|  |  |  |
| --- | --- | --- |
| 文件状态：  [ ] 草稿  [√] 正式发布  []正在修改 | 文件标识： | SYZH1104-TYPT -03 |
| 当前版本： | V1.0 |
| 作者： | 邓虹玉 |
| 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

总体设计说明书

**北京慧博科技有限公司**

**2018年1月**

修订历史版本

|  |  |  |
| --- | --- | --- |
| **日期** | **修订内容** | **修订人** |
| 2018-01-20 | 创建文档 | 邓虹玉 |
| 2016-01-21 | 审核文档 | 张亚情 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

目录

第一章 引言 1

1.1 文档目的 1

1.2 阅读对象 1

1.3 范围 1

1.4 参考资料 1

第二章 系统架构设计 2

2.1 系统业务架构 2

2.2系统数据架构 2

2.3系统应用结构 3

第三章 系统功能设计 6

3.1 系统组件设计 6

3.1.1 系统逻辑分层 6

3.1.2 组件关联设计 10

3.1.3 组件时序设计 11

3.1.4 功能组件设计 12

3.1.5 接口组件设计 13

3.1.6 公共组件设计 14

3.2 系统界面设计 15

3.2.1 界面框架设计 15

3.2.2 系统空控件选择 16

第四章 系统运行环境设计 18

4.1 硬件环境 18

4.2 软件环境 18

1. 引言
   1. 文档目的

该文档旨在使项目管理人员、设计人员和开发人员充分了解公交通用地理信息平台的总体设计思想，使设计人员、开发人员能参照本文档中的技术总体设计，进行系统各功能的详细设计与开发。

* 1. 阅读对象
* 研发组、测试组、设计组等相关项目干系人；
* 基于语义网的站内搜索引擎模块改造项目相关成员。
  1. 范围
  2. 参考资料
* 《基于语义网的站内搜索引擎模块改造项目工作说明书》；
* 《基于语义网的站内搜索引擎模块改造项目需求分析说明书》。

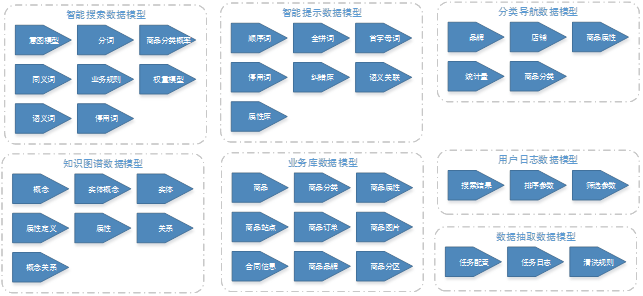
1. 系统架构设计
   1. 系统业务架构

基于语义网的站内搜索引擎模块改造项目系统在业务架构设计方面，强调以数据驱动为前提，以业务支持服务为目的，通过建设智能化电商搜索引擎，提升用户搜索体验，促进商城效益。如下图所示，业务框架主要包括资源抽取、资源加工、数据管理以及智能应用四大板块。



2.2系统数据架构

如下图所示，基于语义网的站内搜索引擎系统主要设计数据模型包括智能搜索数据模型、智能提示数据模型、分类导航数据模型、知识图谱数据模型、业务库数据模型和以及用户日志数据模型。

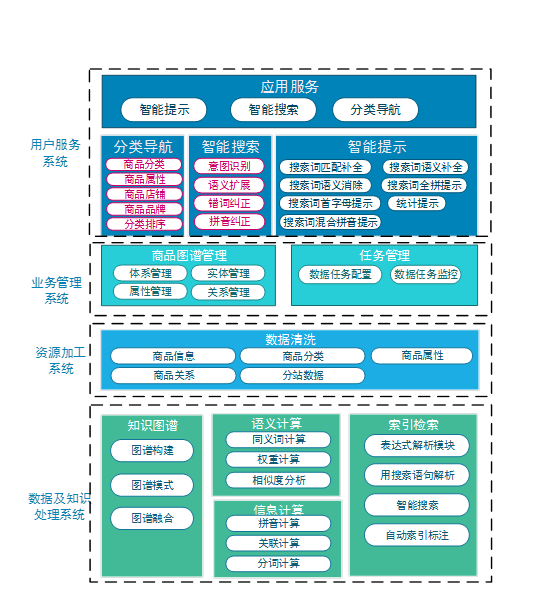


2.3系统应用结构

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。

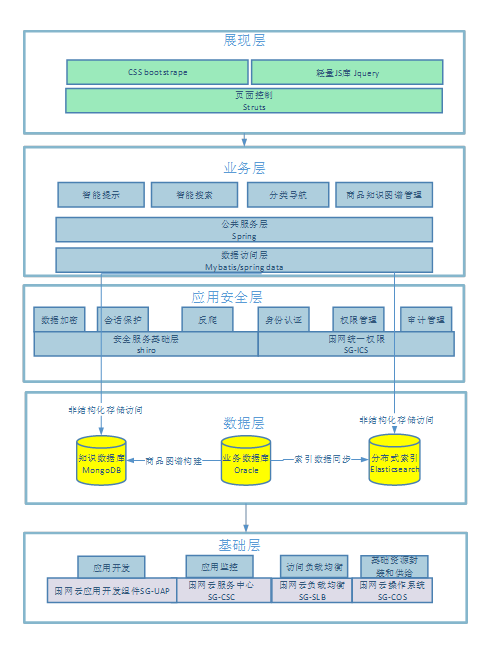
基于语义网的站内搜索引擎系统的应用架构如下图：



业务架构

技术架构实现上，遵循SG-EA技术架构设计规范，采用组件化、动态化的软件技术，利用一致的可共享的数据模型，按照展现层、业务逻辑层、安全层、数据服务层、基础架构服务层进行划分，实现多层技术体系设计。通过集成平台进行界面，数据和服务集成，实现科技资源数据的各接口组件能够为内外部用户提供多层次上集成，实现重用，以满足公司对知识访问的服务需求，为各单位提供高效便捷的知识访问支撑。

技术架构如下图：



1. 系统功能设计
   1. 系统组件设计
      1. 系统逻辑分层

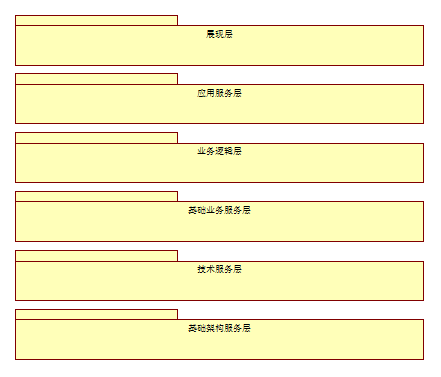
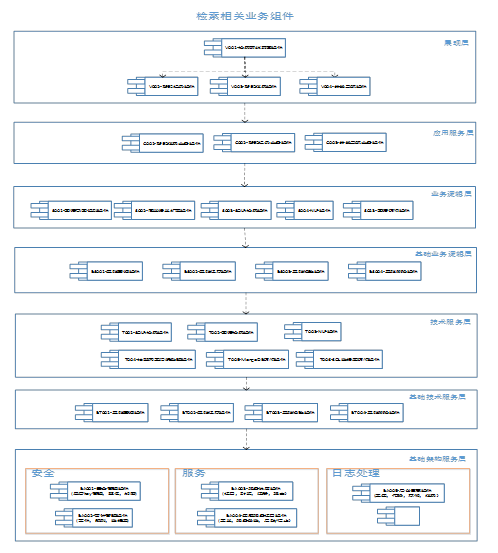


图 系统逻辑分层图

系统逻辑分层说明表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **逻辑层次** | **职责描述** | **技术实现** | **逻辑层次依赖** | **层间通信** |
| 展现层 | * 显示用户界面 * 接受用户输入 * 数据验证 * 呈现数据 * 向应用服务层发送请求 | * 使用JSP技术 * 使用JQUERY技术 * 使用AJAX技术使用 * 使用HTML5/DHTML技术 * 使用XML、JSON技术 * 使用Flex技术 * 使用WEBSOCKET技术 * 使用EASYUI技术 * 使用BOOTSTRAP技术 * 使用JQUERY.NICESCROLL.MIN.JS技术 * 使用ECHARTS3.0技术 * 使用JQGRID技术 * 使用JQUERY-SWIPE技术   使用JQUERY JBOX技术   * 使用MY97DATEPICKER技术 * 使用JQUERY SELECT2技术 * 使用JQUERY ZTREE技术 * 使用JQUERY.UPLOADIFY技术 * 使用JQUERYVALIDATION PLUGIN 1.11技术 | * 依赖应用服务层 | * 采用XML、JSON等数据载体 * 采用HTTP(S)协议进行通信 |
| 应用  服务层 | * 提供完整核心应用的服务 * 接收并处理展现层的数据请求 * 分析用户输入的请求数据 * 向业务服务层发起请求 | * 使用JAVA开发技术 * 使用SPRING进行IOC控制 * 使用SPRING MVC技术 | * 依赖业务逻辑层 | * 采用JAVA调用 * 采用共享缓存数据交换 * 采用HTTP(S)协议进行通信 |
| 业务  逻辑层 | * 接收应用服务层的请求 * 提供基础业务服务 | * 使用SPRING技术 * 使用LOG4J技术 * 使用JAVA开发技术 | * 依赖基础业务服务层 | * 采用JAVA调用 * 采用共享缓存数据交换 * 采用HTTP(S)协议进行通信 |
| 基础业务  逻辑层 | * 封装上层的公共基础业务逻辑处理服务 * 向技术服务层发起请求 * 模块服务调用 | * 使用SPRING技术 * 使用LOG4J技术 * 使用JAVA开发技术 | * 依赖技术服务层 | * 采用JAVA调用 * 采用共享缓存数据交换 * 采用HTTP(S)协议进行通信 |
| 技术服务层 | * 数据持久化操作 * 系统监控服务 * 基本接口服务 | * 使用SPRING技术 * 使用LOG4J技术 * 使用JAVA开发技术 | * 依赖基础技术服务层 | * 采用JAVA调用 * 采用共享缓存数据交换 * 采用HTTP(S)协议进行通信 |
| 基础技术  服务层 | * 封装上层服务使用的公共技术服务功能 * 数据持久化操作 * 数据缓存操作 | * 使用SPRING技术 * 使用STRUTS2技术 * 使用XFIRE技术 * 使用LOG4J技术 * 使用SPRINGDATA技术 * 使用MYBATIES/JDBC技术 * 使用MONGODB访问API技术国网大数据使用 * 使用Elasticsearch访问API技术 | * 依赖基础框架服务层 | * 采用Java调用 * 采用共享缓存数据交换 * 采用http(s)协议进行通信 |
| 基础架构  服务层 | * 支持系统整体基础框架的服务 * 服务器分发，请求服务级别分配 * 应用请求反向代理服务 * 静态元素请求响应服务 | * 使用nginx技术 |  |  |

* + 1. 组件关联设计



组件关联图

组件关联清单表

|  |  |  |  |
| --- | --- | --- | --- |
| 组件编号-名称 | 组件的方法 | 关联的组件编号-名称 | 关联的组件的方法 |
| V002-智能提示组件 | 搜索关联词 | C002-智能提示业务组件 | 提示关联词 |
| V003-智能搜索组件 | 搜索商品、店铺 | C001-智能搜索业务组件 | 搜索资源 |
| V004-分类导航组件 | 加载品类数据 | C003-分类导航业务组件 | 加载品类数据 |
| C002-智能提示业务组件 | 提示关联词 | S003-Solr检索组件 | 查询Solr |
| C001-智能搜索业务组件 | 搜索商品、店铺 | S001-图谱意图识别组件 | 意图识别 |
| S002-同义语义扩展组件 | 关键词语义扩展 |
| S004-NLP组件 | 分词、文本抽取、语义计算等 |
| S003-ES检索组件 | 查询Solr |
| C003-分类导航业务组件 | 加载品类数据 | S005-图谱访问组件 | 查看图谱 |

* + 1. 组件时序设计

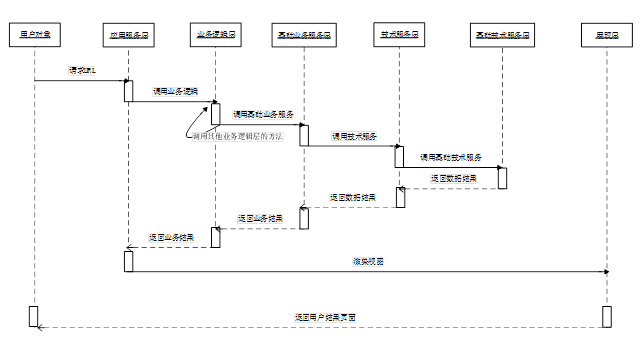


图 组件时序设计图

组件时序设计图的执行步骤如下：

1、用户通过浏览器向服务器发出请求

2、服务器将请求转发给应用服务层，服务层根据客户的请求方法名和参数，调用相应的业务逻辑层

3、业务逻辑层根据传过来的参数进行处理，调用相应的技术服务层方法

4、技术服务层返回给业务逻辑层处理结果

5、业务逻辑层将最终的处理结果返回给应用服务层

6、应用服务层将业务逻辑层返回的结果返回给展示层

7、展示层通过渲染视图将结果返回给用户

8、用户看到想要的结果页面

* + 1. 功能组件设计

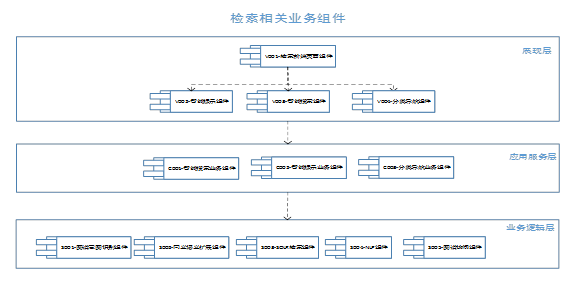


图 功能组件图

功能组件清单表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **功能组件编号-名称** | **功能编号-名称** | **系统逻辑分层** | **包含的组件编号-名称** | **是否接口组件** | **非功能性要求** |
| FUN15-信息检索管理 | FUN1501-信息检索管理 | 展现层 | V002-智能提示组件 | 否 | 易操作性 |
| V003-智能搜索组件 | 否 |
| V004-分类导航组件 | 否 |
| 应用服务层 | C002-智能提示业务组件 | 是 | 容错性  并发性  稳定性 |
| C001-智能搜索业务组件 | 是 |
| C003-分类导航业务组件 | 是 |
| 业务逻辑层 | S001-图谱意图识别组件 | 否 | 无 |
| S003-Solr检索组件 | 否 | 无 |
| S004-NLP组件 | 否 | 无 |
| S005-图谱访问组件 | 否 | 无 |

* + 1. 接口组件设计



图 接口组件图

接口组件清单

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 分类 | 接口组件编号-名称 | 职责描述 | 来源 | 系统逻辑分层 |
| 本系统接口组件 | C001-智能搜索业务组件 | 与知识图谱和Solr交互，提供商品、店铺搜索等服务 | 外部系统接口 | 技术服务层 |
| C002-智能提示业务组件 | 与知识图谱和Solr交互，提供用户搜索关键词联想词等服务 | 外部系统接口 | 技术服务层 |
| C003-分类导航业务组件 | 提供商品过滤筛选 | 外部系统接口 | 技术服务层 |

* + 1. 公共组件设计

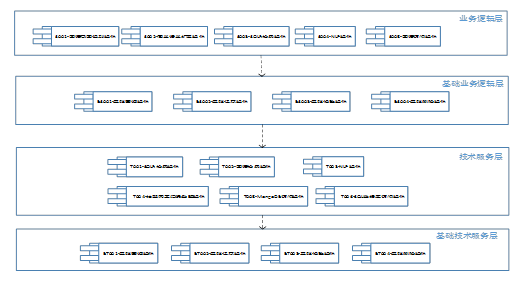


图 公共组件图

公共组件清单

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 分类 | 公共组件编号-名称 | 职责描述 | 来源 | 系统逻辑分层 |
| 本系统重用功能组件 | S001-图谱意图识别组件 | 负责用户搜索意图识别 | 外部系统接口 | 业务逻辑层 |
| S002-同义语义扩展组件 | 负责关键词同义词扩展 | 外部系统接口 | 业务逻辑层 |
| S003-SOLR检索组件 | 负责资源检索 | 外部系统接口 | 业务逻辑层 |
| S004-NLP组件 | 负责文本抽取、分词等NLP功能 | 外部系统接口 | 业务逻辑层 |
| S005-图谱访问组件 | 负责访问知识图谱 | 外部系统接口 | 业务逻辑层 |

* 1. 系统界面设计
     1. 界面框架设计

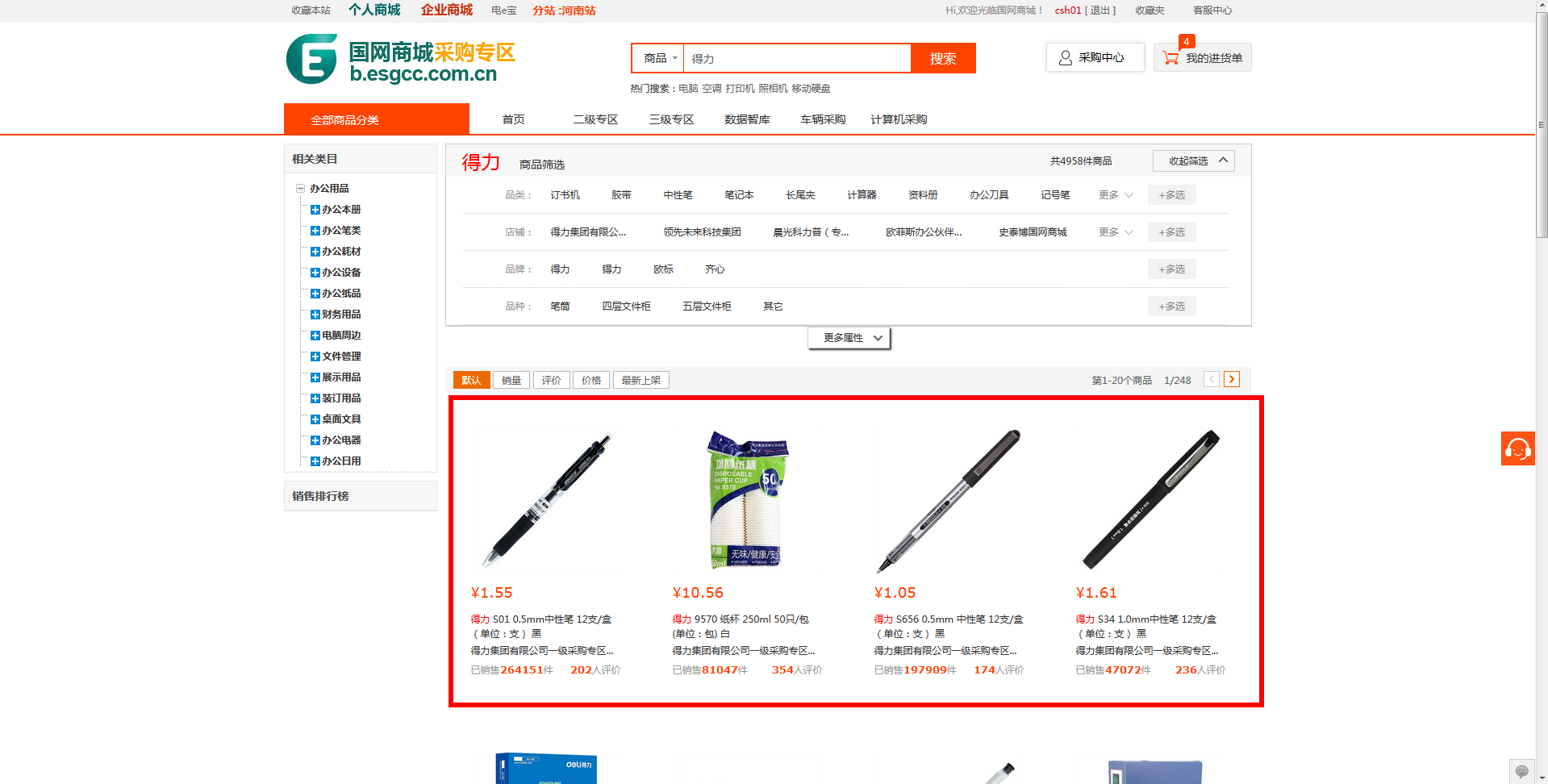
系统登录页设计：



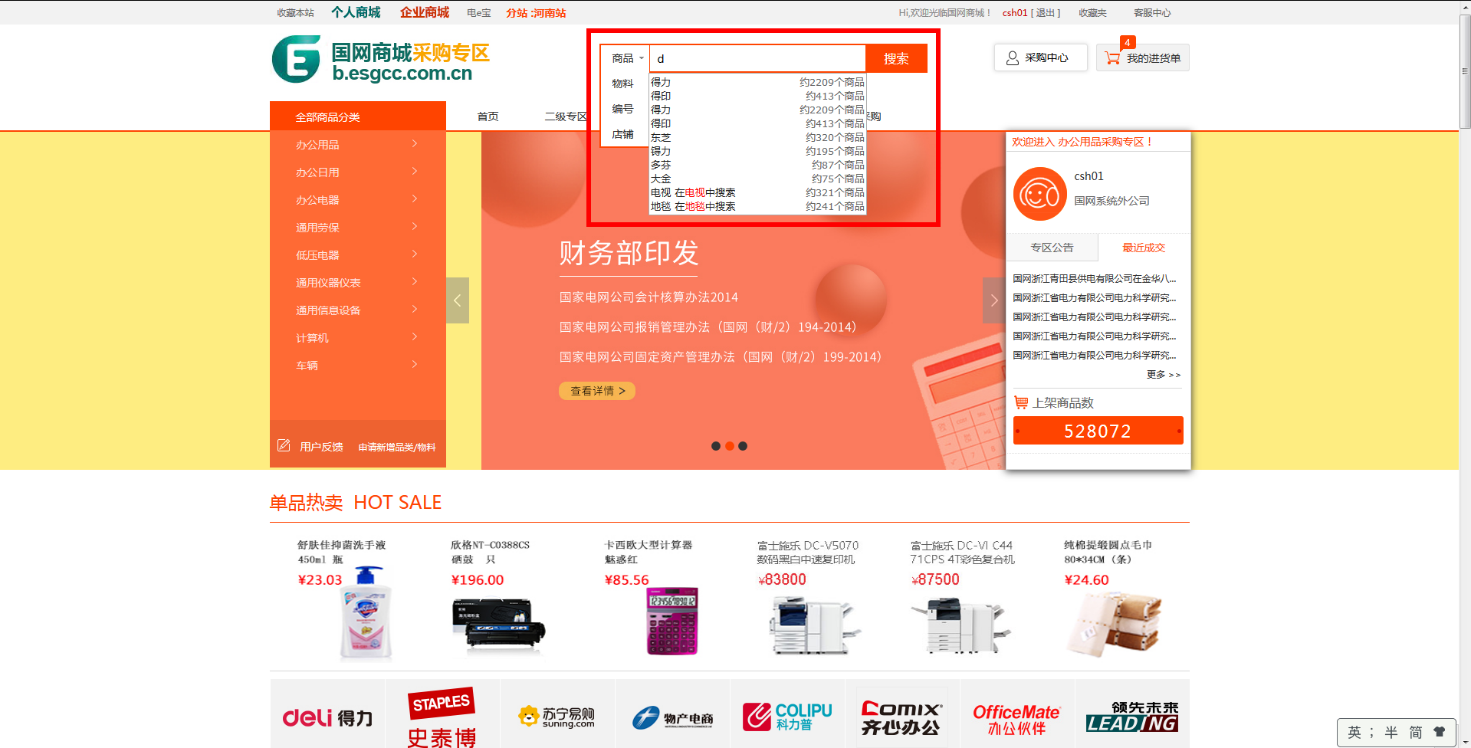
采购专区首页：



表格控件：



下拉控件：



* + 1. 系统空控件选择

系统控件清单

|  |  |  |  |
| --- | --- | --- | --- |
| **业务场景** | **功能页面名称** | **选择控件** | **控件用途** |
| 商品/店铺搜索 | 首页及搜索列表页 | 下拉控件 | 通过下拉控件，以列表的形式展示商品提示、商品数量、意图识别提示。 |
| 表格控件 | 通过表格控件，以列表形式展示单据序号、单据状态、单据号单据类型、提交时间、单据金额、单据事由。其中单据状态、单据类型支持筛选，提交时间、单据金额支持排序功能。 |

1. 系统运行环境设计
   1. 硬件环境

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **设备名称** | **硬件类型** | **用途** | **配置** | **数量** |
| 知识图谱服务器 | 虚拟机 | 用于存储商品-分类语义三元组数据，并提供搜索语义等服务 | CPU:4C  内存：16G以上  硬盘：300GB  网卡：1GB\*2  操作系统：REDHAT LINUX 6.5 | 2 |
| 检索索引服务器 | 虚拟机 | 用于存储并索引商品、店铺等数据，并提供搜索服务 | CPU:4C  内存：16G以上  硬盘：300GB  网卡：1GB\*2  操作系统：REDHAT LINUX 6.5 | 2 |
| WEB服务器 | 虚拟机 | 用于商品、店铺搜索提示及信息检索服务 | CPU:4C  内存：16G以上  硬盘：300GB  网卡：1GB\*2  操作系统：REDHAT LINUX 6.5 | 2 |
| 日志服务器 | 虚拟机 | 用于收集搜索、数据同步等日志数据 | CPU:4C  内存：16G以上  硬盘：300GB  网卡：1GB\*2  操作系统：REDHAT LINUX 6.5 | 1 |

* 1. 软件环境

|  |  |  |  |
| --- | --- | --- | --- |
| **设备名称** | **软件名称** | **说明** | **版本** |
| 知识图谱服务器 | Tokumx | 存储语义三元组数据 | 2.0.2 |
| 知识图谱服务器 | Redis | 数据缓存 | 3.2.8 |
| WEB服务器 | Tomcat | 应用访问及搜索 | 8 |
| WEB服务器 | Nginx | 应用服务反向代理 | 1.10 |
| 知识图谱服务器 | Mysql | 存储图谱配置、同义词等信息 | 5.6 |
| 检索索引服务器 | ElasticSearch | 存储和索引商品、店铺等数据 | 5.5 |
| 日志服务器 | ElasticSearch | 存储和搜索日志数据 | 5.5 |
| 知识图谱服务器 | Tokumx | 存储语义三元组数据 | 2.0.2 |