版本：1.0.0

# 1、功能：

# 2、程序模块划分：

* Eureka注册中心
* Zuul网关
* 数据字典管理模块
* 索引管理模块
* 数据同步模块
* 搜索模块
* 用户权限模块
* 中文navicat模块
* 数据统计模块

# 3、各模块功能

## 3.1 Eureka注册中心

注册中心。各模块将其服务注册到上面，供其他模块发现及调用

## 3.2 网关服务

… …

## 3.3 数据字典管理模块

管理库、表、字段的中英对照信息

## 3.4 索引管理模块

管理索引的创建、修改、删除

## 3.5 数据同步模块

所有创建索引的表均在数据同步队列中，定期对数据进行同步。

## 3.6 搜索模块

管理数据的查询、查询结果的展示

## 3.7 用户权限模块

控制数据字典管理模块和索引管理模块

## 3.8 中文navicat模块

利用数据字典管理模块的信息，呈现一版中文版的数据库信息。

## 3.9 数据统计模块

统计每个库表的记录数、更新时间等相关数据。

# 4. 各模块详述

## 4.1 Eureka注册中心

## 4.2 服务网关

## 4.3 数据字典管理模块

数据字典管理模块包括库信息管理、表信息管理、字段信息管理、外键信息管理。

### 4.3.1 库信息管理

dict\_database表的表结构如下



#### 4.3.1.1 库信息增加

库信息增加有两种方式：

* 手动输入

首先选择数据库（必须是数据库中存在的），然后填入中文库名和库详情即可。

* Excel导入

表头为英文库名、中文库名、库详情。后台解析Excel并作校验，校验无误后入库，校验规则（同时满足下面所有条件）：

（1）英文库名必须为数据库中存在的库

（2）必须为dict\_database中不存在的库

（3）中文库名不得为空

（4）数据不能出现重复

#### 4.3.1.2 库信息修改

可修改库的中文名、库详情，不许修改英文名。

#### 4.3.1.3 库信息删除

库信息删除是一个需要非常谨慎的操作。

若dict\_table中存在属于该库的表，那么该库信息不能被删除。

### 4.3.2 表信息管理

dict\_table的表结构如下



#### 4.3.2.1 表信息增加

表信息增加有两种方式：

* 手动输入

通过下拉方式选择数据库后，填入英文表名（支持模糊匹配）、中文表名，提交即可。

* Excel导入

表头为英文表名、中文表名。后台解析Excel并作校验，校验无误后入库，校验规则（同时满足下面所有条件）：

（1）英文表名必须是当前库下的表

（2）dict\_table中不存在该表

（3）中文表名不得为空

（4）数据不能出现重复

#### 4.3.2.2 表信息修改

可修改英文表名、中文表名，英文表名的修改会触发修改数据库的表名。

#### 4.3.2.3 表信息删除

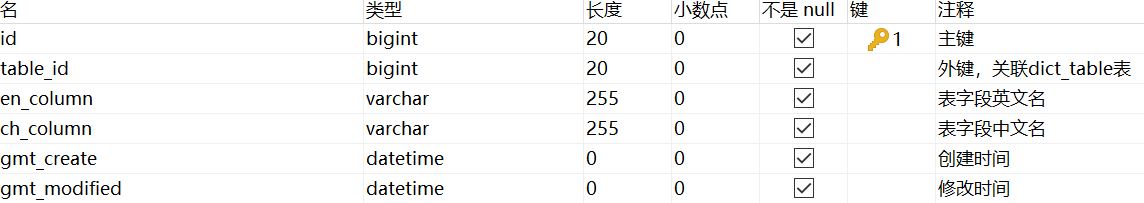
表信息删除代表该表无用了，那么删除该表信息意味着要删除业务系统中所有与该表有关的信息。

为了保证数据字典模块与索引管理模块的尽可能解耦合，同时保证业务的一致性。该步骤若想删除表信息，前提为删除搜索引擎表及搜索引擎中与该表相关的所有信息。

该表信息的删除还会级联删除该表的字段信息与外键信息。

### 4.3.3 表字段信息管理

dict\_column表的表结构如下



#### 4.3.3.1 字段信息增加

字段信息增加有两种方式：

* 手动导入

选择数据库及数据表后，系统加载该表的所有英文字段，用户填写中文字段名，删去无用的字段。

* Excel导入

1. 以库为单位导入某库多个表的字段

表头为英文表名、英文字段名、中文字段名。后台解析Excel并作校验，校验无误后入库，校验规则（同时满足下面所有条件）：

（1）表必须是当前库的表。

（2）所有的英文字段必须在表中真实存在。

（3）这些字段中的任何一个在dict\_column中不存在。

（4）中文字段名均不为空。

（5）数据不能重复。

2. 以表为单位导入当前表的字段

表头为英文字段名、中文字段名。后台解析Excel并作校验，校验无误后入库，校验规则（同时满足下面所有条件）：

（1）所有的英文字段必须在表中真实存在。

（2）这些字段在dict\_column中不存在。

（3）中文字段名不能为空。

（4）数据不能重复。

#### 4.3.3.2 字段信息修改

可修改字段英文名、中文名。其中英文字段名的修改会触发修改表的字段名。

#### 4.3.3.3 字段信息删除

若搜索引擎中存在该表，则该表的字段不支持删除。否则可以删除。

## 4.4 索引管理模块

负责大索引的创建，新表的加入。

### 4.4.1 大索引创建

因为每个数据表加到搜索引擎的某个索引中，所以将表加入搜索引擎之前需要现指定该表所对应的索引名称。若一个索引都没有，则需要创建索引。

创建索引在前端仅需指定索引的名称、索引的主分片数即可，后端完成相关元数据字段的添加，如（elastic\_table\_name, elastic\_database\_name等等）。大索引只有超级管理员可以操作，暂不支持删除、修改操作。

### 4.4.2 表管理

#### 4.4.2.1 新表加入

新表加入涉及5个业务表的操作。这5个插入操作应该在一个事务下完成。

对于一个新表，首先选择其对应的索引。然后，系统展示出该表的中英字段信息，并略去空字段。用户需要对每个字段选择“是否展示”、“是否搜索”、“是否分词”，同时填写该字段的权重。这三个选项是递进的关系，即“需要展示”才涉及到“是否搜索、分词”，“需要搜索”才涉及到“是否分词”。

前端提交的表单格式为

{

“databaseName”: “zfsj\_ajj”,

“tableName”: “tb\_test”,

“indexName”: “elasticIndexOne”,

“columns”: [

{

“enColumn”: “name”,

“displayed”: “true”,

“searched”: “false”,

“analyzed”: “fallse”

}

]

}

后端接到数据后，执行以下操作：

* es\_column表：根据字段的三种类型（搜索不分词、搜索且分词、不搜索）尝试将columns中的enColumn映射到上面三种类型的字段中，若字段不够，添加新的字段。
* mapping\_column表：因为es\_column已经创建了足够多的字段，所以这里将columns中的enColumn的各个字段与es\_column中的字段作映射。将映射结果填入column\_mapping表中。
* mapping\_table表：填入该表及其对应的索引名即可。
* sync\_table表：将该表加入同步队列。

以上四步操作在一个事务中执行。

#### 4.4.2.1 表修改

（1）字段修改

* 字段名的修改不在此进行，数据字典管理模块中的字段修改会触发更新这里的字段名
* 可更改字段的权重（只有被搜索的字段才能改）、是否展示（只有不被搜索的字段才能更改）。不可更改是否搜索、是否分词两个属性。
* 不支持字段删除
* 不支持字段添加

（2）表修改

* 表名在字典管理模块修改。因为为外键关联，所以不必修改这里。
* 不支持修改表所属的索引。

（3）表删除

表删除会触发删除mapping\_column、mapping\_table、sync\_table中的所有相关数据。并删除搜索引擎中该表的相关数据。

## 4.5 数据同步模块

数据同步模块定期扫描sync\_table表，表结构如下



查询es中的数据条数修改index\_records字段；查询数据库中表的记录条数，若table\_records与数据库中查得的该表的实际记录不相符时，证明该表发生了更新，由此启动该表的全量导入流程。

全量导入流程如下：

* 发送需要同步的库名和表名到消息队列kafka，
* **全量读入子模块**解析kafka消息队列中的信息，获取到库名、表名，执行全量读入操作，对读出的每条记录作全字段哈希（生成“hash\_id”字段作为ES文档的id），将数据投递至kafka。{“databaseId”:”1”, ”tableId”:”2”, “data”:{}}

读入完毕后更新sync\_table表，将table\_records字段改为读取的记录条数。

* **过滤和转换子模块**解析kafka消息，对于每条记录，查询redis，将表字段转化为ES中对应的字段，添加“表名”和“库名”两个字段，取值为其对应的主键id。删除没有对照信息的字段。将数据发至批量插入子模块，数据结构为{“index”:”zfsj\_ajj”, data : {“id”:”1”}}
* **批量插入子模块**

使用Elasticsearch的批量插入模块，将数据写入Elasticsearch,记录下出错信息。

## 4.6 搜索模块

搜索分为两部分。首页的全库聚合和以及详情页的单表查询。

### 4.6.1 全库聚合

全库聚合对搜索词进行全索引搜索，并将搜索结果按照库、表二级聚合。

查询串构造与之前基本相同。

返回结果需要对库名、表名作中英映射。因为搜索引擎中存的是库的主键、表的主键，所以直接根据主键从redis中查询即可。

### 4.6.2 单表查询

单表查询需要传入表的id，根据表id过滤出需要该表的数据。

从redis中根据table\_id查询出mapping\_column的相关数据，将字段转换成中文字段，然后返给前端即可。

## 4.7 用户权限模块

暂不添加

## 4.8 中文navicat模块

实现一个中文版的数据库视图，数据字典来自dict\_database, dict\_table, dict\_column。

## 4.9 数据统计模块

数据统计包含库信息统计、表信息统计。

### 4.9.1 表信息统计

表信息统计表结构如下（stat\_table）



每天扫描stat\_table下各个表。对于每个表，去数据库里使用select count(\*)查得该表的真实数据条数，若该值不等于total\_records，则认为该表发生了更新，那么更新total\_records的值，并更新update\_date为当前时间。

定期扫描dict\_table表，保证dict\_table中的表与stat\_table中的表一致，有新增的要及时加入，有删除的要及时删除。

### 4.9.2 库信息统计

库信息统计表结构如下（stat\_database）



每天扫描stat\_database下各个库。对于每个库，从dict\_table中算出该库对应的表的数目，查询该库的所有表的记录和，查询各表的最近更新时间作为该库的更新时间。

定期扫描dict\_database下各个库，保证dict\_database与stat\_database数据的统一。dict\_database表中的增删改查及时更新到stat\_database.