

Esearch布署手册

zxb123

Version 1.0, 2016-02-21

Table of Contents

1. 获取程序包	2
1.1. svn库上下载源码并编译	2
1.2. 从开发人员处获取	2
2. 环境准备	3
2.1. jdk准备	3
2.2. tomcat准备	3
2.3. 初始化sql	3
3. 部署程序包	4
4. 索引抽取	5
4.1. 添加es连接节点	5
4.2. 添加索引表	6
4.3. 配置索引字段信息	7
4.4. 创建索引	9
4.4.1. 指定创建索引sql	9
4.4.2. 直接创建索引	10
4.4.3. 加入队列后创建索引	10
5. 检索测试	12
6. 一键搜索	13
6.1. 配置查询方案	13
6.2. 一键搜索页面	15
6.2.1. 直接搜索	15
6.2.2. 资源展示顺序	16
6.2.3. 照片展示	17
6.2.4. 二次查询	18
6.2.5. 筛选查询	21
6.2.6. 搜索小提示	21
6.2.7. 其它定制化项	22
7. 定时同步索引	23

ESearch工程为elasticsearch插件的java客户端，通过本工程可完成以下操作。

- 将数据库表数据导入到elasticsearch中
- 基于已有的elasticsearch提供全文检索功能
- 管理现有的elasticsearch索引资源



布署本工程前，请先确保已经安装了 [elasticsearch](#) 插件。
elasticsearch安装请参考相关文档。本工程目前对应的elasticsearch版本为 **version1.4.4**，请安装对应版本的elasticsearch插件。



本工程部分功能依赖于metadata及wmf，请确保线上已经布署了这两个工程。

Chapter 1. 获取程序包

最新版本的esearch2程序源码在svn库上
[CODE/product/dmga-esearch](http://svn.dameng.com:2688/code/DM-GAHYB-CODE/product/dmga-esearch)

[http://svn.dameng.com:2688/code/DM-GAHYB-](http://svn.dameng.com:2688/code/DM-GAHYB-CODE/product/dmga-esearch)

1.1. svn库上下载源码并编译

1. 从上面的 [svn](#) 地址下载源代码。
2. 导入eclipse
3. 编译并使用maven打成war包

1.2. 从开发人员处获取

- 联系开发人员获取该工程war包

Chapter 2. 环境准备

在进行该工程部署前，需要在服务器上做好以下准备工作。

2.1. jdk准备

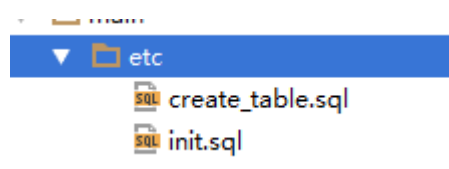
请自行百度或点此 [参考](#)

2.2. tomcat准备

请自行百度或点此 [参考](#)

2.3. 初始化sql

在文档目录下找到/etc目录



1. 执行create_table.sql
2. 执行init.sql

Chapter 3. 部署程序包

1. war包部署

如果是war包，则直接将war包拷贝到tomcat根目录的webapps下，然后启动tomcat，war包会自动解压。

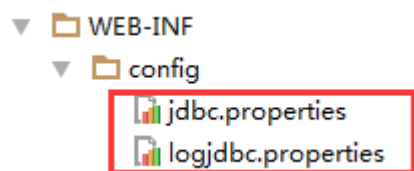
2. 非war包方式部署

非war包方式，需要在conf/server.xml中手动配置程序包所有位置。

配置示例

```
<Context path="/eSearch2" docBase="C:\zxbProgramFiles\workdir\project\code\DM-  
GAHYB-CODE\product\dmga-eSearch\1.0.1\trunk\src\main\webapp" >  
</Context>
```

配置完成后需要手动修改管理库连接地址及密码等



配置文件位置

Chapter 4. 索引抽取

索引抽取在这里指的是从数据库表中将数据抽取到全文库中（即插件处）。
程序实现了索引的全量抽取与增量抽取以及简单的清洗转换功能(数据最好ETL提前清洗好)。

4.1. 添加es连接节点



进入索引管理界面



索引节点管理界面



说明

名称	解释
节点名称	指定一个节点名称
节点ip	如 10.75.198.114:9300
集群名称	必须如插件所配置的 <code>cluster.name</code> 保持一致
是否集群节点	建议默认选择 是
状态	表示是否启用该节点

可以先点击右上角的 **测试** 按钮进行连接测试。

4.2. 添加索引表

1. 进入索引表管理界面，点击 **新增** 按钮

索引表管理							
 显示查询  新增  删除  信息同步  设置同步信息  加入队列							
<input type="checkbox"/>	相关操作				表名称	表别名	
<input type="checkbox"/>	索引字段配置	修改表	删除表	删除索引	同步索引	测试查询	加
<input type="checkbox"/>	索引字段配置	修改表	删除表	删除索引	同步索引	测试查询	加
<input type="checkbox"/>	索引字段配置	修改表	测试查询	自定义sql	RYS_ZDF	重点人员电子	

进入索引表管理界面

2. 选择数据表所属数据源

数据源信息							
 增加  删除选中  刷新							
相关操作	数据源类型	数据源名称	数据源分类	主机名称IP	端口	状态	备注
详细 修改 删除 数据集	DM7	共享库	综合库	10.75.198.93	5236	正常	
详细 修改 删除 数据集	DM7	缓冲库	缓冲库	10.75.198.92	5236	正常	
详细 修改 删除 数据集	DM7	历史库	缓冲库	10.75.198.92	5236	正常	
详细 修改 删除 数据集	Oracle	人口库Oracle	综合库	10.75.198.203	1521	正常	
详细 修改 删除 数据集	DM7	专题库	专题库	10.75.198.92	5236	正常	
详细 修改 删除 数据集	DM7	字典库	综合库	10.75.198.93	5236	正常	
详细 修改 删除 数据集	DM7	综合基础资源库	综合库	10.75.198.93	5236	正常	

进入数据源选择界面

3. 进入数据集选择界面

数据集 综合基础资源库(DM7)					
 显示查询  添加到索引管理表  采集元数据  同步数据列元数据  删除选中  刷新					
<input type="checkbox"/>	模式名	数据集名	注释	描述	关联的资源
<input type="checkbox"/>	WGHXX	RYS_WGH_RYJBXX	网格化人员基本信息		孝感网格化人口基本信息
<input type="checkbox"/>	ZHJCZYK	JGYS_CYFWDW	餐饮服务单位信息		
<input type="checkbox"/>	ZHJCZYK	JGYS_KDZDXX	孝感快递站点信息		孝感孝感快递点信息
<input type="checkbox"/>	ZHJCZYK	JGYS_LXSXX	旅行社信息		孝感全省旅行社信息
<input type="checkbox"/>	ZHJCZYK	JGYS_WLGSXX	孝感物流企业信息		孝感物流信息
<input type="checkbox"/>	ZHJCZYK	RYS_DW_JYXTDTXNLR	孝感教育系统到龄退休教		孝感到龄退休教职工信息
<input type="checkbox"/>	ZHJCZYK	RYS_GZDX_YW_JBXX	工作对象基本信息		
<input type="checkbox"/>	ZHJCZYK	RYS_RYDHXX	运营商客户信息		孝感运营商客户信息
<input type="checkbox"/>	ZHJCZYK	RYS_XGXSBJXX	孝感学生报考信息		孝感学生报考信息
<input type="checkbox"/>	ZHJCZYK	SJYS_XG_TBW	淘宝网订单信息		
<input type="checkbox"/>	ZHJCZYK	SJYS_YM_YCTYGRXX	房产权属登记信息		
<input type="checkbox"/>	ZHJCZYK	WPYS_XG_FCJXX	孝感房管局房产信息		

进入数据集选择界面

选择要添加的数据集，点击 **添加到索引管理表** 即可。

4.3. 配置索引字段信息



在添加索引表后，需要配置索引字段信息后，方可进行索引的创建

1. 在索引表管理界面，点击 索引字段配置

<div>显示查询</div> <div>新增</div> <div>删除</div> <div>信息同步</div> <div>设置同步信息</div> <div>加入队列</div> <div>移出队列</div> <div>查看队列</div> <div>清空缓存</div> <div>手动删除索引</div> <div>中断创建</div> <div>刷新</div>															
<input type="checkbox"/>	相关操作				表名称	表别名	表状态	索引名称	类型名称	更新方式	总数据量	追加数据量	更新数据量	排序号	创建时间
<input type="checkbox"/>	索引字段配置 修改表 删除表 创建索引 加入队列 自定义sql				JGYS_C1	餐饮服务单位	未创建	zhjczyk	jgys_cyfw	自动	0	0	0	999	2016-03-07 16:05

索引字段配置

2. 进入索引字段配置后，可进行字段的配置

索引字段配置													
保存 同步索引 新增索引 刷新 关闭 改后，建议点击清空缓存按钮！													
操作	列名	列注释	数据类型	属性设置	日期格式	权重值	索引分词	搜索分词	是否索引	数据字典	排序号	创建时间	更新时间
删除	BIRTHDAY	出生日期	VARCHAR2			1			分词索引	设置 清除	999	2016-03-07 16:05:55	2016-03-07 16:06:11
删除	CARDNO	身份证号	VARCHAR2	身份证号二次查		1			分词索引	设置 清除	999	2016-03-07 16:05:55	2016-03-07 16:06:11
删除	CARDTYPEID	证件类型	VARCHAR2			1			分词索引	网格化证件类型 设置 清除	999	2016-03-07 16:05:55	2016-03-07 16:06:11
删除	CERTIFICATENUM	证件号码	VARCHAR2			1			分词索引	设置 清除	999	2016-03-07 16:05:55	2016-03-07 16:06:11

索引列编辑界面

配置说明

在上面的配置项中，可编辑的配置项依次为。

列注释

列注释是确保在没有数据列注释的情况下，可以手动添加该列的注释信息。

属性设置

属性设置指配置该列的属性信息。目前包括如下可配置的属性字段。

1. asc:升序排序

结果展示时将以该列进行升序排序

2. desc:降序排序

结果展示时将以该列进行降序排序

3. csrq:标识该字段为出生日期字段

备用属性

4. date:标识该列为日期列

标识该字段为日期列，配置该项后，可提供类似 张三 25 ,张三 1990 这种搜索。

5. gxsj:更新时间

指定该字段为更新时间，同步索引 时将根据该字段进行索引更新。

6. rksj:入库时间

指定该字段为更新时间，**同步索引**时将根据该字段进行索引追加。

7. is_notAnalyzed:不分词处理

不分词处理，配置该项后，可支持 **张? 丰**，**张三*** 这种搜索。
不建议多数字段配该项，会影响性能。对关键字段如XM可以配置该项。

8. is_store:存储索引

保留属性

9. key:关键列

云搜索中使用

10. not_query:非查询列

指定该列在数据检索时，不检索该列内容。

11. not_result:非结果列

指定该列在数据检索完成后，不获取该列内容。

12. not_show:不显示

指定该列在数据展示页面不显示。

13. pk:主键列

指定该列为主键

14. second_query:二次查询列

指定该列为二次查询列（**一键搜索**中使用），配置该项后，在详细页面，点击该列可进行二次查询。

15. sfzh:身份证号列

强烈建议配置此项

一键搜索中照片展示会使用该项信息。

日期格式

配置数据的日期格式

权重值

权重，指在数据检索过程中，该列权重值越高，如果该列有匹配内容，对应的那行数据将会显示越靠前。

强烈建议配置此项，可选值为1~10，值越高表示权重值越大，匹配后越靠前，默认值为1

该项只有在索引创建前配置方生效，如果索引已经创建后再配置此项将不生效。

索引分词

当使用到了分词器时需要配置此项，如ik。未使用分词器时，配置此项将会报错。默认不配置。

搜索分词

当使用到了分词器时需要配置此项，如ik。未使用分词器时，配置此项将会报错。默认不配置。

是否索引

建议默认 **分词索引**，当该列不需要建索引时选择 **不索引**

数据字典

指定后可以 **在创建索引过程中** 对该列进行字典翻译。

排序号

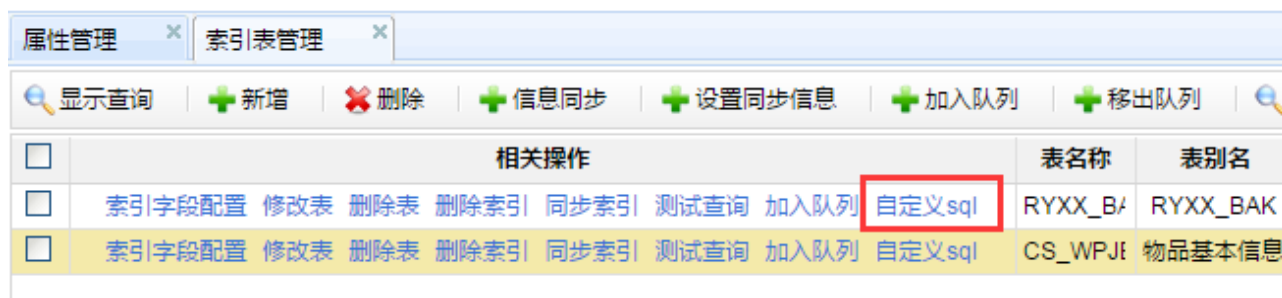
指定后将控制在 **一键搜索** 中列的显示顺序。

4.4. 创建索引

创建索引指的是将数据库表数据抽取到全文资源库。

4.4.1. 指定创建索引sql

1. 选择自定义sql



自定义sql

2. 指定sql语句

自定义SQL

保存

重置

关闭

删除

创建索引SQL :

更新索引SQL :

追加索引SQL :

指定sql语句

创建索引sql

指定用来创建索引的sql语句

更新索引sql

指定同步索引时用来更新索引的sql语句

追加索引sql

指定同步索引时用来追加新索引数据的sql语句



指定索引sql时，上面三者可指定其中一到多个

4.4.2. 直接创建索引

当指定或者没有指定创建索引sql语句后，点击创建索引即可开始索引的创建。



如果没有手动指定创建索引sql，则会以下面的sql默认创建索引。`select * from schema.table`

当需要同时创建多个索引时，可以在多个索引表上点击创建索引。



注意，此时是并行创建，不推荐使用这种方式同时创建多索引。可查看

[加入队列](#)

4.4.3. 加入队列后创建索引

当需要为多个索引表创建索引时，可将待创建的索引表 [加入队列](#)，然后在 [查看队列](#) 中点击 [创建索引](#) 完成索引的创建。队列中的索引表创建任务是串行的，即按顺序依次创建。

1. 将待创建索引表加入队列

属性管理		索引表管理	
显示查询		新增	删除
信息同步		设置同步信息	加入队列
移出			
相关操作			表名称
索引字段配置 修改表 删除表 删除索引 同步索引 测试查询 加入队列 自定义sql			RYXX_BA
索引字段配置 修改表 删除表 删除索引 同步索引 测试查询 加入队列 自定义sql			CS_WPJ

2. 点击查看队列

设置同步信息		加入队列	移出队列	查看队列
		表名称	表别名	表状态
测试查询 加入队列 自定义sql		CS_WPJ	物品基本信息	创建完成
		RYXX_BA	RYXX_BAK	队列中

3. 队列任务中点击创建索引

查看队列		
创建索引 同步索引 停止任务 清空队列 关闭窗口		
序号	任务名称	操作
1	RYXX_BAK	移出

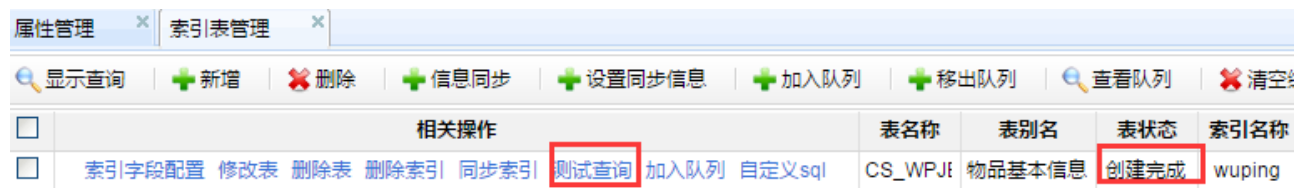


1. 队列中 正在被创建的索引任务 不可以移出队列，而后面 尚未开始创建的索引任务可以被移出队列。
2. 队列任务执行过程中，可随时添加新的索引任务。

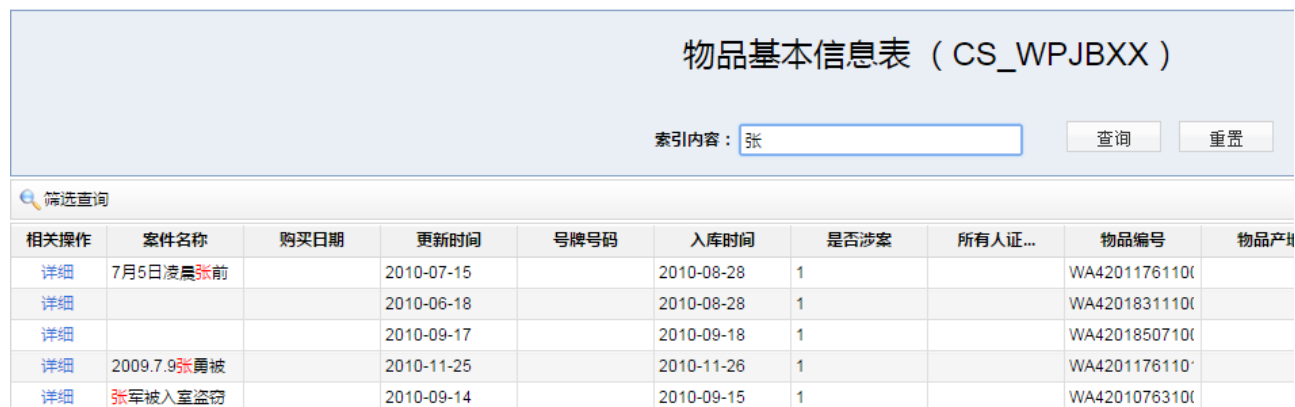
Chapter 5. 检索测试

当索引创建完成后，可以直接在这个管理界面，点击 **测试查询** 进行查询测试。

1. 进入测试查询界面



2. 点击测试查询



Chapter 6. 一键搜索

一键搜索是基于已经整合的索引表进行综合检索的功能。见下图：



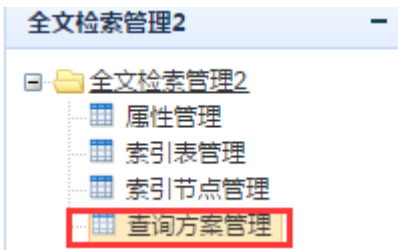
完成上述功能需要完成以下几步



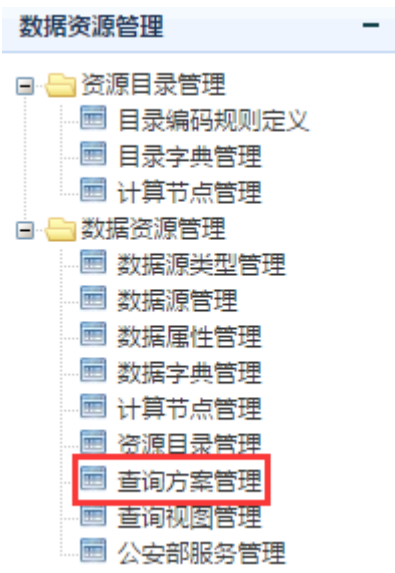
- 索引表对应的数据集进行资源编目
未编目的资源不会显示在一键搜索中
- 后台管理中配置查询方案
- 挂接索引表到查询方案对应的分类下

6.1. 配置查询方案

- 点击查询方案



此处的查询方案同 数据资源>查询方案 相同，只是挂了个链接。



- 2. 手动添加 查询方案，进入查询方案后，手动添加 分类
- 3. 依次点击 设置分类，设置索引



- 4. 将左侧的资源添加至右侧即可





- 左侧的资源为索引表中的数据，且必须为已经_编目_过的。
- 资源名称分为三种颜色。
 - 红色 红色表示该资源未创建索引。
 - 紫色 紫色表示该资源被其它分类已经挂接过。
 - 绿色 绿色表示该资源已创建索引且未被其它分类挂接过。

上述步骤完成后，整个一键搜索即配置完成。



为提高检索速度，分类信息均缓存。配置成功后，请点击 **清除缓存**

6.2. 一键搜索页面

上述工作完成后，可按以下地址进入一键搜索界面。

http://host:port/esearch2/elasticsearch/DefaultQuery_init.do?scheme_id=scheme_id

host

你部署esearch2的机器ip

port

你的tomcat的端口号

scheme_id

你刚才配的查询方案的ID（点击查询方案详细可看到）

示例

http://localhost:8080/esearch2/elasticsearch/DefaultQuery_init.do?scheme_id=ejc

6.2.1. 直接搜索

主关键字框中输入关键字，点击搜索直接开始搜索。

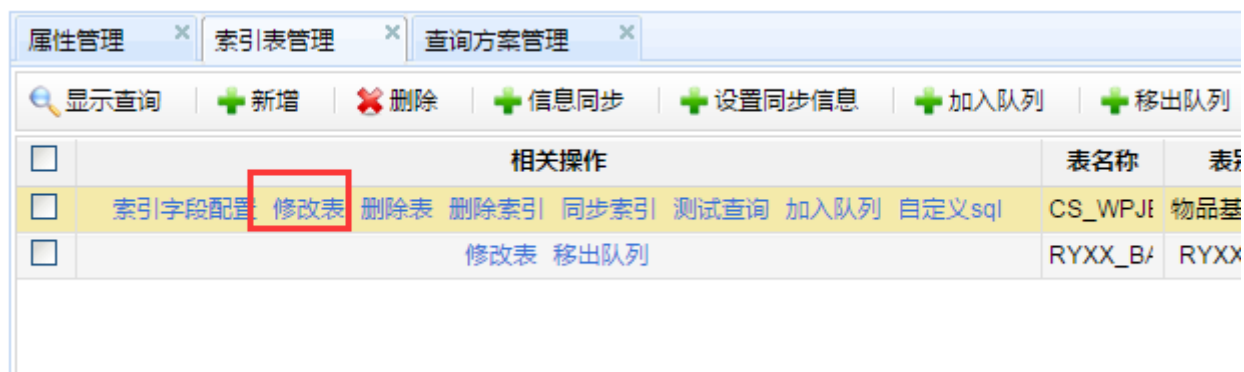


6.2.2. 资源展示顺序

左侧资源展示顺序，可在后台管理中进行配置。



1. 点击修改表



2. 修改排序号并保存

修改索引表

节点名称：

localhost

(不超过32字)

索引名称：

wuping

(不超过32字)

索引类型：

cs_wpjbx

(不超过64字)

*状态：

创建完成

*更新类型：

自动

*创建索引Handler：

InsertHandlerDefaultImpl

*mapping设置：

生成mapping

数据集ID：

dataset_3706BE0C3519F7EE1DFD

(不可编辑)

*排序号：

2

6.2.3. 照片展示

人物照片展示会在如下两个位置进行展示。

操作	相片	更新时间	婚姻状况	主键ID	籍贯	民族	入库时间	证件号码	学历	性别	姓名
加入比对 详细 户成员			未婚	0f37394e-97ee-4dca-9a86-eeb675b9	湖北省黄梅县	汉族		42213019	文盲或半文盲	女	张三桃
加入比对 详细 户成员			已婚	d8bd-4154-9734-	湖北省浠水县	汉族		42112519	文盲或半文盲	女	张三芳

数据列表处展示

查看详细信息

更新时间	undefined
婚姻状况	未婚
主键ID	0f37394e-97ee-4dca
籍贯	湖北省黄梅县
民族	汉族
入库时间	undefined
相片	
证件号码	42213019420217562

数据详细页面展示

如上所示，图片均挂了，图片展示功能不属于一键搜索的范围。需要其它系统提供一个获取照片的接口。例如：

```
http://localhost:8080/esearch2/elasticsearch/GetImg_getImg.do?sfzh=422130194202175629
```

需要添加照片展示功能，需要具备如下条件。

- 1. 具备一个照片获取接口，可根据身份证号获取的。
- 2. 该索引表已经指定了身份证号列。



字段配置

保存

+ 同步索引列

+ 新增索引列

刷新

关闭

操作	列名	列注释	数据类型	属性设置	日期格式
删除	MZ	民族	VARCHAR		
删除	RKSJ	入库时间	TIMESTAMP	非查询列,非结果	
删除	SFZH	证件号码	VARCHAR	身份证号	

6.2.4. 二次查询

二次查询，在此处是指通过在详细页面点击某列信息，然后基于该列值再次进行检索。

功能介绍

- 1. 点击详细页面中的二次查询列后，将以该列值进行二次检索。

查看详细信息

姓名	张雄彪	<div>关联资源信息</div> <div>孝感车辆违法信息 (1)</div> <div>孝感户籍变动信息 (1)</div> <div>孝感驾驶员从业资格信息 (1)</div>
----	-----	--

性别	男
出生日期	1968-10-27 12:00:00
相片	户成员
身份证号	422228196810274076
民族	汉
籍贯	湖北省汉川市
婚姻状况	已婚
人员属性	
户编号	7332008
身高	160
文化程度	初中
所属派出所名	汉川市公安局杨林水陆派出所
服务处所	市杨林沟镇豪富村
地址	湖北省汉川市杨林沟镇豪富村71号

2. 点击检索后的结果行，将跳转到对应资源页面

孝感车辆违法信息

比对

查看比对

违法行为

相似

+

×

查询

重置

操作	违法行为	违法时间	违法地址	违法地点	违法编号	当事人	号牌号码	号牌种类	机动车所有人	决定书类别
加入比对 详细	1109	2007-05-28	环东路(四号圆	324870000000	450331200710	张雄彪	桂HK0335	07	李远辉	1

20

1 / 1

显示从1到1，总 1 条。每页显示：20

配置说明

1. 配置二次查询列

字段配置				
保存 + 同步索引 + 新增索引 刷新 关闭				
操作	列名	列注释	数据类型	属性设置
删除	NAME	姓名	VARCHAR2	二次查询列
删除	GENDER	性别	VARCHAR2	
删除	DOB	出生日期	DATE	
删除	PID	身份证号	VARCHAR2	二次查询列 身份
删除	NATION	民族	VARCHAR2	
删除	NATAL PLACE	籍贯	VARCHAR2	

2. 配置可被二次查询检索的资源

a. 点击查询方案，添加方案，如 关联查询分类

系统管理 +

资源目录管理 +

数据资源管理 +

数据质量管理 +

信息化建设与应用通报管理 +

全文检索管理 -

全文检索管理

索引表管理

索引节点管理

查询方案管理

查询方案管理

显示查询 | + 增加 | x 删除选中 | 刷新

<input type="checkbox"/>	相关操作	方案名称	
<input type="checkbox"/>	详细 修改 删除 设置分类 查询	案件专题	
<input type="checkbox"/>	详细 修改 删除 设置分类 查询	电话号码查询	如1
<input type="checkbox"/>	详细 修改 删除 设置分类 查询	关联查询分类	
<input type="checkbox"/>	详细 修改 删除 设置分类 查询	人员专题	
<input type="checkbox"/>	详细 修改 删除 设置分类 查询	五查一联系	
<input type="checkbox"/>	详细 修改 删除 设置分类 查询	一键搜索	
<input type="checkbox"/>	详细 修改 删除 设置分类 查询	一键搜索BAK	

b. 点击详细，此处scheme_id需要修改为[big red]renyuan

详细	
关闭	
方案名称:	关联查询分类
备注:	
最后更新用户:	superadmin
最后更新时间:	2015-07-02 17:31:36.033
查询URL:	http://10.75.198.96:80/metadata/query/dataQuery_comQuery.do? scheme_id= renyuan

c. 依次点击设置分类，设置索引。同上面挂接资源一样即可。

查询方案-人员专题								
显示查询 + 增加 ✖ 删除选中 🔄 刷新								
<input type="checkbox"/>	相关操作	分类名称	是否启用	排序号	备注	最后更新用户	最后更新时间	
<input type="checkbox"/>	详细 修改 删除 设置条件 设置资源 设置索引	人员信息	启用	1		superadmin	2016-01-19	

6.2.5. 筛选查询

筛选查询指的是在已经查询后的结果集中进行筛选，支持多个条件同时筛选。

比对

查看比对

姓名

相似

张三桃

+

✖

查询

重置

操作	相片	更新时间	婚姻状况	主键ID	籍贯	民族	入库时间	证件号码	学历	性别	姓名
加入比对 详细 户成员			未婚	0f37394e-97ee-4dca-9e86-	湖北省黄梅县	汉族		42213019	文盲或半文盲	女	张三桃

5

⏮

⏪

1

/ 1

⏩

⏭

🔄

显示从1到1，总 1 条。每页显示：5

6.2.6. 搜索小提示

搜索按钮右侧可以点击提示按钮获取搜索语法。

1. 点击搜索提示按钮

一键搜索

主关键字

张三

搜索

2. 查看搜索提示

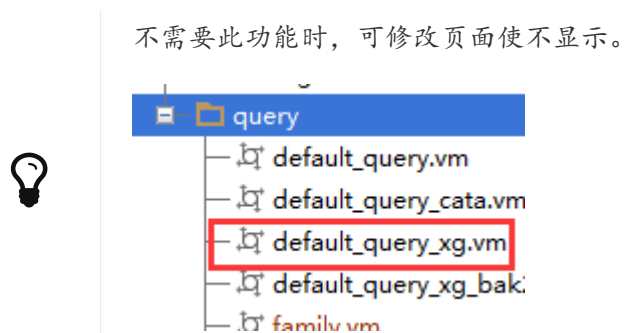


6.2.7. 其它定制化项

在孝感部门共享中，还添加了其它定制化开发的一些功能。

户成员

户成员即展示该用户的户籍变动信息。点击后会跳到对应页面根据身份证号进行查询。

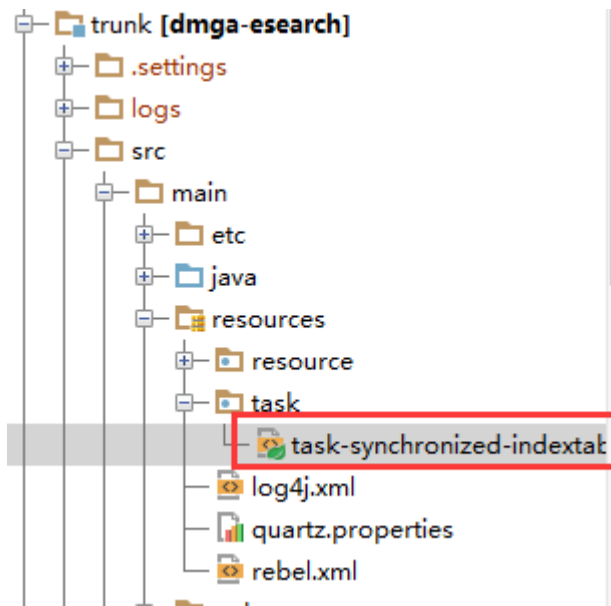


比对

比对，指将同一资源下的多条数据放在一起进行比对显示。 不需要此功能时，可修改页面使不显示。

Chapter 7. 定时同步索引

程序实现了定时执行索引同步任务。需要在如下配置文件中配置。



下图则是已经配置好了同步。



```
<property name="cronExpression" value = "0 28 10 * * ?"></property> ①  
  
<!-- 引用trigger -->  
    <bean class="org.springframework.scheduling.quartz.SchedulerFactoryBean"> ②  
        <property name="triggers">  
            <list>  
                <ref local="synchronizedTrigger" />  
            </list>  
        </property>  
    </bean>
```

① 此处配置定时调度频率，为cron表达式。请自行百度如何编写cron表达式。

② 当不需要此定时任务执行时，请将此bean给注释掉即可。