|  |  |  |
| --- | --- | --- |
| 文件状态：  [ ] 草稿  [√] 正式发布  []正在修改 | 文件标识： | SYZH1104-TYPT-01 |
| 当前版本： | V1.0 |
| 作者： | 邓虹玉 |
| 审批： | 张亚情 |
| 完成日期： | 2018/1/20 |

基于语义网的站内搜索引擎模块改造项目工作说明书

**北京慧博科技有限公司**

**2018年2月**

目录

1 前言 1

1.1 目的 1

1.2 术语 1

1.3 参考 1

2 项目概述 1

2.1 项目目标 1

2.2 项目实施前准备工作 1

2.2.1 开发设备 1

2.2.2 其他依赖 2

2.2.3 其他参与方 2

2.3 系统体系结构 2

2.4 项目工作范围 3

2.5 系统建设方案设计 3

2.5.1 工作描述 3

2.5.2 工作内容 4

2.5.3 工作交付物 4

2.5.4 工作完成标志 4

2.6 系统开发实施 4

2.6.1 工作描述 4

2.6.2 工作内容 4

2.6.3 工作交付物 5

2.6.4 工作完成标志 5

2.7 系统测试验证 5

2.7.1 工作描述 5

2.7.2 工作内容 5

2.7.3 工作交付物 5

2.7.4 工作完成标志 5

2.8 系统部署 6

2.8.1 工作描述 6

2.8.2 工作内容 6

2.8.3 工作交付物 6

2.8.4 工作完成标志 6

2.9 系统验收 6

2.9.1 工作描述 6

2.9.2 工作内容 6

2.9.3 工作交付物 7

2.9.4 工作完成标志 7

2.10 项目组织与管理 7

2.10.1 工作描述 7

2.10.2 工作内容 7

2.10.3 工作交付物 7

2.10.4 工作完成标志 7

2.11 系统维护与技术支持 8

2.11.1 工作描述 8

2.11.2 工作内容 8

2.11.3 工作交付物 8

2.11.4 工作完成标志 8

3 系统建设范围 8

3.1 系统功能范围 9

3.1.1 集客CRM 9

3.1.1.1 功能描述 9

3.1.1.2 系统功能 9

3.1.2 积分系统 11

3.1.2.1 功能描述 11

3.1.2.2 模块功能 11

3.2 系统数据建设范围 12

3.2.1 淘宝平台数据处理 12

3.2.2 系统功能数据处理 12

3.2.3 积分系统数据 12

4 项目交付清单 12

5 项目验收标准 13

5.1 软件验收标准 13

5.2 软件过程文档验收标准 14

6 项目实施进度 14

6.1 实施进度计划表 14

6.2 阶段里程碑 17

7 项目组织结构 17

7.1 项目组织结构 18

7.2 角色分派及角色职责 18

8 变更管理流程 18

8.1 变更管理流程 19

8.2 重大变更定义 19

8.3 变更责任 19

1. 前言
   1. 目的

为保证项目的顺利实施，本文档将作为项目执行过程中的基准。

* 1. 术语

无。

* 1. 参考

无。

1. 项目概述
   1. 项目目标

基于知识图谱技术的站内搜索引擎的优化改造，一是对前端体验优化，改善前端搜索功能、提供搜索增项数据、筛选条件排序优化；二是搜索引擎模块功能的改造，在solr的基础上对搜索引擎进行二次改造，研究利用爬虫技术，抓取采购专区商品页面目标数据，建设商品知识图谱，优化中文分词和搜索结果排序算法；三是开发数据存储功能，存储底层的用户数据，以供数据分析和其他项目应用拓展。

通过上述改造，结合爬虫技术、中文分词、反序索引库等技术的引入，基本实现前端体验合理化、系统功能智能化、数据应用简单化的总体目标。对于用户，站内搜索的优化可以改善用户搜索的准确率，快速找到用户想要的商品，避免过多操作，提升体验；对于运营，站内搜索引擎模块的价值巨大，可以了解用户的意图，改善站点的用户体验，优化网站的结构，可以收集用户的画像相关数据，作为用户浏览相关推荐的参数，另外，用户数据还可以应用在大数据和人工智能上，助力其他产品。

* 1. 项目实施前准备工作
     1. 开发设备

软件开发服务器1台，服务器配置不低于以下要求。

* CPU： 4核
* 内存：16G
* 硬盘：300G
* 系统软件：Centos6.5
  + 1. 其他依赖

无。

* + 1. 其他参与方

无

* 1. 系统体系结构

基于语义网的站内搜索引擎系统主要包含应用服务、业务管理、资源加工、数据及知识处理以及分析管理5个主要的业务应用。

用户服务主为内外部用户提供各种业务的入口和业务服务访问；业务管理主要为业务管理人员管理平台的各种支撑性服务；资源加工主要为商品及相关数据资源转化到知识图谱及索引库服务，语义计算、信息计算主要负责资源数据抽取后，清洗和计算出理等基础技术。

基于语义网的站内搜索引擎系统的应用架构如下图：



* 1. 项目工作范围

项目工作范围主要描述了在项目实施的各个阶段提供服务的主要内容，明确双方的职责，确定阶段交付成果以及完成标准。同时也说明项目不包含的工作任务。

正常情况下，项目工作任务要求按照如下阶段进行描述：

* 编制系统建设方案
* 项目需求分析
* 项目详细设计开发
* 系统开发及测试
* 系统部署
* 上线试运行
* 系统验收
* 系统维护与技术支持
  1. 系统建设方案设计
     1. 工作描述

系统建设方案设计是本项目前期工作，是项目开展的前提。系统建设方案设计的主要工作包括组织相关人员制定技术路线和解决方案。与业务部门沟通解决方案中业务范围和要求，组织销售代表及产品和开发人员对系统技术方案中的实现部分进行评审确认等。

* + 1. 工作内容

在本阶段的工作内容为：

* 制定总体技术方案思路和解决方案框架；
* 组织相关人员一起编写总体技术方案；
* 与业务部门不断沟通，完善技术方案；
* 组织完成技术方案确认工作。
  + 1. 工作交付物

《系统建设方案》，包括项目背景、需求、解决方案、实施计划等内容。

* + 1. 工作完成标志

系统建设方案通过业务部门、研发部门评审。

* 1. 系统开发实施
     1. 工作描述

系统开发实施阶段的主要工作是软件系统开发执行的主要阶段，本阶段的主要够工作有需求分析、系统详细设计、系统开发。

* + 1. 工作内容

在本阶段的工作内容为：

* 组织项目需求分析；
* 进行系统详细设计；
* 进行系统编码以及其他技术实现；
* 组织设计评审、阶段与用户沟通。
  + 1. 工作交付物
* 《详细设计方案》，系统实现的详细设计信息。
* 《数据库设计方案》，数据库结构及存储过程等设计信息。
* 《系统测试报告》，项目内部功能、性能测试报告。
  + 1. 工作完成标志

内部测试报告完成，得到产品部门确认。

* 1. 系统测试验证
     1. 工作描述

系统测试验证阶段主要从功能上、性能上以及系统部署等方面对系统进行测试、验证及完善工作。本阶段的具体工作包括系统测试、业务用户测试、性能测试。

* + 1. 工作内容

在本阶段的工作内容为：

* 进行系统集成测试；
* 组织用户业务测试；
* 执行系统整体性能测试；
* 对部署流程进行前期验证和测试；
  + 1. 工作交付物
* 《系统集成测试报告》，参与各方系统集成测试后的评价报告。
* 《系统用户手册》，用户操作手册。
* 《系统用户测试报告》，用户测试确认报告。
  + 1. 工作完成标志

集成测试报告、用户测试报告经过产品及业务部门确认。

* 1. 系统部署
     1. 工作描述

系统实施阶段是系统部署的执行阶段，包括业务系统、数据库结构等内容的部署。同时，包括用户培训工作。

* + 1. 工作内容

在本阶段的工作内容为：

* 制定详细的系统实施计划，组织系统实施工作；
* 执行系统部署工作；
* 组织用户培训；
* 上线后在线技术支持。
  + 1. 工作交付物
* 《系统部署方案》，系统最终执行的部署流程、每一步操作说明等。
* 《系统部署详细工作计划》，详细工作步骤和时间、人员安排。
* 《系统维护手册》，系统管理维护说明书。
* 《系统上线确认报告》，系统部署成功确认信息。
  + 1. 工作完成标志

《系统上线确认报告》得到领导确认签字。

* 1. 系统验收
     1. 工作描述

系统运行三月后启动系统验收工作，对项目成果进行确认。

* + 1. 工作内容

在本阶段的工作内容为：

* 配合验收规定提供验收资料；
* 负责项目内容讲解和说明。
* 组织项目验收工作；
* 对项目验收报告进行确认。
  + 1. 工作交付物

验收规定的文档资料。

* + 1. 工作完成标志

验收报告签署。

* 1. 项目组织与管理
     1. 工作描述

承担项目整体方案、详细设计、开发编码、测试等过程实施各方组织和进度管理工作。

* + 1. 工作内容

在本阶段的工作内容为：

* 负责项目参与人员的组织工作；
* 负责整体方案的编写和执行过程进度、质量控制；
* 负责例会组织召开和回报。
  + 1. 工作交付物

《项目总体计划》、《项目周报》、《会议纪要》等。

* + 1. 工作完成标志

项目验收确认后结束。

* 1. 系统维护与技术支持
     1. 工作描述

系统验收后，对线上问题进行维护工作。

* + 1. 工作内容

在本阶段的工作内容为：

* 解决系统Bug缺陷的问题；
* 诊断和排查疑难问题；
* 知识转移和培训。
  + 1. 工作交付物

《系统支持记录》、《系统运行维护报告》。

* + 1. 工作完成标志

维护期结束，本项工作完成。

1. 系统建设范围

本系统建设内容包括系统组件开发、系统数据建设两部分。

* 1. 系统功能范围
     1. 智能搜索

|  |  |  |
| --- | --- | --- |
| **序号** | **功能目录** | **功能点描述** |
| 1 | 智能提示 | 搜索词匹配补全 |
| 2 | 搜索词语义补全 |
| 3 | 搜索词语义消歧 |
| 4 | 搜索词全拼提示 |
| 5 | 搜索词首字母提示 |
| 6 | 搜索词混合拼音提示 |
| 7 | 统计提示 |

* + 1. 智能提示

|  |  |  |
| --- | --- | --- |
| **序号** | **功能目录** | **功能点描述** |
| 1 | 智能搜索 | 搜索意图识别 |
| 2 | 搜索语义拓展 |
| 3 | 商品过滤筛选 |
| 4 | 搜索拼音及错词纠正 |

* 1. 系统数据建设范围

1. Mysql相关

1）brand表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 是 | 品牌id | brand\_id | varchar | 15 | 是 |
| 否 | 规范的品牌名称 | brand\_name\_std | varchar | 255 | 是 |

2）brand\_synonym表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 否 | 规范的品牌名称 | brand\_name\_std | varchar | 255 | 是 |
| 否 | 品牌同义词 | brand\_synonyms | varchar | 255 | 否 |

3）cate\_prop表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 是 | 分类id | cate\_id | varchar | 15 | 是 |
| 否 | 分类名称 | cate\_name | varchar | 255 | 否 |
| 否 | 属性名 | cate\_prop\_name | varchar | 255 | 否 |
| 否 | 规范的属性名 | cate\_prop\_name\_std | varchar | 255 | 否 |
| 否 | 属性英文名 | cate\_prop\_name\_en | varchar | 255 | 否 |
| 否 | 属性值映射 | prop\_value\_mapping | varchar | 5000 | 否 |
| 否 | 属性值排序 | prop\_value\_sort | varchar | 1000 | 否 |

4）cate\_synonym表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 主键 | 名称 | 代码 | 数据类型 | 长度 | 非空 |
| 是 | 分类id | cate\_id | varchar | 15 | 是 |
| 否 | 分类名称 | cate\_name | varchar | 255 | 否 |
| 否 | 属性同义词 | cate\_synonyms | varchar | 255 | 否 |

5）no\_result\_search\_word表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 否 | 搜索词 | sw | varchar | 100 | 是 |

2. Elasticsearch相关

1）gwds\_prompt\_es\_index表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 否 | 名称 | name | text | 100 | 是 |
| 否 | 数量 | count | long |  | 否 |
| 否 | 类型 | type | int |  | 是 |

2）gwds\_prod表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 是 | 商品id | prod\_id | text | 64 | 是 |
| 否 | 商品名称 | prod\_name | text | 100 | 是 |
| 否 | 店铺名称 | prod\_store\_name | text | 100 | 是 |
| 否 | 店铺id | prod\_store\_id | text | 64 | 是 |
| 否 | 品类 | prod\_cat\_name | text | 32 | 是 |
| 否 | 品牌 | prod\_brand\_name | text | 32 | 否 |
| 否 | 站点信息 | prod\_site | object |  | 否 |
| 否 | 商品熟悉 | prod\_props | object |  | 否 |

3. Tokumx/MongoDB相关

1）entity\_id

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 是 | 系统自动生成的ID | \_id | int | 32 | 是 |
| 否 | **名称** | name | varchar | 255 | 是 |
| 否 | ID | id | int | 32 | 是 |
| 否 | 概念iD | concept\_id | int | 32 | 是 |
| 否 | 类型（概念或实体） | type | int | 1 | 是 |

2）entity\_synonym

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 是 | 系统自动生成的ID | \_id | int | 32 | 是 |
| 否 | ID | id | int | 32 | 是 |
| 否 | 同义词 | synonym | int | 32 | 是 |

3）parent\_son

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 是 | 系统自动生成的ID | \_id | int | 32 | 是 |
| 否 | 父概念ID | parent | int | 32 | 是 |
| 否 | 子概念ID | son | int | 32 | 是 |

4）attribute\_object

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 是 | 系统自动生成的ID | \_id | int | 32 | 是 |
| 否 | 实体ID | entity\_id | int | 32 | 是 |
| 否 | 属性ID | attr\_id | int | 32 | 是 |
| 否 | 值实体ID | attr\_value | int | 32 | 是 |
| 否 | 实体ClassID | entity\_type | int | 32 | 是 |
| 否 | 值实体ClassID | attr\_value\_type | int | 32 | 是 |

5）attribute\_string

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **主键** | **名称** | **代码** | **数据类型** | **长度** | **非空** |
| 是 | 系统自动生成的ID | \_id | int | 32 | 是 |
| 否 | 实体ID | entity\_id | int | 32 | 是 |
| 否 | 属性ID | attr\_id | int | 32 | 是 |
| 否 | 属性值 | attr\_value | text |  | 是 |
| 否 | 实体ClassID | entity\_type | int | 32 | 是 |

1. 项目交付清单

|  |  |  |
| --- | --- | --- |
| 交付物名称 | 介质形式 | 备注 |
| 《项目总体计划》 | 电子 |  |
| 《总体技术方案》 | 电子、纸质 |  |
| 《详细设计方案》 | 电子、纸质 |  |
| 《数据库设计方案》 | 电子 |  |
| 《系统测试报告》 | 电子 |  |
| 《系统用户手册》 | 电子 |  |
| 《系统用户测试报告》 | 电子、纸质 |  |
| 《系统维护手册》 | 电子 |  |
| 《系统上线确认报告》 | 纸质 |  |
| 《系统需求规格说明》 | 电子 |  |

1. 项目验收标准
   1. 软件验收标准

* 系统功能验收

以本文第4条系统需求范围为依据，以各部门确认的《项目软件需求规格说明书》所定义的功能为验收的标准。

* 性能验收
  1. 数据加载传送响应时间小于20秒；
  2. 业务功能查询响应时间小于5秒；（与三级调度等系统相集成后，由其提供的功能查询响应时间另计）
  3. 随着业务数据的增加，系统响应时间可能会产生相应的变化，但不影响系统的正常使用。
* 软件缺陷

如果验收测试后，系统的残余缺陷按照严重程度分类，分别满足对应的验收标准则视为验收测试合格，存在问题在约定期间内进行修正。

|  |  |  |
| --- | --- | --- |
| 严重程度 | 描述 | 验收标准 |
| A | 系统由于出现问题不能继续运行 | = 0 |
| B | 系统中某项业务（某一模块）无法进行。 | = 0 |
| C | 系统中单一交易不能继续运转或错误，但不影响业务的进行。（有替代办法） | <= 10 |
| D | 其他：系统在某种状态下产生错误，不影响正常业务。限于那些装饰性且在软件处理上没有影响的问题。有着这些性质的缺陷可能留着未解决并积累直到软件的按期发布。 | <=20 |

* 1. 软件过程文档验收标准

文档验收以抽样方式进行，抽样率为20%。如果在验收的文档中，错误的总字数超过抽样文档总字数的5%，或描述每一独立完整的功能/章节错误，则视为验收失败；否则为验收合格。

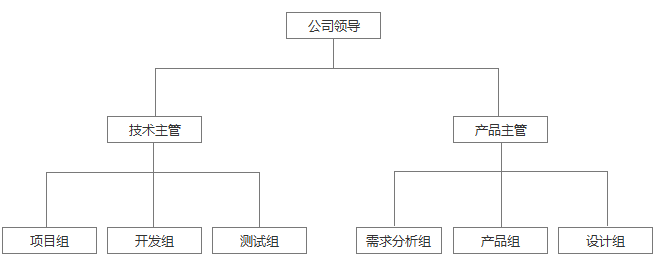
1. 项目实施进度
   1. 实施进度计划表

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 编号 | 活动 | 进度月份表 | | | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | 需求分析 |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | 系统  设计 |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | 系统开发与测试 |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | 项目验收及成果移交 |  |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 系统部署 |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | 用户测试 |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | 试运行 |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | 试运行验收 |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | 竣工验收 |  |  |  |  |  |  |  |  |  |  |  |  |

* 1. 阶段里程碑

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **阶段** | **计划周期** | **工作内容** |
| 1 | 需求分析 | 合同签订日起1个月 | 开展系统、数据的需求调研和分析，编写需求分析报告； |
| 2 | 系统设计 | 合同签订日起2个月 | 完成系统的总体概要设计、详细设计、数据库设计工作； |
| 3 | 系统开发与测试 | 合同签订日起4个月 | 编写代码，完成各个模块、功能的实现；并对系统各模块、功能进行程序测试； |
| 4 | 项目验收及成果移交 | 合同签订日起4个月 | 召开项目验收会；将系统开发的全部程序、相关文档交付给国网公司。 |

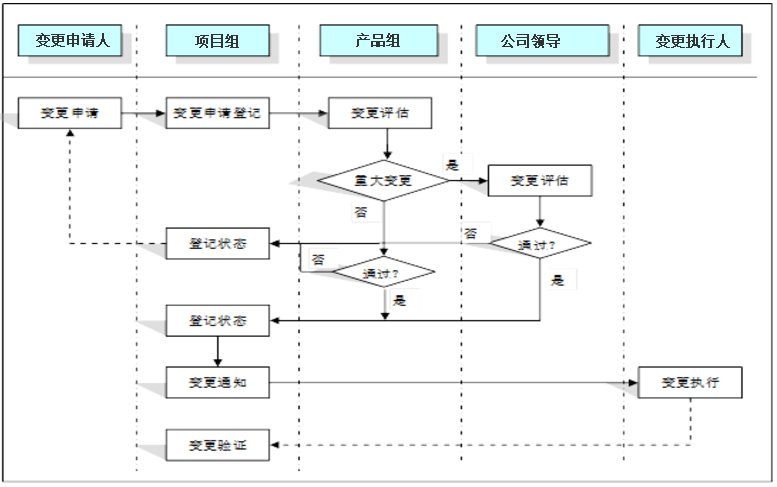
1. 项目组织结构
   1. 项目组织结构



* 1. 角色分派及角色职责

|  |  |
| --- | --- |
| 角色 | 职责 |
| 公司领导 | 负责产品发展方向 |
| 技术主管 | 负责产品开发管理 |
| 产品主管 | 负责产品设计管理 |
| 项目组 | 负责项目协调管理 |
| 开发组 | 负责系统开发、编码、单元测试 |
| 测试组 | 负责系统集成测试 |
| 需求分析组 | 负责业务性能需求分析。 |
| 设计组 | 负责系统设计和技术方案验证。 |
| 产品组 | 负责产品原型设计及产品质量把控 |

1. 变更管理流程
   1. 变更管理流程



* 1. 重大变更定义

重大变更包括：

1、项目进度角度。如果项目变更对项目进度影响超过1个自然月，必须作为重大变更进行评审、评估。

2、项目范围角度。如果每次项目变更影响到项目工作量的±5%，或者累计影响超过±10%，将变更列为重大变更。

* 1. 变更责任

项目发生重大变更后，需要对工作量进行评估，各部门应就该变更处理意见达成一致，并签署相关备忘录。