。

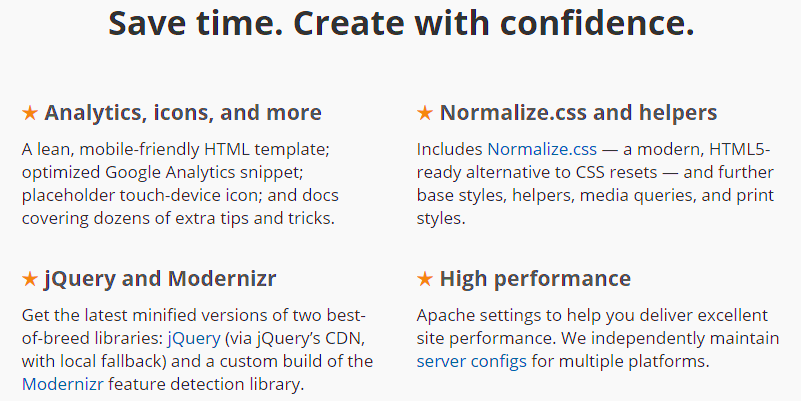


图4 - 1

由于本项目的核心为数据处理的部分，前端仅需展现一些统计图表和处理投诉消息的操作界面即可，业务逻辑相对简单，故采用了jQuery+Bootstrap作为web开发框架，采用Echarts作为可视化库，方便制作高度个性化定制的数据可视化图表。



图4 - 2

考虑到用户通过本项目获得一些统计分析信息后，有分析的需求，增加了一键图表导出为图片、部分图表一键切换类型的功能。

## 4.3开发环境搭建

## 4.4数据库设计

## 4.5服务器设计

## 4.6测试

## 4.7扩展性

# 第五章 成本模型及可行性分析

## 5.1成本模型

但是

## 5.2可行性分析

### 5.2.1 技术可行性

#### 5.2.1.1 核心算法准确度高

本系统经准确性验证表明，实际处理准确度较高，可以精确地把投诉信息进行命名实体识别、情感分析、关键词提取、文本摘要、投诉分类等处理，另外可以将信息分为投诉、询问、无效信息，效果较好。

#### 5.2.1.2 技术选取合理、成熟度高

现阶段的计算机技术、人工智能技术发展极快，我们采用了scrapy爬虫、redis增量式去重+不间断爬取、BiLSTM+crf、lstm、TextRank+Word2Vec、TextCNN、canopy+kmeans等等核心技术，它们充分支持了我们核心算法的实现，取得了不错的成效。

#### 5.2.1.3 开发团队经验丰富

FLY团队目前有开发人员4人，包括1名前端开发人员（包括前端可视化的设计和功能的实现），1名后端设计人员（包括数据库的设计以及后台的数据管理），1名算法人员（进行核心算法的开发），1名系统测试人员（进行系统的测试和提出建议）。以上人员都在自己的开发领域有一定的经验，且学习能力较强，因此有能力在较短时间内完成整个系统。

### 5.2.2 操作可行性

#### 5.2.2.1 界面直观，操作方便

本系统的设计过程中充分考虑到用户的实际使用情况，尽可能直观地提供可视化界面、操作界面。同时我们也有强大的数据分析能力，将数据处理成了多种多样简明直接又不失美观的图表，为相关部门的趋势分析提供方便。

#### 5.2.2.2 拓展性强，维护性好

本项目目前的训练数据全部由团队自己从各省政务网站采集，数据量有限，这也限制了更高的准确性。如果未来获得应用，相关部门可以为本系统提供更加海量的投诉信息数据，我们可以将模型训练得更加精准，更好的帮助相关人员提高工作效率。

本项目当前只针对浙江政务服务网。本项目不需要经过复杂的修改，只需要获得数据，就可以扩展到不同省份地区的政务服务网，为不同地区服务。

### 5.2.3 社会经济可行性

#### 5.2.3.1 成本模型合理

由上述成本模型分析可得，同时通过对系统软件开发成本的估算及可能取得直接或间接效益的评估，本项目在经济上使用成本较低，在收益上具有良好的前瞻性。

#### 5.2.3.2 存在实际需求

近年投诉信息数量越来越多、反应问题越来越细，使用原先的工作人员人工逐条查看分析的模式已经不能满足市民群众快速处理投诉的要求。近年政府的政务服务改革目标是“最多跑一次”，也充分体现了政府部门提高工作效率的祈愿。

本系统可以更好地辅助投诉处理工作人员和相关行政部门，也能从宏观角度、以可视化的统计分析数据来提供投诉信息的各种状况和趋势。