**知识智能感知系统项目计划书**

**学院：计算机学院**

**小组成员：王海 吕鑫 马茼茼 胡曦伟 谢欣旖**

**指导老师：刘小明**

**时间：2019年1月15日**

目录

[1. 引言 3](#_Toc535417761)

[1.1编写目的 3](#_Toc535417762)

[1.2项目简介 3](#_Toc535417763)

[1.3开发背景 3](#_Toc535417764)

[1.4 参考资料 3](#_Toc535417765)

[2. 项目概述 4](#_Toc535417766)

[2.1 项目目标 4](#_Toc535417767)

[2.2 应交付成果 4](#_Toc535417768)

[3. 需求管理 4](#_Toc535417769)

[3.1 需求管理过程 4](#_Toc535417770)

[3.2 任务分解 5](#_Toc535417771)

[3.3变更控制 5](#_Toc535417772)

[4. 项目组织计划 5](#_Toc535417773)

[4.1 组织结构 5](#_Toc535417774)

[4.2 人员分工 6](#_Toc535417775)

[4.3 协作与沟通 7](#_Toc535417776)

[4.3.1 内部协作 7](#_Toc535417777)

[4.3.2 外部沟通与协作 8](#_Toc535417778)

[3.3.3 项目沟通技巧 8](#_Toc535417779)

[4.4项目风险计划 8](#_Toc535417780)

[5.项目时间 10](#_Toc535417781)

[5.1 项目进度计划 10](#_Toc535417782)

[5.2 项目交付期限 12](#_Toc535417783)

# 引言

## 1.1编写目的

为保证整个系统开发项目能够按时且保质保量的完成，使每个人在该项目的开发过程中能够发挥自己的特长及作用，使整个项目的开发过程能够顺利平稳且有序的进行，并为项目开发提供有效且有力的进度参考。

## 1.2项目简介

利用现有的规范化文本处理的理论技术，对互联网信息的采集、存储和检索，以及领域分析，主题发现和文本摘要提取的知识智能感知系统。

信息采集采用爬虫技术利用网页节点进行页面元素配置，实现同一网站的长期时实爬取；信息将存与elasticsearch数据库中，后台搭建elasticsearch高级检索子系统，采用bool嵌套查询的方式对已爬取的数据编辑条件进行数据前端显示；选取所要进行分析的文本信息；选取相应的功能进行算法分析（支持向量机svm算法文本领域分类，隐含狄利克雷分布lda算法将进行主题发现，摘要提取，）

## 1.3开发背景

互联网文本中包含有大量的语义知识信息，然而这些海量文本信息具有容量大、多样性、生产速度快和价值密度低的特征。众多行业领域的研究和应用（如互联舆情分析、网络）都迫切需要一种统一的针对不规范大数据互联网信息的知识库。

## 参考资料

[1]汪云飞：《Java EE开发的颠覆者：Spring Boot 实战》，电子工业出版社2016.3 ；

[2]【Craig Walls】著，丁雪丰 译：《Spring Boot实战》，人民邮电出版社2016.8；

[3]饶琛琳:《ELK Stack权威指南（第2版）》，机械工业出版社2017.5

[4]朴灵：《深入浅出nodejs》，人民邮电出版社2013.12

[5]许家怡，白忠建，吴磊编著：《软件工程：理论与实践》，高等教育出版社2015.4；

# 2. 项目概述

## 2.1 项目目标

产品目标： 构建面向互联网文本数据(新闻、论坛、帖子、微博、微信……)的语义知识智能感知与提取系统。前期基于现有理论技术完成初步实验结果(Baseline)，作为初步的信息化构建方案；后期针对理论模型的改进方案，实现智能化的改进并实验验证。

系统主要包括：

1.互联网信息的智能采集与存储子系统

2.一体化的智能检索子系统

3.面向不同行业领域的主题发现与分析子系统

## 2.2 应交付成果

完整的项目需求书，实训报告统计表。进度安排表，人员分工表。

# 3. 需求管理

## 3.1 需求管理过程

1、组织策划和技术骨干代表编写系统功能描述。

2、调查用户的实际情况，明确用户需求。

3、做好市场调研，总结已有同类系统优势和缺点，针对痛点进行系统设计。

## 3.2 任务分解

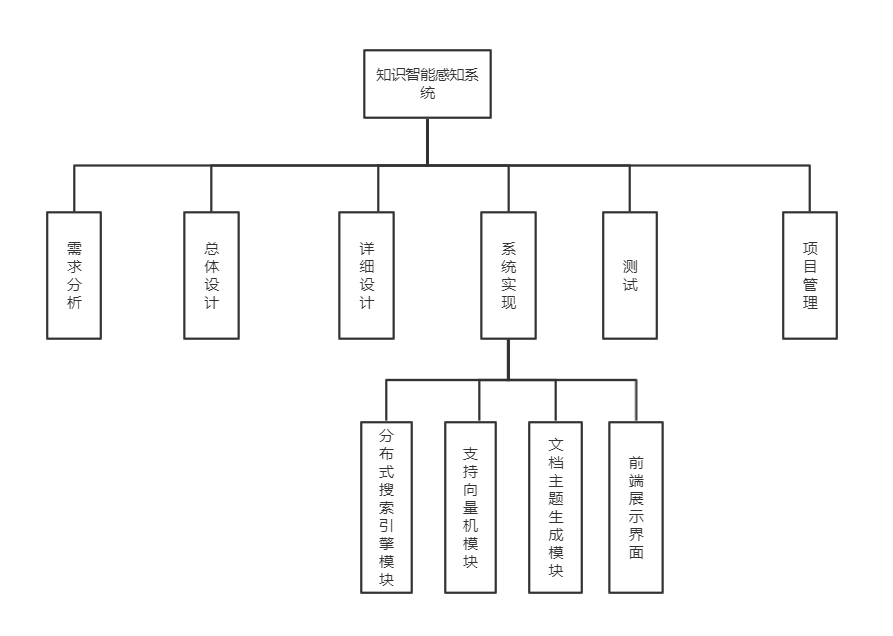


图3-2 工作分解结构图

## 3.3变更控制

1、项目涉及的知识比较复杂，学习起来有一定的困难，各成员需要及时搜索资料，参考文献进行学习。

2、本次开发时间处于学期中，会有其他学科考试备考的冲突。需要留出预留时间以不影响开发进度。

鉴于项目开发可能遇到的问题，我们将项目分成若干个阶段，根据每个阶段的完成情况及时进行汇总调整。

# 4. 项目组织计划

## 4.1 组织结构

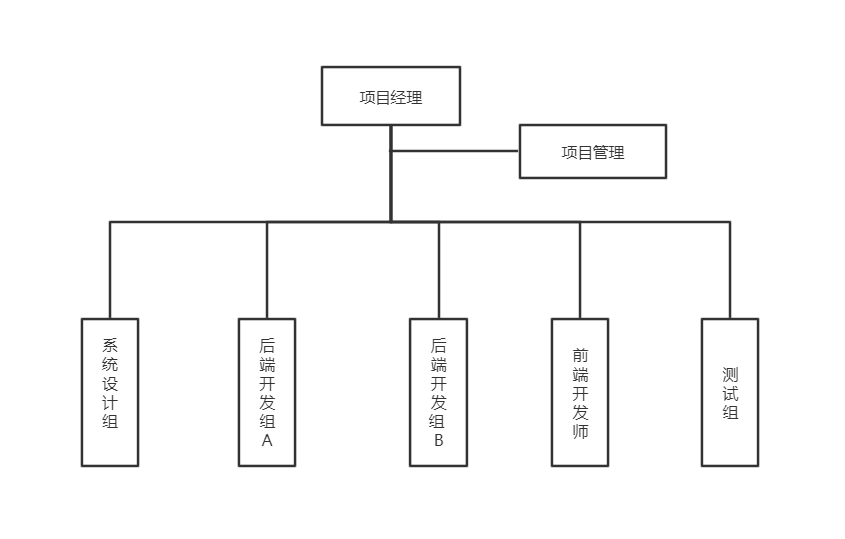


图4-1 软件项目的组织结构

## 4.2 人员分工

|  |  |  |
| --- | --- | --- |
| 姓名 | 角色名称 | 职责描述 |
| 王海 | 项目组长、算法工程师、配置管理员、编码人员 | 1、管理软件项目开发组与项目进度；  2、指导开发人员完成详细设计和代码编写，遵循开发规范，在项目范围内贯彻并监督规范的执行；  3.搭建和管理配置管理库，负责产品的变更管理、配置标识管理、配置审计等工作； 4.负责产品的版本控制、集成构建和发布管理；  5、负责编写系统中的关键模块和关键算法的程序，并进行综合测试、修改、代码走查工作； |
| 吕鑫 | 算法工程师、需求分析员、编码人员 | 1、参与项目的整体规划与实施；  2、负责编写系统中的关键模块和关键算法的程序，并进行综合测试、修改、代码走查工作；  3、负责与项目经理共同进行客户调研、业务流程分析设计及问题沟通协调；  4、负责系统的总体技术方案与系统设计，系统的质量控制。 |
| 马茼茼 | 算法工程师、编码人员 | 1、参与项目的整体规划与实施；  2、负责编写系统中的关键模块和关键算法的程序，并进行综合测试、修改、代码走查工作；  3、负责与项目经理共同进行客户调研、业务流程分析设计及问题沟通协调；  4、负责系统的总体技术方案与系统设计，系统的质量控制。 |
| 胡曦伟 | 算法工程师、编码人员 | 1、参与项目的整体规划与实施；  2、负责编写系统中的关键模块和关键算法的程序，并进行综合测试、修改、代码走查工作；  3、负责与项目经理共同进行客户调研、业务流程分析设计及问题沟通协调；  4、负责系统的总体技术方案与系统设计，系统的质量控制。 |
| 谢欣旖 | 需求分析员、界面设计员 | 1、根据产品规划或者项目要求，对客户需求进行收集、整理、分析，编制需求规格说明书；  2、协助系统架构师、开发工程师对需求进行理解，协助开发人员完成产品的界面、功能、流程设计；  3、基于项目对需求进行管理,负责项目中的需求变更、需求确认、需求控制等需求相关工作，编写各种需求文档；  4、编写对应页面，实现通过接口与后端建立数据交互，解决跨域请求问题。 |

表4-2-1 人员分工表

## 4.3 协作与沟通

项目沟通管理包括为保证及时与合理地生成、收集、分发、储存、提取及最终利用项目信息所需要的各过程。旨在为保证各项目干系人(包括项目团队、关系人、客户及发起人)及时得到信息并对信息做出相应的反应。项目沟通几乎贯穿于项目的每个环节，有效的沟通管理促成项目的成功，各项目干系人都应明白沟通会对项目产生怎样的影响。

### 4.3.1 内部协作

为了保证项目开发过程的顺利进行和信息的有效沟通，特要求如下的沟通计划：

1.每天17:00—17:30项目组成员进行口头交流；

2.每周五的14：00前提交周报告，并制作格式模板；

3.每周一的13：00-14：00召开项目周例会，会后发布会议纪要给相关的项目人员，其中说明项目的进展和存在的问题；及时提交问题报告，问题可以通过网络提交，项目经理会及时获取问题信息。

### 4.3.2 外部沟通与协作

为了满足用户的产品需求，更好更完善产品功能，特要求如下的沟通计划：

1. 项目负责人与指导老师交流时根据指导老师需求进行详细的提问，直至双方达成共识。
2. 与团队成员讨论，划分功能模块，给出工程时间轴，并撰写详细的需求文档，召集客户相关部门的负责人一起商讨需求文档，要听取客户的建议，将任何变动要求都分类纪录。安抚情绪，解答困惑，控制需求变动。
3. 将会以商协中所有分歧性意见都逐条写下，抄送所有相关负责人，并要求他们纠正分歧和确认变动。根据确认过的反馈回复，修改需求文档，知道需求文档定稿。

### 项目沟通技巧

1. 运用正确的表达方式。沟通必须目的明确。在信息交流之前，发送者应考虑好自己将要表达的意图，要力求简明扼要。用简单明了的词句表明自己的意思。
2. 提高倾听技能。沟通不仅仅是说，而是说和听。倾听既是我们取得关于他人第一手信息、正确认识他人的重要途径，也是我们向他人表示尊重的最好方式。
3. 避免无休止的争论。这种无休止的争论带来的结果是没有定论，不仅问题没有解决，而且延误了问题解决的时间。
4. 保持畅通的沟通渠道。沟通固然重要，但如果没有畅通的沟通渠道，组织就必然呈现自发的无组织状态，就无法获得需要的真实的信息，整个组织的运转效能就会下降。
5. 使用高效的沟通工具。在项目组织内，通常会使用相关的成熟的项目管理软件、电子邮件系统、办公自动化系统等工具来支持项目各种信息的生成、传递及存储的要求。这些工具的使用，大大提高了沟通的效率，拉进了沟通双方的距离，减少了不必要的面谈和会议。
6. 把握沟通原则。一是沟通内外有别。即要求团队作为一个整体对外意见要一致，一个团队要用一种声音;二是非正式的沟通又助于关系融洽;三是采用对方能接受的沟通风格;四是沟通的升级原则，即第一步,和对方沟通;第二步,和对方的上级沟通;第三步,和自己的上级沟通;第四步，自己的上级和对方的上级沟通。五是扫除沟通的障碍。

## 4.4项目风险计划

风险是指在项目进行过程中可能发生的事件，这些事件将会对项目按预期时间、资源和预算完成产生重大影响。风险分析的目标是识别这些事件，设法避免这些事件的发生并制定一旦这些事件发生后的处理措施。

下面的这个风险列表就是通过一系列的风险识别、风险评估、风险应对，最后得出项目风险计划表

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 二、项目风险管理 | |  |  |  |  |  |  |
| 风险发生概率的判断准则： | |  |  |  |  |  |  |
| 高风险： >60%发生风险的可能性 中风险： 30-60%发生风险的可能性 低风险：<30%发生防线的可能性 | | | | | | | |
| 序号 | 风险描述 | 发生概率 | 影响程度 | 风险等级 | 风险响应计划 | | |
| 1 | 由于时间紧张最后无法按期完成 | 低 | 极大 | 高 | 充分考虑各种潜在因素，适当留有余地；任务分解要详细，便于考核；如果出现必须延期的情况，组长需及时同相关负责人沟通，并申请延期时间。 | | |
| 2 | 系统没有足够的测试时间 | 高 | 大 | 高 | 持续地监控，项目进度控制随着项目的进行而不断进行的，保证每个环节都有足够的时间。。 | | |
| 3 | 每个人的开发水平不一样，导致工作量完成度出现差错 | 中 | 大 | 高 | 尽可能在一起完成项目，互相帮助一起解决遇到的难题。 | | |
| 4 | 组员成员因意外无法参加设计 | 中 | 中 | 高 | 事先同用户商量解决办法 | | |
| 5 | 开发软件结构体系存在问题，使完成的软件产品未能实现项目预定目标 | 中 | 高 | 中 | 选用正版软件开发 | | |
| 6 | 质量不符合用户要求 | 低 | 高 | 高 | 经常和用户交流工作成果、品牌管理采用符合要求的开发流程、认真组织对产出物的检查和评审、计划和组织严格的独立测试等。 | | |

# 5.项目时间

## 5.1 项目进度计划

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 活动 | 紧前工作 | 工期估计 | 最早 | | 最迟 | | 时差 |
| 开始时间 | 结束时间 | 开始时间 | 结束时间 |
| 1 需求分析 | - | 30 | 0 | 30 | 0 | 30 | 0 |
| 2 总体设计 | 1 | 25 | 25 | 50 | 25 | 50 | 0 |
| 3 分布式搜索引擎子系统详细设计 | 1 | 15 | 50 | 65 | 50 | 65 | 0 |
| 4 支持向量机子系统详细设计 | 1 | 15 | 50 | 65 | 50 | 65 | 0 |
| 5 文档主题分析子系统详细设计 | 1 | 15 | 50 | 65 | 50 | 65 | 0 |
| 6 前端界面详细设计 | 1 | 15 | 50 | 65 | 50 | 65 | 0 |
| 7 数据整理 | 1 | 15 | 60 | 75 | 65 | 80 | 5 |
| 8 分布式搜索引擎子系统编码 | 3,7 | 25 | 75 | 100 | 80 | 105 | 5 |
| 9 SVM模型更改 | 4 | 15 | 100 | 115 | 105 | 120 | 5 |
| 10 支持向量机子系统编码 | 4 | 20 | 125 | 145 | 150 | 170 | 5 |
| 11 SVM模型调优 | 10 | 15 | 145 | 160 | 170 | 185 | 5 |
| 12 LDA模型更改 | 5 | 15 | 100 | 115 | 105 | 120 | 5 |
| 13文档主题分析子系统编码 | 5 | 25 | 115 | 140 | 120 | 135 | 5 |
| 14 LDA模型调优 | 13 | 15 | 140 | 155 | 140 | 155 | 0 |
| 15 前端界面编码 | 6 | 25 | 100 | 125 | 100 | 125 | 0 |
| 16 分布式搜索引擎子系统单元测试 | 8 | 25 | 100 | 125 | 100 | 125 | 0 |
| 17支持向量机子系统单元测试 | 11 | 25 | 140 | 165 | 140 | 165 | 0 |
| 18 文档主题分析子系统单元测试 | 14 | 30 | 140 | 165 | 140 | 165 | 0 |
| 19 前端界面测试 | 25 | 25 | 140 | 165 | 140 | 165 | 0 |
| 20 系统联调与测试 | 16 17 18 19 | 30 | 140 | 170 | 140 | 170 | 0 |
| 21 服务部署 | 20 | 15 | 160 | 175 | 160 | 175 | 0 |
| 22 数据采集 | - | 175 | 0 | 175 | 0 | 175 | 0 |

表5-1-1 项目进度表

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 时间 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| 需求分析 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 总体设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 分布式搜索引擎子系统详细设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 支持向量机子系统详细设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 文档主题分析子系统详细设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 前端界面详细设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 数据整理 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 分布式搜索引擎子系统编码 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SVM模型更改 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 支持向量机子系统编码 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SVM模型调优 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LDA模型更改 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 文档主题分析子系统编码 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LDA模型调优 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 前端界面编码 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 分布式搜索引擎子系统单元测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 支持向量机子系统单元测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 文档主题分析子系统单元测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 前端界面测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 系统联调与测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 服务部署 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 数据采集 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 项目管理 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

图5-1-2 知识智能感知系统项目工作计划甘特图（时间单位：周）

## 5.2 项目交付期限

第20周周四